	<b>UNIVERSIDAD FRANCISCO DE PAULA SANTANDER OCAÑA</b>			
	Documento	Código	Fecha	Revisión
	<b>FORMATO HOJA DE RESUMEN PARA TRABAJO DE GRADO</b>	<b>F-AC-DBL-007</b>	<b>10-04-2012</b>	<b>A</b>
Dependencia	Aprobado		Pág.	
<b>DIVISIÓN DE BIBLIOTECA</b>	<b>SUBDIRECTOR ACADEMICO</b>		<b>1(273)</b>	

## RESUMEN – TRABAJO DE GRADO

AUTORES	<b>BRAYAN FERNANDO SANCHEZ JULIO JAIRO ALONSO IBARRA GONZALES</b>
FACULTAD	<b>DE INGENIERIAS</b>
PLAN DE ESTUDIOS	<b>INGENIERIA DE SISTEMAS</b>
DIRECTOR	<b>ALVEIRO ALONSO ROSADO GOMEZ</b>
TÍTULO DE LA TESIS	<b>ANALISIS DE LOS ESTUDIOS SOBRE INTELIGENCIA ARTIFICIAL PRESENTES EN REPOSITARIOS DE TESIS DE GRADO EN UNIVERSIDADES COLOMBIANAS DEL 2015 AL 2020</b>

### RESUMEN

(70 palabras aproximadamente)

EN LA REALIZACION DE ESTE TRABAJO SE PRETENDE BUSCAR, ESTUDIAR Y ANALIZAR, HACIENDO AGRUPACIONES POR CATEGORIAS DE LAS DIFERENTES VARIABLES DEFINIDAS DURANTE EL ESTUDIO A LOS TRABAJOS DE GRADO Y TESIS RELACIONADOS CON LA INTELIGENCIA ARTIFICIAL SE ENCUENTRAN EN LOS RESPOSITARIOS INSTITUCIONALES DE LAS UNIVERSIDADES QUE TENGAN PROGRAMAS DE ESTUDIO RELACIONADOS CON LAS TECNOLOGIAS DE LA INFORMACION Y LA COMUNICACIÓN, ESTAS INSTITUCIONES SERAN SELECCIONADAS DE LOS RESULTADOS MOSTRADOS POR LAS BASE DE DATOS DEL SISTEMA NACIONAL DE INFORMACION DE LA EDUCACION SUPERIOR.

### CARACTERÍSTICAS

PÁGINAS: 274	PLANOS:	ILUSTRACIONES:	CD-ROM: 1
--------------	---------	----------------	-----------



Vía Acolsure, Sede el Algodonal, Ocaña, Colombia - Código postal: 546552  
 Línea gratuita nacional: 01 8000 121 022 - PBX: (+57) (7) 569 00 88 - Fax: Ext. 104  
[info@ufpso.edu.co](mailto:info@ufpso.edu.co) - [www.ufpso.edu.co](http://www.ufpso.edu.co)

**ANALISIS DE LOS ESTUDIOS SOBRE INTELIGENCIA ARTIFICIAL PRESENTES  
EN REPOSITARIOS DE TESIS DE GRADO EN UNIVERSIDADES COLOMBIANAS  
DEL 2015 AL 2020**

**AUTORES**

**BRAYAN FERNANDO SANCHEZ JULIO**

**JAIRO ALONSO IBARRA GONZALES**

**Proyecto de grado presentado como requisito para optar el título de Ingeniero de  
Sistemas**

**Director**

**ALVEIRO ALONSO ROSADO GOMEZ**

**Magister en Gestión aplicación y desarrollo de software**

**UNIVERSIDAD FRANCISCO DE PAULA SANTANDER OCAÑA  
FACULTAD DE INGENIERIAS  
PLAN DE ESTUDIOS DE INGENIERIA DE SISTEMAS**

**Ocaña, Colombia**

**Septiembre, 2020**



## **Agradecimientos**

*Yo, Brayan Fernando Sanchez Julio, quiero agradecer a mi madre Marilse Julio Mora quien siempre ha querido que sea feliz dándome su apoyo y motivación, quien gracias a su esfuerzo y lucha hoy puedo tener un proyecto de vida, también quiero agradecer a mi abuela Dignora Quintero de Sanchez quien me acogió desde muy temprana edad y me ha apoyado como una segunda madre en todo mi proceso de crecimiento, las quiero mucho a ambas. También quiero agradecer a mi hermana, tios, y primos quienes me brindaron su apoyo y guía durante diferentes etapas de mi vida.*

*Yo, Jairo Alonso Ibarra Gonzales, quiero principalmente agradezco a Dios quien me permitió culminar mis estudios, a mi abuela María del Carmen Quintero por su esfuerzo y apoyo durante todo el proceso de mi carrera; por su compromiso, confianza y amor incondicional, también a mi madre Mariela Ojeda por su acompañamiento, motivación y entrega, a mis hermanos, a mis tíos y demás familiares quienes fueron muy importantes para alcanzar mi meta.*

*Por último, queremos dar nuestros agradecimientos a la Universidad Francisco de Paula Santander Seccional Ocaña por brindarnos la oportunidad de formarnos como profesionales. A los diferentes profesores que durante toda nuestra trayectoria educativa nos guiaron, nos motivaron e incentivaron a aprender.*

*Gracias a todos por permitirnos alcanzar nuestra meta.*

## Tabla de Contenido

Capítulo 1. Análisis de los Estudios Sobre Inteligencia Artificial Presentes en Repositorios de Tesis de Grado en Universidades Colombianas del 2015 al 2020.....	1
1.1. Planteamiento del Problema.....	1
1.2. Formulación del Problema .....	5
1.3. Objetivos .....	5
1.3.1. Objetivo general. ....	5
1.3.2. Objetivos específicos. ....	6
1.4. Justificación.....	6
1.5. Delimitaciones.....	8
1.5.1. Delimitación operativa. ....	8
1.5.2. Delimitación conceptual.....	9
1.5.3. Delimitación geográfica. ....	9
1.5.4. Delimitación temporal.....	10
Capítulo 2. Marco Referencial .....	10
2.1. Marco Histórico .....	10
2.1.1. Antecedentes históricos internacionales. ....	10
2.2. Marco Contextual.....	15
2.2.1. Creación del mercado de inteligencia artificial:.....	16
2.2.2. Priorización de las innovaciones creadoras de mercado. ....	17
2.2.3. Políticas basadas en evidencia y métricas de impacto para la regulación.....	17
2.2.4. Experimentación regulatoria. ....	17
2.2.5. Infraestructura de datos de fácil acceso. ....	18
2.2.6. Mercado de IA como generador de equidad e inclusión.....	18
2.2.7. Marco ético para la IA y seguridad. ....	19
2.2.8. Compromisos creíbles y producto de consensos.....	19
2.2.9. Ambiente de experimentación para desarrollar políticas de talento. ....	19
2.2.10. El rol estratégico de las universidades y la investigación académica en la creación del mercado de inteligencia artificial. ....	20
2.2.11. Atracción de talento internacional. ....	20
2.2.12. Políticas sobre futuro del trabajo basadas en evidencia. ....	21

2.2.13. El estado como facilitador y usuario de la IA. ....	21
2.2.14. Acceso continuo a conocimiento de la comunidad internacional. ....	21
2.3. Marco Conceptual .....	22
2.3.1. Inteligencia artificial. ....	22
2.3.2. Machine learning.....	23
2.3.3. Minería de datos.....	23
2.3.4. Sistemas expertos.....	24
2.3.5. Redes neuronales.....	24
2.3.6. Algoritmos geneticos. ....	25
2.3.7. Big Data. ....	25
2.3.8. Internet de las cosas (IoT).....	25
2.3.9. Chatbots. ....	26
2.3.10. Robótica. ....	26
2.3.11. Deep Learning.....	27
2.4. Marco Teórico.....	27
2.4.1. Revisión sistemática y metaanálisis. ....	28
2.4.2. Fases para escribir una revisión sistemática de literatura. ....	28
2.5. Marco Legal .....	33
2.5.1. Ley No. 1341 del 30 de julio de 2009.....	33
2.5.2. Ley estatutaria 1581 de 2012. ....	34
2.5.3. Ley 1273 de 2009.....	36
Capítulo 3. Diseño Metodológico .....	36
3.1. Tipo de Investigación.....	36
3.2. Población y Muestra.....	37
3.2.1. Población.....	37
3.2.1. Muestra.....	37
Capítulo 4. Resultados .....	42
4.1 Análisis general de la información.....	50
4.1.1. Ramas de la inteligencia artificial más utilizadas por las universidades .....	57
4.1.2. Ramas de la inteligencia artificial por proyecto. ....	58
4.1. Categorización por rama de la inteligencia artificial. ....	60

4.1.1. Minería de datos .....	60
4.1.2 Sistemas multi agentes. ....	101
4.1.3 Visión Artificial. ....	103
4.1.4 Inteligencia computacional. ....	110
4.1.5. Internet de las cosas IoT.....	118
4.1.6. Algoritmos genéticos. ....	136
4.1.7. Robótica. ....	141
4.1.8. Sistemas expertos. ....	144
4.1.9. Redes neuronales artificiales.....	151
4.1.9. Procesamiento de lenguaje Natural.....	175
4.1.10. Big Data. ....	180
4.1.11. Machine learning.....	200
Conclusiones. ....	225
Recomendaciones.....	
Referencias.....	230
Apéndices.....	260

## Tablas

Tabla 1 Ecuación de la muestra. ....	38
Tabla 2 Calculo de la muestra para las tesis y trabajos de grado. ....	38
Tabla 3 Formula del muestro aleatorio estratificado de forma aleatoria proporcional. ...	39
Tabla 4 Calculo de la proporción entre el tamaño de la muestra y la población de estudio .....	40
Tabla 5 Calculo del número de tesis y trabajos de grados a estudiar por universidad.....	40
Tabla 6. Universidades que no tienen repositorio. ....	44
Tabla 7. Universidades que no tienen tesis en sus repositorios. ....	46
Tabla 8. Universidades que solicitan acceso de usuario y contraseña. ....	48
Tabla 9. Universidades a las que no les funciona su repositorio. ....	49

## Figuras

Figura 1: Niveles y cantidades de programas educativos que se encuentran ofertados dentro del campo de las TIC. ....	3
Figura 2: Análisis de los repositorios de las 212 Instituciones de Educación Superior que ofrecen programas relacionados con las Tecnologías de la Información y Comunicación (TIC)..	5
Figura 3. Categorización de universidades por las razones que no hicieron parte de este estudio. ....	43
Figura 4. Numero de tesis relacionadas con la inteligencia artificial por año. ....	51
Figura 5. Numero de universidades por departamento. ....	52
Figura 6. Numero de tesis por departamentos. ....	53
Figura 7. Numero de tesis de trabajos de grados por universidad. ....	54
Figura 8. Número de estudiantes que trabajaron en las diferentes áreas de la inteligencia artificial e IoT. ....	55
Figura 9. Numero de tesis y trabajos de grados por universidad enfocados en el departamento de Bogotá. ....	56
Figura 10. Numero de las ramas de la Inteligencia Artificial e IOT que fueron aplicadas en las tesis de investigación de las universidades del departamento de Bogotá. ....	57
Figura 11. Ramas de la inteligencia artificial <i>más</i> aplicadas del 2015 al 2020. ....	58
Figura 12. Numero de tesis aplicadas por área de profesión. ....	59
Figura 13. Numero de tesis y tiempo de duración de su desarrollo. ....	60
Figura 14. Grafica que representa el número de veces que se usaron determinados algoritmos en las tesis de la rama de minería de datos. ....	99
Figura 15. Grafica que representa las veces que se usaron determinadas metodologías para el desarrollo de las tesis. ....	100
Figura 16. Grafica que representa el número de veces que se usaron los algoritmos dentro de los proyectos de investigación. ....	103
Figura 17. Grafica que presenta el número de veces que usaron determinadas metodologías de desarrollo dentro de la rama de Sistemas Multi Agentes. ....	103
Figura 18. Grafica que representa el número de veces que fue usado un algoritmo en los proyectos de investigación. ....	110
Figura 19. Grafica que representa el número de veces que fue usada una metodología en los proyectos de investigación relacionados con la Visión Artificial. ....	110
Figura 20. Grafica que representa el número de veces que se usaron determinados algoritmos en las tesis relacionadas con la Inteligencia Computacional. ....	117
Figura 21. Grafica que representa el número de veces que se usaron determinadas metodologías dentro de la rama de Inteligencia Computacional. ....	117
Figura 22. Grafica que representa el número de veces que se usaron determinadas metodologías en la ramas de algoritmos genéticos. ....	140

Figura 23. Grafica que representa el número de veces que se usaron determinados algoritmos en las tesis relacionadas con algoritmos genéticos. ....	141
Figura 24. Grafica que representa el número de veces que se usaron determinadas metodologías dentro de la rama de Sistemas Expertos. ....	150
Figura 25. Grafica que representa el número de veces que se usaron determinados algoritmos en las tesis relacionadas con Sistemas Expertos. ....	151
Figura 26. Grafica que representa el número de veces que se usaron determinadas metodologías dentro de la rama de las Redes Neuronales. ....	174
Figura 27. Grafica que representa el número de veces que se usaron determinados algoritmos en las tesis relacionadas con las Redes Neuronales. ....	174
Figura 28. Grafica que representa el número de veces que se usaron determinadas metodologías dentro de la rama de Procesamiento de Lenguaje Natural. ....	179
Figura 29. Grafica que representa el número de veces que se usaron determinadas metodologías dentro de la rama de la Big Data. ....	199
Figura 30. Grafica que representa el número de veces que se usaron determinadas metodologías dentro de la rama de la Big Data. ....	200
Figura 31. Grafica que representa el número de veces que se usaron determinadas metodologías dentro de la rama de Machine Learning. ....	224
Figura 32. Grafica que representa el número de veces que se usaron determinados algoritmos en las tesis relacionadas con del Machine Learning. ....	224

# **Capítulo 1. Análisis de los Estudios Sobre Inteligencia Artificial Presentes en Repositorios de Tesis de Grado en Universidades Colombianas del 2015 al 2020**

## **1.1. Planteamiento del Problema**

Colombia no se puede quedar atrás frente a los avances de las tecnologías digitales y, es por eso que actualmente enfrenta el gran reto de prepararse y comenzar a abrirle las puertas a la cuarta revolución industrial y a la transformación económica y social que esta pueda traer al país, pero, para que este cambio empiece a suceder se deben estudiar y solucionar varios problemas, entre los cuales se encontró uno muy importante, y es el de fortalecer el talento y la capacidad de los profesionales en tecnologías de la información, pero para lograrlo se debe brindar un espacio donde puedan innovar tanto en el sector público, como privado, con el fin de preparar al país para la transformación que conlleva la industria 4.0, por lo cual, es necesaria la participación activa de diferentes entidades estatales como por ejemplo, el departamento nacional de planeación y consejería presidencial para asuntos económicos y transformación digital (Consejo Nacional de Política Económica y Social, 2019).

Además de la escasez de personal calificado en inteligencia artificial (IA), existen otros problemas que hacen más difícil la adopción de esta tecnología en Colombia y demás países latinoamericanos, entre los cuales se puede mencionar la falta de conocimiento sobre este tipo de soluciones tecnológicas por parte de las empresas, lo que las lleva a ser más cautelosas y prefieran optar por otras soluciones no tan innovadoras, pero ya conocidas (Rueda, 1993). Este rechazo a trabajar sobre inteligencia artificial genera una ausencia de datos, los cuales son un



requisito fundamental al momento de realizar desarrollos sobre alguna de las diferentes técnicas que ofrece (ENDEAVOR, 2018), lo que hace más compleja su implementación, acompañado de la lenta asimilación de esta nueva tecnología por parte de la sociedad, lo cual, al ser la base fundamental de la cuarta revolución industrial, provoca un impacto económico que no permite una mejora en la calidad de vida productiva de las personas, como sucede en otros países que ya han asimilado esta tecnología (Albrieu, Rapetti, López Brest, Larroulet, & Sorrentino, 2018).

Haciendo una búsqueda de las universidades que ofrecen programas activos relacionados con las tecnologías de la información y comunicación que se encuentra en la página web del sistema nacional de información de la educación superior (SNIES<sup>1</sup>), como se muestra en la figura 1, donde se observan el número de niveles educativos (Doctorados, maestrías, especializaciones, ingenierías, técnica, tecnologías) ofertados por las universidades que tienen programas dentro del campo de las tecnologías de la información y comunicación (TIC).

---

<sup>1</sup> <https://hecaa.mineducacion.gov.co/consultaspublicas/programas>

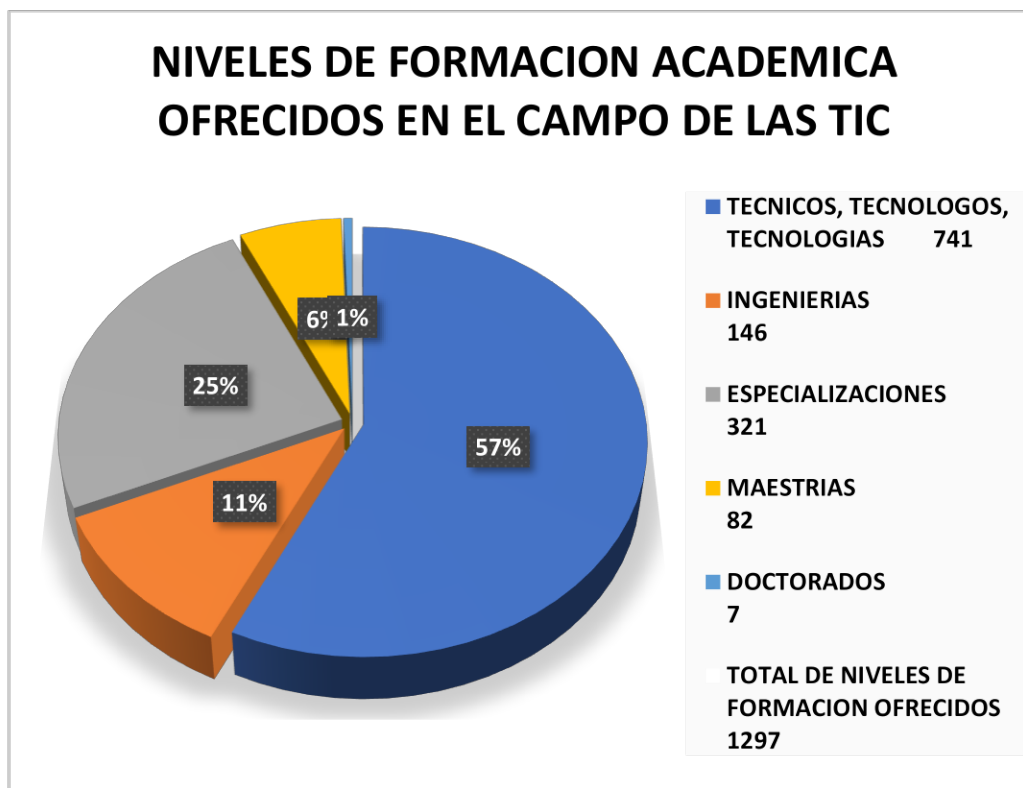


Figura 1: Niveles y cantidades de programas educativos que se encuentran ofertados dentro del campo de las TIC.

Como lo muestra la figura anterior existen en el país 1297, programas académicos activos que buscan la formación en diferentes niveles educativos en el campo de las TIC. Esto sugiere la siguiente reflexión; están estos programas fomentando el aprendizaje de la inteligencia artificial en sus estudiantes, una de las respuestas, podría ser que sí, al contar con cursos que la imparten; sin embargo, se generaría el cuestionamiento, si en realidad se está apropiando y se fomenta el desarrollo de competencias adecuadas para suplir las necesidades que sobre esta área requiere el país (Moreno, 2019). Desde este punto de vista el proyecto de grado o tesis de grado se puede convertir en una herramienta para no solo fomentar, si no que también apropiarse las competencias que demanda una nueva sociedad cada vez dependiente a la ayuda de la tecnología (Galindo, Ruiz, & Ruiz, 2017).

Para lograr tener un estimativo de que tanto están trabajando las universidades sobre inteligencia artificial, en sus proyectos de grado, se debe hacer una revisión de los repositorios institucionales, para estimar por ejemplo que temas dentro del área de la inteligencia artificial, son los más desarrollados y que cantidad de todos los proyectos de un programa son de inteligencia artificial, para esta manera poder hacer un análisis del aporte que están generando las universidades a esa necesidad que como lo señala la revista (Dinero, 2020), en su artículo publicado bajo el título “¿Cuáles son las empresas más innovadoras del país?”, viene en aumento la tendencia de contar con personal dentro de las empresas que tengan dominio de IA, eso también coincide con las declaraciones del departamento nacional de planeación (Consejo Nacional de Política Económica y Social, 2019). En la figura 2, corresponde al análisis de los repositorios donde se almacenan las tesis y trabajos de grado de los diferentes niveles educativos ofertados por las 212 universidades que desarrollan programas dentro del campo de las TIC según el SNIES.

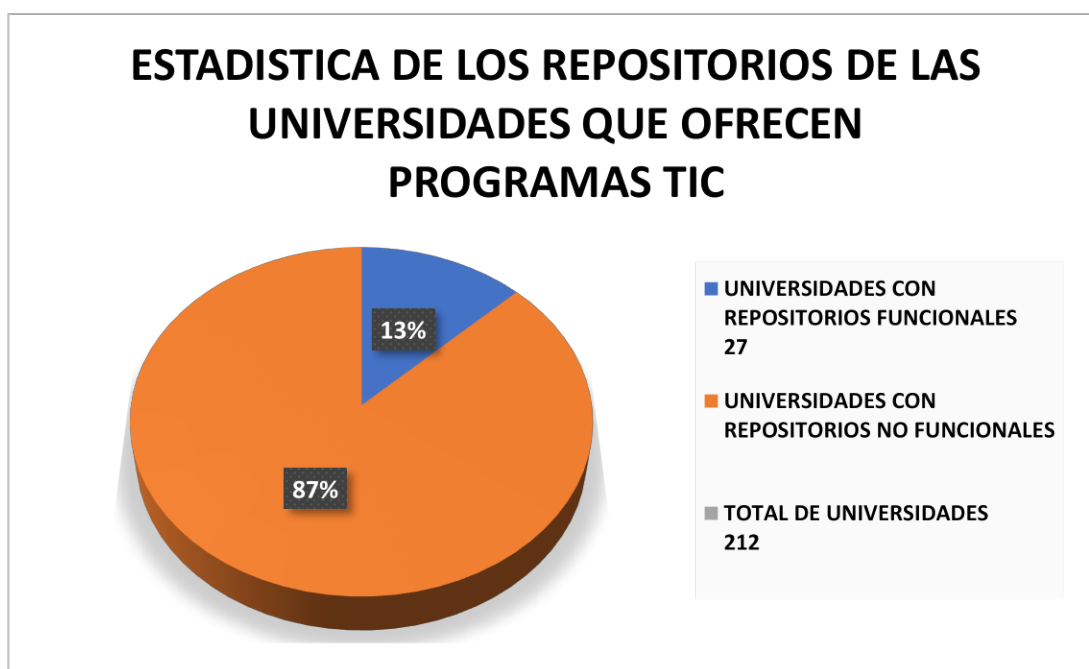


Figura 2: Análisis de los repositorios de las 212 Instituciones de Educación Superior que ofrecen programas relacionados con las Tecnologías de la Información y Comunicación (TIC).

Si bien se conoce la falta de profesionales en el área de inteligencia artificial de los países latinoamericanos (ENDEAVOR, 2018), se hace complicado que, en Colombia, las universidades no estén estimulando la apropiación de la IA como una de las ventajas diferenciadoras de sus egresados les puede servir para acceder más rápido a la vida laboral, por estas razones, se hace necesario adelantar estudios profundos que encuentren un comportamiento y permitan sugerir una evolución de IA en el país a nivel académico, específicamente en los proyectos de final de carrera en educación formal e independiente de su nivel de formación.

## **1. 2. Formulación del Problema**

¿Cuáles son las características que generalizan los proyectos de grado en educación superior sobre inteligencia artificial en Colombia comprendidos entre los años 2015 al 2020?

## **1.3. Objetivos**

### **1.3.1. Objetivo general.**

Analizar los estudios sobre inteligencia artificial presentes en repositorios de trabajos y tesis de grados de universidades colombianas del 2015 al 2020.

### **1.3.2. Objetivos específicos.**

Seleccionar los repositorios institucionales y los programas académicos que contengan proyectos de grado relacionados con la inteligencia artificial para delimitar el ámbito de búsqueda.

Definir las categorías en las cuales se organizará la revisión crítica de los documentos para orientar la sistematización de la información.

Analizar la información recolectada haciendo agrupaciones y recuentos de tendencias para hacer conclusiones de la evolución de la inteligencia artificial a nivel de las universidades de colombianas.

### **1.4. Justificación**

Los países latinoamericanos tienen la desventaja de no encontrarse tan evolucionados en la cuarta revolución industrial como Estados Unidos o China, a pesar de que este cambio ha empezado hace poco tiempo (Albrieu, Rapetti, López Brest, Larroulet, & Sorrentino, 2018), pero esto a la vez sirve como punto de referencia porque indica que tienen mucho terreno tecnológico que recorrer y mejorar. En el caso de Colombia, una vez se haya analizado su punto inicial, es decir, cuando se haya determinado la caracterización, se puede establecer un número estimado de profesionales actuales enfocados y las técnicas más empleadas en inteligencia artificial; obtenido estos datos, también se hace posible prever el número de profesionales que hacen falta para evitar un posible rezago en el avance hacia la industria 4.0, y así obtener un enfoque más preciso

sobre los puntos más débiles que se deben de estimular desde diferentes entidades nacionales como el Departamento Nacional de Planeación y la Consejería Presidencial para Asuntos Económicos y Transformación Digital que con su aporte, posibilitarían aumentar el número de desarrollos que se realizan en esta área y por consiguiente mejorar los conocimientos sobre las ventajas y, beneficios que esta nueva tecnología puede aportar en los diferentes ámbitos laborales (Consejo Nacional de Política Económica y Social, 2019).

Además, la economía colombiana se beneficiará al fortalecer los puntos frágiles en los sectores que carecen de desarrollo a nivel tecnológico, y que se evidencia que son el impedimento para que se genere una evolución constante y notable en todos los diferentes proyectos que se realizan y son involucrados en la cuarta revolución industrial (Rueda, 1993), por esto, un método efectivo para mitigar estos flagelos es aplicar la IA en todos los campos; ya que es una tecnología que puede ser aplicada desde la agricultura hasta brindar respaldo al sistema de salud por medio de herramientas de diagnóstico basadas en sistemas expertos, lo que indica que las startups al aplicar estas tecnologías se beneficiaran de los datos arrojados para dar inicio o fortalecer los servicios ofrecidos por sus emprendimientos a los diferentes sectores aplicando técnicas de soluciones y transformación social por medio de la innovación (Dinero, 2020). Uno de los puntos importantes de esta tecnología es crear un incremento positivo en el Producto Interno Bruto (PIB) por persona, por la ventaja principal de optimización de trabajo y aumento en la producción, lo que indica un crecimiento saludable de la economía y por consiguiente una mejor calidad de vida para las personas (Albrieu, Rapetti, López Brest, Larroulet, & Sorrentino, 2018).

Esta investigación pretende encontrar las características generales que tienen los proyectos de grado sobre IA en Colombia, para tratar de describir un comportamiento y tendencia sobre la evolución de este tema en los últimos cinco años y establecer un panorama de cuanto viene la academia aportando en la formación de profesionales del sector de las TIC con capacidades demostradas o validadas para trabajar en IA y aprovechar las oportunidades que trae la cuarta revolución industrial. El desarrollo de esta tesis proporcionara resultados novedosos y rigurosos desde el punto de vista de la información recolectada y del aprendizaje generado, para los siguientes interesados:

- Para la comunidad en general, brindara un consolidado de como se viene desarrollando la IA en las universidades de Colombia, proporcionando un documento de consulta de interés nacional.
- Para el programa de ingeniería de sistemas, las tendencias aquí descritas pueden ser utilizadas como elementos de reflexión sobre la pertinencia de los proyectos de grado que sobre IA se desarrollan dentro del programa, con relación al panorama nacional.
- Para los autores, proporciona una oportunidad de profundizar a nivel técnico sobre los procesos y elementos que componen una solución de IA. Adicionalmente, a nivel investigativo desarrollaran habilidades en la realización de revisiones sistemáticas de literatura específicamente en esta área.

## **1.5. Delimitaciones**

### **1.5.1. Delimitación operativa.**

Se pueden presentar diferentes tipos de inconvenientes al momento de indagar en el Sistema Nacional de Instituciones de Educación Superior; algunos repositorios no están organizados de manera que se pueda realizar una búsqueda válida para IA, algunas universidades no cuentan con un repositorio activo o la página web se encuentra mal desarrollada y no permite realizar búsquedas, otras no permiten su acceso a la plataforma ya que se debe ingresar por medio de un usuario y contraseña, en otras ocasiones, cuando se localizaban los trabajos de grado o tesis de determinado programa de algunas universidades, estas no permitían un filtro para clasificar esos trabajos según su enfoque o tema de investigación por lo cual se debía revisar todas las tesis para determinar cuáles podrían aplicar a IA.

### **1.5.2. Delimitación conceptual.**

Los términos que enmarcan este proyecto son los siguientes: Industria 4.0, Inteligencia artificial, Machine Learning, Minería de datos, Big Data, algoritmos genéticos, sistemas expertos, redes neuronales, Internet de las cosas (IoT).

### **1.5.3. Delimitación geográfica.**

Este trabajo se desarrollará en la Universidad Francisco de Paula Santander Ocaña, en el programa de Ingeniería de Sistemas, basados en los trabajos de grado y tesis sobre inteligencia artificial que se encontraron en diferentes universidades que muestra el Sistema Nacional de Información de la Educación Superior al aplicárseles un filtro por Programas que se encuentran



Activos a nivel nacional y que están catalogados en el campo de las Tecnologías de la Información y la Comunicación (TIC).

#### **1.5.4. Delimitación temporal.**

Para el desarrollo de la investigación se estimó un tiempo determinado de 4 meses a partir de la aprobación del proyecto.

## **Capítulo 2. Marco Referencial**

### **2.1. Marco Histórico**

Aplicar una revisión sistemática de literatura permite que en el desarrollo de las investigaciones en las cuales la fuente de información proviene de datos abiertos, y que estos son de acceso permitido a los desarrolladores de proyectos de investigación, les permita diferenciar, clasificar o agrupar cantidades de información, de manera que se evalúen y se logren interpretar el valor de las variables asignadas al comportamiento de dicha información, y se obtenga un valor mayor de los datos, por lo tanto los tanto se hace referencia a los siguientes estudios en los cuales se aplicaron una revisión sistemática, que por medio de técnicas de la inteligencia artificial, analizan y generan una evaluación detallado de dichos datos.

#### **2.1.1. Antecedentes históricos internacionales.**

##### **2.1.1.1. Revisión sistemática aplicada al diagnóstico de fallas en el sector industrial.**

El diagnóstico de fallas de maquinaria rotativa juega un papel importante para la confiabilidad y seguridad de sistemas industriales modernos. Como campo emergente en aplicaciones industriales y efectiva solución para el reconocimiento de fallas, las técnicas de inteligencia artificial (IA) han estado recibiendo creciente atención de la academia y la industria. Sin embargo, los grandes desafíos se enfrentan con los métodos de IA bajo las diferentes condiciones reales de funcionamiento (Liu, Zio, Chen, & Yang, 2018).

#### **2.1.1.2. Revisión sistemática de visión por computadora e inteligencia artificial en agricultura de precisión para cultivos de granos.**

Revisión sistemática que tiene como objetivo identificar la aplicabilidad de la visión por computadora en la agricultura de precisión para la producción de cinco granos más producidos en el mundo: maíz, arroz, trigo, soja y cebada. En este sentido, presentamos 25 artículos seleccionados en los últimos cinco años con diferentes enfoques para tratar aspectos relacionados con la detección de enfermedades, grano calidad y fenotipado. A partir de los resultados de la revisión sistemática, es posible identificar grandes oportunidades, como la explotación de GPU (Unidad de procesamiento de gráficos) y técnicas avanzadas de inteligencia artificial, como DBN (Deep Belief Networks) en la construcción de métodos robustos de visión por computadora aplicados a la precisión en la agricultura (Rieder & Patrício, 2018).

#### **2.1.1.3 Revisión sistemática de la dermatoscopia y la dermatoscopia digital en inteligencia artificial para el diagnóstico de melanoma.**

Se realizó una búsqueda bibliográfica sobre la dermatoscopia y la dermatoscopia digital aplicando inteligencia artificial para el diagnóstico de melanoma utilizando varias bases de datos. Títulos y Los resúmenes de los artículos recuperados se examinaron mediante un formulario de evaluación de la literatura.

Se desarrolló un formulario de evaluación de calidad para evaluar la calidad de los estudios. En la actualidad no hay evidencia clara sobre la precisión diagnóstica de la dermatoscopia en comparación con inteligencia artificial. El objetivo es Evaluar la precisión diagnóstica de la dermatoscopia y la dermatoscopia digital utilizando inteligencia artificial para el diagnóstico de melanoma y comparar la precisión diagnóstica de los diferentes algoritmos dermatoscópicos entre sí y con dermatoscopia digital / inteligencia artificial para la detección de melanoma (SM, AP, J, & AD., 2009).

#### **2.1.1.4. Revisión sistemática de mapeo Inteligencia Artificial para Mama IRM en 2008–2018.**

Búsqueda sistemática de la literatura realizada para identificar artículos sobre el uso de IA en la RM de mama. Para cada artículo identificado, el apellido del primer autor, año de publicación, revista de publicación, Web of Science Core Categoría de la revista de colección, país de afiliación del primer autor, diseño del estudio, conjunto de datos, objetivo (s) del estudio, métodos de IA utilizados y, cuando estén disponibles, se registraron los resultados del diagnóstico, Las técnicas más utilizadas se analizaron de acuerdo con objetivos clínicos específicos. Los algoritmos de aprendizaje supervisados se utilizaron principalmente para la

caracterización de lesiones, con el valor AUC del análisis ROC que varía de 0,74 a 0,98 (mediana, 0,87) y con el de imágenes de pronóstico que varía de 0,62 a 0,88 (mediana, 0,80), mientras que el aprendizaje no supervisado se utilizó principalmente para fines de procesamiento de imágenes (Codari, Schiaffino, Sardanelli, & Trimboli, 2019).

#### ***2.1.1.5. Una revisión sistemática de la inteligencia artificial y los robots en la co-creación de valor estado actual y vías de investigación futuras.***

Revisión sistemática de la literatura sobre el tema para avanzar teóricamente Análisis de IA y robots en co-creación de valor. Al revisar sistemáticamente 61 artículos sobre AI y de robótica, que han sido publicados en los mejores servicios y marketing revistas de investigación, identificamos cuatro temas en literatura, a saber, campo genérico avance, proveedores de servicios de apoyo, permitiendo la integración de recursos entre proveedores de servicios y beneficiarios, y apoyando el bienestar de los beneficiarios. Con la identificación del primer conjunto de literatura sobre IA y robots en co-creación de valor, impulsamos un subcampo importante de literatura de co-creación de valor. Adicionalmente, Para avanzar en el campo, sugerimos construir sobre la teoría y la ciencia actor-red y Estudios tecnológicos para comprender la agencia tecnológica en la co-creación de valor.

Teniendo en cuenta que la tecnología tiene agencia, abre nuevas vías de investigación interesantes en torno a los bots de compras y la interacción de primera línea entre humanos que probablemente para influir en la integración de recursos, el compromiso del cliente y la co-creación de valor en el futuro (Kaartemo & Helkkula, 2018)

#### **2.1.1.6. Inteligencia artificial aplicada en sistemas expertos.**

En los sistemas expertos, como técnica aplicada perteneciente al ámbito de la Inteligencia Artificial, presentan características que los identifican como eficaces procesadores de conocimiento. En este sentido, disponen de una base de conocimiento que incorpora representaciones de hechos, heurísticos y esquemas de razonamiento, utilizando estructuras formales de representación próximas a cómo las entiende el usuario conocedor del problema a resolver. El proceso de inferencia, a partir del conocimiento contenido en la base, incluye, además, procesos de aprendizaje, facilidades de comunicación con el usuario, y facilidades de modificación y explicación del conocimiento utilizado (Martinez, 2011).

#### **2.1.1.7 Inteligencia artificial y computación avanzada.**

Inteligencia Artificial (I.A.) se puede definir como aquella inteligencia o artefactos científicos contruidos por humanos, o sea que dice que un sistema artificial posee inteligencia cuando es capaz de llevar a cabo tareas que, si fuesen realizadas por un humano, se diría de este que es inteligente. Dentro de las ciencias de la computación, la rama de la I.A. se basa en intentar dotar al funcionamiento de las aplicaciones informáticas de un comportamiento inteligente similar al humano para la toma de decisiones. La I.A. dentro de la computación no es algo reciente ya que sus albores se pueden referir al comienzo de la Cibernética, hace más de 60 años, pero en los últimos años su avance se ha hecho más notable y se ha extendido a numerosas áreas. Así, ha sido llevada a intentar simular comportamientos humanos, no sólo razonamiento, sino también aspectos más difíciles de medir como es el caso de la creatividad o arte artificial. Una

prueba de ello es la imagen de fondo que ilustra la portada de este libro, generada por un algoritmo genético (Romero, Dafonte, Gomez, & Penousal, 2007).

## **2.2. Marco Contextual**

El contexto se haya conformado por un conjunto de situaciones, fenómenos y circunstancias que se combinan en un momento y lugar específico de la historia y que tienen evidentes consecuencias sobre los sucesos que toman lugar dentro de sus límites espaciotemporales, además de permitirnos la comprensión eficaz del mensaje. Si bien su definición puede tornarse complicada por tratarse de un concepto extremadamente abstracto, lo central de esta palabra es que supone la especificidad de los fenómenos ya que los mismos se han combinado de modo único e irrepetible para tener influencia en lo que en él pasa (Bembibre , 2009).

Las revoluciones industriales no sucedieron de un día para otro, les llevo tanto tiempo, al punto que pasaron desapercibidas por la sociedad y los países que las lideraron tuvieron grandes cambios positivos a nivel de economía e industria convirtiéndose en grandes potencias, mientras que los países que no se acoplaron a estos nuevos cambios se enfrentaron a un rezago en la evolución de la infraestructura industrial y un crecimiento más lento de la economía. Este ciclo se repite nuevamente con el surgimiento de la cuarta revolución industrial, la cual implica que en los países se mejoren factores como: la demanda interna de esta nueva tecnología, creación de

políticas y estrategias en pro de la IA, mejoramiento de la infraestructura tecnológica e innovación sobre la misma entre otros.

Latinoamérica presentan bastante rezago frente a la cuarta revolución industrial encontrándose fuera de los top 30 de los países que van a la vanguardia de esta tecnología, según el estudio A Government Artificial Intelligence Readiness Index en la cual se le asigna un puntaje a cada país teniendo en cuenta diferentes factores como los mencionados, México lidera a los países latinoamericanos ocupando el puesto 32 con un puntaje de 6.673, y por debajo de países como Uruguay, Chile y Brasil se encuentra Colombia en el puesto 44 con un puntaje de 5.945, todos estos países fuera del top 30, lo cual indica un tardanza en la adaptación y asimilación de las nuevas tecnologías imprescindibles para lograr un avance notorio en la industria 4.0.

Es por esto que el Consejo Nacional de Política Económica define 14 principios estratégicos enfocados en el desarrollo y asimilación de la inteligencia artificial en el país posibles y con los cuales también se busca reducir el posible impacto negativo en la profundización de diferencias sociales e inequidad que esta tecnología pueda generar.

### **2.2.1. Creación del mercado de inteligencia artificial:**

Colombia se convertirá en un laboratorio para la creación de un mercado propio de IA en el que interactúen diseñadores, proveedores, intermediarios y consumidores de esta tecnología.

Este mercado se convertirá a su vez en el centro dinamizador del naciente ecosistema de 4 revolución industrial en el país (Consejo Nacional de Política Económica y Social, 2019).

### **2.2.2. Priorización de las innovaciones creadoras de mercado.**

Para la creación de un mercado de IA en Colombia, el Gobierno nacional generará condiciones propicias para el desarrollo de innovaciones creadoras de mercado con el potencial de jalonar el desarrollo de talento y los cambios regulatorios requeridos (Consejo Nacional de Política Económica y Social, 2019).

### **2.2.3. Políticas basadas en evidencia y métricas de impacto para la regulación.**

Toda regulación que impacte el desarrollo de IA en Colombia debe ir acompañada de evidencia que soporte la necesidad de intervención del Estado, con el fin de evitar convertirse en una limitante injustificada para el desarrollo de un mercado competitivo de IA. Así mismo, con este enfoque se estimulará el análisis costo-beneficio en la creación de la nueva regulación y en la imposición de medidas administrativas aplicables a la IA (Consejo Nacional de Política Económica y Social, 2019).

### **2.2.4. Experimentación regulatoria.**

Colombia debe adoptar modelos de regulación novedosos que permitan a las entidades regulatorias y de supervisión conocer las últimas tecnologías y a los emprendedores



experimentar nuevos modelos. Así mismo, es esencial considerar un enfoque de gobernanza de la IA basado en riesgos que debe ser promovido dentro de las entidades del Gobierno y el sector privado (Consejo Nacional de Política Económica y Social, 2019).

#### **2.2.5. Infraestructura de datos de fácil acceso.**

Colombia debe seguir desarrollando una infraestructura de datos completa y que permita el diseño e implementación de sistemas de IA en el país, priorizando la creación e identificación de bases de datos masivos que sean interoperables y contengan información estructurada, así como disminuyendo barreras innecesarias e injustificadas al acceso a datos, para los desarrolladores de esta tecnología. También se deben generar modelos para que entidades de diversos sectores localizadas en Colombia o en el exterior puedan compartir e intercambiar información de forma rápida y sencilla. Esto significa que la regulación de protección de datos debe garantizar los derechos de los ciudadanos bajo un marco de gestión del riesgo (Consejo Nacional de Política Económica y Social, 2019).

#### **2.2.6. Mercado de IA como generador de equidad e inclusión.**

Colombia debe desarrollar condiciones específicas que permitan que la IA pueda ser dirigida a consumidores no tradicionales y a mejorar la calidad de vida de los colombianos, en especial a aquella población pobre y vulnerable, así como para cerrar brechas entre zonas urbanas y rurales, entre otros ejemplos (Consejo Nacional de Política Económica y Social, 2019).

### **2.2.7. Marco ético para la IA y seguridad.**

El gobierno nacional reconoce que el uso de IA conlleva una serie de desafíos éticos que deben ser considerados y atendidos por el Estado, tales como, justicia, libertad, no discriminación, transparencia, diseño responsable, seguridad, privacidad y el rol de los derechos humanos, entre muchos otros. Estos principios deben ser discutidos y construidos con el apoyo del sector privado y la comunidad científica y académica del país experta en la materia (Consejo Nacional de Política Económica y Social, 2019).

### **2.2.8. Compromisos creíbles y producto de consensos.**

Colombia debe implementar una política creíble y que sea reconocida a nivel internacional que genere seguridad jurídica para la inversión. Esto también requiere la creación de consensos entre diversos sectores que permita un efecto duradero de esta política en el país y con efecto transversal. Así mismo, deben definirse mecanismos y respuestas institucionales que faciliten la coordinación entre distintas entidades del Estado (Consejo Nacional de Política Económica y Social, 2019).

### **2.2.9. Ambiente de experimentación para desarrollar políticas de talento.**

El Gobierno nacional debe fomentar las habilidades y las metodologías más efectivas para que los colombianos puedan adoptar la IA, reconociendo los efectos que esta tecnología

también puede tener entre la población adulta, facilitando el aprendizaje a cualquier edad y la posibilidad de implementar programas de entrenamiento intensivo y rápido (Consejo Nacional de Política Económica y Social, 2019).

#### **2.2.10. El rol estratégico de las universidades y la investigación académica en la creación del mercado de inteligencia artificial.**

Colombia debe fomentar proyectos académicos de inteligencia artificial que se desarrollen en los centros de educación superior del país, para que puedan servir a la construcción del mercado de IA. Para esto, es vital que investigadores y estudiantes tengan los elementos y el acompañamiento suficiente para emprender, generando puntos de interacción con la empresa privada (Consejo Nacional de Política Económica y Social, 2019).

#### **2.2.11. Atracción de talento internacional.**

El gobierno nacional debe estimular la creación de un mercado en IA, por medio de la creación de programas que permitan que expertos en la materia de todo el mundo vean en Colombia un mercado atractivo en el cual trabajar. Con esto se busca que Colombia pueda convertirse en un referente mundial en temas de inteligencia artificial, generando alianzas internacionales que permitan este reconocimiento (Consejo Nacional de Política Económica y Social, 2019).

### **2.2.12. Políticas sobre futuro del trabajo basadas en evidencia.**

El gobierno nacional debe realizar un seguimiento al mercado laboral para obtener evidencia oportuna sobre los efectos propios de esta tecnología en el mercado laboral. Esta evidencia permitirá diseñar medidas para enfrentar los posibles efectos negativos de esta tecnología en el empleo e implementar medidas que den una protección real al trabajador (Consejo Nacional de Política Económica y Social, 2019).

### **2.2.13. El estado como facilitador y usuario de la IA.**

El Estado debe brindar todas las condiciones para que distintos sectores desarrollen y participen de forma competitiva en el mercado de IA. El Estado debe convertirse en uno de los grandes usuarios de esta tecnología para dar solución a los distintos retos que enfrenta el sector público. De esta forma, el mercado de IA debe servir como una fuente de soluciones para el sector público impulsado por emprendedores (Consejo Nacional de Política Económica y Social, 2019).

### **2.2.14. Acceso continuo a conocimiento de la comunidad internacional.**

Colombia reconoce el liderazgo internacional de ciertos países e instituciones a en el desarrollo de la IA. Por lo tanto, es necesario tener un intercambio con entidades internacionales líderes en el tema y generar estrategias para tener un intercambio constante, mitigando las asimetrías de información a través de una interacción permanente con la comunidad

internacional, especialmente de la región (Consejo Nacional de Política Económica y Social, 2019).

### **2.3. Marco Conceptual**

Es importante tener en cuenta, al momento de realizar una revisión sistemática de literatura que en la cual se busca analizar por medio de técnicas de análisis de documentación web, y que el campo en el cual se va desarrollar la investigación es la inteligencia artificial , manejar el concepto de las principales ramas que abarcan esta ciencia, y de esta manera conocer a fondo como está integrada, sus áreas, las técnicas y algoritmos que utilizan para la recolección de información de datos abiertos y así lograr un buen estudio y llevar a cabo la debida caracterización de los datos.

#### **2.3.1. Inteligencia artificial.**

Es una disciplina del área de la tecnología, uno de los principales objetivos de la inteligencia artificial es lograr emular los pensamientos y actitudes racionalistas de los seres humanos, para lograr el funcionamiento de cada una de las ramas que la componen se aplican métodos probabilísticos, que por medio de algoritmos precisos, y técnicas de agrupamientos de grandes cantidades de datos, permiten que un sistema se considere inteligente, y logre reaccionar con similitud al comportamiento de los seres humanos, por medio de reconocimiento de patrones (Aravena Garrido, y otros, 2018).

### **2.3.2. Machine learning.**

Machine Learning es una rama de la inteligencia artificial que consiste en crear sistemas capaces de aprender por sí mismos. Las decisiones se toman en base a datos recogidos del entorno pretende estudiar el reconocimiento de patrones y el aprendizaje de máquinas. Según Arthur Samuel, pionero en el campo de la inteligencia artificial y creador de uno de los primeros programas de “autoaprendizaje” del mundo, este campo de la ciencia le da a las máquinas la capacidad de aprender sin ser programadas explícitamente. Dicho de otra manera, son algoritmos que proporcionan conclusiones relevantes a partir de un conjunto de datos obtenido, sin necesidad de escribir instrucciones específicas para esto. Un algoritmo de Machine Learning está compuesto por unos datos recibidos por la máquina, una función de coste que se pretende minimizar, una función de optimización a maximizar y un modelo (Shai Shalev & Shai Ben, 2014).

### **2.3.3. Minería de datos.**

Se define la minería de datos como una rama de la inteligencia artificial, la cual consiste en aplicar un conjunto de técnicas que tienen como objetivo descubrir información contenida en un gran conjunto de datos, la mayoría de las búsquedas se realizan en bases de datos que contienen un volumen de datos muy grandes, y que son de tipo abierto, para lograr el análisis de

patrones y realizar una evaluación del comportamiento de estos mismos, caracterizados , y agrupados en grandes grupos y sub grupos (Grandez Marquez, 2017).

#### **2.3.4. Sistemas expertos.**

Los sistemas expertos son programas que capturan el conocimiento de un experto e imitan sus procesos de razonamiento cuando resuelven los problemas en un determinado dominio. Los sistemas expertos son un subconjunto especial dentro de los sistemas basados en el conocimiento, que incorporan en la base de conocimiento (KDD) del sistema el conocimiento de un experto (Becerra , 2011).

#### **2.3.5. Redes neuronales.**

Las redes neuronales artificiales son modelos que intentan reproducir el comportamiento del cerebro. Como tal modelo, realiza una simplificación, averiguando cuales son los elementos relevantes del sistema, bien sea por que la cantidad de información de la que se dispone es excesiva o bien porque es redundante, Una elección adecuada de sus características más una estructura conveniente, es el procedimiento convencional utilizado para construir redes capaces de realizar determinada tarea (Zapata Garrido & Diaz Mojica, 2007).

### **2.3.6. Algoritmos genéticos.**

Los algoritmos genéticos evolutivos son técnicas de optimización que trabajan sobre poblaciones de soluciones y que están diseñadas para buscar valores óptimos en espacios complejos. Están basados en procesos biológicos que se pueden apreciar en la naturaleza, como la selección natural o la herencia genética. Parte de la evolución está determinada por la selección natural de individuos diferentes compitiendo por recursos en su entorno (Dorronsoro Díaz , 2006).

### **2.3.7. Big Data.**

Es definido como un enorme contenido de información, que puede ser generalizada y agrupada en datos más representativos, los cuales por su gran tamaño tienden a no poder ser analizados por el mismo software que los contiene, para esto se crean algoritmos de análisis de información que permiten analizar el grande volumen de datos, en el ámbito del negocio el big data puede ser apreciado como una oportunidad de desarrollo, y un valor agregado mayor a los ingresos, y al generar un mayor grado de competencia (Romero Castillo, 2018).

### **2.3.8. Internet de las cosas (IoT).**

Es una tecnología que consiste en la conexión de los usuarios de diferentes con las cosas del vivir cotidiano, También denominado internet de los objetos, la expansión de esta área de la



tecnología va mucho más allá que realizar la comunicación entre dos personas, o en la relación que hay entre los usuarios y los dispositivos digitales, es uno de los agentes más importantes y influyentes en la revolución industrial 4.0 (Barrios Moises, 2018).

### **2.3.9. Chatbots.**

El termino de chatbots es definido como un programa informático, que hace referencia a un rama de la inteligencia artificial, la cual se basa en el aprendizaje automático para lograr sus funciones, por medio del procesamiento de lenguaje natural que permite realizar una simulación que conecta a dos usuarios, uno de origen con otro de destino, de tal manera que se logre una conversación, con respuestas razonables, que logre llegar a simular una conversación entre dos personas (Garcia Reina, 2018).

### **2.3.10. Robótica.**

La robótica es una de las ramas de la tecnología, que complementa un área de la inteligencia artificial que se base en el estudio y realización de máquinas que tengan el conocimiento programable y ajustado a las necesidades, que les permita desempeñar tareas físicas, que sean monitoreadas o no por el ser humano, en el área de la robótica se integran diferentes disciplinas como lo son, mecánica, informática y electrónica (Arroyave Zuluaga & Mosquera Renteria, 2015).

### **2.3.11. Deep Learning.**

Es el análisis al contenido de información que puede ser suministrada por una plataforma web, o por libros en la web, pero su principal enfoque está adaptado a la investigación de datos que no se encuentran organizados ni provienen de bases de datos relacionales y organizadas, su tarea es agrupar y caracterizar datos sin etiquetas por medio de algoritmos creados con la tarea de analizar datos abiertos, que pueden provenir de imágenes, correos electrónicos, conferencias y videos, los cuales deben ser etiquetados, por medio de estas etiquetas el aprendizaje profundo es más efectivo, y estos datos están destinados a complementar el aprendizaje automático (Romero Castillo, 2018).

## **2.4. Marco Teórico**

La revisión sistemática de literatura. Estudio que se realiza sobre otros estudios, que determina los aspectos cualitativos y cuantitativos de investigaciones primarias con el objetivo de resumir información, y caracterizarla, de tal manera que sea posible dividirla en grupos, y evaluarla por medio de herramientas metodológicas y matemáticas que permitan combinar la información, y así mantener el grado de importancia de la descripción de datos.

Ventaja de la revisión sistemática de literatura. Estudio eficiente, que permite analizar y combinar la información de estudios primarios, obteniendo la consistencia de los resultados, para realizar una evaluación completa a la información que se genera partiendo de las diferentes categorías, grupos, de manera que el tamaño de muestra del estudio sea mayor, de tal manera que

implementando métodos matemáticos es considerado como un método estadístico calificado (Manterola, Astudillo, Arias, Claros, & MINCIR, 2011).

#### **2.4.1. Revisión sistemática y metaanálisis.**

La revisión sistemática de literatura en un campo de amplio contenido de estudios, permite a un grupo de investigación indagar y obtener resultados más precisos que respondan a los objetivos de las preguntas de interés del estudio, de tal manera que aplicar el Metaanálisis en una investigación de un tema complejo donde para los investigadores es importante lograr dividir en grupos y subgrupos la información, para lograr de manera detallada, obtener resultados y generar las conclusiones validas, resulta efectiva por que el objetivo principal del Metaanálisis es evaluar la precisión de cada grupo y subgrupo, y obtener la relevancia de cada información requerida para el cumplimiento de cada de uno de los objetivos (Manterola, Astudillo, Arias, Claros, & MINCIR, 2011).

#### **2.4.2. Fases para escribir una revisión sistemática de literatura.**

Las revisiones sistemáticas de literatura permiten a los investigadores, hacer un método de búsqueda en bancos de datos, que se encuentran organizados por su fechas de publicación, o por el tipo de contenido que tienen, la información que se obtiene permite hacer un análisis cualitativo o cuantitativo a las variables declaradas que contienen el valor agregado final, por lo

tanto es importante realizar un procedimiento detallado a la clasificación de estos datos, para esto la revisión sistemática debe pasar por estas fases (Rory J, 2013).

#### **2.4.2.1. Fase 1: Búsqueda de la literatura**

El funcionamiento de los motores de revisión de literatura, es monitoreado por algoritmos programados con el objetivo de hacerle un seguimiento detallado a los datos, por lo tanto es importante elegir una fuente de información y un problema objetivo claro, que permita en la primera fase de reconocimiento de la información, declarar variables precisas que son las más importantes ya que deben ir radicadas en el manuscrito, que va a contener los resultados finales, o las predicciones esperadas por parte de la revisión, se debe plantear un seguimiento a información reciente dependiendo la naturaleza del estudio, si la revisión sistemática es aplicada a una gran cantidad de estudios, y debe considerarse cualitativa (Rory J, 2013).

#### **2.4.2.2. Fase 2: Búsqueda de información para realizar la revisión sistemática.**

Se debe elegir un motor de búsqueda efectivo, que se realice un seguimiento constante a la información, ya que si se presenta una búsqueda donde no se incluyan ciertos estudios relacionados al tema que se está aplicando, se considera como un estudio de menor valor, por lo tanto la cadena de búsqueda que se está generando para la recolección de los datos, y la asignación de variables, debe ser estricta, por lo general con el crecimiento de investigadores que

aplican revisiones sistemáticas de literatura, se presentaran nuevas investigaciones que mejoraran o devaluaran el poder de predicción por medio de los algoritmos de seguimiento, por esto el motor de búsqueda y asignador de valor a los datos, debe incluir la nueva información, se debe adaptar a las actualizaciones de los estudios que se presentaran (Rory J, 2013).

#### **2.4.2.3 Fase 3: Gestionar hallazgos de la revisión sistemática.**

La recolección de los hallazgos a partir del estudio que se seleccionó para aplicar la revisión sistemática de literatura, debe ser precisa, por eso al momento de definir las variables de agrupamiento de datos se debe ser claro y objetivo, partiendo de esto todo tipo de estudio que fue descartado, por diferentes razones, que deben ser descartados, y ya no deben ser incluidos en la siguiente fase de estudio, debido a los criterios de selección de datos definido, también juegan un papel importante en la caracterización de los datos, ya que descartan cualquier tipo de sesgo que evite que el estudio sea riguroso, por lo tanto la elección de los datos y la información detallada del agrupamiento y comportamiento de estos, permiten que los hallazgos tengan valor y logren obtener resultados y conclusiones válidas (Rory J, 2013).

#### **2.4.2.4 Fase 4: Interpretar Hallazgos de la revisión sistemática.**

La interpretación de los hallazgos generados a partir de la recolección de los datos primarios, deben ser analizados, agrupados, y evaluados, de manera que se analicen cada uno de los posibles resultados, y variaciones que estos presentaran, después estos datos ya identificados y caracterizados completamente, tomaran un valor agregado y sustentado por un método probabilístico, el cual será el encargado de darle un peso significativo a las variables, y de

permitir definir el estudio como cuantitativo, para luego poder ser analizado rigurosamente por una técnica de metaanálisis que dará una descripción a fondo de cada grupo de datos (Rory J, 2013).

#### **2.4.2.5 Fase 5: Estructurar la revisión sistemática.**

La estructuración de la revisión sistemática, parte desde la participación de los investigadores, el objetivo que cada uno plantea en la recolección de datos primarios, y los resultados que se obtendrán a partir de los criterios de aceptación que les aplicaran a los mismos, de ahí parte la clasificación de los hallazgos y de resultados, por otra parte un investigador se puede centrar en el peso o valor que toma un grupo de datos que se caracterizaron y darle una conclusión por medio de valores numéricos y porcentajes de aceptación, por otra parte el otro integrante de la investigación puede fijar sus conclusiones en base las características de los datos, esto quiere decir que sus objetivos y conclusiones se reflejaran a partir de cualidades, y así el trabajo de investigación se estructura de una manera cualitativa, y la otra cuantitativa, para conformar una revisión sistemática de literatura completa, y que obtiene resultados, estos toman valor cualquier sea el método de estudio de los datos (Rory J, 2013).

#### **2.4.2.6 Fase 6: Introducción a la revisión sistemática de literatura.**

Para aplicar una revisión sistemática de literatura que su estudio plante la investigación del comportamiento de grandes volúmenes de datos , y a partir de estos generar soluciones a la problemática estudiada, se debe enfatizar en el objetivo de su estudio y los grupos de

información que se van a analizar, lo que quiere decir que la investigación debe ser clara en el por qué se realiza y en que campos se verá reflejado el estudio, los resultados que se esperan de este, y por qué es necesario implementar la revisión sistemática en esta área (Rory J, 2013).

#### **2.4.2.7 Fase 7: Descripción de los métodos de revisión sistemática de literatura.**

Cuando la elección de la fuente de los datos ya está definida, y la recolección de los mismos está organizada y caracterizada, lo siguiente es aplicar un método que permita dar una explicación y justificación detallada, de cómo está organizada la entrada de los datos a los algoritmos de precisión, de igual manera se debe definir los criterios que definen la aceptación y no aceptación de fuentes de información que se estudian, la estrategia de búsqueda que se aplica para realizar el análisis a todos los documentos, y el tipo de seguimiento y calificación que se les van a dar a cada grupo a evaluar, como también debe establecerse si la revisión y sus métodos de obtención de datos se escogió de alguna fuente ya organizada (Rory J, 2013).

#### **2.4.2.8 Fase 8: Resultados de la revisión sistemática.**

Los resultados, pruebas y explicaciones que se generan a partir de concluir el estudio por medio de una revisión sistemática, deben ser implementados por medio de métodos cuantitativos, que permitan ser demostrados a través de diagramas que representen efectivamente el funcionamiento y el valor de la información caracterizada, y definida por criterios de aceptación, toda esta información de los resultados deben ser demostrados y evaluados por gráficas y métodos que ilustren el resultado y conclusiones de la revisión sistemática, pero lo que es

previamente obligatorio es establece de manera textual, la definición de la información, la estrategia de recolección de los datos primarios y los patrones de estudio que se aplican (Rory J, 2013).

#### **2.4.2.9. Fase 9: Discusión de la revisión sistemática de literatura.**

Después de que el manuscrito se ha realizado completamente, en la fase de discusión se implementa una estrategia que permita enumerar la cantidad de revisiones sistemáticas que se relacionan en cuanto al tema de investigación, y a partir de esto, generar detalles de datos que ofrecen diferentes comportamientos, datos que son contradictorios en otras investigaciones y en la que se plasmó muestra resultados más positivos, o de diferente efecto, lo cual refleja en las fases anteriores el acertamiento de los hallazgos y la utilidad que se le dio a los datos primarios, por medio de la discusión se puede realizar un metaanálisis, que permite estudiar a fondos, los resultados obtenidos, tanto con el mismo resultado en las conclusiones, como aquellas revisiones que aportaron diferentes resultados, y de ahí sacar nuevos valores agregados a información, que no había pasado por la aceptación de los criterios principales (Rory J, 2013).

## **2.5. Marco Legal**

### **2.5.1. Ley No. 1341 del 30 de julio de 2009.**

Por la cual se definen principios y conceptos sobre la sociedad de la información y la organización de las tecnologías de la información y las comunicaciones (TIC), se crea la agencia nacional de espectro y se dictan otras disposiciones.



Artículo 17.- Objetivos del ministerio. Los objetivos del ministerio de tecnologías de la información y las comunicaciones.

1. Diseñar, formular, adaptar y promover las políticas, planes, programas y proyectos del sector de tecnologías de la información y las comunicaciones, en correspondencia con la constitución política y la ley, con el fin de contribuir al desarrollo económica, social y política de la nación, y elevar el bienestar de las colombianas.
2. Promover el uso y apropiación de las tecnologías de la información y las comunicaciones entre las ciudadanas, las empresas, el gobierno y demás instancias nacionales como soporte del desarrollo social, económico y político de la nación.
3. Impulsar el desarrollo y fortalecimiento del sector de las tecnologías de la información y las comunicaciones, promover la investigación e innovación buscando su competitividad y avance tecnológico conforme al entorno nacional e internacional.
4. Definir la política y ejercer la gestión, planeación y administración del espectro radioeléctrico y de los servicios postales y relacionados, con excepción de lo dispuesto en artículo 76 de la constitución política. (Congreso de Colombia, 2019)

#### **2.5.2. Ley estatutaria 1581 de 2012.**

Por la cual se dictan disposiciones generales para la protección de datos personales.  
Reglamentada parcialmente por el Decreto Nacional 1377 de 2013.

Artículo 1°. Objeto. La presente ley tiene por objeto desarrollar el derecho constitucional que tienen todas las personas a conocer, actualizar y rectificar las informaciones que se hayan recogido sobre ellas en bases de datos o archivos, y los demás derechos, libertades y garantías constitucionales a que se refiere el artículo 15 de la Constitución Política; así como el derecho a la información consagrado en el artículo 20 de la misma.

Artículo 2°. Ámbito de aplicación. Los principios y disposiciones contenidas en la presente ley serán aplicables a los datos personales registrados en cualquier base de datos que los haga susceptibles de tratamiento por entidades de naturaleza pública o privada.

La presente ley aplicará al tratamiento de datos personales efectuado en territorio colombiano o cuando al responsable del Tratamiento o Encargado del Tratamiento no establecido en territorio nacional le sea aplicable la legislación colombiana en virtud de normas y tratados internacionales.

El régimen de protección de datos personales que se establece en la presente ley no será de aplicación:

- a) A las bases de datos o archivos mantenidos en un ámbito exclusivamente personal o doméstico. Cuando estas bases de datos o archivos vayan a ser suministrados a terceros se deberá, de manera previa, informar al Titular y solicitar su autorización. En este caso los responsables y Encargados de las bases de datos y archivos quedarán sujetos a las disposiciones contenidas en la presente ley;
- b) A las bases de datos y archivos que tengan por finalidad la seguridad y defensa nacional, así como la prevención, detección, monitoreo y control del lavado de activos y el financiamiento del terrorismo;

- c) A las Bases de datos que tengan como fin y contengan información de inteligencia y contrainteligencia;
- d) A las bases de datos y archivos de información periodística y otros contenidos editoriales;
- e) A las bases de datos y archivos regulados por la Ley 1266 de 2008;
- f) A las bases de datos y archivos regulados por la Ley 79 de 1993. (Ministerio de comercio, industria y turismo, 2013)

### **2.5.3. Ley 1273 de 2009.**

Por medio de la cual se modifica el Código Penal, se crea un nuevo bien jurídico tutelado - denominado “de la protección de la información y de los datos”- y se preservan integralmente los sistemas que utilicen las tecnologías de la información y las comunicaciones, entre otras disposiciones. (Congreso de Colombia, 2009)

## **Capítulo 3. Diseño Metodológico**

### **3.1. Tipo de Investigación**

Se utilizo la investigación cuantitativa ya que según Hueso & Cascant (2012) afirman que “es un conjunto de técnicas matemáticas que se utilizan para estudiar variables de interés de una determinada población” (p. 3) lo cual nos permitirá obtener resultados solidos sobre la información analizada.

También se considera de tipo descriptivo debido a que en esta se involucran los métodos de organización, descripción, visualización y resumen de datos originados a partir de la recolección de la información. Los datos se pueden resumir numéricamente mediante estadísticos como la media o gráficos (Hueso & Cascant, 2012).

## **3.2. Población y Muestra**

### **3.2.1. Población.**

Se aplico un filtro a la información registrada en el SNIES con el cual se puedan obtener los nombres de las universidades a nivel nacional que tengan carreras dentro del campo de las TIC, basados en estas condiciones se obtuvieron un total de 1297 programas de educación superior de diferentes universidades del país, a las cuales se hizo una búsqueda en sus repositorios para dar con un total de 167 tesis y trabajos de grado que cumplieran con el requisito de estar relacionadas con las TIC, con la inteligencia artificial y haber sido desarrolladas entre los años 2015-2020.

### **3.2.1. Muestra.**

Teniendo en cuenta que para la investigación se conoce la población de estudio se opta trabajar con el muestreo probabilístico y posteriormente a esto se aplica la técnica del muestreo aleatorio estratificado empleando la forma de la afijación proporcional para tener un dato

recomendado del número de tesis y trabajos de grado a elegir de cada repositorio universitario (Otzen & Manterola, 2017).

Para llevar a cabo esto primero se debe calcular el tamaño de la muestra:

Tabla 1

*Ecuación de la muestra.*

---

**ECUACIÓN DE LA MUESTRA APLICANDO MUESTREO  
PROBABILÍSTICO**

---

$$M = \frac{K^2 * (p * q) * N}{((e)^2 * (N - 1)) + (K)^2 * (p * q)}$$


---

*Nota Fuente:* Otzen & Manterola, (2017). Obtenido de:  
<https://scielo.conicyt.cl/pdf/ijmorphol/v35n1/art37.pdf>

Donde:

N: Representa la población total de tesis y trabajos de grado válidos para el estudio.

K: Porcentaje de la certeza de los resultados de la investigación 95%.

p: Diferencia entre la probabilidad de ocurrencia y lo real 50%.

q: Proporción de individuos que poseen las características de estudio 50%

e: Diferencia entre los resultados obtenidos y los posibles resultados si analizamos todas las tesis y trabajos de grado 5%.

Teniendo en cuenta lo mencionado se procede a realizar el cálculo de la muestra:

Tabla 2

*Calculo de la muestra para las tesis y trabajos de grado.*

---

**MUESTRA DE TESIS Y TRABAJOS DE GRADO**

---

---


$$m = \frac{((1.96)^2 * (0.5 * 0.5) * 162)}{((0.05)^2 * (162 - 1)) + ((1.96)^2 * (0.5 * 0.5))}$$

$$m = \frac{3.8416 * 40.5}{0,4025 + 0.9604}$$

$$m = \frac{155,5848}{1,3629}$$

$$m = 114,1571 \approx 114$$


---

*Nota Fuente:* Autores del proyecto.

Una vez calculado el tamaño de la muestra se procede a aplicar la formula del muestreo aleatorio estratificado en su forma de fijación proporcional a las universidades que cumplieron con los requisitos de selección, debido a que con esta técnica se logra obtener que el tamaño de la muestra de cada estrato sea proporcional al tamaño del estrato que le dio origen, respecto a la población total (Otzen & Manterola, 2017).

Tabla 3

*Formula del muestro aleatorio estratificado de forma aleatoria proporcional.*

---

**FORMULA DEL MUESTRO ALEATORIO ESTRATIFICADO DE  
FORMA AFIJACION PROPORCIONAL**

---

Paso 1.  $P = \frac{m}{N}$

Paso 2.  $MP = P * (\text{Numero de tesis o trabajo de grado de cada repositorio})$

---

*Nota Fuente:* Otzen & Manterola, (2017). Obtenido de:  
<https://scielo.conicyt.cl/pdf/ijmorphol/v35n1/art37.pdf>

Donde:

m: Es la muestra calculada con anterioridad.

N: Representa la población total de tesis y trabajos de grado válidos para el estudio.

P: Proporción entre el tamaño de la muestra y la población de estudio.

MP: Es el número de tesis de y trabajo de grado validos por universidad a estudiar.

Tabla 4

*Calculo de la proporción entre el tamaño de la muestra y la población de estudio*

**CÁLCULO DE LA PROPORCIÓN ENTRE EL TAMAÑO DE LA MUESTRA Y LA POBLACIÓN DE ESTUDIO**

$$P = \frac{114}{162}$$

$$P = 0,7037037$$

*Nota Fuente:* Autores del proyecto.

Tabla 5

*Calculo del número de tesis y trabajos de grados a estudiar por universidad*

<b>UNIVERSIDAD OBJETO DE ESTUDIO</b>	<b>CALCULO MP</b>	<b>NUMERO DE TRABAJOS DE GRADO Y TESIS A ELEGIR POR UNIVERSIDAD</b>
Escuela Colombiana De Ingeniería Julio Garavito.	$MP = 0,7037037 * 16$ $MP = 11,2592 \approx 11$	11
Universidad Católica De Colombia	$MP = 0,7037037 * 21$ $MP = 14,7777 \approx 15$	15
Universidad Cooperativa De Colombia	$MP = 0,7037037 * 1$ $MP = 0,7037037 \approx 1$	1
Universidad Del Norte	$MP = 0,7037037 * 10$ $MP = 7,037037 \approx 7$	7
Universidad De Los Andes	$MP = 0,7037037 * 19$ $MP = 13,3703 \approx 13$	13
Universidad Externado De Colombia	$MP = 0,7037037 * 2$ $MP = 1,4074 \approx 1$	1
Universidad Libre	$MP = 0,7037037 * 11$ $MP = 7,7407 \approx 8$	8
Universidad Tecnológica De Bolívar	$MP = 0,7037037 * 1$ $MP = 0,7037037 \approx 1$	1

Universitaria Agustiniiana-Uniagustiniana	$MP = 0,7037037 * 3$ $MP = 2,1111 \approx 2$	2
Universidad de los Llanos	$MP = 0,7037037 * 1$ $MP = 0,7037037 \approx 1$	1
Universidad Francisco de Paula Santander Ocaña	$MP = 0,7037037 * 2$ $MP = 1,4074 \approx 1$	1
Universidad tecnológica de Pereira – UTP	$MP = 0,7037037 * 9$ $MP = 6,3333 \approx 6$	6
Universidad De Cundinamarca-UDEC	$MP = 0,7037037 * 5$ $MP = 3,5185 \approx 4$	4
Corporación Universitaria La Sallista	$MP = 0,7037037 * 1$ $MP = 0,7037037 \approx 1$	1
Institución Universitaria De Envigado	$MP = 0,7037037 * 3$ $MP = 2,1111 \approx 2$	2
Universidad EIA	$MP = 0,7037037 * 2$ $MP = 1,4074 \approx 1$	1
Corporación Universidad Piloto De Colombia	$MP = 0,7037037 * 8$ $MP = 5,6296 \approx 6$	6
Fundación Universitaria Internacional De La Rioja – UNIR	$MP = 0,7037037 * 2$ $MP = 1,4074 \approx 1$	1
Pontificia Universidad Javeriana	$MP = 0,7037037 * 2$ $MP = 1,4074 \approx 1$	1
Universidad Distrital-Francisco José De Caldas	$MP = 0,7037037 * 11$ $MP = 7,7407 \approx 8$	8
Universidad De Cartagena	$MP = 0,7037037 * 7$ $MP = 4,9212 \approx 5$	5
Universidad Autónoma De Manizales	$MP = 0,7037037 * 7$ $MP = 4,9259 \approx 5$	5
Universidad Del Cauca	$MP = 0,7037037 * 6$ $MP = 4,2222 \approx 4$	4



Universidad Autónoma De Bucaramanga - UNAB	$MP = 0,7037037 * 4$ $MP = 2,8148 \approx 3$	3
Universidad Autónoma De Occidente	$MP = 0,7037037 * 3$ $MP = 2,1111 \approx 2$	2
Universidad De San Buenaventura	$MP = 0,7037037 * 7$ $MP = 4,9259 \approx 5$	5
Universidad Santiago De Cali	$MP = 0,7037037 * 1$ $MP = 0,7037037 \approx 1$	1
<b>TOTAL</b>		116

*Nota Fuente:* Autores del proyecto.

## Capítulo 4. Resultados

El número de universidades totales que ofrecen programas en el campo de las TIC según la base de datos del SNIES, son 212, sin embargo, solo 27 de estas fueron tenidas en cuenta para el estudio; dejando por fuera a 185 instituciones. En la figura XX se consolidan el total de las razones por las cuales no fueron tenidas en cuenta varias instituciones.

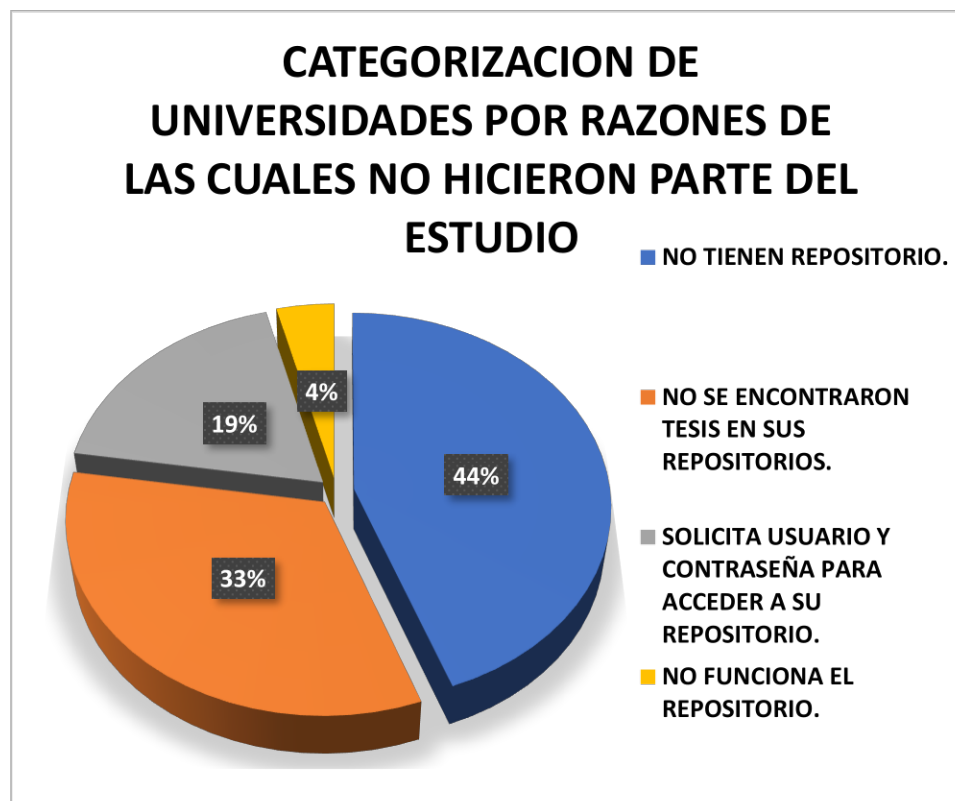


Figura 3. Categorización de universidades por las razones que no hicieron parte de este estudio.

A continuación, se describe con mayor detalle las razones por las cuales no fue posible acceder a la información de las tesis de algunas universidades:

- La universidad no tiene repositorio: De las 185 universidades, el 44% que equivale a 67 de estas instituciones, no tuvieron resultados satisfactorios al momento de buscar sus repositorios, debido a que, aunque muchas contaban con páginas web oficiales, estas no contaban con la categoría de “repositorio” o algún termino similar que permitiera identificar una base de datos donde se encontrarán guardadas las tesis y trabajos de grado de sus estudiantes. En la tabla 6, se hace un listado de las instituciones que tienen este comportamiento.

Tabla 6.  
Universidades que no tienen repositorio.

	<b>NOMBRE DE LA UNIVERSIDAD</b>
1	Instituto Nacional De Formación Técnica Profesional De San Andrés
2	Corporación Universitaria Americana
3	Corporación Educativa Del Litoral
4	Corporación Politécnico De La Costa Atlántica
5	Corporación Universitaria Empresarial De Salamanca
6	Corporación Universitaria Reformada - CUR -
7	Institución Universitaria ITSA
8	Universidad Sergio Arboleda
9	Fundación Centro De Investigación Docencia Y Consultoría Administrativa -F-CIDCA-
10	Fundación Universitaria Para El Desarrollo Humano - UNINPAHU
11	Politécnico Indoamericano
12	Corporación De Educación Tecnológica Colsubsidio- Airbus Group
13	Corporación Instituto Superior De Educación Social -ISES-
14	Corporación Internacional Para El Desarrollo Educativo -CIDE-
15	Corporación Politécnico Colombo Andino
16	Corporación Tecnológica Industrial Colombiana - TEINCO
17	Corporación Unificada Nacional De Educación Superior -CUN-
18	Corporación Universitaria De Cataluña
19	Fundación Universitaria Compensar
20	Fundación De Educación Superior Nueva América
21	Fundación Universidad Autónoma De Colombia -FUAC-
22	Fundación Universitaria Antonio De Arévalo - UNITECNAR
23	Fundación Universitaria Colombo Internacional - UNICOLOMBO
24	Fundación Universitaria Cafam -UNICAFAM
25	Fundación Universitaria Horizonte
26	Fundación Universitaria Salesiana
27	Politécnico Internacional Institución De Educación Superior
28	Universidad Antonio Nariño

29	Universidad INCCA De Colombia
30	Universidad Pontificia Bolivariana
31	Universidad Católica De Oriente -UCO
32	Institución Universitaria Digital De Antioquia -Iu. Digital
33	Institución Universitaria Pascual Bravo
34	Corporación De Estudios Tecnológicos Del Norte Del Valle
35	Corporación Escuela Superior De Administración Y Estudios Tecnológicos- EAE
36	Corporación Regional De Educación Superior-Cres-De Cali
37	Corporación Universitaria Centro Superior – UNICUCES Cali
38	Corporación Universitaria Minuto De Dios -UNIMINUTO
39	Corporación Universitaria Para El Desarrollo Empresarial Y Social - CUDES
40	Escuela Militar De Aviación Marco Fidel Suarez Cudes
41	Fundación Academia De Dibujo Profesional
42	Fundación Centro Colombiano De Estudios Profesionales, - F.C.E.C.E.P.
43	Fundación Tecnológica Autónoma Del Pacifico
44	Politécnico Colombiano Jaime Isaza Cadavid
45	Politécnico Grancolombiano
46	Fundación Escuela Colombiana De Mercadotecnia -ESCOLME-
47	Fundación Universitaria Autónoma De Las Américas
48	Corporación Tecnológica Católica De Occidente - TECOC -
49	Corporación Universitaria De Sabaneta - UNISABANETA
50	Universidad Del Valle Gentium
51	Universidad De La Amazonia
52	Universidad Del Sinú - Elías Bechara Zainum - UNISINU -
53	Colegio Mayor Del Cauca
54	Corporación Universitaria Autónoma Del Cauca
55	Corporación Universitaria Comfacauca - UNICOMFACAUCA
56	Corporación Universitaria De Santa Rosa De Cabal-UNISARC
57	Corporación Instituto De Administración Y Finanzas – CIAF
58	Corporación De Educación Del Norte Del Tolima
59	Instituto Tolimense De Formación Técnica Profesional
60	Institución Universitaria Antonio José Camacho

61	Fundación Escuela Tecnológica De Neiva - Jesús Oviedo Pérez - FET
62	Instituto Universitario De La Paz
63	Universidad Del Tolima
64	Corporación Universitaria Antonio José De Sucre – CORPOSUCRE UD
65	Corporación Interamericana De Educación Superior- CORPOCIDES
66	Instituto Tecnológico Del Putumayo
67	Politécnico Santafé De Bogotá

- No se encontraron tesis en sus repositorios: De las 185 universidades, el 33% que equivale a 50 de estas instituciones, no tuvieron resultados al momento de buscar dentro de sus repositorios (ver tabla 7), debido a que, aunque muchas contaban con la sección de repositorio o base de datos destinadas para almacenar los trabajos de investigación, estos se encontraban con un numero demasiado pequeño de tesis o los años de publicación correspondían a los de hace una década o en su defecto las secciones por carrera de sus repositorios marcaba un total de cero (0) en sus resultados, con los criterios de búsqueda utilizados.

Tabla 7.

*Universidades que no tienen tesis en sus repositorios.*

	<b>NOMBRE DE LA UNIVERSIDAD</b>
1	Corporación Universidad De La Costa -CUC
2	Universidad Simón Bolívar
3	Corporación Universitaria Iberoamericana
4	Fundación Universitaria Empresarial De La Cámara De Comercio De Bogotá
5	Corporación Universitaria Rafael Núñez
6	Universidad Tecnológica De Bolívar
7	Fundación Universitaria Konrad Lorenz

8	Fundación Universitaria Los Libertadores
9	Pontificia Universidad Javeriana
10	Universidad Central
11	Universidad EAN
12	Universidad ECCI
13	Universidad Externado De Colombia
14	Universidad Pedagógica Y Tecnológica De Colombia - UPTC
15	Universidad Santo Tomas
16	Universidad Autónoma Latinoamericana-UNAULA-
17	Universidad Católica Luis Amigo
18	Universidad De Medellín
19	Universidad De San Buenaventura
20	Fundación Universitaria Católica Del Norte
21	Universidad Católica Luis Amigo
22	Universidad De Caldas
23	Universidad De Manizales
24	Fundación Universitaria De San Gil - Unisangil -
25	Fundación Universitaria Juan De Castellanos
26	Universidad De Boyacá UNIBOYACA
27	Fundación Universitaria De Popayán
28	Tecnológico De Antioquia
29	Fundación Universitaria Internacional Del Trópico Americano
30	Instituto Tecnológico Metropolitano
31	Servicio Nacional De Aprendizaje-Sena-
32	Fundación Universitaria María Cano
33	Institución Universitaria De Envigado
34	Universidad Icesi
35	Corporación De Estudios Superiores Salamandra
36	Fundación Universitaria Católica Lumen Gentium
37	Universidad Católica De Pereira
38	Fundación Universitaria Del Área Andina -
39	Fundación Universitaria Comfamiliar Risaralda
40	Universidad Cooperativa De Colombia
41	Universidad Autónoma De Bucaramanga -UNAB
42	Unidades Tecnológicas De Santander
43	Universidad De La Guajira

44	Universidad Del Valle
45	Universidad De La Sabana
46	Fundación Universidad De Bogotá - Jorge Tadeo Lozano
47	Universidad Del Magdalena - UNIMAGDALENA
48	Universidad De Nariño
49	Universidad Mariana
50	Fundación De Estudios Superiores Comfanorte - F.E.S.C

- Solicita usuario y contraseña para acceder a su repositorio: Como se lista en la tabla 8, de las 185 universidades, el 19% que equivale a 28 de estas instituciones, no fueron tenidas en cuenta para el estudio debido a que para acceder a sus repositorios solicitaban un usuario y contraseña que por lo general son asignados a sus estudiantes. Estas universidades son las siguientes:

Tabla 8.

*Universidades que solicitan acceso de usuario y contraseña.*

	<b>NOMBRE DE LA UNIVERSIDAD</b>
1	Corporación Universitaria Antonio José De Sucre - CORPOSUCRE
2	Fundación Universitaria San Mateo - San Mateo Educación Superior
3	Corporación Universitaria CENDA
4	Corporación Universitaria De Ciencia Y Desarrollo - UNICIENCIA
5	Elite- Escuela Latinoamericana De Ingenieros, Tecnólogos Y Empresarios
6	Escuela Tecnológica Instituto Técnico Central
7	Fundación Tecnológica Autónoma De Bogotá-Faba-
8	Institución Universitaria Marco Fidel Suarez – IUMAFIS
9	Institución Universitaria Salazar Y Herrera
10	Corporación Universitaria Remington
11	Universidad De Ibagué

12	Corporación Universitaria Del Caribe – Cesar
13	Tecnológica FITEC
14	Universidad De Santander – UDES
15	Universidad La Gran Colombia
16	Universidad Del Quindío
17	Institución Universitaria EAM
18	Instituto Superior De Educación Rural-ISER-
19	Universidad Francisco De Paula Santander
20	Corporación Universitaria Autónoma De Nariño - Aunar- Compensar
21	Corporación Universitaria Del Meta
22	Fundación Universitaria Compensar
23	Instituto Nacional De Formación Técnica Profesional - Humberto Velásquez García
24	Instituto Nacional De Formación Técnica Profesional De San Juan Del Cesar
25	Universidad Sur colombiana
26	Universidad Popular Del Cesar
27	Universidad Industrial De Santander
28	Universidad Nacional De Colombia

- No funciona su repositorio: De las 185 universidades, el 4% que equivale a 6 de estas instituciones, no fueron tenidas en cuenta para el estudio a pesar de que la universidad tenía un repositorio, al acceder a este se presentaron inconvenientes como que al estar en la página web del repositorio y querer seleccionar una categoría la página no mostraba nada más, o mostraba un mensaje como si esa sección no estuviera existiera. La tabla 9, consolida las universidades con este comportamiento:

Tabla 9.

*Universidades a las que no les funciona su repositorio.*

	<b>NOMBRE DE LA UNIVERSIDAD</b>
1	Corporación Universitaria Latinoamericana – CUL



2	Dirección Nacional De Escuelas
3	Escuela Superior De Guerra General Rafael Reyes Prieto
4	Universidad Manuela Beltrán -UMB-
5	Universidad Católica De Manizales
6	Universidad Nacional Abierta Y A Distancia - UNAD

#### **4.1 Análisis general de la información**

En las siguientes figuras, se muestran los resultados obtenidos luego de categorizar los resultados de las búsquedas y realizar el análisis a los diferentes proyectos de investigación que se encuentran presentes en los repositorios de las diferentes universidades del país, a estas se les aplicó un filtro para lograr una categorización con el fin de obtener la mayor información posible y detallada que permitiera ver el comportamiento de los estudios realizados y de esta manera obtener datos precisos sobre qué tan popular ha sido la inteligencia artificial a nivel nacional para investigaciones durante el periodo de años del 2015 al 2020.

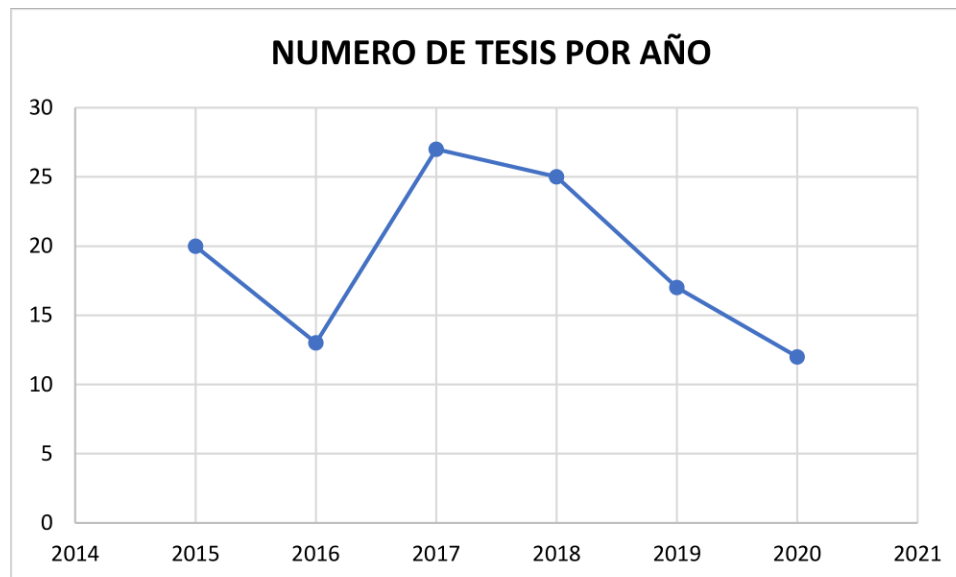


Figura 4. Numero de tesis relacionadas con la inteligencia artificial por año.

En la figura 5, se observa de manera gráfica el número de universidades por departamento que aportaron trabajos de grado a este estudio y que han desarrollado proyectos de investigación usando alguna de las ramas de la inteligencia artificial.

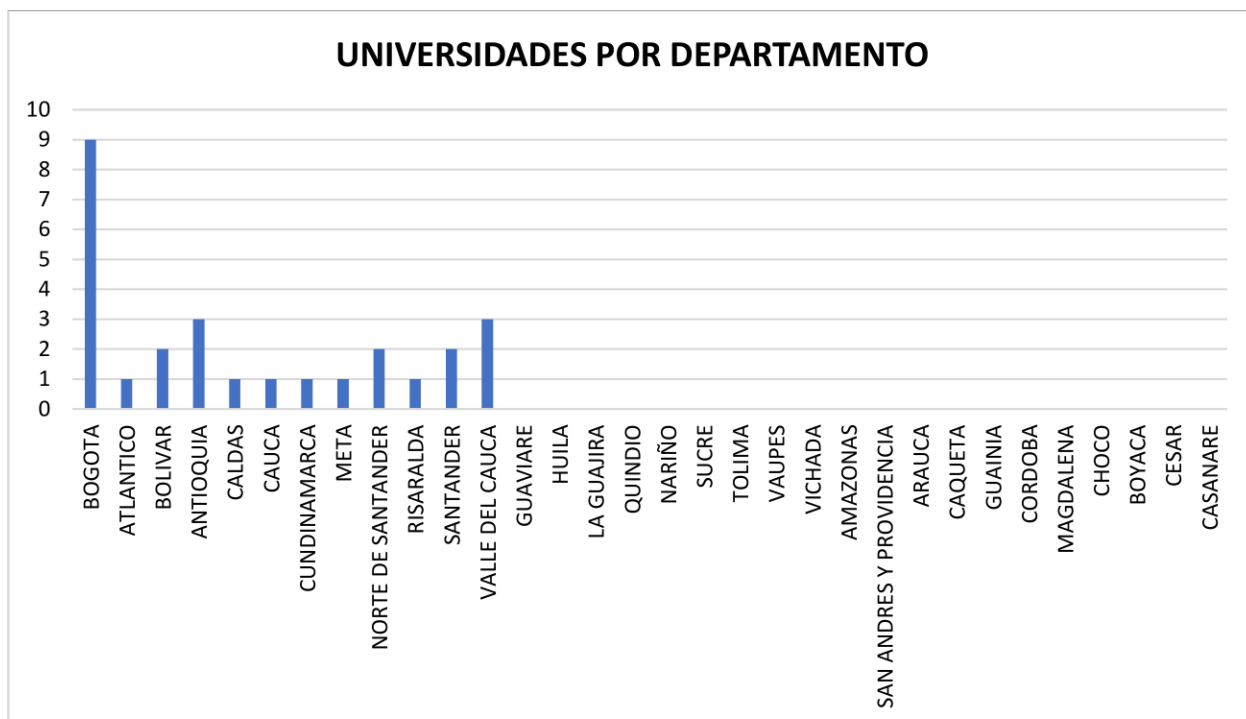


Figura 5. Numero de universidades por departamento.

En dieciocho (18) de los veinte (20) departamentos como son, el Guaviare, Huila, La Guajira, Quindío, Nariño, Sucre, Tolima, Vichada, , San Andrés y Providencia, Arauca, Caquetá, Guainía, Guajira, Córdoba, Magdalena, Choco, Boyacá, Cesar y Casanare la razón por la cual el número de universidades es cero (0) se debe a que al momento de aplicar los filtros necesarios en la base de datos del SNIES y realizar la búsqueda en cada uno de los resultados, las instituciones mostradas no cumplían con los requisitos suficientes para ser incluidas dentro del estudio, y, en los dos (2) departamentos restantes, Vaupés y Amazonas, no se tuvieron en cuenta debido a que el SNIES no mostro ningún resultado.

En la Figura 6; se muestran las cifras que se obtuvieron a partir del estudio realizado a cada una de las universidades mostradas por la base de datos del SNIES y que así mismo

cumplían con todos los requisitos en cuanto al aporte de proyectos basados en inteligencia artificial e Iot, a partir de esto se logró concluir que las ramas más representativas fueron la minería de datos influyendo en un 25.86% de los estudios, las redes neuronales artificiales con un porcentaje del 18.10%, el Iot con un 15.51% y el aprendizaje automático (Machine learning) con un porcentaje del 15.65%, Con estas cifras se determinó que la minería de datos es el área más fuerte de estudio de inteligencia artificial que las universidades del país utilizan para hacerle frente a los nuevos retos tecnológicos y así estar preparados para la revolución industrial.



Figura 6. Numero de tesis por departamentos.



Figura 7. Numero de tesis de trabajos de grados por universidad.

Esta gráfica representa el estudio realizado a las universidades que presentan más tesis y las cuales son más influyentes en formar estudiantes que se enfocan en realizar los proyectos que permitan generar soluciones en diferentes sectores sociales en los diferentes departamentos del país, aplicando diferentes áreas de la inteligencia artificial e IOT, y de esta manera se logra

concluir que la universidad más preparada es la Universidad Católica de Colombia que en todos los programas que ofrece como formación profesional, anualmente se están realizando proyectos de investigación importantes ya que son enfocados muchos al desarrollo de nuevas técnicas que permitan mejorar los servicios de la salud, e influir notablemente en la mejora que se obtiene con modelos de predicción a partir del análisis de gran volumen de datos.

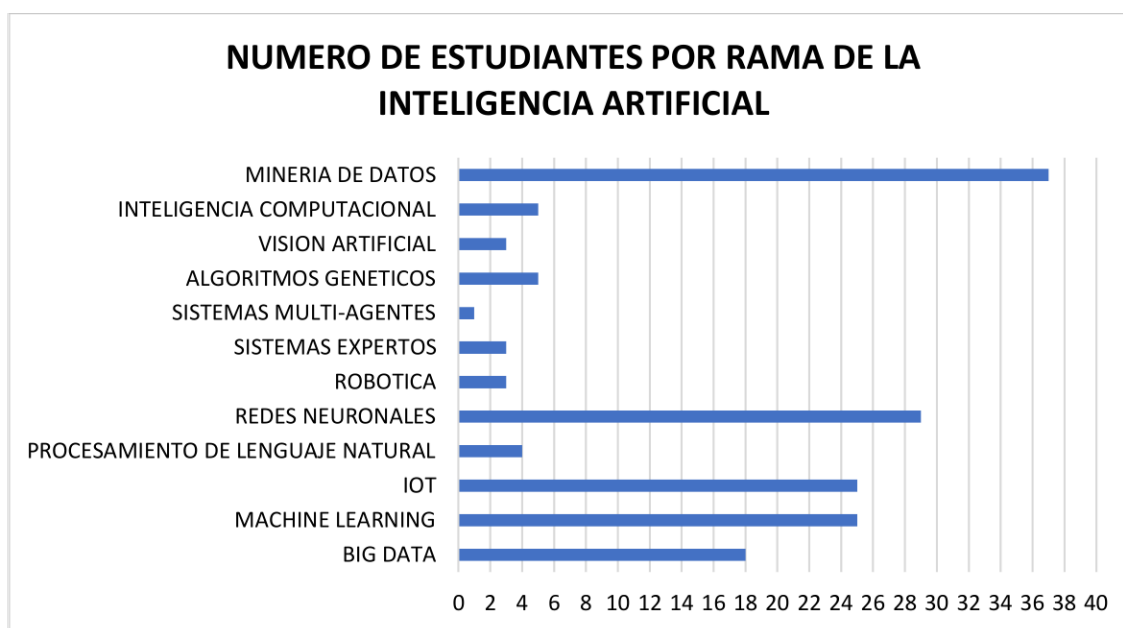


Figura 8. Número de estudiantes que trabajaron en las diferentes áreas de la inteligencia artificial e Iot.

En esta sección del documento se muestra la gráfica de información que contiene el comportamiento de la cantidad de estudiantes y el enfoque que dieron sus proyectos de investigación con el fin de dar soluciones en el ámbito social aplicando las diferentes ramas de la inteligencia artificial, de tal manera que obtuvimos 158 profesionales en diferentes programas de las TIC fueron los que trabajaron con esta nuevas tecnologías emergentes, y que un porcentaje del 23.41% aplicaron la minería de datos como la herramienta de desarrollo de sus

investigaciones, las redes neuronales también fueron incluidas en las investigaciones en un 18.35%, el IOT y aprendizaje automático (machine learning) también se trabajaron en un 15.82% cada una, de tal manera que las otras 8 ramas fueron trabajadas en un 27.6 %, así logrando obtener un 100% de los resultados en técnicas y algoritmos que fueron utilizados en cada de las tesis, en los diferentes programas ofertados por las universidades de Colombia.

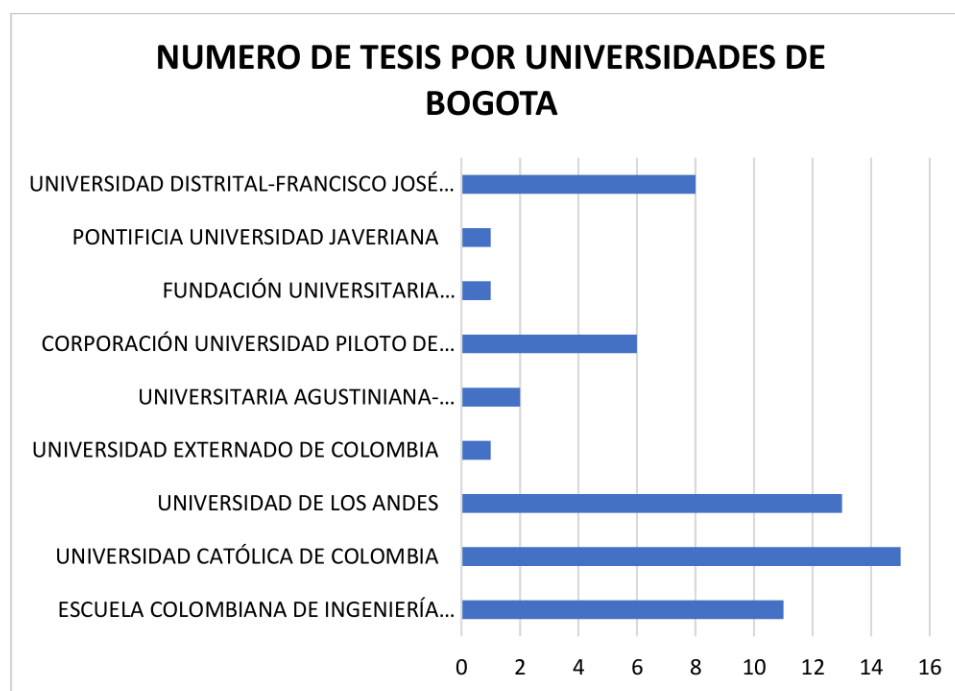


Figura 9. Numero de tesis y trabajos de grados por universidad enfocados en Bogotá.

A continuación realizamos la descripción de los datos obtenidos, al hacerle un estudio profundo a las universidades de Bogotá, el cual fue el más importante desde el 2015 al 2020 por el gran aporte de profesionales y proyectos enfocados en la inteligencia artificial que buscan solucionar y generar nuevas técnicas más eficientes en cada sector del desarrollo del país, por lo que es importante resaltar que la universidad Católica de Colombia es la más preparada en el departamento, ya que sus proyectos de investigación semestrales buscan obtener aumentar sus

competencias y por consiguiente su conocimiento y los beneficios que las nuevas tecnologías emergentes le ofrecen a la vida cotidiana, enfocándose en la solución de los diferentes retos que se van presentado en la economía, salud, y demás áreas que hacen crecer y actualizar a una sociedad, que constantemente está sufriendo cambios a los cuales los profesionales deben buscar una solución eficiente.

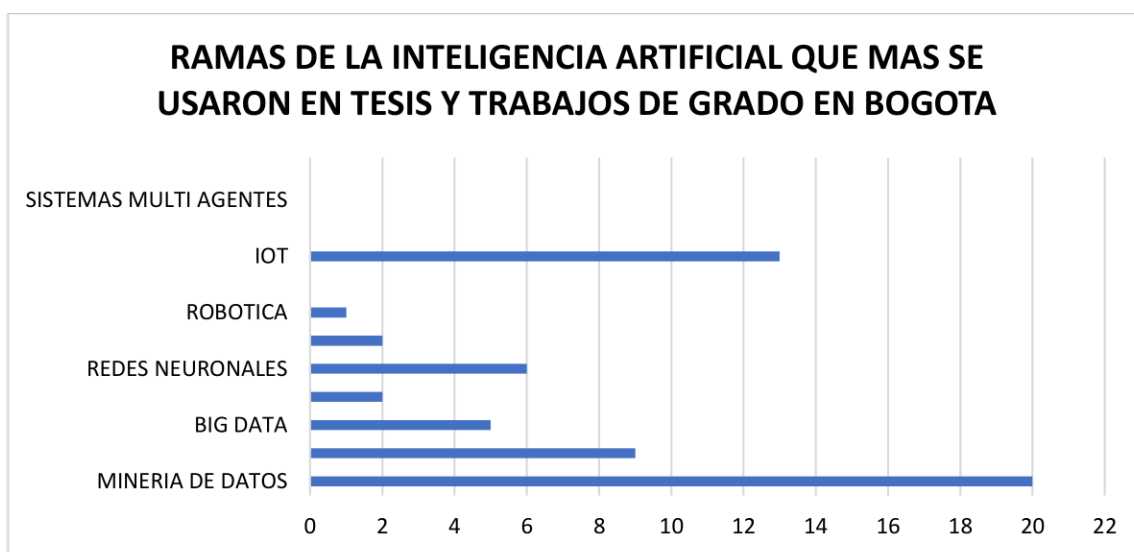


Figura 10. Numero de las ramas de la Inteligencia Artificial e IOT que fueron aplicadas en las tesis de investigación de las universidades de Bogotá.

#### **4.1.1. Ramas de la inteligencia artificial más utilizadas por las universidades**

En esta sección del documento se exponen las cifras concretas de las ramas más utilizadas en las universidades del departamento de Bogotá, haciendo un análisis específico y logrando obtener la información precisa de la que la minería de datos y el internet de las cosas (IOT) fueron las más influyentes ya que su objetivo fue concentrado en diferentes sectores a los cuales se buscaba aplicarlas y mostrar el avance que estas mismas ofrecen, y los beneficios que se obtuvieron en diferentes aspectos, como lo es el crear modelos de predicción que permiten que los profesionales de la salud obtengan una herramienta más eficiente a partir del estudio de la



información con la técnica de la minería de datos creando una tecnología que le permite diagnosticar enfermedades más técnicamente y tener un paso adelante en las soluciones que se le pueden brindar a los pacientes en un tiempo más efectivo, así como también el internet de las cosas mejora el campo laboral creando maquinas inteligentes que permiten optimizar trabajos y generar producción con menos personal de trabajo, y con mejores resultados.

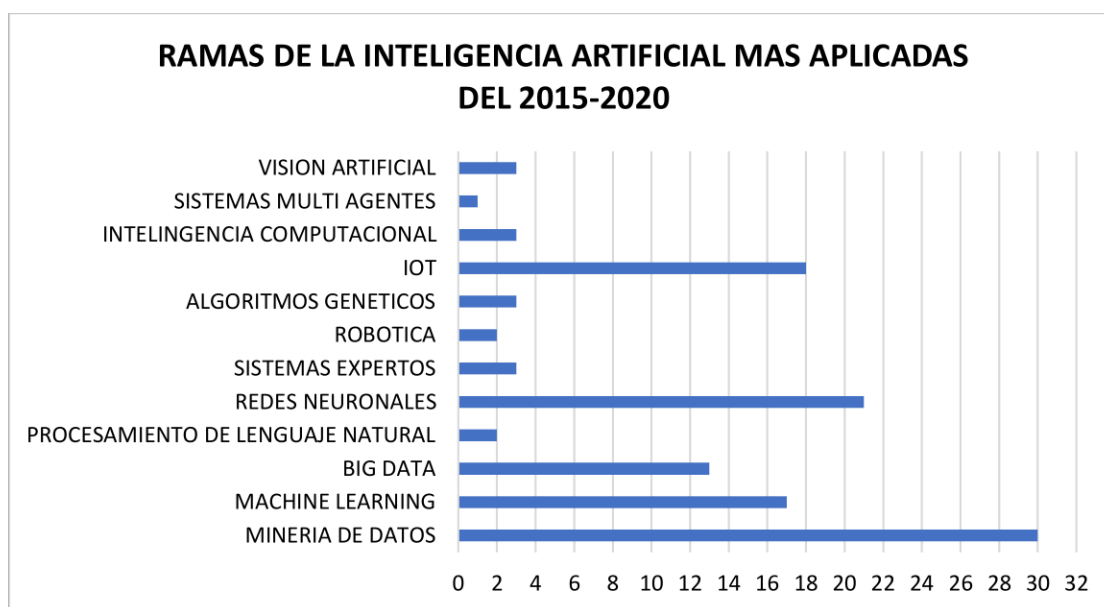


Figura 11. Ramas de la inteligencia artificial **más** aplicadas del 2015 al 2020.

#### 4.1.2. Ramas de la inteligencia artificial por proyecto.

En esta sección la gráfica representa la cantidad encontrada en cuanto a las ramas de la Inteligencia artificial encontradas en cada uno de los proyectos de Investigación alojados en los repositorios de las universidades de educación superior de Colombia.

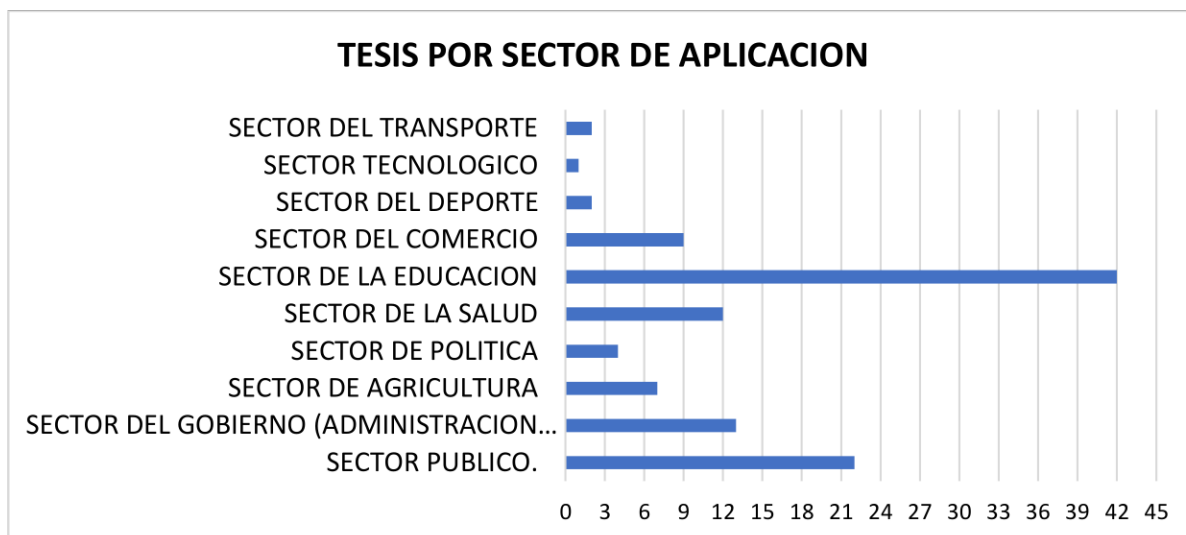


Figura 12. Numero de tesis aplicadas por área de profesión.

En los estudios de los proyectos de investigación se definen un sector de aplicación donde se verá reflejado el trabajo realizado a partir de diferentes técnicas y algoritmos de desarrollo de inteligencia artificial e iot, esto quiere decir el área que se utilizó para llegar a generar una solución en sectores en específicos buscando automatizar y crear modelos de predicción que solucionen diferentes situaciones de la vida cotidiana, por lo tanto en la gráfica se muestran los diferentes sectores donde se aplicaron los trabajos de grado, y la cantidad de proyectos que se realizaron a cada sector.

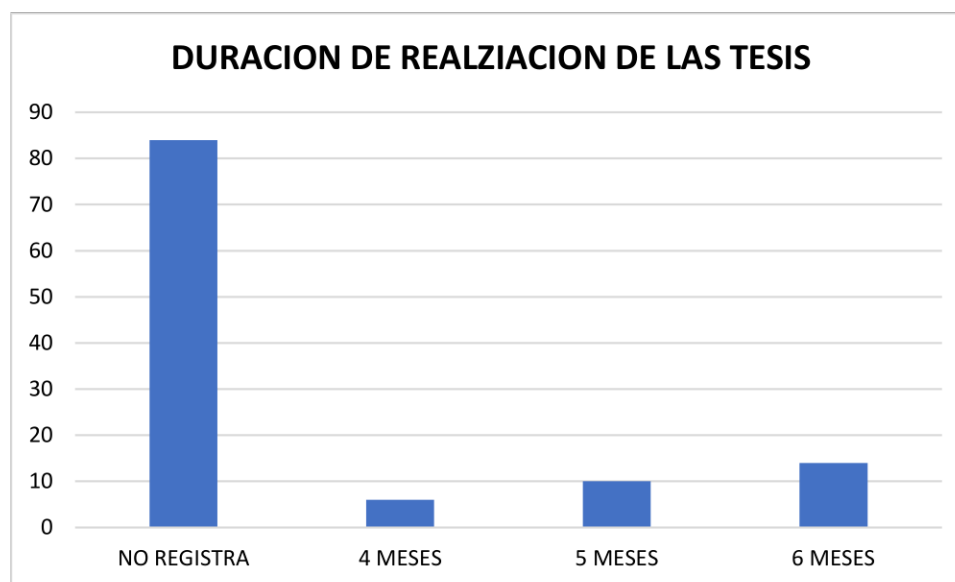


Figura 13. Numero de tesis y tiempo de duración de su desarrollo.

En esta grafica se muestra un intervalo de tiempo en el cual se plantea el desarrollo de los proyectos de grado, los cuales en la presentación de los documentos de las tesis se establecen un periodo calculado en meses, definido muchas veces en el cronograma de actividades, y otras en las delimitaciones temporales de cada proyecto de investigación.

#### **4.1. Categorización por rama de la inteligencia artificial.**

##### **4.1.1. Minería de datos**

Se define la minería de datos como una rama de la inteligencia artificial, la cual consiste en aplicar un conjunto de técnicas que tienen como objetivo descubrir información contenida en un gran conjunto de datos, la mayoría de las búsquedas se realizan en bases de datos que contienen un volumen de datos muy grandes, y que son de tipo abierto, para lograr el análisis de patrones y realizar una evaluación del comportamiento de estos mismos, caracterizados , y agrupados en grandes grupos y sub grupos (Grandez Marquez, 2017).

#### **4.1.1.1. Metodología crisp-dm.**

Es la guía de referencia más utilizada en el desarrollo de proyectos de minería de datos, esta metodología está formada por 4 niveles de abstracción con una organización jerárquica dividida en tareas que van desde el nivel de complejidad más general, y así continuando en el desarrollo de los casos más específicos. (Torres Samboni, 2015)

Estructura el proceso en seis fases:

- Fase 1: Comprensión del negocio

Esta fase inicial se enfoca en la comprensión de los objetivos del proyecto y las exigencias desde una perspectiva de negocio o de la definición del problema, luego se define un problema de minería de datos tomando el conocimiento de los datos y se hace una planeación preliminar para lograr dichos objetivos. (Torres Samboni, 2015)

- Fase 2: Comprensión de los datos.

La fase de entendimiento de datos comienza con la recolección de datos inicial y continua con las actividades que permiten familiarizarse con los datos, identificar los problemas de calidad de datos, descubrir los primeros conocimientos en los datos o descubrir subconjuntos interesantes para formar hipótesis en cuanto a la información oculta. (Torres Samboni, 2015)

- Fase 3: Preparación de los datos.

La fase de preparación de datos cubre todas las actividades necesarias para construir el conjunto de datos final (los datos que serán provistos en las herramientas de modelado) de los iniciales datos en bruto. Las tareas de preparación de datos probablemente van a tener varias iteraciones y no tienen un orden establecido. Las tareas incluyen la selección de tablas, registros y atributos, así como la transformación y la limpieza de datos para las herramientas que los modelan. (Torres Samboni, 2015)

- Fase 4: Modelado.

En esta fase, se selecciona la técnica de modelado que se va a utilizar, aunque es posible que en la fase de comprensión del negocio ya se haya seleccionado. Sin embargo, existen varias técnicas de modelado para el mismo problema, sino que cada una depende de unos requerimientos específicos sobre cómo deben estar preparados los datos. (Torres Samboni, 2015)

- Fase 5: Evaluación.

En esta etapa del proyecto, se ha construido un modelo (o modelos) que parece tener la alta calidad de una perspectiva de análisis de datos. Antes de proceder al despliegue final del modelo, es importante evaluar y revisar las actividades que se hicieron para crearlo, y determinar si se están consiguiendo los objetivos de negocio evaluados en la primera etapa. En el final de esta fase, se debe decidir el uso que se le dará a los resultados obtenidos de la minería de datos. (Torres Samboni, 2015)

- Fase 6: Despliegue

La creación del modelo no es generalmente el final del proyecto. Incluso si el objetivo del modelo es de aumentar el conocimiento de los datos, el conocimiento adquirido tendrá que ser organizado y presentado de tal forma que el cliente pueda usarlo. Dependiendo de los requerimientos, la fase de desarrollo puede ser tan simple como la generación de un informe o tan compleja como la realización repetida de un proceso cruzado de minería de datos a través de la empresa. En muchos casos, es el cliente, no el analista de datos, quien lleva el paso de desarrollo. (Torres Samboni, 2015)

La sucesión de fases, no es necesariamente rígida. Cada fase se descompone en varias tareas generales de segundo nivel. Crisp.dm establece un conjunto de tareas y actividades para

cada fase del proyecto, pero no especifica cómo llevarlas a cabo. (Moine, Gordillo, & Haedo, 2011)

Ámbitos en el que se aplicaron tesis relacionadas con la inteligencia artificial y metodología CRISP-DM:

### **1) Problemas sociales de Colombia.**

#### **a. Análisis de sentimientos sobre el posconflicto colombiano utilizando herramientas de minería de texto.**

- **Problema abordado.**

El problema abordado en la investigación es enfocado a realizar un análisis sobre todo lo que sucede en el posconflicto colombiano, las razones, diferentes sucesos, y afectaciones que ocurren en el país, a través de la técnica de inteligencia artificial conocida como minería de datos, partiendo del análisis que se va realizar a la red social twitter, y el estudio detallado de los tweets. (Torres Samboni, 2015)

- **Datos utilizados:**

- Opiniones de personajes del gobierno
- Opiniones de medios de comunicación.
- Opiniones de grupos o personas de algún grupo político.
- Opiniones de personas independientes.

- **Herramientas de software:**

- Knime: plataforma de minería de datos.

#### **b. Determinación y post-procesamiento de perfiles con técnicas no supervisadas de minería de datos como soporte a la toma de decisiones en la estrategia unidos para la superación de la pobreza en Colombia.**

- **Problema abordado.**

El problema abordado en el proyecto de investigación fue enfocado a procesos de acompañamiento y ejecución de planes de ayuda ofrecidos a las familias más vulnerables y pobres en Colombia, a partir de un estudio de información de datos no ordenados, aplicando un modelado de datos por medio de data mining., aplicando un algoritmo clustering que permite agrupar un gran comportamiento u opiniones en diferentes perfiles, y así clasificar esta información y darle un valor. (Silva Monsalve, 2018)

- **Datos utilizados.**

- Tweets en la red social Twitter.

- **Herramientas de software.**

- Weka es una plataforma web, utilizada para procesos de datos de aprendizaje automático y minería de datos.

**c. Ciudades inteligentes y datos abiertos: Un dashboard basado en minería de datos.**

- **Problema abordado.**

En el problema abordado se busca predecir el comportamiento que tienen los habitantes de una ciudad partiendo del análisis de unos datos ya propuestos, Investigando los datos Abiertos de la fuente datos.gov.co permitiendo la toma de decisiones con la información producto del método de minería de datos utilizado., en este caso aprendizaje supervisado, para el apoyo en procesos de dirección política. (Estevez Blanco, 2017)

- **Datos utilizados.**

- 50000 registros, de las observaciones de los usuarios fuente de datos datos.gov.co.

- **Herramientas de software.**

- Rapidminer es una herramienta perfecta para crear modelos y a posterior la realización de análisis predictivos de grandes volúmenes de datos.
- Mongo db es una base de datos de documentos que ofrece una gran escalabilidad y flexibilidad, así como un modelo de consultas e indexación avanzado.
- Tableau es una plataforma de análisis visual que transforma la manera en que usamos los datos para resolver problemas.

## 2) Sector Educativo

### a. Mapa de la situación académica colombiana a través del análisis de las bases de datos del ICFES.

- **Problema abordado.**

El problema abordado se enfoca en realizar un análisis detallado que permita obtener el resultado de un estudio de la situación académica colombiana basado en el análisis de los resultados de las pruebas estandarizadas SABER 11 que son presentadas por los estudiantes del grado final de educación intermedia, buscando por medio de Técnicas de Minería de Datos lograr tener un conjunto de datos completo de las calificaciones. (Sánchez Pérez, 2020).

- **Datos utilizados.**

- Información obtenida de las bases de datos del ICFES, donde se encuentran publicados los resultados históricos de las pruebas que se han aplicado a los estudiantes colombianos de último grado de formación media en el período 2015-2019.



- **Herramientas de software.**
  - Tableau Archivo de Datos: consolidado.csv Exportado del software Knime.
  - Knime plataforma de minería de datos.

**b. Diseño e implementación de una plataforma de análisis de datos en el sector educativo.**

- **Problema abordado.**

El problema abordado en la investigación consiste en implementar una plataforma que contenga métodos de análisis de grandes volúmenes de datos que en este caso permitan evaluar el desempeño en las pruebas estudiantiles a partir de la técnica de la minería de datos partiendo del análisis de datos de los resultados que obtienen cada estudiante, y así registrados y evaluados en una plataforma web inteligente. (Osorio Salcedo, 2016)

- **Datos utilizados.**
  - Pruebas realizadas por estudiantes de los colegios partiendo del análisis por grupos, Materias, Notas.
- **Herramientas de software.**
  - Plataforma apache spark es un motor de procesamiento de clústeres rápido de uso general que puede implementarse en un clúster de hadoop o en modo independiente.
  - Rapidminer es un programa informático para el análisis y minería de datos.
  - Weka plataforma web, utilizada para procesos de datos de aprendizaje.

**c. Modelo de evaluación de impacto de los programas de pregrado y postgrado de la facultad de ingeniería de la universidad católica de Colombia.**

• **Problema abordado.**

En el desarrollo de este proyecto se investigó cómo otras instituciones de educación superior (IES) llevaron a cabo las mediciones de impacto para evaluar distintos procesos en las funciones sustantivas de docencia, investigación, extensión e internacionalización. En consonancia con esto, para establecer los factores claves que están a la vanguardia de la Educación Superior, se realizó un estudio de minería de datos. Para identificar las variables que se deben contrastar para establecer los indicadores que se utilizarán para medir el impacto. **Fuente especificada no válida.**

• **Datos utilizados.**

- Número de equipos con software para la investigación: cantidad de programas para la investigación.
- Número total de equipos: cantidad total de equipos en la Facultad de ingeniería,
- Número de prácticas realizadas en las asignaturas: cantidad de prácticas durante cada asignatura.
- Número total de asignaturas: cantidad total de asignaturas facultad de ingeniería, 100 equipos utilizados.

• **Herramientas de software.**

- Software vantage point es una herramienta de escritorio para la minería de textos que ayuda a navegar rápidamente por grandes volúmenes de texto estructurado con el fin de visualizar patrones y relaciones.

- Balanced ScoreCard es una herramienta de gestión que permite implementar la estrategia de una empresa a partir de una serie de medidas de actuación, permitiendo un control permanente sobre todos los factores de la organización.

**d. Uso de regresión logística para predecir deserción estudiantil temprana**

- **Problema abordado.**

En el problema de investigación abordado se realiza un análisis a las instituciones educativas de nivel superior tienen un flagelo de deserción estudiantil, que se presenta en diferentes niveles de estudio desde el comienzo de la carrera hasta el final de la misma, por lo tanto el problema abordado por el tema de investigación es de usar el área de la inteligencia artificial conocida como minería de datos para crear un sistema que permita predecir la deserción estudiantil a través del análisis de grandes volúmenes de datos. (Prieto Arciniegas, 2015)

- **Datos utilizados.**

- Datos de los estudiantes.
- Semestre de estudiantes
- Materias por semestre

- **Herramientas de software.**

- Pandas es un paquete de python que proporciona estructuras de datos similares a los dataframes de R.
- Python para la construcción de la data set.
- Jupyter Software de cálculos matemáticos para python e inteligencia artificial.
- Scikit learn es una herramienta realmente útil en máster en data science.

**e. Caracterización y segmentación de público objetivo para scouting universitario a partir de minería de datos.**

- **Problema abordado.**

El problema por resolver en este proyecto es como mejorar la atención brindada por la oficina de scouting por medio de desarrollos de software como, chats más intuitivos para los nuevos interesados en entrar a la universidad de los Andes, creando así una herramienta que permita por medio de técnicas de minería de datos organizar los datos y crear una elección de estudiantes mucho más efectiva. (Morales Guerrero, 2015)

- **Datos utilizados.**

- Los datos con los que funcionara esta tecnología serian datos propios del interesado que llenar un formulario para obtener información del programa interesado.
- Puntaje registrado en el ICFES.
- Datos de las dudas preguntadas por el aspirante al correo de la oficina de scouting.

- **Herramientas de software**

- Bases de datos utilizadas para la extracción de los datos.

### 3) Comercio electrónico

#### a. Predicción de compra sobre consultas de información en comercio electrónico.

- **Problema abordado.**

En el proyecto de investigación en el problema abordado se busca encontrar un modelo que permita realizar la clasificación de consultas de información como potenciales o no compras. Se ha definido que una consulta es buena cuando la probabilidad de que resulte en una compra es mayor a 0.5. Se deben utilizar las variables capturadas en el formulario llenado por el cliente y se

debe extraer información del campo de texto de comentarios de la consulta para que esta sirva de entrada para el modelo predictivo. (Serrano Gonzales, 2015).

- **Datos utilizados.**
  - Datos de consultas de información realizadas desde el 1 de enero de 2002 hasta el 12 de agosto de 2015.
  - Datos de las consultas realizadas desde el 1 de enero de 2012.
  - Consultas de información del pasado introducidas de forma manual.
- **Herramientas de software.**
  - Sistemas de información herramienta para la extracción de los datos.
  - Bases de datos sistemas de alojamiento de información.
  - Weka plataforma web, utilizada para procesos de datos de aprendizaje.

#### 4) Sector salud

##### **a. Predicción del tiempo de la estancia hospitalaria de pacientes con trastorno afectivo bipolar en la clínica san juan de dios de Manizales por medio de técnicas de minería de datos.**

- **Problema abordado.**

En el problema abordado se propone una investigación relacionada con la inteligencia artificial, específicamente con el componente de búsqueda de conocimiento en bases de datos con la intención de construir un posible modelo de predicción de estancia hospitalaria para pacientes con TAB, en donde a partir de factores demográficos se permita conocer el tiempo que el paciente estará hospitalizado durante el tratamiento. La investigación fue desarrollada bajo la metodología CRISP-DM, que es una metodología utilizada para los proyectos de minería de

datos, para la aplicación de las técnicas multivariadas se utilizó el programa estadístico informático SPSS. (Zuluaga Valencia, 2017)

- **Datos utilizados**
  - Base de datos: Hoja de cálculo con los pacientes objeto de estudio
- **Herramientas de software.**
  - Bases de datos transaccionales (SQL),
  - Software de análisis estadístico IBM SPSS statistics versión 20.0 y SAS en su versión 9.1.

**b. Aplicación software para la clasificación de pacientes con Alzheimer basado en técnicas de data mining.**

- **Problema abordado.**

El problema abordado se enfoca en el sector de la salud, se busca realizar una herramienta a partir de técnicas de minería de datos ya que ha sido muy utilizada para crear modelos de predicción, clasificación y selección de información detallada la cual permita obtener de manera automatizada, diagnósticos adelantados, por medio de información y análisis de grandes volúmenes de datos que son organizados en diferentes categorías a las cuales según el tipo de reincidencia de los datos se creen estadísticas que sirven para ser tomadas como datos de valor agregado, en el presente estudio la aplicación de software sirve para clasificar el nivel de la enfermedad de Alzheimer que se presenta en adultos mayores utilizando minería de datos. (Álvarez, Padilla, & Sepúlveda, 2020)

- **Datos utilizados.**
  - Serie de estudios de imágenes de acceso abierto (OASIS).

- **Herramientas de software.**

- Azure VM es una máquina virtual que permite la implementación de una serie de aplicaciones y programas informáticos en un sistema virtualizado.
- Django es un framework web de alto nivel que permite el desarrollo rápido de sitios web seguros y mantenibles.
- GraphQL es Un lenguaje de consultas que permite hacer consultas y esperar una respuesta predecible.
- Tableau es una plataforma de análisis visual que transforma la manera en que usamos los datos para resolver problemas.
- Python ide es una herramienta de código abierto multiplataforma que viene por defecto al descargar el software.

- c. **Perfil de identificación de dolor en neonatos a partir de variables fisiológicas.**

- **Problema abordado**

El problema abordado a solucionar es reducir los índices de mortalidad de las ucis debido a las falencias en cuanto a disponibilidad de talento humano por medio de la información que se genera de los diferentes dispositivos de monitoreo de pacientes internados en estas unidades basados en el comportamiento de variables fisiológicas como lo son el ritmo cardiaco, el ritmo respiratorio, la concentración de oxígeno en la sangre y la presión arterial. (Ramírez Rincón, 2015)

- **Datos utilizados**

- Se utilizó la información de una base de datos desarrollada por el MIT lab for computational physiology, que contiene información de más de 40 mil pacientes hospitalizados en unidades de cuidados intensivos entre 2001 y 2012.

- **Herramientas de software.**
  - MMIC sistema embebido para conexiones de circuitos integrados.
  - Bases de datos sistemas de alojamiento de información.
- d. Esquematización de un sistema para la ayuda en la toma de decisiones medicas al momento de identificar dolor en pacientes dentro la unidad de cuidados intensivos neonatal.**

- **Problema abordado.**

La finalidad del proyecto es desarrollar un marco de referencia metodológico para el desarrollo e implementación de un sistema para la ayuda en la toma de decisiones en el momento de la identificación del dolor en pacientes neonatos en la unidad de cuidados intensivos neonatal. Se considera de gran importancia el desarrollo de dicho marco de referencia, aplicando la minería de datos como la herramienta principal para desarrollar la propuesta de investigación. (Durán Gómez, 2016)

- **Datos utilizados.**
  - Datos de pacientes en homig.
  - Base de datos mimic-III.
  - Datos obtenidos de la plataforma de Physionet.
- Herramientas de software.
  - Weka plataforma web, utilizada para procesos de datos de aprendizaje

## **5) Sector político.**

- a. Las redes sociales como factor de predicción de resultados electorales en campañas presidenciales.**
- **Problema abordado.**



El problema abordado se enfoca en el sector de la política a partir de generar estrategias que permitan cada vez aumentar la posibilidad de posesión de los candidatos a los diferentes cargos, y de esta manera se buscaba implementar una campaña que fuera apoyada desde los resultados en opiniones y estados que se frecuentan en la red social Twitter, analizando diferentes puntos de vista, estados, opiniones, a través de minería de datos se logra el objetivo, partiendo de técnicas de minería de opinión para la recolección de la información y así mismo su clasificación. (Romero Moreno, 2019)

- **Datos utilizados**
  - 6.057.839 tweets en español que fueron extraídos de la red social Twitter (emociones, sentimientos, menciones, entre otros).
- **Herramientas de software.**
  - Lenguaje de programación R aplicado al entorno computacional y precisamente dedicado a la estadística.
  - Software Arima software utilizado para crear modelos de predicción a partir de datos analizados estadísticamente.

## 6) Transporte

### a. Tendencias y características de los viajeros que visitan la ciudad de Pereira por medio de técnica de minería de datos.

- **Problema abordado.**

Se busca implementar una herramienta de minería de datos en el sector del turismo en una área comercial que le da muchas ganancias a los que ofrecen trabajo, y a los trabajadores por lo que es un gran aporte en la economía de la ciudad, país o donde se centren estos negocios, por lo que se implementó una herramienta que a partir de la minería de datos les permite obtener un

modelado de datos que generen valor agregado, para crear estrategias de nivel comercial a partir del análisis de estos grandes volúmenes de Datos. (Salcedo Toro, 2015)

- **Datos utilizados.**
  - Encuesta online.
  - Ingresos utilizados por los viajeros, analizando desde el tiempo de estadía, gastos, gustos, alojamientos.
- **Herramientas de software.**
  - MySQL como motor de base de datos.
  - Rapidminer es un programa informático para el análisis y minería de datos.

**b. Análisis y predicción de accidentes de tránsito en barranquilla mediante técnicas de minería de datos.**

- **Problema abordado.**

El problema abordado en la investigación radica en que la ciudad de Barranquilla existe grandes datos de accidentes que se provocan de distintos modos, como en lugares por lo que los investigadores plantearon una solución a partir de la técnica de minería de datos, haciendo un análisis a todos los registros de accidentes que existen, y dando paso a un modelo de predicción que permita mitigar estos accidentes. (Martinez Bielostotzky & Rueda Iglesias, 2019)

- **Datos utilizados.**
  - Registros proporcionados por el instituto nacional de medicina legal y ciencias forenses (INMLCF) categorizadas por: vehículos, víctimas y momento del accidente.
- **Herramientas de software.**

DSaaP data science as a platform es una plataforma para el análisis de datos en la gestión de riesgos en accidentes en la ciudad de barranquilla.

## 7) Ingeniería de software

### a. Descubrimiento de conocimiento en una empresa de outsourcing de TI de la ciudad de Medellín aplicando técnicas de minería de datos que permita identificar potencialidades en el éxito de los proyectos de desarrollo de software.

- **Problema abordado.**

El problema abordado es una solución que permita identificar potencialidades del éxito de proyectos de software, por lo que surgió la necesidad de poder generar conocimiento útil a partir del análisis de datos históricos de proyectos exitosos y fallidos de una empresa de outsourcing de TI aplicando técnicas de minería de datos, con el fin de obtener patrones que permitieran a la organización tomar decisiones y guiar los proyectos de software a la consecución del éxito.

(Gallego Gallego, 2015)

- **Datos utilizados.**

- Datos históricos de proyectos exitosos y fallidos de una empresa de outsourcing de TI.

- **Herramientas de software.**

- Software SPSS: Sirve para obtener estadísticas avanzadas.
- Software HUGIN. Herramienta de apoyo.

## 8) Sector público gubernamental

**a. Búsqueda de patrones en el comportamiento de los visitantes de la plataforma "oferto" de la cámara de comercio de Armenia y Quindío, a través de la aplicación de minería web.**

- **Problema abordado.**

El problema abordado se enfoca en dar una solución tecnológica a la cámara de comercio de la ciudad de Armenia y Quindío que es una entidad privada sin ánimo de lucro, que requiere de un sistema avanzado que les permita ver la interacción que tienen los visitantes de su plataforma Oferto, haciendo una recolección de diferentes aspectos de los visitantes, como lo es el tiempo, cualidades, búsquedas y otras actividades, para esto se planteó una herramienta que utilizaba minería de datos, aplicando la técnica de minería web, extrayendo contenidos y cualidades de la plataforma, así por medio de la identificación de patrones. (Angarita García & Muñoz Franco, 2016)

- **Datos utilizados.**

- Muestra de las actividades de los usuarios de los días entre el 8 de marzo y el 21 de mayo del 2015.

- **Herramientas de software.**

- MySQL Workbench software que permite desarrollar de forma gráfica bases de datos.
- Weka es una plataforma web, utilizada para procesos de datos de aprendizaje automático y minería de datos.
- Rapidminer es una herramienta perfecta para crear modelos y a posterior la realización de análisis predictivos de grandes volúmenes de datos.

- Watson Analytics es una plataforma intuitiva en la nube que le permitirá descubrir fácilmente el valor de sus datos, incluso si no cuenta con conocimientos técnicos o analíticos.
- C#: Lenguaje de programación.

#### **4.1.1.2. Metodología xp extreme programming.**

La metodología es caracterizada por proponer como factor primario la adaptabilidad antes que la previsibilidad, es por esto que posee un conjunto de valores que la identifican además de un conjunto de etapas definidas, Estos valores como la simplicidad a la hora de agilizar el desarrollo, comunicación en cuanto a la forma de codificar, retroalimentación establecida por el cliente al dar su opinión de los avances del proyecto. (Barriga Mariño, 2017)

Consiste en:

- Planificación del proyecto

Esta etapa comprende la definición de historias de usuario con el cliente, las historias tienen una gran similitud a los casos de uso relacionándose con el levantamiento de requerimientos.

- Diseño.

Consta del planteamiento de diseños simples y sencillos, pero de calidad, también promueve la utilización de glosarios de términos para facilitar el entendimiento del diseño y abrir paso a la reutilización de código por ejemplo en posteriores actualizaciones, adicionalmente, plantea el tratamiento de riesgo para su reducción de forma rápida.

- Codificación.

Esta fase comprende la programación o codificación de las historias o requerimientos, esta programación debe hacerse controlada por un estándar definido esto mantiene el código consistente y facilita su comprensión y escalabilidad.

- **Pruebas**

Como fase final, implica el testeo del código para verificar su funcionalidad, sometiendo a estas pruebas distintas clases del sistema omitiendo los métodos más triviales, estos test no deben tener ninguna dependencia del código que en un futuro evaluará.

Ámbitos en el que se aplicaron tesis relacionadas con la inteligencia artificial y

### **1) Sector político.**

#### **a. Desarrollo y aplicación de una herramienta de extracción y almacenamiento de datos de Twitter a un contexto social de violencia política.**

- **Problema abordado.**

En el problema abordado en la investigación de los antecedentes del proyecto, se identifica que la mayoría de herramientas de extracción y almacenamiento de datos de redes sociales pertenecientes al mercado actual, se centran en un contexto específico, generalmente enfocado al marketing y los negocios, Luego de identificar estos factores y si se pretendiera realizar un tipo de análisis, estadístico, de varianza con enfoque al análisis de sentimientos y con carácter predictivo, que involucre los datos generados por diferentes cuentas en una red social tan influyente como lo es twitter, teniendo en cuenta la necesidad de que la obtención de estos datos pueda generarse de una forma sencilla e implique bajos costos monetarios, además de tomar como contexto para la aplicación funcional de la herramienta la problemática social referente a la violencia política que se genera a través de este medio de comunicación. (Barriga Mariño, 2017).

- **Datos utilizados.**

- Datos analizados en los tweets: reconocimiento del conflicto, situaciones pre-traumáticas y creencias identitarias.

- **Herramientas de software.**

- Plataforma GitHub Es una plataforma de desarrollo colaborativo) para alojar proyectos utilizando el sistema de control de versiones Git. Se utiliza principalmente para la creación de código fuente de programas de ordenador.
- Start uml: Diagramas de componentes, Diagrama de despliegues,
- Liferay plataforma gestora de contenidos

**b. Modelo conceptual para determinar el próximo presidente de estado a través de data mining y Twitter.**

- **Problema abordado**

Se quiere determinar cuál será el próximo presidente de Colombia realizando un estudio y aplicando las técnicas de data minning en los comentarios realizados por una cantidad de usuarios en la red social Twitter, enfocado el estudio de datos en el comportamiento que tienen en diferentes características de los tweets y de esta manera predecir que preferencias existen en cuanto a política en elecciones del presidente actual. (Corzo Sosa, 2018)

- **Datos utilizados.**

- Los comentarios realizados por usuarios en la red social de Twitter.

- **Herramientas de software.**

- Angular es un framework open source desarrollado por Google para facilitar la creación y programación de aplicaciones web de una sola página.
- JavaScript es un lenguaje de programación interpretado ligero.
- Flask es un micro framework de python que permite crear aplicaciones web con pocas líneas.

#### **4.1.1.3 Metodología entrenamiento y validación**

La metodología se caracteriza por realizar una validación cruzada de k-veces, las cuales permiten que todos los datos iniciales sean divididos de manera aleatoria creando subconjuntos en el desarrollo de las fases y etapas de construcción de proyectos a partir del entrenamiento que es realizado una misma cantidad de veces a diferencia de las pruebas para la clasificación, la estimación de precisión en el número total de clasificaciones correctas. (Han & Kamber)

Consiste en:

- Recolección de documentos.

La fuente de información es una combinación entre documentos almacenados en el sistema de archivos y el registro histórico de correspondencias electrónicas en el Sistema de Información de gestión documental Orfeo.

- Preprocesamiento de datos.

Los textos que conforman el corpus de documentos deben ser limpiados de caracteres que son de poca utilidad en la minería de textos para entender el sentido de este, tales como palabras vacías, puntuación, números y fechas.

- Aprendizaje.



Las operaciones de acá en adelante se ejecutaron en R-server y los modelos fueron diseñados para operar en clúster.

- Evaluación y selección.

El resultado de la serie de experimentos se consigna en la tabla 13 utilizando la prueba según Friedman, en paréntesis se le asignó un ranking a cada exactitud obtenida, siendo (1) el algoritmo con más alta exactitud y (3) el que obtuvo la peor. Y al final se promedian los resultados del ranking.

- Prueba de hipótesis.

Con el modelo GBT obtenido durante de la serie de iteraciones, se probó la exactitud utilizando el conjunto de datos de comprobación, el cual conforme se explica en la metodología se compone de 1000 documentos no vistos anteriormente, en la siguiente matriz de confusión se muestran las predicciones por dependencia.

Ámbitos en el que se aplicaron tesis relacionadas con la inteligencia artificial y metodología de entrenamiento y validación:

### **1) Sector público gubernamental**

#### **a. Modelado basado en datos para la clasificación semiautomática de correspondencia electrónica: caso de estudio para la administración pública colombiana.**

- **Problema abordado.**

La solución al problema consiste en crear un modelo de clasificación de datos para las empresas que manejan las direcciones de correo electrónico como base fundamental de relación con sus clientes, por lo tal se implementa una herramienta que de manera semiautomática logre clasificar y dirigir los correos que diariamente se reciben en la empresa de administración

pública, por medio de algoritmos de clasificación de textos, aplicando minería de datos. (Vargas Antolínez, 2018).

- **Datos utilizados.**
  - Correos electrónicos clasificado en las siguientes categorías:
  - Empleo
  - Desarrollo
  - Mérito
  - Municipio
  
- **Herramientas de software.**
  - SHUF sistema operativo Linux.
  - Base de datos sistema Orfeo es Oracle data base,
  - SQL par Oracle denominado TOAD.

A continuación, se muestran los pasos que se tuvieron en cuenta para el desarrollo de este proyecto:

- Umbralizacion.

Es uno de los más importantes métodos de segmentación. El objetivo es convertir una imagen en escala de grises a una nueva con sólo dos niveles, de manera que los objetos queden separados del fondo.

- Segmentación.

Subdivide una imagen en sus partes constituyentes u objetos, con el fin de separar las partes de interés del resto de la imagen, por lo tanto, el nivel al que se lleva a cabo esta subdivisión depende del problema a resolver.

- Binarización.

Este proceso transforma la imagen en escala de grises, a una imagen de blanco y negro (unos y ceros) mediante una comparación o umbral.

- Máscara.

Es una imagen que contiene únicamente dos colores, blanco y negro, adopta la función de selección según el rango que sea asignado a esta, pintando de color blanco los pixeles que están contenidos dentro del rango especificado a la máscara, y descarta todo lo que es de diferente valor, pintándolo de negro, haciendo posible aislar elementos indeseados.

- Discriminador de áreas.

Esta función evalúa la cantidad de pixeles conectados, permitiendo descartar áreas de diferentes tamaños, haciendo un filtrado de la imagen y resaltando los objetos de interés, esto se implementa para tener en cuenta los objetos de un tamaño determinado, poder etiquetarlos como deseados y descartar el resto, con el fin de eliminar ruido.

- Opening.

Este proceso está compuesto por una operación en conjunto en donde interactúan en orden el proceso de erosión y posteriormente el de dilatación aplicada a la imagen con el mismo elemento estructural.

- Detección de bordes.

Es la técnica más común para detectar discontinuidades significativas en el nivel de gris, debido a que son más frecuentes en las aplicaciones prácticas. Los métodos de extracción de bordes de una imagen se basan en la diferencia que experimenta una característica en dos regiones adyacentes y que indican la existencia de un borde. A su vez los bordes pueden

clasificarse por su anchura, ángulo de su pendiente de variación, y las coordenadas de su punto medio.

Ámbitos en el que se aplicaron tesis relacionadas con la inteligencia artificial y metodología de reconocimiento de patrones:

**1) Sector agropecuario.**

**a. Diseño e implementación de un sistema de detección de malezas en cultivos cundiboyacenses.**

• **Problema abordado**

Existe una problemática en el sector agroindustrial en lo referente a la aplicación de agroquímicos, debido a que las malas técnicas de implementación de estos químicos presentan efectos secundarios negativos sobre el suelo y los cultivos. El sector agricultor colombiano en cuanto a aplicación e innovación tecnificada, automatización y agricultura de precisión no es sobresaliente debido a que existe carencia de investigación automática e implementación de la misma, por lo que se pretende abordar de forma contundente esta situación, brindando alternativas y basadas en técnicas de inteligencia artificial en este caso minería de datos.

(Montenegro Bermudez & Parada Rojas, 2015)

• **Datos utilizados.**

- Datos de las características del Cultivo.
- Datos de los químicos a Utilizar.

• **Herramientas de software.**

- Weka plataforma de software para el aprendizaje automático.
- Eclipse entorno de desarrollo con programación Java.
- Matlab Software para operaciones matemáticas.

#### 4.1.1.4 Metodología de 7 fases.

La metodología que consta de 7 fases implementada en el trabajo de grado permite tener un orden y un procedimiento adecuado en el desarrollo del proyecto de investigación que fue enfocados a la gran extracción de información, para que así lograr obtener los resultados esperados. (Rojas Gutierrez & Aguilar, 2017)

Consiste en:

- Obtener las bases de datos.

Se buscan las fuentes de extracción, posteriormente se califican las fuentes de acuerdo a algunos criterios de selección que están asociados a calidad de datos, para la finalmente realizar la definición de las variables.

- Obtener la herramienta de análisis de datos a utilizar.

Se busca la herramienta de análisis de datos con la cual se va a trabajar, para el caso se escogió la herramienta RapidMiner.

- Aplicar las bases de datos en la herramienta de análisis.

Aplicar las bases de datos en la herramienta de análisis: Se aplican los datos obtenidos de las fuentes de extracción dentro de la herramienta de análisis de datos.

- Aplicar los algoritmos de búsqueda.

Con el objetivo que permita identificar patrones de comportamiento dentro de las bases de datos aplicadas en la herramienta de análisis, por tal motivo, se selecciona un algoritmo de clustering para ejecutarla en RapidMiner.

- Realizar la búsqueda de datos que permitan identificar patrones de comportamiento en dichas enfermedades.

Se ejecuta el algoritmo sobre los datos a analizar dentro de la herramienta de análisis de datos.

- Obtener los resultados de las búsquedas.

Obtener los resultados de las búsquedas: Se extraen los resultados obtenidos de la búsqueda en la herramienta de análisis de datos para los patrones identificados.

- Realizar la documentación sobre resultados obtenidos.

Realizar la documentación sobre resultados obtenidos: Se realiza la documentación de acuerdo con los resultados obtenidos de la herramienta de análisis de datos.

Ámbitos en el que se aplicaron tesis relacionadas con la inteligencia artificial y metodología de las 7 fases:

## 1) Sector salud

### a. Minería de datos para el descubrimiento de patrones en enfermedades respiratorias en Bogotá, Colombia.

- **Problema abordado.**

El problema abordado se enfoca en crear un modelo de predicción que cree alternativas de solución temprana a las enfermedades respiratorias que son una de las principales causas de mortalidad en Bogotá, especialmente en menores de cinco años y adultos mayores, y de acuerdo con la secretaria general de la salud, se implementa una herramienta de minería de datos que permita por medio de un modelo de predicción de síntomas de enfermedades respiratorias que han sido tomada de un banco de datos, lograr dar una solución efectiva a los trabajadores de la salud. (Rojas Gutierrez & Aguilar, 2017).

- **Datos utilizados.**

- Datos Obtenidos de una base de datos que contiene los diagnósticos clínicos de los pacientes. (CLUSTERES).
- **Herramientas de software.**
  - RapidMiner Herramienta de análisis de Grandes volúmenes de Datos.

#### **4.1.1.5 Metodología descriptiva**

La metodología descriptiva aplicada a la realización de revisiones sistemáticas, principalmente orientadas al área de la salud, es la adecuada según la naturaleza del presente estudio donde se tomó la metodología propuesta por Kitchenham68, en la cual se presenta un método para la realización de revisiones sistemáticas en el contexto de ingeniería del software y el desarrollo de proyectos. (Cifuentes Cortes & Martinez Naizaque, 2018)

Consiste en:

- Definición correcta del tema.

Se en define el tema de investigación correcto y los datos que se van a estudiar para poder realizar el estudio.

- Elaboración de un plan de trabajo.

El plan de trabajo consiste en las tareas predefinidas en el desarrollo del proyecto a partir de la elección del tema de investigación y las fases de desarrollo.

- Búsqueda bibliográfica.

La selección para realizar la búsqueda bibliográfica de los contenidos a evaluar acerca de los datos ya definidos como la base fuente de la información y del estudio.

- Selección y acceso a los documentos.

La selección de los estudios donde se encuentran los datos es una etapa donde se definen los documentos y si el acceso a estos mismos es posible.

- Enriquecimiento de la documentación mediante intercambios personales.

Los cambios de información son provenientes del estudio individual que realizan los investigadores para llegar a obtener todos los documentos necesarios para la realización del estudio y de la utilidad de los datos.

- Análisis de los documentos.

El análisis de la documentación se genera a partir de los datos, lo que quiere decir que las categorías definidas para la evaluación de la información permiten obtener el estudio detallado de los sistemas.

Ámbitos en el que se aplicaron tesis relacionadas con la inteligencia artificial y metodología descriptiva:

### **1) Sector del Internet.**

#### **a. Revisión sistemática para las técnicas de minería web de contenido.**

- **Problema abordado.**

El problema que radica es la dificultad que se presenta para acceder y analizar los grandes volúmenes de información que se almacena en la web, y todos lo que se le presenta a los usuarios que requieren encontrar ciertos temas que le ayudan a definir categorías, y agregar valor a estas grandes cantidades de datos. (Cifuentes Cortes & Martinez Naizaque, 2018).

- **Datos utilizados.**
  - Se manejaron fuentes de información de: Redalyc, Scielo, scopus, IEEE Xplore, Google Scholar y Web of Science.
- **Herramientas de software.**



- Navegadores web.

#### **4.1.1.6 Metodología de método científico.**

En la metodología de método científico se deben seguir unos pasos como lo es identificar cual es el tipo de investigación para lograr obtener el paradigma positivista que se basa en el método científico, que logre dar una solución que sea objetiva con base en un enfoque centrado en el análisis de datos y la aplicación de algoritmos de aprendizaje. (Arias Roa, 2019)

Consiste en:

- **Conjunto de datos.**

Los datos por estudiar son la fuente principal del estudio de información que se realiza para conocer las fuentes de acceso a la misma, a partir de repositorios de grandes volúmenes de información.

- **Procesamiento y limpieza de datos.**

En esta etapa se realiza la selección de datos más menos importantes para el proceso de clasificación con el objetivo de conservar solo las características o los atributos más relevantes, se calcula la matriz de dispersión para observar cuánto afecta un atributo a otro conjunto y sus correlaciones.

- **Identificación de características.**

La literatura en el estado del arte sugiere usar las características de si un equipo juega en casa o fuera de casa, estas características se transforman en una codificación en una variable discreta para su procesamiento posterior.

- **Muestreo.**

En esta etapa se va a clasificar el conjunto de datos en dos de manera que una sea para realizar el entrenamiento y la segunda sea para realizar las pruebas.

- **Construcción del modelo de predicción.**

En esta etapa, se aplica los clasificadores de aprendizaje automático necesarios para realizar predicción del ganador de un partido de fútbol de la categoría A del fútbol profesional colombiano.

- **Salida de datos.**

En esta etapa se va a obtener el resultado que genera el modelo predictivo donde se indica el resultado del modelo predictivo.

Ámbitos en el que se aplicaron tesis relacionadas con la inteligencia artificial y metodología del método científico:

- 1) Sector deportivo**

- a. Modelamiento predictivo del ganador de un partido de fútbol de la categoría A del fútbol profesional colombiano usando aprendizaje de máquina.**

- **Problema abordado.**

El problema abordado en este proyecto se investigó cómo otras instituciones de educación superior (IES) llevaron a cabo las mediciones de impacto para evaluar distintos procesos en las funciones sustantivas de docencia, investigación, extensión e internacionalización. En consonancia con esto, para establecer los factores claves que están a la vanguardia de la educación Superior, se realizó un estudio de minería de datos para identificar las variables que se deben contrastar para establecer los indicadores que se utilizarán para medir el impacto. (Prieto Alvarez, 2018)

- **Datos utilizados.**

- Número de equipos con software para la investigación: Cantidad de programas para la investigación.
- Número total de equipos: Cantidad total de equipos en la Facultad de Ingeniería.
- Número de prácticas realizadas en las asignaturas: Cantidad de prácticas durante cada asignatura.
- Número total de asignaturas: Cantidad total de asignaturas Facultad de Ingeniería, 100 Equipos utilizados.
- **Herramientas de software.**
  - Software Vantage Point es herramienta de escritorio para la minería de textos que ayuda a navegar rápidamente por grandes volúmenes de texto.
  - Balanced Score Card es herramienta propone un balance entre los indicadores de cada una de estas perspectivas.

#### **4.1.1.7 Metodología Open Up.**

La metodología Open Up fue utilizada para crear la herramienta de predicción de estudiantes desertores de la universidad Francisco de paula Santander Ocaña en el programa ingeniería de sistemas, es un proceso unificado mínimamente suficiente, lo que significa que solo se incluye contenido fundamental capaz de adaptarse al contenido o proceso que sea necesario, teniendo como centro la arquitectura para el impulso en el desarrollo. (Díaz Pedroza & Chindoy Chasoy, 2019)

Consiste en:

- Inicio fase iteración.

Esta fase se encarga de la comprensión del problema y de esta forma responder a la pregunta: ¿Por qué es necesario el desarrollo de un software que permita la automatización de los procedimientos manuales dentro del proceso KDD?

- Fase de elaboración iteración.

Esta fase se encarga de detallar los requisitos del sistema, la creación de una arquitectura, el diseño de los casos de uso.

- Fase de construcción iteración.

En esta fase se desarrollan los diagramas del sistema y el producto final logrando los objetivos planteados.

Ámbitos en el que se aplicaron tesis relacionadas con la inteligencia artificial y metodología Open Up:

### **1) Sector educativo**

**a. Desarrollo de un software que permita la detección de estudiantes desertores del programa de ingeniería de sistemas utilizando procesos y elementos de minería de datos.**

- **Problema abordado.**

El objetivo del funcionamiento del software es que por medio de herramientas de minería de Datos se detecte la cantidad de estudiantes desertores en el programa de ingeniería de Sistemas de la universidad Francisco de Paula Santander Ocaña, evaluando diferentes aspectos desde el rendimiento académico, y diferentes factores que afectan el normal curso de la carrera. (Diaz Pedroza & Chindoy Chasoy, 2019).

- **Datos utilizados**
  - Datos Obtenido de la Base de Datos de la UFPSO.

- Clase: Estado académico, código del estudiante, año de admisión a la carrera, promedio acumulado, promedio por semestre, calificaciones por semestre para cada asignatura, materias perdidas, cantidad de veces que repite cada materia, materias cursadas.
- **Herramientas de software.**
  - WEKA Herramienta computacional, El método de búsqueda (search method).

#### **4.1.1.8 Metodología Alan Porter.**

La metodología establecida por Alan Porter consiste en crear una propuesta de desarrollo de proyectos que está basada en unas fases de responsabilidad denominadas de inteligencia en todos sus procesos, y es muy representativa a nivel de análisis e interpretación de los subprocesos que se presentan en las diferentes etapas, partiendo desde la extracción de datos de los documentos principales. (Giraldo Zuluaga, 2019)

Consiste en:

- Fase 1: Inteligencia
  - a. Profundización del problema.

Definir las preguntas guías del proceso.

- b. Selección de las fuentes de información.

Hacer explícitas las fuentes de información que se utilizarán para recuperar los metadatos y datos a analizar.

- c. Refinamiento de la búsqueda y recuperación de la información.

Establecer la(s) ecuación(es) a utilizar para recuperar la información. Con base en la revisión de los resultados se entra en un proceso iterativo de calibración de la ecuación, con base en los resultados.

- Fase 2: Análisis y diseño

- a. Depuración de los datos.

Eliminar la redundancia y las variaciones innecesarias en los datos. Para esto se depuran los metadatos y datos, dejando estos en un formato apropiado para ser utilizado por la herramienta de análisis.

- b. Análisis básico.

Empezar el análisis de los datos a través de un análisis exploratorio, la limpieza de datos (adicional al paso dos y como resultado del análisis exploratorio), el análisis de primer orden (producir listas) y el análisis de segundo orden (producir matrices).

- c. Análisis avanzado.

Analizar el grado de conexión entre los diferentes elementos.

- Fase 3: Interpretación.

- a. Representación.

Escoger la forma de presentar los datos dependiendo de la población objetivo del estudio.

- b. Interpretación.

Hacer inferencias con base

en los resultados obtenidos.

- c. Construcción del perfil de la investigación.

Usar los resultados y plasmar las inferencias en informes, documentos técnicos, etc. Hecho esto, queda que se utilicen por parte de los usuarios como soporte o insumo de sus decisiones.

Ámbitos en el que se aplicaron tesis relacionadas con la inteligencia artificial y metodología de Alan Porter:

### **1) Sector comercial.**

#### **a. La minería de texto como herramienta para la identificación del perfil**

##### **investigativo: caso gestión del conocimiento en América latina 2012 a 2018**

- **Problema abordado.**

El problema abordado en esta investigación es sobre gestión del conocimiento en América latina de 2012 a 2018, objetivamente buscando crear una herramienta de perfil investigativo en diferentes áreas del sector comercial, pero buscando todo lo que tenga que ver con trabajos donde se aplique la minería de texto como el principal método de recolección de información, y a partir de la misma crear un modelo de solución mayor por el valor agregados a los datos.

(Giraldo Zuluaga, 2019)

- **Datos utilizados.**

- La base de datos bibliográfica de resúmenes y citas de artículos de revistas científicas llamada SCOPUS.

- **Herramienta de software.**

- Knime es una plataforma de minería de datos.

#### **4.1.1.9 Metodología proceso de ingeniería**

Consiste en:

La metodología conocida como el proceso realizado de ingeniería, consiste en llevar a a cabo el análisis de cada una de las etapas de desarrollo de proyectos que se involucran en las

siguientes fases, para la realización del proyecto, siendo estos los pasos considerados más importantes para la realización del análisis de sentimientos en el entorno de mercadeo móvil. (Mina Trujillo, 2017)

- Fase de investigación.

En esta fase se obtuvo conocimiento sobre el concepto de aprendizaje automático, su aplicabilidad, su relación con el análisis de sentimientos, cuáles son los antecedentes, y los procesos que se llevan a cabo la construcción de un modelo para análisis de sentimientos, desde el preprocesamiento hacia la fase de entrenamiento.

- Fase de educación.

En esta fase se llevó a cabo un levantamiento de requerimientos en el cual se tuvieron en cuenta los objetivos planteados, y las funcionalidades principales que esperaban los dueños de Grasshoppers del aplicativo móvil.

- Fase de análisis y diseño.

En esta fase se definió la arquitectura de la aplicación y el diseño detallado de la misma en base a los requisitos que se levantaron durante la fase de educación.

- Fase de implementación de modelo en el API Predicción:

En esta fase se llevó a cabo el entrenamiento del modelo de Google Predicción, utilizando como datos de entrenamiento una recopilación de comentarios de páginas de venta en línea colombianas y de la página de Facebook de Grasshoppers store.

- Fase de implementación.

En esta fase se llevó a cabo la implementación de la aplicación de servidor, y de los aplicativos móviles que consumirían los servicios Rest.

- Fase de integración:



En esta fase se integraron al aplicativo móvil de ventas los servicios de Firebase Analytics y Google predicción API. Con el propósito de aplicar el análisis de sentimientos sobre los comentarios que los clientes hacen y generar los eventos para así posteriormente visualizar los reportes en la consola de Firebase.

- Fase de pruebas:

En la fase de pruebas se llevó a cabo un diseñado de casos de prueba basado en casos de uso, para lo cual se definieron escenarios de casos de uso y a partir de allí se construyeron los casos de prueba, finalmente se aplicaron las respectivas pruebas y se registraron los resultados en las fichas de casos de prueba.

- Fase de implementación de modelo en el api predicción.

Ámbitos en el que se aplicaron tesis relacionadas con la inteligencia artificial y Metodología de proceso de ingeniería:

### **1) Sector comercial.**

#### **a. Análisis de sentimiento en entornos de mercadeo móvil.**

- **Problema abordado.**

El mercadeo en línea es implementada por empresas pequeñas, medianas y grandes alrededor del mundo, muchos negocios migran o complementan su base de venta con el mercadeo móvil en línea dados los beneficios de este y el aumento del uso por parte de los consumidores de tecnologías como las redes sociales o el microblogging; llegando a este punto la cuestión a tratar es que tanto se está aprovechando el potencial del mercadeo móvil en línea, y consecuentemente de qué manera se pueden potencializar los beneficios del uso de esta técnica, concretamente registrando ciertas acciones ejecutadas por los clientes y analizando la

retroalimentación generada por los usuarios o consumidores de estos servicios de mercadeo móvil en línea. (Mina Trujillo, 2017)

- Datos utilizados.
  - Comportamiento en plataformas de servicio web, y tiendas online por cada Usuario.
- Herramientas de software.
  - Hibernate puente para conectar dos modelos de datos diferentes.
  - Spring MVC es una implementación del patrón de diseño "front controller".
  - Android studio es el entorno de desarrollo integrado (IDE) oficial para el desarrollo de apps.
  - Firebase analytics es una herramienta de analítica mobile desarrollada por Google.

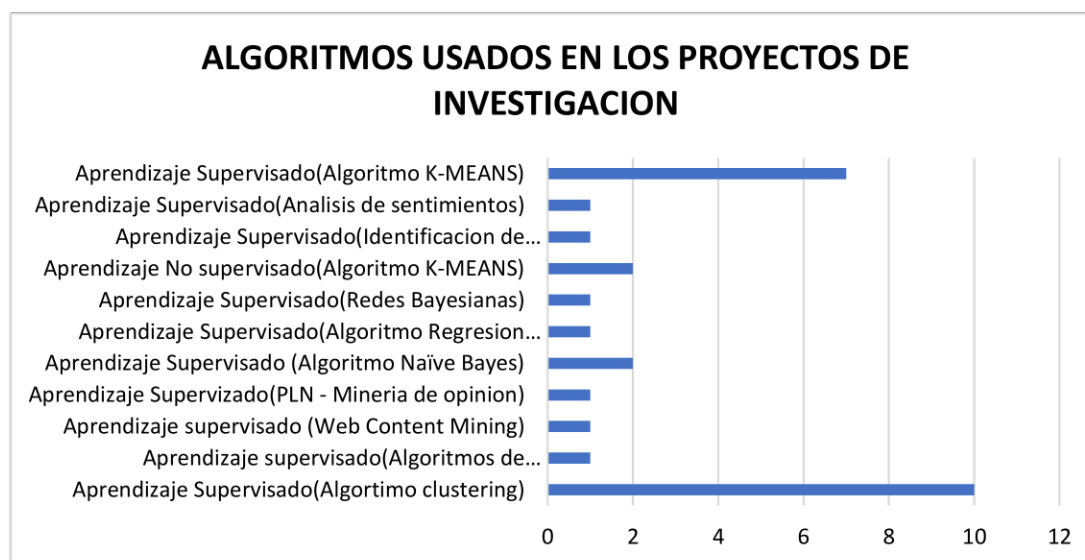


Figura 14. Grafica que representa el número de veces que se usaron determinados algoritmos en las tesis de la rama de minería de datos.

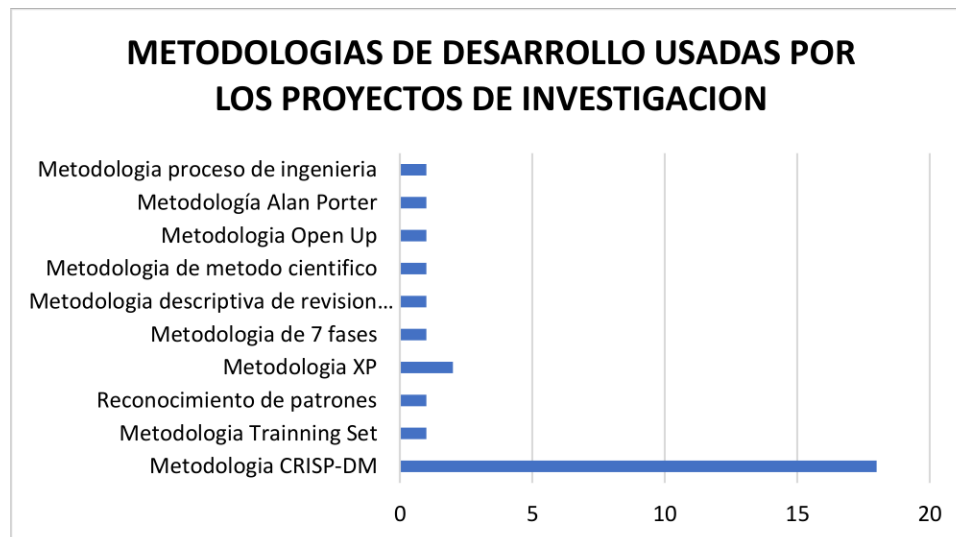


Figura 15. Grafica que representa las veces que se usaron determinadas metodologías para el desarrollo de las tesis.

En los trabajos de grado estudiados donde se encuentra aplicada el área de minería de datos, se obtuvo que los sectores más representantes en los estudios fueron el sector público, de educación, de la salud y comercial, luego del análisis detallado de cuáles fueron las metodologías de desarrollo que más se aplicaron para la solución de los diferentes problemas abordados, encontramos que la metodología más representante fue la metodología Crisp-dm, y en unos cuantos casos también la metodología Xp, logrando así resaltar que los algoritmos supervisados y clustering, fueron los más utilizados para obtener los datos y llevarlos a los diferentes software para el ordenamiento de los mismos, donde las plataformas Weka y knime son las más representantes para el alojamiento y clasificación de la información a partir de modelos de predicción.

### **4.1.2 Sistemas multi agentes.**

Los sistemas multi agentes se basan en unas características principales las cuales se enfocan en que cada agente del sistema tiene un punto de vista limitado (no tiene la información completa); no existe un control global para todo el sistema; los datos están descentralizados; la computación es asíncrona y por último permite su interoperación con sistemas existentes.

(Giraldo Giraldo, Guzmán Luna, & Ovalle Carranza, 2005)

#### **4.1.2.1 Metodología PSP.**

La metodología PSP consiste en revisar los procesos de los datos que las personas quieren organizar para su uso, en los cuales se tiende a cometer errores, por siguiente se sigue a analizar los datos de los defectos que son analizados por una lista de chequeo que permite a los ingenieros encontrar las fallas en la organización de la información y así darle paso a la realización de los proyectos. (Larco Ampudia, 2015)

Consiste en:

- Plan: producir un plan para desarrollar el programa definido en los requerimientos.
- Design: producir una especificación del diseño del programa.
- Code: transformar el diseño en instrucciones de lenguaje de programación.
- Compile: trasladar las instrucciones en código fuente al código ejecutable.
- Test: verificar que el código ejecutable satisfaga los requerimientos.
- Postmortem: resumir y analizar los datos del proyecto.

(Larco Ampudia, 2015)

Ámbitos en el que se aplicaron tesis relacionadas con la inteligencia artificial y metodología de PSP:

## 1) Seguridad informática.

### a. Sistema multi-agente deliberativo para la obtención y análisis de datos en honeynets.

- **Problema abordado.**

El propósito de esta tesis es recolectar e interpretar los datos de ataques, identificados por una red honeyNet, bajo un esquema distribuido y autónomo para el apoyo a la toma de decisiones, e identificación de comportamientos de ataques. En la actualidad, los sistemas de misión crítica y de núcleo de negocio, son expuestos al mundo a través de internet, lo que los convierte en blanco de diferentes ataques para la obtención de acceso o de datos internos de cada organización.

- **Datos utilizados.**

- Registros en la plataforma.
- Registro de actividades dentro la plataforma.
- Identificación de firmas.

- **Herramientas de software.**

- IDS (intrusion detección system): Software detección de intrusos java:  
Lenguaje. compilado orientado a objetos.
- Java fx: Plataforma de software.
- Hibernate: Herramienta de mapeo objeto relacional.
- MySql: Motor de base de datos.
- Gson: Librería gratuita liberada por Google.



Figura 16. Grafica que representa el número de veces que se usaron los algoritmos dentro de los proyectos de investigación.

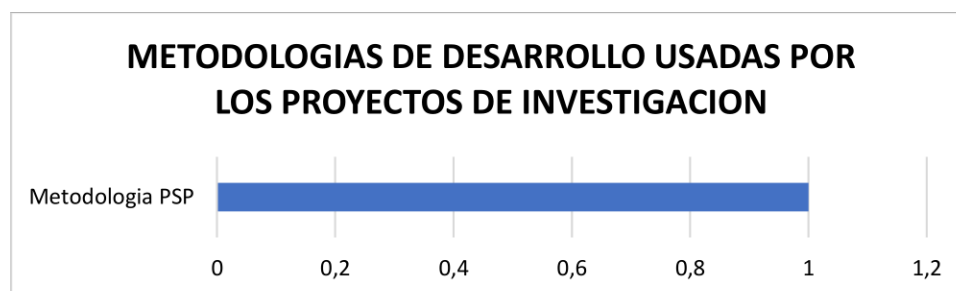


Figura 17. Grafica que presenta el número de veces que usaron determinadas metodologías de desarrollo dentro de la rama de sistemas multi agentes.

En los trabajos de grado estudiados donde se encuentra aplicada el área sistemas multi agentes, se logró identificar que el sector de la tecnología fue donde el trabajo de grado propone realizar un estudio para la solución de un problema basado en datos de ataques, el algoritmo utilizado fue agentes deliberativos por medio del se realiza un manejo de los registros en la plataforma, y por medio de la metodología psp se organiza el funcionamiento del proyecto.

#### 4.1.3 Visión Artificial.

El sistema de visión artificial (SVA) nace con el desarrollo de la inteligencia artificial con el fin que una máquina pueda asimilar todos los elementos de una imagen, así se concibe la visión artificial como el “proceso de extracción de información del mundo físico a partir de imágenes utilizando para ello un computador (Cha K. et al. 1992). La visión artificial o

comprensión de imágenes describe la deducción automática de la estructura y propiedades de un mundo tridimensional, tanto estático como dinámico, a partir de una o varias imágenes bidimensionales de ese mundo (Nalwa.1993). (Miquelarena, Cerrolazaa, Febresa, & Prietoa, 2010)

#### **4.1.3.1. Metodología de cascada**

La metodología de cascada es la utilizada para la realización del proyecto la cual es caracterizada por que su modelo de ejecución de tareas es descendente, cada etapa debe cumplirse de manera total, para así poder darle el paso a la siguiente, esto es una ventaja que tiene esta metodología ya que mitiga la posibilidad de error presentado simultáneamente entre una o más etapas.

Consiste en:

- Planeación del proyecto: en esta fase se determinarán objetivos que se cubrirán en el transcurso del proyecto.
- Análisis: se analiza toda la información obtenida en la investigación previa y se evalúan los requerimientos mínimos que se necesitan para el desarrollo del proyecto, con los requerimientos se puede seleccionar el hardware y software requerido por el proyecto.
- Desarrollo del prototipo: se desarrollará un prototipo que cumpla lo requerimientos del proyecto, estas son algunas etapas del desarrollo del prototipo.
- Prueba de errores: el algoritmo ya desarrollado se somete a unas pruebas exhaustivas para determinar si en el sistema no falla por algún error.
- Verificación: se verifica el algoritmo ya corregido por los anteriores fallos y buscar nuevas vulnerabilidades que genere el algoritmo.

- Finalización del proyecto: finalización del prototipo funcional y la documentación del proyecto.

Ámbitos en el que se aplicaron tesis relacionadas con la inteligencia artificial y metodología de cascada:

### **1) Sector de la Tecnología.**

#### **a. Elaborar un prototipo de escritorio para reconocer imágenes de vista frontal por medio de inteligencia artificial con python openCV.**

- **Problema abordado**

El problema abordado en el proyecto de investigación fue crear modelos de funcionamiento a partir del reconocimiento de imágenes, esta tecnología se ha vuelto bastante influyente en el mercado tecnológico ya que se generan procesos autónomos que permiten por medio de un sistema generar resultados, tanto como de acceso, reconocimiento y en muchos casos de seguridad, Este proyecto es enfocado a crear un prototipo que realice un reconocimiento de imágenes por medio de la inteligencia artificial y su área que es visión artificial. (Acuña Ocampo, 2017)

- **Datos utilizados.**

- Imágenes Guardadas en las bases de datos.
- Herramientas de software.
- Linux debían Raspbian Jessie es una distribución del sistema operativo GNU/Linux basado en Debian.
- Video Core IV 3D graphics core es un procesador multimedia de bajo consumo desarrollado originalmente por alphasosaic ltd que es ahora propiedad de broadcom. Es un procesador de tipo DSP de dos dimensiones



con capacidad para codificar y decodificar archivos multimedia manteniendo el bajo consumo.

## 2) Metodología eigenfaces.

Esta metodología toma un enfoque holístico hacia el reconocimiento facial: Una imagen facial es un punto de un espacio de dimensional alto en cuanto a espacio de imagen y una representación de menor dimensión encontrada; donde la clasificación se vuelve más sencilla.

(Hoyos Monroy, 2019)

Consiste en:

Análisis y diseño para el prototipo de software.

En esta etapa se definen todas las características que va utilizar el prototipo de software partiendo desde el análisis de la interfaz hasta el funcionamiento.

- Construcción en desarrollo del prototipo.

Se define que tecnologías van a ser aplicadas para el desarrollo del prototipo en su primer estado que es como se ve y la estructura que compone al sistema de software, lo que quiere decir que es cuando se define la imagen del prototipo.

- Desarrollar y ejecutar plan de pruebas al prototipo de software.

En el plan de pruebas a ejecutar en el prototipo de software permite realizar una validación del funcionamiento eficiente del sistema de software.

- Elaborar documentación de los resultados de las pruebas.

La fase de la elaboración de las pruebas realizadas al prototipo de software es la tarea que permite verificar el funcionamiento real del prototipo por lo cual lo siguiente es hacer pruebas reales y constantes del uso del software.

Sector de la educación.

**a) Prototipo de software para el reconocimiento facial por medio de visión artificial usando la metodología de análisis eigenfaces.**

- **Problema abordado.**

El problema abordado en el proyecto de investigación es crear un prototipo de software que permita obtener el reconocimiento facial de las personas utilizando patrones de reconocimiento de imágenes a partir del uso de visión artificial, de manera que se automatice el proceso de reconocer rostros a partir de técnicas emergentes.

- **Datos utilizados.**

- Detección de Rostros.
- Reconocimiento de rostros.
- Descripción del algoritmo.
- Análisis de componentes principales (ACP).

- **Herramientas de software.**

- Python 3.x con las librerías: o open cv 4.1.0 o pandas 0.24.2 o pytz 2019.1 o six 1.12.0 o python-dateutil 2.8.0 o numpy 1.16.3.
- Sistema operativo mac os, windows o linux.
- Driver o controlador de la cámara web.

### **3) Metodología ágil Kanban.**

La metodología es adaptada para el desarrollo de software, con el propósito de optimizar la secuencia de trabajo, en la cual se representa visualmente las tareas a desarrollar y su estado, esta es la más apropiada para el desarrollo de este proyecto, porque permite ajustar la cantidad de

trabajo a la capacidad de ejecución, y así tener un enfoque claro del estado de cada una de las tareas. (Pachon Espinel, 2018)

Consiste en:

- Recolección de información:

Revisión de la literatura, para establecer herramientas informáticas a emplear.

- Identificación de requerimientos:

Transcripción de entrevista, donde se identifican las especificaciones que necesitaría el prototipo.

- Diseño de software:

Estandarización de residuos a clasificar.

- Diseño BD:

Diagrama de clases.

- Flujo de navegación:

Determinar la navegabilidad.

- Desarrollo:

Selección de algoritmos.

- Pruebas y ajustes

Pruebas de funcionalidad.

- Construcción de manuales técnico y de usuario:

Captura de video del funcionamiento del aplicativo.

- Entrega de proyecto finalizado.

Revisión del director del proyecto.

Problema abordado. desarrollados con la metodología ágil Kanban.

## **1) Sector del Comercio.**

### **a. Prototipo de sistema automatizado con visión artificial para la selección de empaques de plástico, vidrio y lata en el proceso de reciclaje.**

- **Problema abordado.**

El problema abordado en el proyecto de investigación es crear por medio de herramientas tecnológicas basadas en inteligencia artificial un prototipo de un sistema que permita ser automatizado de manera que aplicando visión artificial se logre realizar la selección de materiales de tipo reciclaje, automatizando estos procesos para la mejora del ambiente de trabajo y de la empresa. (Pachon Espinel, 2018)

- **Datos utilizados.**

- Asignación de colores a los residuos, negro, azul, blanco, para la selección de los datos y su clasificación respectiva.

- **Herramientas de software.**

- Visual Studio Integration 9.2 Deep learning institute.
- Microsoft .net es el framework de trabajo para desarrollo.
- Microsoft visual C++, IDE de desarrollo de software.
- Matlab software para cálculos matemáticos.
- Herramienta de hardware: Arduino.



Figura 18. Grafica que representa el número de veces que fue usado un algoritmo en los proyectos de investigación.

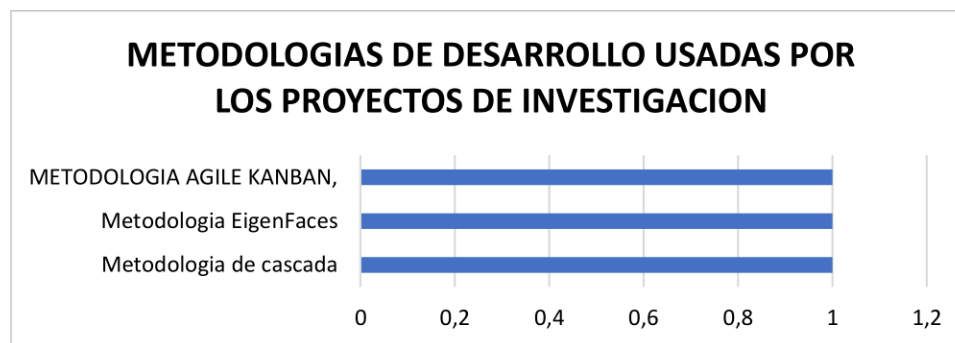


Figura 19. Grafica que representa el número de veces que fue usada una metodología en los proyectos de investigación relacionados con la Visión Artificial.

En los trabajos de grado estudiados donde se encuentra aplicada el área de visión artificial, logramos obtener información relevante que muestra que los sectores más beneficiados con los estudios fueron el sector de la tecnología, de educación y comercial, en estos estudios los datos utilizados son imágenes que se encuentran alojadas en diferentes bases de datos, y el algoritmo que más representa esta área es el algoritmo basado en clustering, a partir de esto se logró concluir que fueron diferentes metodologías las utilizadas para el desarrollo de los proyectos entre las cuales resaltan la metodología de cascada, eigen faces, y Kanban.

#### **4.1.4 Inteligencia computacional.**

Es el estudio de diseño para la creación de agentes inteligentes para realizar acciones en un entorno así obteniendo resultados en los objetivos planteados, El objetivo científico central de

inteligencia computacional, es entender los principios que hacen que el comportamiento inteligente sea posible, en los sistemas naturales o artificiales. (Buendia Robinson & Covilla Yarce, 2015)

### **1) Metodología de desarrollo XP (Xtreme Programming).**

La metodología de desarrollo XP permite plantear claramente los procesos metodológicos que se utilizan para construir software, siendo los requerimientos una de las problemáticas planteadas como necesidad para crear este trabajo de grado, para ayudar a estandarizar el proceso de captura de requerimientos software desde el discurso en el lenguaje natural del cliente.

(Buendia Robinson & Covilla Yarce, 2015)

Consiste en 4 fases:

- Fase de exploración: Hace uso de las historias de usuario para definir el alcance general del proyecto, esta actividad se realiza junto con los clientes, se evalúan los tiempos iniciales de desarrollo, aunque no sean los finales, teniendo en cuenta que lo que se conoce del proyecto en esta instancia es algo muy general y los requerimientos se pueden modificar en cada iteración.
- Fase de planificación: Los requerimientos definidos por el usuario en la fase anterior son utilizados en la fase de planificación en la que se acuerda con los interesados del proyecto el orden en que serán implementados, esta fase es corta a comparación de las demás y el resultado de esta es un plan de entregas.
- Fase de iteraciones: Es quizá la fase más importante en el ciclo de la metodología XP. En esta los requerimientos son desarrollados y se hace la entrega, aquí también el usuario interactúa en la presente fase pues se debe hacer un análisis de

la entrega y de lo que se requiere para así retroalimentar y solicitar cambios si es necesario, además en esta fase también es posible medir el avance del proyecto.

- Fase de puesta en producción: Al finalizar cada iteración se hace entrega de módulos funcionales y sin errores, el cliente puede solicitar no poner el sistema en producción hasta que se obtenga la funcionalidad completa del sistema.

Ámbitos en el que se aplicaron tesis relacionadas con la inteligencia artificial y metodología XP:

### **1) Sector de aplicación desarrollo de software.**

#### **a. Sistema para la captura y análisis de requerimientos de los proyectos basados en inteligencia computacional para la universidad de Cartagena.**

- **Problema abordado.**

En el problema abordado en la investigación se busca agilizar el proceso de captura de requerimientos, optimizando el tiempo de desarrollo en las distintas fases del proceso de software, haciendo uso de los nuevos tics, y al mismo tiempo brindar al cliente una herramienta inteligente, capaz de guiarlo en el momento de explicar sus necesidades. (Buendia Robinson & Covilla Yarce, 2015)

- **Datos utilizados.**

- Se utilizan los datos de una encuesta realizada a profesionales del desarrollo de software

- **Herramientas de software.**

- Corpus.
- Software diseñado para el estudio de documentos relacionados con la lingüística.

- Django: Es un framework web de alto nivel que permite el desarrollo rápido de sitios web seguros y mantenibles.

Es la guía de referencia más utilizada en el desarrollo de proyectos de minería de datos.

Estructura el proceso en seis fases:

- Fase 1: Comprensión del negocio

Esta fase inicial se enfoca en la comprensión de los objetivos del proyecto y las exigencias desde una perspectiva de negocio o de la definición del problema, luego se define un problema de minería de datos tomando el conocimiento de los datos y se hace una planeación preliminar para lograr dichos objetivos. (Torres Samboni, 2015)

- Fase 2: Comprensión de los datos.

La fase de entendimiento de datos comienza con la recolección de datos inicial y continua con las actividades que permiten familiarizarse con los datos, identificar los problemas de calidad de datos, descubrir los primeros conocimientos en los datos o descubrir subconjuntos interesantes para formar hipótesis en cuanto a la información oculta. (Torres Samboni, 2015)

- Fase 3: Preparación de los datos.

La fase de preparación de datos cubre todas las actividades necesarias para construir el conjunto de datos final (los datos que serán provistos en las herramientas de modelado) de los iniciales datos en bruto. Las tareas de preparación de datos probablemente van a tener varias iteraciones y no tienen un orden establecido. Las tareas incluyen la selección de tablas, registros y atributos, así como la transformación y la limpieza de datos para las herramientas que los modelan. (Torres Samboni, 2015)

- Fase 4: Modelado.



En esta fase, se selecciona la técnica de modelado que se va a utilizar, aunque es posible que en la fase de comprensión del negocio ya se haya seleccionado. Sin embargo, existen varias técnicas de modelado para el mismo problema, sino que cada una depende de unos requerimientos específicos sobre cómo deben estar preparados los datos. (Torres Samboni, 2015)

- Fase 5: Evaluación.

En esta etapa del proyecto, se ha construido un modelo (o modelos) que parece tener la alta calidad de una perspectiva de análisis de datos. Antes de proceder al despliegue final del modelo, es importante evaluar y revisar las actividades que se hicieron para crearlo, y determinar si se están consiguiendo los objetivos de negocio evaluados en la primera etapa. En el final de esta fase, se debe decidir el uso que se le dará a los resultados obtenidos de la minería de datos. (Torres Samboni, 2015)

- Fase 6: Despliegue

La creación del modelo no es generalmente el final del proyecto. Incluso si el objetivo del modelo es de aumentar el conocimiento de los datos, el conocimiento adquirido tendrá que ser organizado y presentado de tal forma que el cliente pueda usarlo. Dependiendo de los requerimientos, la fase de desarrollo puede ser tan simple como la generación de un informe o tan compleja como la realización repetida de un proceso cruzado de minería de datos a través de la empresa. En muchos casos, es el cliente, no el analista de datos, quien lleva el paso de desarrollo. (Torres Samboni, 2015)

La sucesión de fases no es necesariamente rígida, cada fase se descompone en varias tareas generales de segundo nivel. CRISP-DM establece un conjunto de tareas y actividades para cada fase del proyecto, pero no especifica cómo llevarlas a cabo. (Moine, Gordillo, & Haedo, 2011)

Ámbitos en el que se aplicaron tesis relacionadas con la inteligencia artificial y metodología CRISP-DM:

Problema abordado. desarrollados con la metodología CRISP – DM.

Sector del Comercio.

**a) Sistema eléctrico de filtrado de información basado en inteligencia computacional para recomendación de atractivos turísticos del caribe colombiano.**

**Problema abordado.**

El sector del turismo es un negocio que aporta mucho dinero a la economía de las ciudades y el país, por lo que se hace importante crear un sistema de filtrado eléctrico que permita atraer en diferentes épocas del año turistas y así las ganancias sean mayores por lo tanto se planteó un sistema eléctrico que se basa en una técnica de inteligencia artificial denominada Inteligencia computacional, que permite crear un modelo de predicción que genera mejor oferta de productos y de más características que hacen que sean mejores ofertas a la hora de las ventas a turistas. (Jiménez García & Montoya Campo, 2017)

- **Datos utilizados.**

- Bases de datos y servicios web de turismo en Colombia, 400 sitios conocidos por el ministerio de industria y turismo.

- **Herramientas de software.**

- Json es un formato ligero de intercambio de datos.
- Start uml Software para desarrollo de diagramas.
- Ramcat una herramienta que permite realizar recomendaciones personalizada.

## **2) Sector ambiental.**

### **a. Predicción adaptativa de la calidad del agua mediante técnicas de inteligencia computacional.**

- **Problema abordado.**

Nuestro país Colombia es el 6 país con más propuestas y mayor oferta hídrica en el mundo, por lo tanto, el ministerio del medio ambiente sabe que hay que adaptarse a los cambios que hacen que deban integrarse a todos los procesos del análisis de calidad del agua, la propuesta de solución es implementar un modelo de predicción adaptativa a la calidad del agua utilizando tecnologías emergentes, que sean adaptadas al área de la inteligencia artificial. (Corrales, 2015)

- **Datos utilizados.**

- 10 mil registros de datos de calidad del agua: piscicola, consumo humano, recreacional.
- Datos provenientes de USGS.

- **Herramientas de software.**

- Weka plataforma web, utilizada para procesos de datos de aprendizaje.
- Start UML software para el desarrollo de diagramas.



Figura 20. Grafica que representa el número de veces que se usaron determinados algoritmos en las tesis relacionadas con la Inteligencia Computacional.

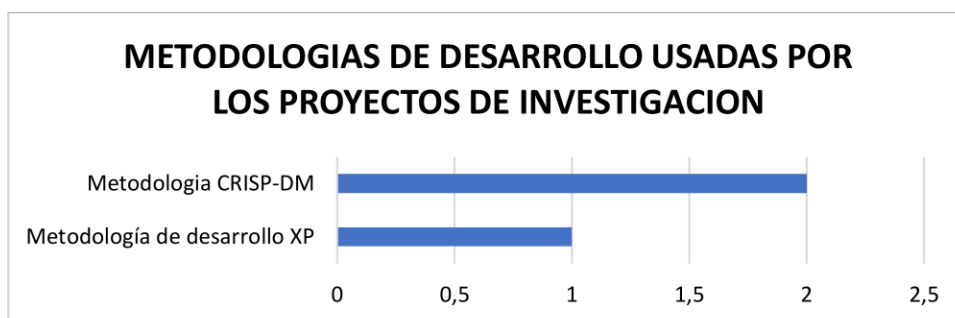


Figura 21. Grafica que representa el número de veces que se usaron determinadas metodologías dentro de la rama de Inteligencia Computacional.

En los trabajos de grado estudiados donde se encuentra aplicada el área inteligencia computacional, se obtuvo que los sectores más representantes en las tesis de grado son el sector ambiental, del comercio y desarrollo de software, para los cuales la información obtenida en los proyectos de investigación fueron datos abiertos, las metodologías más implementada para la realización de los proyectos son Crisp-dm y metodología xp que están basadas la aplicación de fases que permiten el completo desarrollo de los proyectos, y los algoritmos aplicados fueron Algoritmos de lógica difusa, filtrado colaborativo, adaptivos complejos, la plataforma utilizada para el alojamiento y clasificación de datos es Weka.

#### **4.1.5. Internet de las cosas IoT.**

El internet de las cosas es conocido como un área de la inteligencia artificial que permite la creación de sistemas híbridos esto quiere decir que logra la interacción de sistemas programados con utilidad y aplicación en sensores que llevan a cabo la automatización de procesos inteligentes, y así conectar e instrumentalizar cualquier objeto de la vida cotidiana con sistemas con capacidad de decisión propia. (Accenture; Fundacion de la innovacion Bankinter, 2011)

A continuación, se muestran los pasos que se tuvieron en cuenta para el desarrollo de este proyecto:

- Se evalúan las condiciones para saber si pueden ser considerada como fraude electrónico.
- Se determinan las técnicas más utilizadas para realizar fraude electrónico.
- Se determinan las acciones que pueden ayudar a identificar o clasificar una transacción financiera como fraude.
- El problema de los fraudes financieros cae en el ojo internacional con el fin de identificarlo y contra restarlo.
- Se establecen las cifras de fraude electrónico para Colombia.
- Explicación de lo que es un motor de riesgo transaccional.
- Se explica los beneficios de un motor de riesgo tradicional.
- Se explica cómo funciona un motor de riesgo tradicional.
- Se determinan los tipos de controles a llevar a cabo en caso de que el score determinado por el algoritmo de aprendizaje no sea el suficiente para realizar la transacción.

## 1) Sector del comercio.

### a. Motores de riesgo transaccionales ¿solución eficaz para reducir agresivamente el fraude electrónico?

- **Problema abordado.**

En el problema abordado se busca describir como los motores de riesgo tradicional pueden ayudar a reducir las cifras de fraudes electrónicos en el sector financiero a nivel mundial a partir de un algoritmo de aprendizaje que sea capaz de evaluar desde una serie de entradas de variables, un patrón de operación, que permita identificar el perfil de usuario y así validar la autenticidad de quien realiza la transacción. (Sarmiento Martínez, 2015)

- **Datos utilizados.**

- Se utilizaron la puntuación asignada por el motor de búsqueda a cada transacción realizada a partir de variables de entrada que aumentarían con el tiempo.

- **Herramientas de software.**

- Motores de riesgo transaccional.

### b. Gestión de la seguridad en el internet de las cosas.

- **Problema abordado.**

Como en la actualidad se vienen presentado que las empresas y corporaciones que ofrecen sistemas automatizados a partir de temas de internet de las cosas hacen que sea objetivo la razón de crear modelos de gestión que permitan analizar la seguridad que se presenta en los proyectos, basándose en que la información que se produce en cada dispositivo genera confianza, y no permita que sean vulnerados los sistemas y mal uso de informaciones este documento se describen, por medio de un breve resumen, las amenazas de seguridad latentes que se encuentran

en el internet de las cosas. Se describe el estado actual y las amenazas comunes sobre las plataformas de tecnología que se basan en este concepto. (Zabala Jaramillo, 2016)

- **Datos utilizados.**
  - Datos en tiempo real de Sensores distribuidos en la capa de adquisición de datos.
- **Herramientas de software.**
  - Arduino, Librería Arduino, Adaptación (Microcontroladores Arduino).

**c. Gestión de riesgos en el internet de las cosas (IoT).**

Pasos para la realización del proyecto:

- Se da un concepto general de lo que es el IoT y su importancia.
- Se explica la importancia de la gestión de riesgos en el ámbito IoT en el área empresarial basados en las ISO 27001 y ISO 31000.
- En esta parte se evalúa el riesgo teniendo en cuenta las siguientes fases:
  - Identificación del riesgo.
  - Método de análisis del riesgo.
  - Evaluar el riesgo.

Luego de evaluado el riesgo se debe realizar el tratamiento del mismo basándose en dos puntos:

- Acciones de riesgo.
  - Plan de tratamientos de riesgos.
- Luego de tratado el riesgo se debe realizar un monitoreo y revisión de la amenaza y de la forma como fue trata esa amenaza.

- Por último, se lleva un registro de reportes e informes de las amenazas tratadas para realizar las mejoras correspondientes.

- **Problema abordado.**

En el problema abordado del proyecto de investigación se busca establecer una referencia para realizar una gestión de riesgos y adaptarlo al IoT con base en estándares de seguridad, entre los cuales se encuentra iso27001, iso31000 y metodologías magerit. (Gantiva Henao, 2020)

- **Datos utilizados.**

- La información brindada por las iso27001 y iso31000 para definir la referencia para la gestión de riesgos en el internet de las cosas.

- **Herramientas de software.**

- Bases de datos.

**d. Mecanismo de monitoreo para bicicletas en movimiento, usando enfoque de IOT.**

A continuación, se muestran los pasos que se tuvieron en cuenta para el desarrollo de este proyecto:

- Investigación inicial: Búsqueda de la información y adaptarla con teorías.
- Inclusión en procesos: Necesidades de la sociedad, teniendo en cuenta las características del usuario final.
- Proceso investigativo de maduración transferencia de la tecnología: Permite crear prototipos, materializarlos, transferirlos a la sociedad.
- **Problema abordado.**

Este proyecto se desarrolló con el fin de encontrar una solución a la problemática de inseguridad en las vías y ciclo rutas que actualmente tiene la ciudad de Bogotá, Colombia. Gracias a la



acogida de los miles de usuarios que a diario transitan en su bicicleta por la ciudad y conocen lugares y rutas con problemas de inseguridad, pero no tienen la manera de alertar a otros usuarios para que tomen vías alternas o tengan más precaución al rodar por estos lugares. (Romero Guzman, Gonzalez Castro, & Jaime Fuentes, 2017)

- **Datos utilizados**
  - Información y alarmas enviadas por GPS por medio del uso de Arduino, shield ethernet, módulo sim 808 y GPS.
- **Herramientas de software.**
  - Base de datos de mysql.
  - Google maps.
  - Arduino y librerías: Las cuales son: dfrobot\_sim808 y la librería software serial.

## 2) Sector de la tecnología.

### a. Arquitectura de aplicaciones de software embebido en microcontroladores para tarjetas de captura de datos de la IOT.

- **Problema abordado.**

El problema abordado en el proyecto de investigación busca analizar una arquitectura de software embebido aplicada en microcontroladores que se consideran fundamentales a la hora de crear sistemas de automatización por esta razón el objetivo del proyecto de investigación se enfocó en crear una arquitectura que le permita a los desarrollados tener un lineamiento estandarizado de técnicas y herramientas para el uso de microcontroladores de tarjeta de captura de datos en sistemas IoT . (Mendoza Ibarra J. F., 2018)

- **Datos utilizados.**

- Resultados en proyectos utilizando IoT .
- **Herramientas de software.**
  - Tecnología lora.
- b. El internet de las cosas y sus riesgos para la privacidad.**

A continuación, se muestran los pasos que se tuvieron en cuenta para el desarrollo de este proyecto:

- Se realiza una breve introducción del internet de las cosas y como ha ido aumentando en los últimos años.
  - Se explica que es el internet de las cosas y todo lo que abarca.  
Explica las formas en cómo se comunica los dispositivos en el IoT .
  - Presenta diferentes e importantes empresas que apuestan y usan IoT .
  - Explica los riesgos del internet de las cosas.
- **Problema abordado.**
  - En el problema abordado en el proyecto de investigación se busca exponer y prevenir sobre las vulnerabilidades y riesgos que existen en el IoT debido al aumento destacable en su uso y el constante crecimiento de los sistemas y sus utilidades brindadas en diferentes sectores, permitiendo así automatizar muchas tareas que al día de hoy eran tediosas para el ser humano. (González Larin, 2017)
- **Datos utilizados.**
  - Datos de diferentes fuentes de investigación.
- **Herramientas de software.**
  - Navegadores web.
- c. Seguridad de la información en la internet de las cosas.**

A continuación, se muestran los pasos que se tuvieron en cuenta para el desarrollo de este proyecto:

- Se hace una breve introducción.
- En este punto se comenta la historia del IoT y algunas estadísticas.
- En este punto se comenta algunos antecedentes de violaciones de seguridad a dispositivos que se conectaban conectados a internet.
- **Problema abordado.**

El problema abordado en la investigación es enfocado en realizar un análisis a los sistemas de internet de las cosas, partiendo de que el aumento de estos proyectos va en constante crecimiento, por lo que es importante realizar un estudio que permita verificar la seguridad que existe en estos sistemas y como están siendo controlados. (Cañon Miranda, 2016)

- **Datos utilizados.**
  - Datos de diferentes fuentes de investigación.
- **Herramientas de software.**
  - Navegadores web.

#### **d. La importancia de la gestión de riesgos y seguridad en el internet de las cosas (IoT).**

- **Problema abordado.**

El problema que se busca solucionar en el proyecto de investigación se enfoca en un estudio que permita resaltar la importancia de la gestión de riesgos y la seguridad en el internet de las cosas ya que su constante crecimiento debe justificar los grandes riesgos de información que se presentan. (Molina García, 2019)

- **Datos utilizados.**

- Datos de diferentes fuentes de investigación.

- **Herramientas de software.**

- Navegadores web.

**e. Modelado, optimización y simulación de balanceo de flujo de información entre IoT y FOG Computing.**

- **Problema abordado.**

En el problema que se aborda en la investigación se busca disminuir la cantidad de información perdida entre IoT-FOG y minimizar la utilización de los enlaces. esto, considerando que pocos FOG deben atender a miles de IoT. para este primer estudio se supone: la latencia de todos los FOG es igual, los enlaces son directos entre IoT y FOG . el procesamiento necesario para un flujo de información es igual en todos los FOG y es determinado por el IoT. el tiempo de procesamiento de un flujo depende solo de su tamaño y de la capacidad de procesamiento del FOG. datos utilizados.

- **Datos utilizados.**

- Se utilizaron datos dados durante el curso académico.

- **Herramientas de software.**

- Fog computing.
- Gams.
- Cloud computing.

**3) Sector de la agricultura.**

**a. Un middleware para la gestión labores agrícolas en un ambiente de internet de las cosas.**

- **Problema abordado.**

En el proyecto de investigación se busca suplir la necesidad de monitorear las variables de nuestro entorno, aumenta el número de sensores, estándares, formatos de datos y heterogeneidad de la información disponible para analizar; haciendo que estos sistemas se hagan más complejo y, como consecuencia las soluciones genéricas para gestionarlos y controlarlos alcanzan sus límites. (Ordoñez Erazo, 2018)

- **Datos utilizados.**

- Datos obtenidos del entorno Físico.

- **Herramientas de software.**

- Middlewares es el software que proporciona servicios a las aplicaciones, fuera de lo que ofrece el sistema operativo.

**b. Diseño de un sistema IoT para la medición de temperatura, humedad y luz en huertas urbanas.**

A continuación, se muestran los pasos que se tuvieron en cuenta para el desarrollo de este proyecto:

- Investigar el funcionamiento y uso de los sensores.
  - Realizar un prototipo de medición en un pequeño jardín con los sensores.
  - Familiarizarse en temas del módulo de ethernet que se utilizara para la creación de la red.
  - Hacer la implementación de la primera pequeña red para verificar la conectividad del módulo ethernet.
  - Tener charlas o actividades con los administradores de las huertas para identificar las necesidades que presentan a nivel tecnológico.

- Realizar un sistema básico para evidenciar los datos tomados por los sensores.
- Realizar la conectividad al módulo ethernet con los datos tomados con los sensores o medidores implementados.
- **Problema abordado.**

En este proyecto, se desarrolla un sistema mediante el cual se monitoreará y visualizará la temperatura del ambiente, la intensidad de luz, la humedad relativa del ambiente y humedad del suelo, contenida en una huerta urbana ubicada en la localidad de Bosa, a través de una red IoT que permitirá la parametrización de variables y generación de alertas, esto con el fin de establecer el avance que brinda esta tecnología para el desarrollo de una huerta urbana. (Osma-Molina & Villamizar-Monar, 2020)

- **Datos utilizados.**
  - Variables del ambiente (temperatura, humedad e intensidad de luz).
- **Herramientas de software.**
  - Plataforma Ubidots: servicio en la nube
  - IDE desarrollo del microcontrolador Arduino.

#### 4) Sector desarrollo de software.

##### a. Arquitectura de aplicaciones de software embebido en microcontroladores para tarjetas de captura de datos de la IoT.

- **Problema abordado.**

En el problema abordado en el proyecto de investigación se propone un modelo arquitectónico de referencia, desde dispositivos físicos hasta el consumo y análisis de datos en las capas superiores (aplicación), es decir en las aplicaciones que usaran los datos, modelo mediante el cual a través de algunos requerimientos se puede obtener una arquitectura concreta.

Sin embargo, si bien los autores identifican a los dispositivos como sistemas embebidos que interactúan entre el mundo digital y el físico, conectando entidades físicas del mundo real a internet, el modelo no contempla ni define qué arquitectura o qué componentes debe tener el software de un sistema embebido para que pueda formar parte de un ecosistema del Internet de las cosas. (Mendoza Ibarra J. F., 2018)

- **Datos utilizados.**
  - Datos obtenidos del entorno físico.
- **Herramientas de software.**
  - Protocolo http.
  - Notación JSON.

## **5) Sector público.**

### **a. Desarrollo de una aplicación basada en la plataforma OpenMTC para la interconexión y administración de dispositivos funcionales empleando el concepto de internet de las cosas.**

A continuación, se muestran los pasos que se tuvieron en cuenta para el desarrollo de este proyecto:

- Se define, especifica y se explican las restricciones del problema.
- Se comienza el desarrollo del diseño de la arquitectura global del programa.
- Se realiza la recolección de la información para la arquitectura del proyecto, y se establecen las alternativas al diseño propuesto.
- En este punto se realiza la implementación se describe como se realiza la misma y se explican los resultados obtenidos.

- **Problema abordado.**

En el problema abordado se busca lograr una mejor integración de la tecnología con la vida cotidiana de las personas esto con el fin de poder conocer fácilmente el entorno en el que se encuentran y poder interactuar y poder interactuar con este, sumado a esto obtener todos los beneficios posibles que ofrece el IoT en las casas inteligentes como por ejemplo llevar un mejor control de lo que comemos, la cantidad de agua, luz, gas, gasolina, y otros recursos que utilizamos a diario.

- **Datos utilizados.**

- Documentación de la plataforma MTC.
- Documentación titulada “demonstration of openmtc – m2m solutions for smart cities and the internet of things”.
- Otra fuente utilizada es “m2m solutions for smart cities and the internet of things”.
- “Openmtc python sdk documentación”.

- **Herramientas de software.**

- Freedom board house.
- Aplicación móvil.

**b. Integración de internet de las cosas en sistema embebido system on chip, con aplicación a domótica.**

A continuación, se muestran los pasos que se tuvieron en cuenta para el desarrollo de este proyecto:



- Se hizo énfasis en cada uno de los elementos que hicieron parte del comportamiento del sistema y su estructura y se realizan pruebas de funcionamiento y depuración en cada etapa.
- Se inicia realizando un modelamiento básico del sistema, mediante el uso de algunos diagramas de UML, con el objeto de tener una guía a seguir.
- Se comienza con el Diagrama de Casos de Uso.

- **Problema abordado.**

El problema abordado en la investigación se enfoca en dar una solución a través de un sistema embebido de conexiones entre tecnología y objetos, procesos que hacen parte de la domótica, automatización a las casas.\_(Valiente Crisancho, 2017)

- **Datos utilizados.**
  - Investigaciones de sistemas embebidos ya creados.
- **Herramientas de software.**
  - Start UML.
  - Base de datos mysql.

## **6) Sector de la educación.**

### **a. Modelo de gestión estratégica para el desarrollo de un campus inteligente basado en conceptos de smart city en la universidad de Cartagena - campus piedra de Bolívar.**

#### **Metodología de tipo mixta.**

La metodología de desarrollo mixta requiere de una investigación descriptiva y documental para la indagación de información del área de estudio sobre conceptos que engloben

los temas de investigación, obteniendo el conocimiento a través de artículos, revistas, trabajos electrónicos y consultas en internet que posteriormente se analizó y con base en los objetivos planteados. (Ospino Pinedo, 2017)

Consiste en:

Etapa 1: Generación un libro blanco para identificar, profundizar y afianzar la filosofía, la tecnología y los servicios de un smart campus. inicialmente se realizó una revisión de la literatura y proyectos afines, con el fin de comprender los conceptos necesarios para la puesta en marcha del proyecto de smart campus.

Etapa 2: identificación de los conceptos de smart city que se puedan aplicar al campus universitario. se estableció cuáles son los principales criterios que definen una smart city y posteriormente se hizo una clasificación para determinar cuáles de esos criterios pueden ser aplicados dentro de un campus universitario.

Etapa 3: Realización de un estudio diagnóstico en cuanto a infraestructura, servicios, aplicaciones y usuarios del campus piedra de bolívar de la universidad de Cartagena con relación a los conceptos de smart city.

Etapa 4: Generar una propuesta de carácter integral y sostenible que permita definir un marco genérico de acción, en el que se amparen las soluciones ti actuales y futuras. posterior al estudio diagnóstico.

Etapa 5: Validación el modelo garantizando el carácter sostenible de la propuesta para el campus piedra de bolívar de la universidad de Cartagena. como etapa final se validó el modelo de gestión estratégica desarrollado en el objetivo general para garantizar un modelo eficiente y sostenible de gestión de la universidad de Cartagena, campus piedra de bolívar.

- **Problema abordado.**

En el problema abordado en el proyecto de investigación se planteó en la universidad de Cartagena plantear la automatización de varios procesos dentro del campus, aplicando conceptos de smart city en las tareas, funciones y utilidades dentro del campo de estudio de la universidad dándole uso a las herramientas de inteligencia artificial. (Maza Figueroa & Orozco Acevedo, 2017)

- **Datos utilizados.**

- Artículos.
- Revistas.
- Trabajos electrónicos y consultas en internet.

- **Herramientas de software.**

- “Smart cards” o tarjetas inteligentes, portal multi acceso (web, tv, internet móvil, canal telefónico, etc.)

## 7) Sector de la salud.

### a. **Arquitectura de un sistema de tele monitorización para hospitalización domiciliaria de adultos mayores apoyada en tecnologías de Internet de las cosas (IoT).**

Metodología aplicada con tres tipos de carácter:

- Investigativo

Consiste en realizar una investigación detallada que permite sustentar de donde se obtiene la información, tener fuentes validas de los datos que se van a utilizar para obtener un sistema valido.

- Analítico.

Toda la información que se obtiene desde todas las fuentes de investigación, debe pasar por una fase de análisis que consiste en evaluar cada dato, partiendo desde el grupo categoría donde se va alojar, y el resultado que se esperan de ellos. (Chagui Romero & Martinez Franco, Prototipo de sistema de monitoreo no invasivo de pacientes cronicos que padecen hipertension, asma y/o EPOC. [Trabajo de grado de ingenieria, Universidad de Cartagena], 2018)

Practico. después de obtener los datos y organizarlos se deben poner en prácticas los datos logrando así concluir el desarrollo del proyecto a través de pruebas funcionales.

- **Problema abordado.**

En el problema abordado se busca diseñar una arquitectura para la tele monitorización del adulto mayor en hospitalización domiciliaria utilizando tecnologías de internet de las cosas con el fin de llevar un seguimiento constante a los adultos mayores debido a que por su avanzada edad son propensos a sufrir de enfermedades que requieren de un cuidado especial y constante. (Segura Gazcón & Velásquez Hernández, 2017)

- **Datos utilizados.**

- Una encuesta realizada a 10 casas de familia donde se encontraban viviendo personas mayores.

- **Herramientas de software.**

- Aplicación móvil.
- Servidor de Google.

**b. Prototipo de sistema de monitoreo no invasivo de pacientes crónicos que padecen hipertensión, asma y / o epoc.**

### Metodología de desarrollo RUP:

La metodología de desarrollo RUP es un tipo de investigación aplicada dado que se utilizó un contexto real que se presenta entre las instituciones prestadoras de salud y el seguimiento de las variables de cambio de los pacientes con enfermedades crónicas, más específicamente, los que padecen hipertensión.

Consiste en:

#### 1. Tipo de investigación y enfoque.

En esta etapa se define el tema de investigación y el enfoque al cual se verá orientada la investigación.

#### 2. Población y muestra

En la etapa de población y muestra se define la cantidad de población a estudiar, y a partir de esto se obtiene una muestra real de los datos.

#### 3. Fuentes de información.

En la etapa de fuentes de información se define de donde se obtendrán los datos.

#### 4. Instrumentos de recolección de información.

Los instrumentos de recolección de información permiten obtener detalla la información y a partir de esta obtener resultados y porcentajes en coincidencias de información.

#### 5. Análisis e interpretación de resultados.

En la etapa de interpretación se le da un valor a los datos y a partir de los mismos se crea una modelo de predicción de resultados.

#### 6. Desarrollo generado por objetivos.

En la etapa final del proyecto se generan unos resultados que van definidos a partir de los objetivos que se plantean para realizar el proyecto.

- **Problema abordado.**

El presente proyecto fue realizado con el objetivo de desarrollar el prototipo de un sistema de monitoreo remoto para el seguimiento de la presión arterial y saturación de oxígeno en sangre de pacientes crónicos que padecen de hipertensión, asma y/o epoc en Colombia, haciendo uso de software libre y tecnologías de Internet de las cosas. (Chagui Romero, Prototipo de sistema de monitoreo no invasivo de pacientes crónicos que padecen hipertensión, asma y / o epoc. [Trabajo de grado de ingeniería, Universidad de Cartagena], 2018)

- **Datos utilizados.**

- Bases de datos indexadas:
- Dispositivos.
- Sensores.
- Funciones.

- **Herramientas de software.**

- Node js es un entorno en tiempo de ejecución multiplataforma, de código abierto.
- Programación JavaScript.
- Express JS framework de aplicaciones web minimalista y flexible.
- MongoDB y por ello se usó el ODM.

En los trabajos de grado estudiados donde se encuentra aplicada el área internet de las cosas IOT, se obtuvo que los sectores más representantes en las tesis de grado son el sector comercial, y el sector de la tecnología, para los cuales los datos que se utilizaron para las pruebas y el funcionamiento de los dispositivos automatizados son obtenidos en tiempo real, a través de sensores y dispositivos inteligentes, que permiten optimizar trabajos, producción, y eficacia en

los problemas que se enfocan a solucionar, a partir de unas metodologías que se basan en fases, que deben cumplirse una a una completamente para el desarrollo del proyecto de investigación.

#### **4.1.6. Algoritmos genéticos.**

El área de la inteligencia artificial conocida como algoritmos se ven inspirados en la selección natural que han logrado todo el proceso evolutivo de los vivos para resolver problemas de optimización, y aprendizaje en las máquinas. (Alvarez V., Hurtado G., & Trujillo M., 1995)

##### **1) Sector de la educación.**

###### **a. Generación automática de resúmenes de múltiples documentos.**

A continuación, se muestran los pasos que se tuvieron en cuenta para el desarrollo de este proyecto:

- Buscar la forma de hibridar los algoritmos base.

En esta fase busca como por medio de programación sincronizar el funcionamiento de los algoritmos en conjunto, que las funciones de uno complementen los resultados de las salidas del otro.

- Selección de la versión del algoritmo.

La selección de la versión del algoritmo es importante ya que la actualización de la función de cada uno va variando, por lo tanto, hay que escoger la que más se ajusta al sistema.

- Proceso de diseño de la función objetivo.

La función objetivo del proyecto es la etapa donde se define que tipo de funcionamiento tiene el proyecto y los resultados que se obtendrán.

- Proceso de configuración.

En el proceso de configuración se adaptan los sistemas a los algoritmos a través de funciones que permiten el uso adecuado de los mismos.

- **Afinación de los pesos asociados.**

La afinación de los pesos asociados en esta etapa se evalúa el funcionamiento de los algoritmos, y cómo funcionan en conjunto.

- **Problema abordado.**

Debido al crecimiento exponencial de documentos de textos digitales en internet y en la organizaciones cada día se incrementa la necesidad de que los sistemas de búsqueda sean más efectivos y ofrezcan a sus usuarios resúmenes de unos documentos de una misma temática, permitiéndoles poder decir más rápidamente cuales libros se adaptan y responden más a su necesidad de acceso a la información , Buscando de esta manera por medio de la técnica de Inteligencia artificial en esta caso algoritmos genéticos , se busca reducir el tiempo de las búsquedas y satisfacción de usuarios en el uso de las consultas en las web. (Cuellar Chacon, 2018)

- **Datos utilizados**
  - Documentos: duc2005. duc2006.
  - Encontramos 4 categorías que son:
    - Numero de tópicos.
    - Numero de documentos.
    - Fuente de datos.
    - Longitud del resumen.
- **Herramientas de software.**



- Visual estudio .net. entorno de programación de Microsoft que permite desarrollar aplicaciones para diferentes plataformas.
- Software apache splitta herramienta de código abierto.

## **2) Sector del transporte.**

### **a. Optimización en la distribución de rutas de colegios privados en Pereira con ayuda de algoritmos de inteligencia artificial.**

Los pasos que se tuvieron en cuenta para desarrollar este proyecto fueron:

- Población:  
Colegios privados de Pereira que presten el servicio de transporte a sus estudiantes.
- Unidad de análisis:  
Base de datos local, generada por el software, donde se tiene información sobre las pruebas realizadas en cada mapa. Mostrando el algoritmo utilizado, el id del mapa, el valor de la función objetivo encontrada, el tiempo de ejecución, la configuración aplicada al algoritmo, numero de rutas, capacidad total y porcentaje de ejecución del algoritmo (si se guardó para una prueba determinada).
- Muestra:  
Colegio Liceo Taller San Miguel, el colegio cuenta un total aproximado de 750 estudiantes, de los cuales 630 pertenecen a la sede campestre ubicada en el kilómetro 8 de la vía a Armenia, y 120 al jardín infantil que se encuentre en el barrio de Alamos.

- **Variables:**

Las variables que alimentaran el programa para encontrar una solución.

- **Problema abordado.**

En la actualidad los colegios manejan diferentes rutas de llegada a sus instalaciones educativas, muchas veces partiendo del uso de un sistema transporte del colegio, y otras veces conductores contratados por padres de familia para hacer seguro el traslado a clases de los estudiantes, por tal razón es importante destacar que muchos de los estudiantes viven en diferentes lados de la ciudad, y esto conlleva a que el recorrido muchas veces retrasen sus llegadas oportunas a clases, realiza un proyecto de análisis y diseño de un sistema que por medio de Inteligencia artificial y algoritmos que se basan en Algoritmos genéticos lograr crear una solución de las rutas más cercanas para el traslado de los estudiantes a sus colegios.\_(Pinea Cortes & Pinzon Rivera, 2015)

- **Datos utilizados.**

- Datos organizados.
- Numero de rutas y capacidad de cada una.
- Número de estudiantes y ubicación de cada uno.
- Ubicación del colegio.

- **Herramientas de software.**

- Google maps software con ubicación en tiempo real.
- Start UML software para la creación de diagramas de despliegues.
- Software bugs: para algoritmos genéticos.

### **3) Sector público.**

**a. Diseño de un sistema de soporte de decisiones para control de inundaciones en áreas urbanas mediante algoritmos genéticos NSGA-II. Evolutivos basado en preferencias y redes neuronales.**

- **Problema abordado.**

A través de los eventos de catástrofes naturales que se presenta, se plantea en la solución del problema abordado diseñar un sistema de soporte de decisiones para control de inundaciones en áreas urbanas. (Toledo Cubillos, 2016)

- **Datos utilizados.**

- Utiliza diagramas de Pareto con dos funciones objetivo:
- Costos de daños por inundación
- Costos de inversión.

- **Herramientas de software.**

- Aplicación desarrollada en Java denominada Optimizador de Sistemas de Drenaje (OSD).

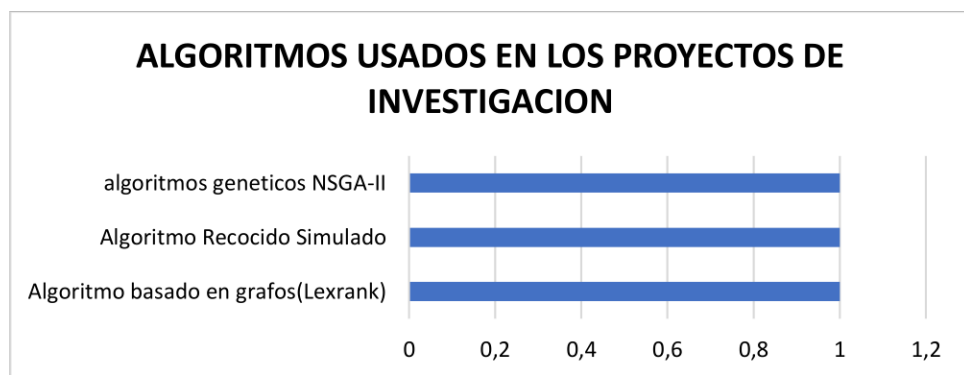


Figura 22. Grafica que representa el número de veces que se usaron determinadas metodologías en la rama de algoritmos genéticos.

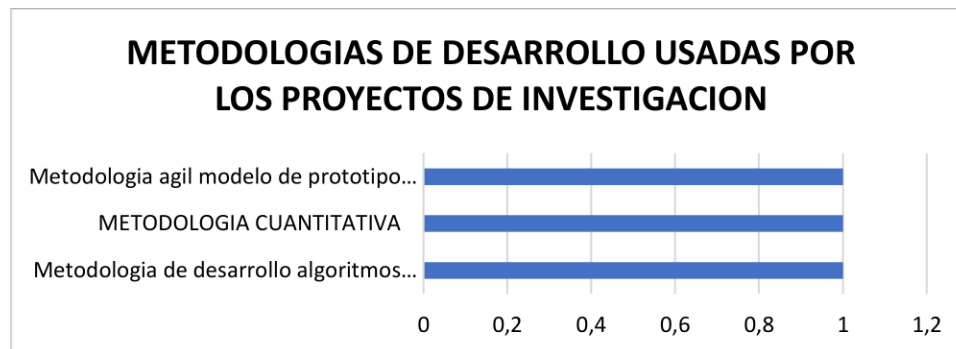


Figura 23. Grafica que representa el número de veces que se usaron determinados algoritmos en las tesis relacionadas con algoritmos genéticos.

En los trabajos de grado estudiados donde se encuentra aplicada el área de algoritmos genéticos, se obtuvo que los sectores más representantes en las tesis de grado son el sector de la educación, sector del transporte y el sector público, para los cuales los datos que se utilizaron fueron predefinidos según la funcionalidad y el diseño del proyecto de investigación por los desarrolladores de las tesis, las metodologías implementadas fueron, cuantitativa, ágil, y la metodología de desarrollo de algoritmos híbridos, para estos trabajos de grado los algoritmos que fueron utilizados para la ejecución de los programas son, NGS – II, algoritmo basado en grafos (lexrank), algoritmo recocido simulado.

#### **4.1.7. Robótica.**

El área de la inteligencia artificial conocida como robótica, basada en los avances en mecánica, electrónica e informática. El desarrollo de robots en el terreno industrial, bélico y aeroespacial durante el siglo es una rama que permite el avance y mejora de muchos sectores al construir un modelo de proyectos funcionales que logran optimizar trabajos en diferentes sectores. (Sánchez, y otros, 2007)

##### **1) Sector del comercio.**

###### **a. ECIBOT: Agente inteligente para la interacción en lenguaje natural.**

- **Problema abordado.**

En el problema abordado en el proyecto de investigación se busca mostrar al lector de forma detallada cada uno de los aspectos del proyecto ecibot, un agente conversacional con el que se quiere dar soporte al centro de atención telefónica y crear un canal adicional de comunicación donde la comunidad (estudiantes, administrativos e interesados en los servicios que ofrece la escuela) puedan resolver sus inquietudes de forma rápida y concisa. (Sánchez Prieto & Alvarado Pérez, s.f.)

- **Datos utilizados.**

- Información de la Escuela Colombiana de Ingeniería Julio Garavito.

- **Herramientas de software.**

- Bot Framework frameworks para el uso y automatización de chatbots.

## **2) Sector de la educación.**

### **a. Propuesta pedagógica para incentivar el pensamiento científico y creativo a través de la robótica en niños de 3 a 6 años del colegio Carlos Castro Saavedra.**

Metodología cualitativa-descriptiva.

La metodología cualitativa y descriptiva consiste en la extracción de grandes cantidades de datos que pasan por una fase de descripción esto se genera a partir de categorías y grupos establecidos según la necesidad y el uso adecuada a la información, para luego así pasar por una fase de análisis de cualidades que obtienen en detalle el valor de los datos y de esta manera obtener los resultados esperados en los proyectos de investigación. (Zuluaga Agudelo, 2019)

Consiste en:

1. Método de investigación.

Por medio de la observación se determina la importancia de la implementación de los planes de aula para potenciar el pensamiento científico y creativo en niños y niñas de la primera infancia. dicha observación se hace de manera participativa por medio de la planeación de clases didácticas de robótica básica (juegos, videos, fotos, preguntas y respuestas, manualidades).

## 2. Técnicas para la recolección de la información.

En esta fase se definen las técnicas y métodos con los que se recogerá la información y se analizar.

## 3. Fuentes primarias.

Reuniones: se dialoga con los directivos, docentes y estudiantes de la institución Carlos castro Saavedra antes de la ejecución de las actividades. Fuentes secundarias: se toma como referencia el proyecto de investigación - definición de un ambiente lúdico y creativo de aprendizaje fundamentado en la robótica educativa.

- **Problema abordado.**

En la mayoría de las instituciones educativas del país se observa que la enseñanza de la tecnología va dirigida principalmente a los alumnos de grados superiores pasando por alto que los pequeños son más receptivos, adaptándose mejor y más fácilmente, La robótica como ciencia participe y transversalizada por las tic tiene como ventaja en que, al ser aprovechada en los niños desde los primeros años de su etapa escolar, permiten afrontar la interacción con la tecnología como algo natural y que forma parte de su rutina escolar, por lo tanto el proyecto abarca una propuesta pedagógica que parte a través de incentivar a través del pensamiento científico y creativo acceder a conocimiento acerca de la Robótica. (Zuluaga Agudelo, 2019)

- **Datos utilizados**
  - Identificar temas.
  - Delimitar la temática.
  - Estandarizar y documentar los procesos.
  
- **Herramientas de software.**
  - No se utilizaron.

#### **4.1.8. Sistemas expertos.**

Son programas de computación que se derivan de una rama de la investigación informática llamada Inteligencia Artificial. El objetivo científico de la inteligencia artificial es entender la inteligencia. Está referida a los conceptos y a los métodos de inferencia simbólica, o de razonamiento por computadora, y cómo el conocimiento usado para hacer esas inferencias será representado dentro de la máquina. (Pignani, s.f.)

##### **1) Sector político.**

##### **a. Desarrollo y aplicación de una herramienta de extracción y almacenamiento de datos de Twitter a un contexto social de violencia política.**

Metodología XP.

Es una metodología de desarrollo de software planteada por Kent Beck, Este tipo de programación se diferencia de las metodologías comunes principalmente en que propone como factor primario la adaptabilidad antes que la previsibilidad. Esta propuesta metodológica posee un conjunto de valores que la identifican además de un conjunto de etapas definidas. (Barriga Mariño, 2017)

Consiste en:

### 1. Planificación del proyecto

Esta etapa comprende la definición de historias de usuario con el cliente, las historias tienen una gran similitud a los casos de uso relacionándose con el levantamiento de requerimientos, pero con algunas diferencias, estas historias son líneas escritas por el cliente en lenguaje no técnico y son usadas para la estimación de tiempos, también se definen las iteraciones y tiempos en cuanto a la velocidad del proyecto.

### 2. Diseño.

Consta del planteamiento de diseños simples y sencillos, pero de calidad, también promueve la utilización de glosarios de términos para facilitar el entendimiento del diseño y abrir paso a la reutilización de código por ejemplo en posteriores actualizaciones, adicionalmente, plantea el tratamiento de riesgo para su reducción de forma rápida, el descarte de la funcionalidad extra y la refactorización de código sin alterar su funcionalidad con el fin de optimizar el funcionamiento.

### 3. Codificación.

Esta fase comprende la programación o codificación de las historias o requerimientos, esta programación debe hacerse controlada por un estándar definido esto mantiene el código consistente y facilita su comprensión y escalabilidad.

### 4. Pruebas.

Como fase final, implica el testeado del código para verificar su funcionalidad, sometiendo a estas pruebas distintas clases del sistema omitiendo los métodos más triviales, estos test no deben tener ninguna dependencia del código que en un futuro evaluará.

- **Problema abordado.**

Como producto de la investigación de los antecedentes del proyecto, se identifica que la mayoría de herramientas de extracción y almacenamiento de datos de redes sociales



pertenecientes al mercado actual, se centran en un contexto específico, generalmente enfocado al marketing y los negocios, luego de identificar estos factores y si se pretendiera realizar un tipo de análisis, estadístico, de varianza con enfoque al análisis de sentimientos y con carácter predictivo, que involucre los datos generados por diferentes cuentas en una red social tan influyente como lo es Twitter, teniendo en cuenta la necesidad de que la obtención de estos datos pueda generarse de una forma sencilla e implique bajos costos monetarios, además de tomar como contexto para la aplicación funcional de la herramienta la problemática social referente a la violencia política que se genera a través de este medio de comunicación. (Barriga Mariño, 2017)

- **Datos utilizados.**

- Tweets de la red social Twitter.

- **Herramientas de software.**

- Plataforma GitHub.
- UML: Diagramas de componentes diagrama de despliegues.
- Liferay plataforma gestora de contenidos.

## **2) Sector de la educación.**

### **a. Prototipo de tutor inteligente para el aprendizaje de la programación de computadores.**

Metodología de desarrollo, Reconocimiento de patrones.

La técnica de reconocimiento de patrones se define como el campo o área en la cual se obtienen razones del conocimiento de los datos en grandes volúmenes a través del uso de algoritmos y así obtener el método en el cual se logre definir las categorías agrupadas para utilizar la información de una manera más precisa. (Quiroga Saavedra, 2016)

#### **1. Reconocimiento de patrones**

El reconocimiento de patrones es un campo que tiene que ver con el descubrimiento de regularidades en la información a través del uso de algoritmos y valerse de estas regularidades para tomar acciones como clasificar la información en categorías.

## 2. Algoritmos de clasificación.

Cuando el conjunto de entrenamiento contiene vectores de entrada con su correspondiente vector objetivo se habla de una rama del reconocimiento de patrones llamada aprendizaje supervisado y en particular, cuando el objetivo del problema es asignar cada vector de entrada a una de un conjunto discreto de categorías se habla de problemas de clasificación.

### 2.1 K Nearest Neighbors

En la fase de entrenamiento de este método se almacenan todos los ejemplos de prueba, con su etiqueta y luego, cuando se va a hacer la predicción se calcula la distancia del elemento de prueba a cada ejemplo de entrenamiento.

### 2.2 Random forest.

Esté método se basa en arboles de decisión, en los cuales se crean árboles para predecir el valor de una variable con respecto a variables de entrada y donde cada nodo del árbol corresponde con una variable y el valor de la variable de destino esa representado por el camino desde la raíz a la hoja.

## 3. Medición de desempeño

### 3.1 Matriz de confusión.

La matriz de confusión contiene información sobre la clasificación real y la realizada por un sistema de predicción.

### 3.2 Precisión y recall.

Precisión es la fracción de casos obtenidos que son relevantes, mientras que recall hace referencia al porcentaje de casos relevantes que son obtenidos.

- **Problema abordado.**

Se busca reducir el déficit de profesionales en el campo de la programación y mejorar la calidad profesional de los mismos desarrollando un tutor inteligente para apoyar el proceso de aprendizaje de programación de computadores usando técnicas de Aprendizaje de Maquina.

(Quiroga Saavedra, 2016)

- **Datos utilizados.**

- Datos relevantes por estudiante: Género, edad, estrato socioeconómico, composición familiar, colegio de grado, estudiantes trabajando, estudiantes con hijos, desempeño académico.

- **Herramientas de software.**

- No registra.

### **3) Sector de la tecnología.**

#### **a. Diseño e implementación de agentes para la plataforma juego del agua.**

A continuación, se muestran los pasos que se tuvieron en cuenta para el desarrollo de este proyecto:

##### **1. Apropiación de la tecnología.**

Solicitud y acceso al código de la plataforma actual, aprendizaje de tecnologías de desarrollo, instalación de herramientas de desarrollo, lectura de código y búsqueda de interfaces de entrada para componente de software a implementar.

##### **2. Definición del agente.**

Definición de entras del agente y de las reglas de comportamiento.

### 3. Implementación.

En esta etapa de desarrollo se toman todas las definiciones realizadas en la etapa anterior (diseño de agente) y se implementan dentro del componente software.

### 4. Integración de la documentación

Integración con la plataforma web del juego del agua: en esta tarea se realizan las reuniones para la integración del código y la verificación del funcionamiento en la plataforma web. Documentación: en esta tarea se termina la documentación del código generado, la creación del documento de descripción de la arquitectura y el manual de usuario.

### 5. Validación

Validación de todas las etapas anteriores, desde el funcionamiento hasta la recolección de datos.

#### 5.1 Pruebas de funcionamiento.

Esta tarea incluye la creación de diferentes configuraciones (sesiones) para validar el agente implementado, la identificación de errores y las oportunidades de mejora.

#### 5.2 Pruebas de motor de reglas.

El funcionamiento de esta tarea se separa, porque en estas pruebas se pretende comprobar el funcionamiento de cada regla.

- **Problema abordado.**

El problema abordado se enfoca en crear un modelo de adaptación al juego del agua, donde a partir de 4 agentes ya declarados, el administrador de la plataforma por medio del razonamiento deductivo permita cambiar el estado de dificultad de las áreas, que el contenido de preguntas que se le realizan a cada jugador vaya en acorde al grado de dificultad en el que se

encuentra, todo esto a partir de los sistemas expertos y agentes que se identificaron anteriormente. (Tavera Sanchez, 2015)

- **Datos utilizados.**
  - Base de datos: con todos los datos de juegos antes realizados.
- **Herramientas de software.**
  - Apache Tomcat.
  - Mysql motor de bases de datos.
  - Eclipse IDE de programación java.

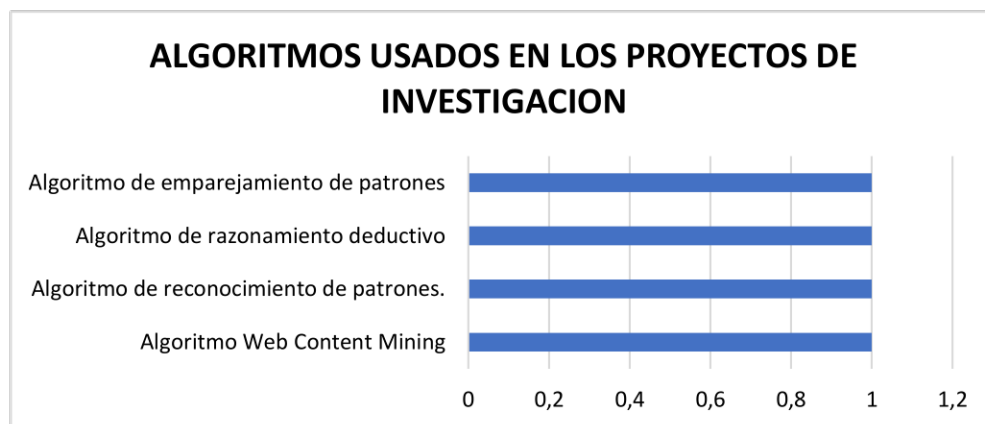


Figura 24. Grafica que representa el número de veces que se usaron determinadas metodologías dentro de la rama de sistemas expertos.

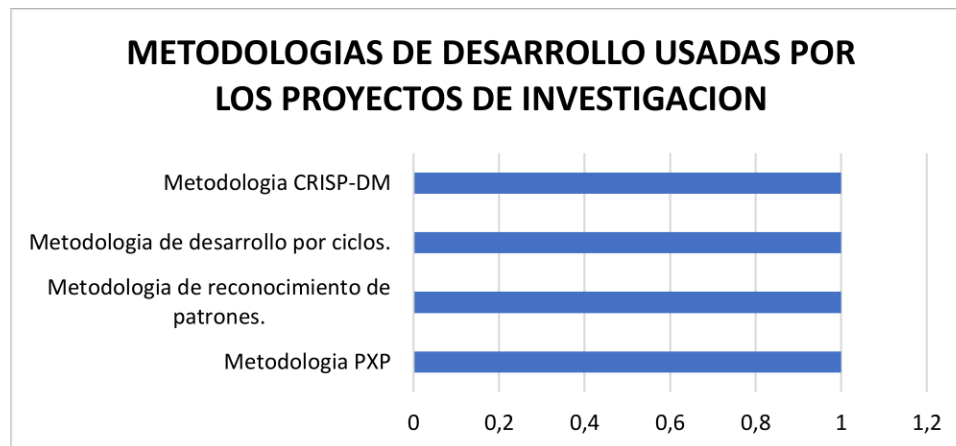


Figura 25. Grafica que representa el número de veces que se usaron determinados algoritmos en las tesis relacionadas con sistemas expertos.

En los trabajos de grado estudiados donde se encuentra aplicada el área de la robótica, se obtuvo que los sectores más representantes en las tesis de grado son el sector del comercio y sector de la educación, los algoritmos utilizados en los proyectos fueron, algoritmos aplicados a chabots, y de precisión, entre las metodologías que representan estas investigaciones está la cualitativa, descriptiva, y los datos que se utilizan para la validación del correcto funcionamiento de los sistemas inteligentes, son definidos por los investigadores, basándose en modelos de la vida real, y adaptando esta información a sistemas programables.

#### **4.1.9. Redes neuronales artificiales.**

##### **1) Sector de la salud.**

##### **a. Un modelo de credit scoring utilizando inteligencia artificial para el fondo de empleados de salud en Risaralda – fesar.**

A continuación, se muestran los pasos que se tuvieron en cuenta para el desarrollo de este proyecto:

##### **1. Hipótesis.**

De lo expuesto previamente se deriva la siguiente hipótesis: La utilización de un modelo computacional basado en inteligencia artificial para credit scoring durante el estudio de un crédito en el Fondo de empleados de Salud en Risaralda permitirá al analista conocer la predicción del riesgo de incurrir en mora para un crédito solicitado mejorando la cartera y reduciendo sus costos administrativos.

## 2. Variables.

Cuando un asociado solicita un crédito en fesar debe diligenciar todos los datos requeridos en el respectivo formato denominado “formato de solicitud de crédito” (ver anexo b), varios de ellos son transferidos a otro formato denominado “formato estudio de crédito” (ver anexo c), en el cual se recopilan las variables empleadas por el comité de crédito para realizar el respectivo análisis de la solicitud y decidir si es aprobado o no.

## 3. Instrumentos.

La información fue recolectada directamente desde la base de datos de fesar, son datos reales y no se realizó ningún tipo de ajuste o cambio, así como tampoco se afectó la protección de los datos personales de los asociados.

- **Problema abordado.**

Los fondos de empleados invierten parte de su presupuesto en la sistematización de los procesos contables y crediticios de manera que logran optimizar en cierto grado la operación diaria en el análisis de las solicitudes de crédito. Dicha sistematización no contempla el análisis del riesgo en la colocación de un crédito empleando métodos o modelos computacionales modernos, A partir de esto aplicando inteligencia artificial se crea un modelo basado en modelo scoring, y a través de estadísticas, matemáticas econométricas e inteligencia artificial, lograr

obtener un resultado efectivo de los dineros guardados en el fondo de empleados individualmente, a cada cartera de asociado. (Vargas Orozco, 2015)

- **Datos utilizados.**
  - Clasificación de categorías denominadas:
  - 30 días de mora respectivamente.
  - 60 días de mora respectivamente.
  - 90 días de mora respectivamente.
  - 120 días de mora respectivamente.
- **Herramientas de software.**
  - Excel software de hojas de cálculo para operaciones matemáticas.
  - RapidMinner herramienta de consultas.
  - Sql Server Sprees lenguaje de consultas.

**b. Implementación de un sistema de captura de gestos usando un leap motion y redes neuronales para su clasificación.**

Metodología ágil Kanban.

La metodología es adaptada para el desarrollo de software, con el propósito de optimizar la secuencia de trabajo, en la cual se representa visualmente las tareas a desarrollar y su estado, esta es la más apropiada para el desarrollo de este proyecto, porque permite ajustar la cantidad de trabajo a la capacidad de ejecución, y así tener un enfoque claro del estado de cada una de las tareas. (Pachon Espinel, 2018)

Consiste en:

1. Recolección de información.

Revisión de la literatura, para establecer herramientas informáticas a emplear.



## 2. Identificación de requerimientos.

Transcripción de entrevista, donde se identifican las especificaciones que necesitaría el prototipo.

## 3. Diseño de software.

Estandarización de residuos a clasificar.

## 4. Diseño base de datos.

Diagrama de clases.

## 5. Flujo de navegación

Determinar la navegabilidad.

## 6. Desarrollo.

Selección de algoritmos.

## 7. Pruebas y ajustes

Pruebas de funcionalidad.

## 8. Construcción de manuales técnico y de usuario.

Captura de video del funcionamiento del aplicativo.

## 9. Entrega de proyecto finalizado.

Revisión del director del proyecto.

- **Problema abordado.**

Se creo un prototipo funcional el cual permite por medio de una interfaz, realizar la captura de los datos para el entrenamiento del sistema de una manera sencilla ofreciendo libertad al usuario de crear gestos predeterminados. Durante el proceso se evaluaron las capacidades del Leap motion para reconocer las falanges de las manos en diferentes ocasiones y la eficiencia de

las técnicas clasificación del aprendizaje de máquinas usadas para reconocer gestos del alfabeto dactilológico de la lengua de señas colombiano. (Muñoz Garcia, 2017)

- **Datos utilizados.**
  - Datos a 00evaluar.
  - Tiempo de Entrenamiento.
  - Error aproximado.
  - Cantidad de Datos.
- **Herramientas de software.**
  - Angular.js framework para desarrollo en la web.
  - Base de datos: MongoDB.

### **c. Red neuronal para diagnóstico de hipertensión arterial.**

A continuación, se muestran los pasos que se tuvieron en cuenta para el desarrollo de este proyecto:

#### 1. Selección de datos en la web.

En esta etapa se eligen los sistemas de bases de datos, o contenidos de donde se extraerá la información detallada.

#### 2. Clasificación de datos.

La clasificación de los datos se realiza para permitir que cada dato obtenga el valor oportuno asignado por variables.

#### 3. Resultados de la recopilación de datos.

Los resultados obtenidos a través del análisis del gran volumen de información sirven para crear el modelo de predicción.

- **Problema abordado.**

Según la organización mundial de la salud la hipertensión o presión arterial alta es un trastorno en el que los vasos sanguíneos tienen una tensión habitual alta donde se pueden dañar, Por lo tanto, Por medio de este proyecto de investigación Se busca predecir con una alta probabilidad la enfermedad de la hipertensión evitando hacer pruebas diagnósticas utilizando técnicas de inteligencia artificial primordialmente las redes neuronales basadas en arboles de decisión, y aprendizaje supervisado. (García Montero, 2018)

- **Datos utilizados.**

- Se utilizan datos relacionados con la enfermedad almacenados en center for disease contro an prevention e información relacionados con la presión arterial, sus factores de riesgo y condiciones relacionadas donde los usuarios pueden ver mapas de enfermedades cardiacas y accidentes cerebrovasculares a nivel del condado y sus factores de riesgo por grupo racial/étnico, junto con mapas de condiciones ambientales sociales y servicios de salud, para todo estados unidos o para un estado o territorio elegido.

- **Herramientas de software.**

- Node js es un entorno en tiempo de ejecución multiplataforma, de código abierto.
- Visual Studio Code Insiders IDE de desarrollo de programación.
- Weka plataforma de minería de datos.
- RapidMinner herramienta de consultas.

**d. Diseño de un clasificador de señales EEG provenientes de una interfaz BCI utilizando redes neuronales para controlar un robot hexápodo.**

A continuación, se muestran los pasos que se tuvieron en cuenta para el desarrollo de este proyecto:

1. Adquisición de la señal.

En esta etapa se realiza el registro de todas las señales que tiene el cerebro por medio de dispositivos especializados.

2. Entrenamiento de las personas para el control de los estados mentales.

Se enseñan a las personas las señales provenientes del dispositivo Emotiv de forma continua.

3. Reconocimiento de patrones.

En esta etapa se utiliza el software MATLAB se extraen las propiedades más relevantes, eliminando las posibles características que afecten el clasificador.

4. Diseño y fabricación del robot hexápodo.

La fabricación del robot hexápodo fue necesario establecer el diseño o modelo más adecuado al entorno e interacción con el paciente.

5. Sistema de interacción.

Se utilizó una cámara web como dispositivo externo de acompañamiento al robot hexápodo.

6. Pruebas de paciente con condición de cuadriplejía.

Se realizaron todas las pruebas del funcionamiento efectivo de la interfaz cerebro computador logrando así controlar los movimientos del robot.

- **Problema abordado.**

El problema abordado en el proyecto de investigación busca mejorar calidad de vida de las personas que se encuentran en estado de parálisis parapléjico de completamente todos los

movimientos del cuerpo aplicando técnicas de la inteligencia artificial para mejorar la salud.

(Escobar Ortiz, 2018)

- **Datos utilizados.**
  - Tarjeta de adquisición de datos llamada beaglebone back.
- **Herramientas de software.**
  - Matlab software para cálculos matemáticos.

## 2) Sector del comercio.

### a) Sistema inteligente para apoyar el proceso de microinyección a través de técnicas artificiales inteligentes y modelo cae integración.

A continuación, se muestran los pasos que se tuvieron en cuenta para el desarrollo de este proyecto:

#### 1. Diseño y análisis geométrico de las micro partes.

En este campo la geometría el análisis se llevará a cabo mediante modelado cad, que proporcionará las principales variables geométricas que afectan la calidad de la pieza.

#### 2. Análisis de procesos por simulación cae.

Simulaciones de la inyección Se realizará el proceso para analizar el comportamiento del material al transportar fuera del proceso

#### 3. Creación de base de datos.

Basada en los análisis geométricos y de proceso proporcionados. por los simuladores CAD y CAE, se establecerá una base de datos donde se registrarán los resultados y variaciones obtenidos

#### 4. Diseño de sistemas predictivos basados en inteligencia artificial.

Esta etapa busca para integrar todos los resultados obtenidos y almacenados en la base de datos con un sistema experto en inteligencia, que es capaz de determinar el más apropiado variables para obtener micro partes con índices de alta calidad.

- **Problema abordado.**

En la actualidad, la necesidad de producir micro partes de plástico ha aumentado considerablemente debido al progreso que han hecho los dispositivos tecnológicos y al deseo de hacerlos más pequeño y con mayor número de funcionalidades. Por lo tanto, el proyecto de investigación está enfocado en crear un sistema inteligente que permita crear productos a pequeña escala aplicando un modelo de integración cae. (Barriga Mariño, 2017)

- **Datos utilizados**

- Variables:
  - Temperatura,
  - Presión.
  - Geometría.
  - Vértices.
  - Tiempo de fabricación.
  - Flujo de entrada.
  - Fuerza de cierre.

- **Herramientas de software.**

- Cae Ingeniería asistida por ordenador.

**b. Reconocimiento y conteo de monedas colombianas mediante procesamiento digital de imágenes y redes neuronales convolucionales.**

A continuación, se muestran los pasos que se tuvieron en cuenta para el desarrollo de este proyecto:

1. Categorizar monedas.

En esta etapa se asigna un código de identificación única a cada moneda que permita identificarla en los procesos.

2. Entrada de Monedas.

A través de las características ya organizadas se aplican las entradas de las monedas, de la manera en que serán ingresadas y la cantidad mínima.

3. Clasificación de monedas.

Las categorías de las monedas son definidas previamente a partir del valor y de más características.

4. Resultados de reconocimiento.

El resultado del reconocimiento de cada moneda se genera a partir del código unitario que se le asigna independiente a cada una de ellas.

5. Sumatorias

Las operaciones serían los resultados del funcionamiento del sistema.

- **Problema abordado.**

Las pequeñas y medianas empresas también deben ir enfrentados a los cambios que se presentan frecuentemente en los sistemas de conteo de dinero y demás áreas, por lo tanto, el problema abordado fue crear un sistema que por medio de técnicas de inteligencia artificial se logre realizar el conteo de monedas es las pymes. (Carrillo Lindado & Llanos Terán, 2017)

- **Datos utilizados.**

- Monedas de diferentes valores, clasificadas por sus atributos.

- **Herramientas de software.**

- Cámaras de procesamiento digital de Imágenes

**c. Sistema para el reclutamiento y selección de personal.**

Metodología de desarrollo scrum e incremental.

1. Recolección de información.

En esta etapa se llevará a cabo es estudio de mercado y la recolección de la información necesaria para tener la base del proyecto

2. Análisis de la información.

Una vez recolectada la información se pasará a realizar el respectivo análisis de los datos obtenidos.

3. Análisis de viabilidad.

Una vez concluido el análisis, se procederá a tener una reunión en la cual se plantearán los posibles riesgos que pueda presentar el proyecto.

4. Diseñar el diagrama relacional de la base de datos

En esta etapa se pretende realizar el diagrama de base de datos.

5. Realización de los sketches.

Los sketches son una visión gráfica y primitiva de cómo será la solución.

6. Realización de los wireframes.

Los wireframes son una visión más estructurada de los sketches, son las guías visuales que representarán el esqueleto de la web, contando además con el flujo de información (entradas) y los procesos que serán llevados a cabo por el aplicativo.

7. Creación de los mockups.

Los mockups también son representaciones graficas del aplicativo



- **Problema abordado.**

El problema abordado trata sobre un sistema de reclutamiento y selección de personas, el cual tiene como principal función servir como apoyo al proceso de contratación de una empresa, optimizando los tiempos que conlleva el mismo y pudiendo generar beneficios para la organización. Contando además con un diseño agradable e intuitivo, tanto para el aspirante como para el reclutador y con una infraestructura escalable y de óptimo rendimiento. (Viviescas Sandoval., 2019)

- **Datos utilizados.**

- Documentación en base de datos.
- Variables Tiempos, costos, precios, satisfacción del usuario.

- **Herramientas de software.**

- No registra.

### **3) Sector de la educación.**

**a. Modelo predictivo para la determinación del nivel de riesgo de deserción estudiantil en tecnología en electrónica de La Universidad distrital empleando redes neuronales artificiales.**

- **Problema abordado.**

La deserción estudiantil universitaria sigue siendo una de las mayores problemáticas en la Universidad distrital, debido a que generan impactos negativos en el ámbito social y económico. Los recientes estudios de la universidad distrital sobre retención de estudiantes muestran que la tasa de graduación en promedio fue de 42% y la tasa de deserción estudiantil fue del 58% para el periodo 1992-2010, por lo tanto, se planteó un modelo predictivo para la determinación del nivel

de riesgo estudiantil en tecnología electrónica en la universidad distrital empleando redes neuronales. (Tello Castañeda, 2016)

- **Datos utilizados.**
  - Base de Datos de la universidad Distrital.
  - Información del estudiante, calificaciones, promedios.
- **Herramientas de software.**
  - Bases de datos: Sistemas de información.

**b. Plataforma adaptativa de refuerzo de suma utilizando redes neurales artificiales y sistema de reputación bayesiana.**

A continuación, se muestran los pasos que se tuvieron en cuenta para el desarrollo de este proyecto:

Adaptación de plataforma de juego.

1. El jugador ingresa (Unity).

Acceso del jugador a la plataforma del juego.

2. Activa la red neuronal.

En esta etapa se activa la red neuronal para identificar el estado del jugador.

3. Activación del modelo de reputación bayesiano.

Los siguientes niveles del juego se ven activados por el algoritmo de reputación que automatiza las preguntas a continuar.

4. Base de datos.

Se organizan las preguntas en las bases de datos.

5. Procesamiento de Plataforma web.

El funcionamiento efectivo de la plataforma web permite ver cómo están interactuando los datos.

#### 6. Responsable de calificaciones

Las calificaciones son obtenidas después de evaluar cada pregunta y archivadas en la base de datos.

- **Problema abordado.**

El problema abordado en el proyecto de investigación es el de crear una plataforma que se adapte al nivel de capacidad de solucionar ejercicios matemáticos en los estudiantes de 1 año escolar según su avance aplicando redes neuronales como solución tecnológica. (Iglesias Martinez & Ilias del Rio, 2019)

- **Datos utilizados.**

- Estudiantes utilizando el juego, A partir de sus capacidades se generan la activación de los Modelos de dificultad.

- **Herramientas de software.**

- Sistema de reputación bayesiano (brs).
- Firebase plataforma móvil creada por Google, cuya principal función es desarrollar y facilitar la creación de apps.
- Angular framework para desarrollo en la web.
- Node .js es un entorno JavaScript que nos permite ejecutar en el servidor.

### **c. Sistema de reconocimiento óptico para mapeo conversión braille- español en el INCI.**

Metodología Xtreme Programming (XP).

Es una metodología que identifica de manera formal y operacional el desarrollo de cada una de las etapas de análisis de información, marcando así una diferencia entre los diferentes niveles lógicos de acción que se obtienen, a partir de los cuales se va construir una arquitectura ingenieril de una solución efectiva para darle solución a los problemas presentados y obtener los resultados esperado. (Avellaneda Díaz, 2015)

#### 1. Planificación.

Permite cualificar la arquitectura del negocio, para modelar prospectivamente las preguntas pertinentes a la formalización técnica ingenieril y económica del proyecto, definiendo prioridades y valorando acciones sustantivas de trabajo.

#### 2. Diseño.

Categoriza como entregable la llamada maqueta de configuración de la solución, proyectando con calidad las interfaces, los modelos y estructuras de navegación e interacción y definiendo las clases, modelos de datos y arquitectura semántica a nivel modular con la que se desarrollaran los aplicativos.

#### 3. Desarrollo.

Define las categorías de refactorización de los aplicativos, junto con el patronato de calidad y el proceso de socialización con el usuario final, identificando, las características de los entregables que cada grupo del proyecto general.

#### 4. Pruebas.

En este escenario xp, define, normatiza y regula los procedimientos y principios funcionales que regularan y aseguraran la calidad del grupo de arquitectura de software,

estableciendo para ello las fases convencionales relacionadas en ingeniería de software como pruebas de caja blanca y caja negra.

- **Problema abordado.**

El problema abordado en la investigación es aplicar una solución de software que reconozca y convierta caracteres braille a textos en español que sean editables que permitan administrar desde la aplicación los resultados, esto se realiza a partir del análisis de imágenes que son obtenidas del escaneo de documentos que se encuentran en el sistema braille, sirviendo como una herramienta de apoyo a los estudiantes que padecen de discapacidades visuales, y logren tener una comunicación efectiva con sus docentes. (Avellaneda Díaz, 2015)

- **Datos utilizados**

- Imágenes en alfabeto braille con extensiones tiff, png y jpg.

- **Herramientas de software.**

- Entorno de desarrollo netbeans (version 7.4).
- BraullieApp app que contiene las letras en formato braile.

#### **4)Sector de las Telecomunicaciones.**

##### **a. Modelo inteligente de decisión de espectro que mejora el desempeño en redes de radio cognitiva.**

A continuación, se muestran los pasos que se tuvieron en cuenta para el desarrollo de este proyecto:

1) Modelo inteligente de decisión de espectro propuesto. En este punto se realiza el procesamiento de los datos obtenidos en la medición.

2) En este punto se analizan los datos de los tipos de señales que se obtuvieron con un analizador de espectro en la facultad de ingeniería de la Universidad distrital Francisco José de Caldas.

3) En este punto se desarrolla el procesamiento de los datos obtenidos en la medición, se describe el proceso de normalización y luego se procede con el cálculo de los tres parámetros requeridos como entradas al algoritmo GRA.

4) En este punto se diseña la etapa de clasificación por medio del GRA que es una técnica de ponderación de atributos efectiva de MADM que analiza la mejor alternativa con el mayor valor del coeficiente de relación gris.

5) Diseño de la etapa de predicción, una vez se termina la fase anterior se inicia un proceso de predicción basado en las estadísticas tomadas en la medición para bajo tráfico y alto tráfico. Para la su esta predicción le permitirá identificar instantes de tiempo en los que puede existir una llegada del pu al canal y pueda tomar decisiones eficaces.

6) En este punto se define el definitivo modelo inteligente de decisión de espectro.

7) En este punto se realiza el procedimiento de evaluación del desempeño del modelo propuesto analizando por partes los algoritmos de clasificación y predicción.

8) Se crea un simulador que pueda reflejar mediante un experimento el comportamiento de tráfico real de una red inalámbrica de tal manera que se pueda ejecutar el modelo diseñado. En este simulador el SU puede realizar su proceso de ejecución.

- **Problema abordado.**

Se presenta el diseño de un modelo de decisión espectral dinámico para redes de radio cognitiva que permite a los usuarios secundarios acceder al espectro de manera oportunista y utilizar el canal sin afectar el tráfico de los usuarios primarios. El objetivo de este trabajo es

utilizar el recurso del espectro de manera eficiente eligiendo el canal apropiado en un instante de tiempo y reduciendo la cantidad de Ansoff a realizar por el usuario secundario. (Bernal Ariza, 2018)

- **Datos utilizados.**

- Obtenidos en la medición en la banda de frecuencia 824mhz - 879 mhz.

- **Herramientas de software.**

Start UML diagramas.

VM modelo de clasificación

## **5) Sector público.**

### **a. Evaluación de una metodología para la generación de modelos digitales del terreno a partir de estereofotogrametría digital aplicando el modelo de correlación por mínimos cuadrados mediante una red neuronal artificial.**

A continuación, se muestran los pasos que se tuvieron en cuenta para el desarrollo de este proyecto:

#### 1. Etapa de conceptualización y diseño:

Inicialmente en esta etapa se establecen los requisitos y restricciones que en el modelo evaluado se deben tener en cuenta para su construcción y correcto funcionamiento con el fin de darle solución al problema inicialmente planteado.

#### 2. Etapa de desarrollo:

En esta etapa se materializa el procedimiento, que de manera general consiste en seleccionar el modelo estereoscópico para someterlo al proceso de correlación y posteriormente su resultado es agregado en una red neuronal artificial.

3. Etapa de análisis de resultados Inicialmente se procesa el grupo de datos con las 5 clasificaciones y posteriormente con 10, de la misma forma se cambian los valores especificados en los parámetros de la red neuronal; estos experimentos se realizan con un grupo de datos denominado datos de entrenamiento, que corresponden al 20% del total de los datos trabajados. (Cajamarca Montoya, 2017)

- **Problema abordado.**

Actualmente en el desarrollo de diferentes proyectos de gran magnitud, se hace necesario en el campo de la ingeniería y más específicamente en el área de la Geomática, la automatización de diversos procesos con altos niveles de precisión ya que de esto depende que se modelen todos los factores presentes en la zona de la forma más real posible. En este documento se hace un acercamiento general a tres temas fundamentales a trabajar en un proyecto cuyo objetivo general es evaluar una metodología para la elaboración de modelos digitales del terreno (MDT) introduciendo la teoría de redes neuronales artificiales<sup>1</sup> junto al modelo de correlación por mínimos cuadrados.

- **Datos utilizados.**

- Imágenes Digitales.

- **Herramientas de software.**

- Weka Plataforma de aprendizaje automático.
- Start UML software para crear diagramas de despliegues.

**b. Diseño e implementación de un sistema creador de arte digital basado en el concepto de cadáver exquisito utilizando inteligencia artificial.**

Metodología de desarrollo cadáver exquisito.

Consiste en:



1. Reunión de todas las imágenes.

2. Selección de ellas mismas por medio de píxeles declarados en ciertas posiciones que permiten destacar los aspectos más fundamentales de las imágenes.

El problema de investigación se enfoca en darle una solución o punto de vista a los espectadores acerca de las relaciones profundas que hay entre las máquinas y los seres humanos.

(Salazar Cardenas & Vargas Trujillo, 2017)

- **Datos utilizados.**

- Uso de imágenes cuyos píxeles fueron redistribuidos y agregados a una selección de una capa que interactúa con el usuario.

- **Herramientas de software.**

- Python lenguaje de programación.
- OpenCV librería de python.

## **6) Sector de la tecnología.**

### **a. Sistema de visión artificial para detección y ubicación del golpe de un balón en superficie plana.**

A continuación, se muestran los pasos que se tuvieron en cuenta para el desarrollo de este proyecto:

1. Planeación y recolección de requisitos.

Etapa de levantamiento de los requisitos.

2. Modelado del prototipo.

En esta etapa se crean el modelo del funcionamiento del prototipo.

3. Construcción del prototipo.

En esta etapa se juntan los pasos anteriores y se de paso a crear el prototipo.

#### 4. Validación del prototipo.

Etapas de pruebas del funcionamiento del prototipo.

#### 5. Despliegue final del producto.

Etapas de muestras de utilidad del proyecto.

- **Problema abordado.**

En la actualidad muchas industrias y organizaciones han acogido a la inteligencia artificial (I.a) dentro de sus procesos más importantes, diseñando herramientas que permitan reducir tiempos y costos en los procesos de fabricación de productos, o en la verificación de calidad de los mismos, Otro tipo de enfoque uso que se puede dar a la inteligencia artificial, consiste en la detección e identificación de objetos esto comúnmente se conoce como visión artificial. Esto puede ser utilizado para diferentes sistemas, entre los que se encuentran los sistemas de movilidad, manufactura e incluso, videojuegos. (Murcia Jimenez, 2020)

- **Datos utilizados.**

- Datos estándar de imágenes y un formato de etiquetación.

- **Herramientas de software.**

- Tensor flow y keras librería de inteligencia artificial.

### 7) Sector del transporte.

#### **a. Reconocimiento óptico de caracteres en placas vehiculares haciendo uso de redes neuronales convolucionales.**

A continuación, se muestran los pasos que se tuvieron en cuenta para el desarrollo de este proyecto:

##### 1. Levantamiento de información y dataset.

En esta etapa se hace el levantamiento de la información obteniendo una fuente información de la data set escogido.

## 2. Entrenamiento de la red neuronal.

Con los datos obtenidos y la definición de las categorías se hace el entrenamiento de la red neuronal.

## 3. Implementación del algoritmo.

En esta etapa se hace la implementación del algoritmo partiendo de las categorías definidas y las funciones que se deben ejecutar.

## 4. Inferencia en la red.

En esta etapa se evalúa como actúa la red neuronal según el entrenamiento dado. análisis de los resultados. Los resultados son obtenidos a partir de una serie de pruebas con diferentes datos, hasta llegar a obtener el mínimo porcentaje de error.

- **Problema abordado.**

En la ciudad de Pereira el flujo vehicular es bastante grande, lo que hace que diariamente ocurran muchos accidentes por diferentes infracciones de tránsito, por lo tanto la Universidad tecnológica de Pereira en el programa de ingeniería de sistemas y computación, Se enfocaron en realizar una herramienta basada en inteligencia artificial, utilizando las redes neuronales como la técnica central junto a cámaras de seguridad y así lograr el reconocimiento de placas vehiculares, que son detectadas incumpliendo las diferentes señales de tránsito, y de esta manera lograr reducir el índice de accidentes. (Mosquera Mosquera & Martinez Rendon, 2018)

- **Datos utilizados.**

- Placas Vehiculares.
- Imágenes o fotografías tomadas por cámaras de seguridad.

- **Herramientas de software.**

- Sift (Scale-invariant features transform) y surf (speeded-up robust features).

## **8) Sector del gobierno.**

### **a. Construcción de un prototipo de software para el procesamiento y reconocimiento de las marchas básicas de un caballo en los ejes (x,y,z), con el fin de generar modelos de movimientos aproximados.**

A continuación, se muestran los pasos que se tuvieron en cuenta para el desarrollo de este proyecto:

1. Análisis de requerimientos globales.
2. El diseño de hardware y software que captura un número de muestras de un movimiento por medio de los sensores seleccionados.
3. Crear dos redes neuronales artificiales que reconocen los movimientos capturados bajo un cierto porcentaje de error, así como la visualización de datos mediante gráficas.

- **Problema abordado.**

En el proyecto de investigación en el problema abordado se busca implementar una solución de prevención de accidentes por medio de un simulador en la escuela de carabineros Alfonso López Pumarejo. (Camacho Barragán & Carranza Rivera, 2019)

- **Datos utilizados.**

- Recopilación y procesamiento de la información de datos dados por los acelerómetros mma7455 los cuales están adaptados a un robot equino de 6 grados de libertad, analizando las posiciones (x, y, z) capturando los datos en una interfaz.

- **Herramientas de software.**

- Matlab como herramienta de desarrollo de software.

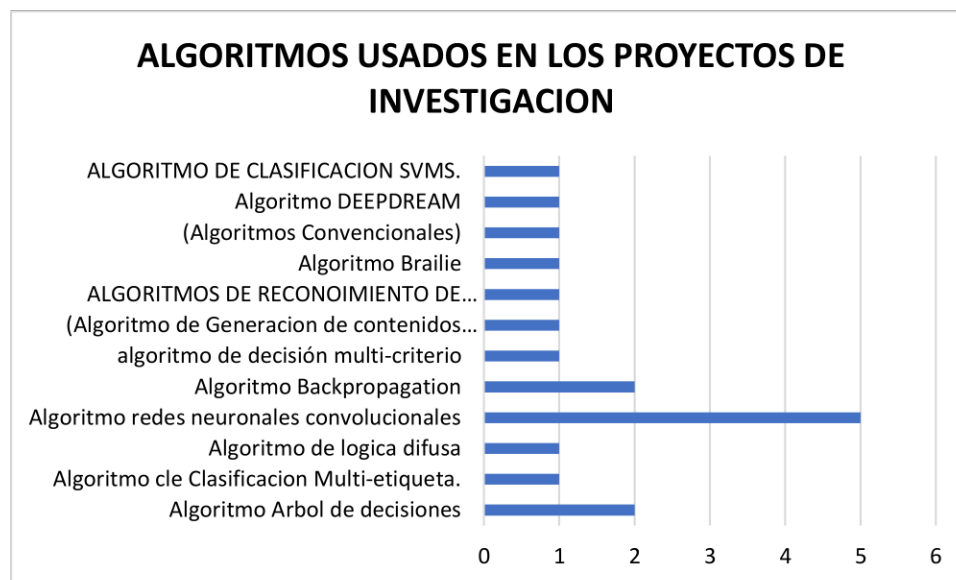


Figura 26. Grafica que representa el número de veces que se usaron determinadas metodologías dentro de la rama de las redes neuronales.

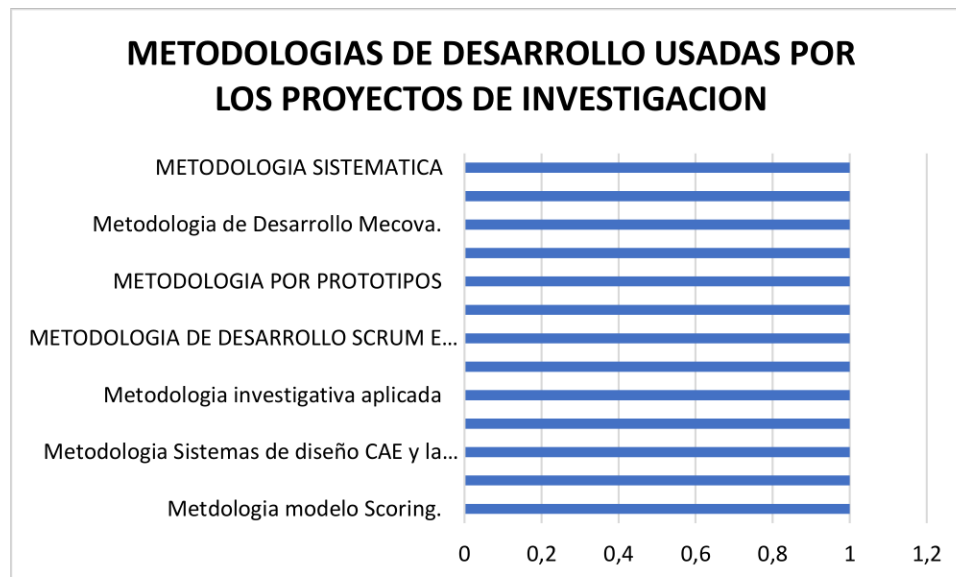


Figura 27. Grafica que representa el número de veces que se usaron determinados algoritmos en las tesis relacionadas con las redes neuronales.

En los trabajos de grado estudiados donde se encuentra aplicada el área de redes neuronales, se obtuvo que los sectores más representantes en las tesis de grado son el sector de la

salud, comercio y educación, donde los algoritmos de aprendizaje supervisado, backpropagation, y redes convolucionales fueron las aplicadas para el desarrollo de los proyectos de investigación, la información extraída para alimentar las redes neuronales son datos abiertos obtenidos de bases de datos elegidas objetivamente en el proceso de inicio de la realización del proyecto.

#### **4.1.9. Procesamiento de lenguaje Natural.**

El procesamiento de lenguaje natural consiste en la utilización de un lenguaje natural para comunicarnos con la computadora, debiendo esta entender las oraciones que le sean proporcionadas. El uso de estos lenguajes naturales facilita el desarrollo de programas que realicen tareas relacionadas con el lenguaje o bien, desarrollar modelos que ayuden a comprender los mecanismos humanos relacionados con el lenguaje. (Cifuentes Cortes & Martinez Naizaque, 2018)

##### **1) Sector público.**

###### **a) Watson analytics una solución cloud para procesos de big data.**

A continuación, se muestran los pasos que se tuvieron en cuenta para el desarrollo de este proyecto:

1. Adquisición de contenido: Es el primer paso de una aplicación de deepQA, la identificación y recopilación del contenido se resumen en 2 pasos, el primero es analizar ejemplos de preguntas del espacio del problema (rama de la ciencia o situacional).

2. Dada una cantidad razonable de datos, deepQA inicia de forma automática un proceso para alimentarse de nuevos datos a partir de la semilla base implantada.

3. Clasificación de las preguntas. Su propósito es identificar si las preguntas se pueden dividir en preguntas más sencillas y a su vez determinar cuáles partes de estas preguntas requieren un procesamiento especial.

4. Enfoque y detección TRL. Un tipo de respuesta de léxico es una palabra o frase en la pregunta que especifica el tipo de respuesta que se debe presentar sin cualquier intento de entender su semántica.

- Detección de relaciones. La mayoría de las preguntas contienen relaciones, si son sujeto-verbo, sintáctico, predicados o relaciones semánticas.
- **Problema abordado.**

En la actualidad las empresas han buscado mecanismos de producción que les permitan generar mayores ingresos, o agregar valor a cada producto o servicios que ofrecen, por lo tanto, se han enfocado en darle valor a los datos que en la fecha se ha convertido en el activo más valioso, por lo tanto se creó un método que por medio de Watson Analytics permita darle valor y crear un modelo de respuesta a grandes volúmenes de datos alojados en la nube, partiendo de mucha información no estructurada. (Garzon Castillo, 2017)

- **Datos utilizados.**
  - Documentos científicos, blogs, tweets e Información de internet.
- **Herramientas de software.**
  - Hadoop es una estructura de software de código abierto para almacenar datos y ejecutar aplicaciones en clústeres de hardware comercial.

**b. Modelo de interacción inteligente del reglamento del congreso de Colombia para dispositivos móviles “MIRC”.**

A continuación, se muestran los pasos que se tuvieron en cuenta para el desarrollo de este proyecto:

1. Análisis y revisión documental.

En esta fase se realizó una recopilación de toda la información necesaria para la actualización del reglamento del congreso, consistente en toda la jurisprudencia que la modificaron.

2. Diseño e implementación.

En esta fase se realizó un diseño del modelo, y se describe los recursos relevantes necesarios para su implementación.

3. Pruebas y validación de resultados.

Una vez implementado el prototipo de software app MIRC con el contenido del reglamento del congreso, que puede ser consultado a través de dispositivos móviles, se realizó una encuesta de satisfacción del uso del prototipo.

- **Problema abordado.**

El problema abordado en la investigación nace desde la inclusión de herramientas en los dispositivos electrónicos que permitan ofrecer y hacer más fácil la relación entre personas con diferentes discapacidades fundamentales, de manera que los beneficiados a través del modelo de integración inteligente puedan tener acercamiento y una relación eficiente con los reglamentos del congreso de la republica así tengan problemas de analfabetismo y de más, aplicando énfasis en calidad de creación de software. (Pinzón Ortiz, 2019)

- **Datos utilizados.**

- Un total de 495 archivos individuales de texto, que contienen el desglose de cada uno de los 393 artículos de la Ley 5 de 1992 (actualizada a diciembre de



2018) (Congreso de Colombia, 2018) y otros archivos anexos usados en la app MIRC.

- **Herramientas de software.**

- Framework de desarrollo app inventor.
- FileZilla cliente ftp.

## **2) Sector del gobierno.**

### **a. Desarrollo de armas cibernéticas utilizando inteligencia artificial (desarrollo de cyber usando armas inteligencia artificial).**

A continuación, se muestran los pasos que se tuvieron en cuenta para el desarrollo de este proyecto:

#### **1. Modelo de predicción de hashtags.**

Este modelo nos permite asignarle una o más etiquetas a un texto. En este caso usaremos los tweets recolectados para identificar en cada uno de ellos etiquetas relacionados al terrorismo.

#### **2. Modelo de identificación de entidades.**

Este módulo usa ner (named-entity recognition) que sus siglas indican reconocimiento de entidades nombradas. Consiste la extracción de un texto en este caso los tweets entidades como personas, organizaciones, eventos, productos entre otras cosas.

#### **3. Módulo de determinación de proximidad.**

El módulo de determinación de proximidad es la etapa donde se busca obtener los números más acertados de proximidad que se presentan en el sistema.

- **Problema abordado.**

El proyecto está enfocado en crear una solución efectiva en el sector de agencias privadas del gobierno, implementando herramientas que permitan lograr avances en los resultados o

estrategias que se utilizan para la detección de enemigos cibernéticos, creando armas tecnológicas en la web de avanzado uso tecnológico como lo es aplicando inteligencia artificial, por medio de análisis de tweets y comportamiento de diferentes usuarios.

- **Datos utilizados.**
  - Información del api de Twitter
- **Herramientas de software.**
  - Logstash es un pipeline de procesamiento de datos del lado del servidor que ingesta datos de una multitud de fuentes simultáneamente, los transforma y luego los envía a un "escondite", como elasticsearch.
  - Elasticsearch es un motor de búsqueda y analítica.

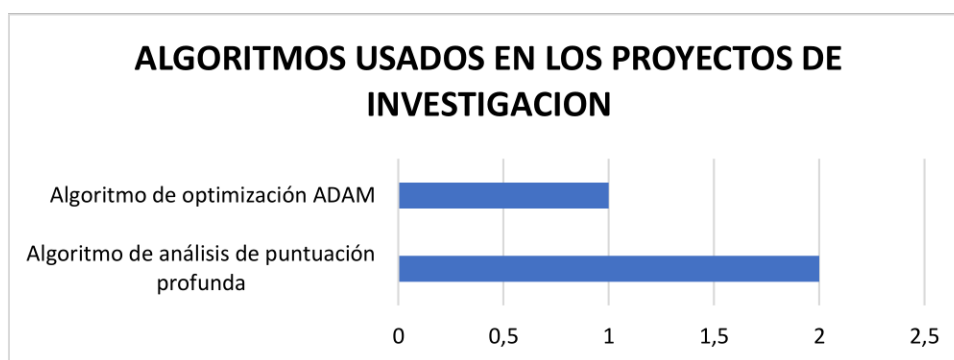


Figura 28. Grafica que representa el número de veces que se usaron determinadas metodologías dentro de la rama de procesamiento de lenguaje natural.

En los trabajos de grado estudiados donde se encuentra aplicada el área de lenguaje de procesamiento de lenguaje natural, se obtuvo que los sectores más representantes en las tesis de grado son el sector del gobierno, y el público, donde los algoritmos que se escogieron para la realización de estos proyectos, los algoritmos utilizados son optimización ADAM y algoritmo de análisis de puntuación profunda, y los datos del procesamiento son imágenes elegidas, a las cuales se les hace un análisis desde su tamaño, y las principales características, sombras, colores, imágenes.

#### **4.1.10. Big Data.**

Big Data se describe a menudo como conjuntos de datos extremadamente grandes que crecido más allá de la capacidad de gestionarlos y analizarlos con métodos tradicionales herramientas de procesamiento de datos. La búsqueda de pistas en la Web revela casi definición universal. (Ohlhorst, 2013)

##### **1) Sector comercial.**

##### **a. Uso de analíticas para predecir los computadores afectados por Malware, en una institución financiera en Colombia, 2017.**

Metodología de Box y Jenkins.

La metodología de Box y Jenkins es basada en la aceptación y análisis de hipótesis, es importante tener claro que al aceptar una hipótesis como cierta no se puede concluir respecto a la veracidad de los resultados obtenidos, sino que solo se aporta evidencia a su favor. De igual forma, menciona que la hipótesis de trabajo es considerada como la inicial que plantea el investigador al dar una respuesta anticipada al problema objeto de la investigación. (Mayorga García, 2017)

Consiste en 4 Fases:

1. La primera fase consistió en identificar el posible modelo ARIMA que sigue la serie, que lo requiere.

2. La segunda fase: Seleccionando provisionalmente un modelo para la serie estacionaria, se pasa a la segunda etapa de estimación, donde los parámetros  $\alpha$  y  $\beta$  del modelo se estiman por máxima verosimilitud y se obtienen sus errores estándar y los residuos del modelo.

3. La tercera fase es el diagnóstico, donde se comprueba que los residuos no tienen estructura de dependencia y siguen un proceso de ruido blanco.

4. La cuarta fase es la predicción, una vez que se ha obtenido un modelo adecuado se realizan las predicciones de este.

- **Problema abordado.**

El problema abordado en el proyecto de investigación consiste en crear un mecanismo de protección para predecir ataques de malware, softwares maliciosos a los equipos de trabajo de las empresas. (Mayorga García, 2017)

- **Datos utilizados.**

- Datos abiertos provenientes de Eventos que se presentan en diferentes fallas ocurridas en los sistemas de la empresa Verizon mediante Ataques por medio de Malware y software maliciosos, haciendo un estudio de:
  - Sectores de Afectación más recurrente.
  - Formas en las que se introdujo el virus.
  - Daños causados.
  - Tiempos establecidos para la eliminación del virus.

- **Herramientas de software.**

- Alterix: Software donde se puede visualizar los resultados correspondientes a los datos procesados.
- Software comercial tableau: Dispone de grandes beneficios para la presentación de datos.

### **b. Implementación de una herramienta de autogestión y autoconfiguración para la implementación de servicios en proyectos de big data.**

Metodología de 4 fases:

La metodología consiste en realizar la definición de todos los servicios a implementar, para llevar a cabo la selección y realizar un análisis comparativo entre en 16 herramientas a partir de las cuales van a ser seleccionadas 8 de ellas, y luego implementar las etapas teniendo en cuenta parámetros de usabilidad, facilidad de implementación y siempre manteniendo el enfoque en herramientas open source. (Torres Robles & Rincon Saavedra, 2017)

1) Levantamiento de la información desde el concepto de big data, abarcando su historia, arquitecturas y herramientas que facilitan su gestión.

2) Se realiza un proceso de evaluación basado en criterios de impacto, recursos y tiempo, que generan sobre las organizaciones al ser implementadas.

3) Se diseña el mecanismo a implementar, basándose en las arquitecturas zeta, kappa y lambda; haciendo uso de ocho herramientas seleccionadas que se administran a través de chef server, por medio de rutinas (scripts) diseñadas para su administración y gestión.

4) Ejecutar pruebas comparativas para medir el grado de optimización generado al implantar la herramienta vs los procesos normales de instalación, al igual que el impacto que esta genera en la administración de ti en las organizaciones.

- **Problema abordado.**

El problema abordado en el proyecto de investigación busca analizar, investigar y facilitar la información a los usuarios sobre las actuales herramientas gratuitas que permiten la autogestión, autoconfiguración de la infraestructura y el manejo de datos en proyectos orientados a big data. (Torres Robles & Rincon Saavedra, 2017)

- **Datos utilizados.**

- Se realiza un estudio a los datos encontrados en la Web sobre las herramientas a utilizar para el estudio.

- **Herramientas de software.**

- Mongo db un sistema de base de datos NoSQL orientado a documentos de código abierto
- Hadoop es una estructura de software de código abierto para almacenar datos y ejecutar aplicaciones en clústeres de hardware comercial.
- Clúster mpi protocolo de comunicación entre ordenadores para computación paralela.
- Apache Spark es una solución Big Data orientada al procesamiento de datos. En este artículo te enseñamos su funcionamiento básico.
- Grafana es un software libre basado en licencia de Apache 2.0, que permite la visualización y el formato de datos métricos.

**c. Plan de implementación de tecnologías big data para la optimización de estrategias comerciales y de segmentación.**

Metodología basada en modelo de prototipos.

La metodología basada en prototipos es basada en una representación de aquellos aspectos que sean visibles para el cliente o usuario final, construyendo prototipos que permiten la retroalimentación de los interesados que facilitan el refinamiento 21 de los requisitos iniciales planteados y a su vez garantiza un entendimiento pleno del caso de estudio. (Carrillo Álvarez, 2016)

Consiste en:

Fase 1: Esta fase no forma parte de la propuesta de esta tesis, pero reconociendo su importancia dentro del proceso de implementación se considera dentro de las fases a desarrollar.

Fase 2: Describe las actividades que permitirán la implementación del escenario realizable planteado en la figura 15 del capítulo 6. El inicio de esta fase depende de tener implementado el crm, haber unificado los clientes en un solo repositorio central, y tener un gobierno de datos implementado y funcional.

Fase 3: Esta fase marca el inicio del escenario deseado descrito en el capítulo 6. Se espera que para ese momento sea necesario replantear algunas actividades de las descritas a continuación, basados en que muy seguramente ya existan nuevas necesidades de analítica, visualización, de integración de información tal vez con móviles e incluso se pueda estar pensando en migración a un modelo cloud.

- **Problema abordado.**

El problema abordado es realizar un análisis que se viene haciendo muy importante en las empresas que ofrecen diferentes servicios en el sector comercial, partiendo de que es bastante amplia la cantidad de información que se manejan en ellas en diferentes sectores y que actualmente los datos es el activo más valioso y costoso en una empresa, tanto así que se hizo primordial crear un plan de implementación de tecnologías big data, que permitieran optimizar estrategias comerciales y de segmentación en los negocios. (Carrillo Álvarez, 2016)

- **Datos utilizados.**

- Encuestas realizadas
- Grupo objetivo: gerentes, subgerentes y coordinadores de las unidades de negocio de publicidad, circulación y administración.

- Tamaño de la Encuesta: 14 encuestas reales, 14 encuestas ponderadas
- Técnica de recolección de datos: diligenciamiento de formulario desarrollado y publicado en google docs.
- Tipo de la muestra: la muestra es dirigida, los encuestados seleccionados son empleados de vanguardia liberal.
- **Herramientas de software.**
  - Hadoop es una estructura de software de código abierto para almacenar datos y ejecutar aplicaciones en clústeres de hardware comercial.
  - Ferramentas business intelligence y business analytics.

**d. La ventaja competitiva del big data en el marketing empresarial y la innovación de productos.**

- **Problema abordado.**

En el desarrollo del documento se centrará en la innovación de productos, como actúan los fabricantes con un modelo tradicional y como el uso del Big Data, como paradigma de soluciones facilitaría y superaría la efectividad del modelo tradicional. Se hablará de las tres etapas por las que debe pasar un producto que se encuentra en lanzamiento al mercado, desde el punto de vista de un área de marketing y se atacaran las principales razones por las cuales fallan los lanzamientos por medio de soluciones contempladas con big data. (Rodriguez Saldaña, 2017)

- **Datos utilizados.**
  - Textos, emails, búsquedas web, contenidos.
- **Herramientas de software.**
  - No registra.



**e. Consideraciones generales para la aplicación del Big Data en negocios de redes de mercado.**

Metodología método clustering.

La metodología basada en técnicas de clustering, la cual sugiere aplicar un método basado en técnicas predictivas, el cual permitirá encontrar los grupos y los valores atípicos de la base de datos con el fin de detectar los usuarios de las redes sociales con similares patrones de comportamiento. (Salgado Jaime, 2017)

Consiste en:

1. emplear el algoritmo de k-means para la detección de los grupos.
2. ubicación de los puntos más lejanos con el propósito de identificar los casos atípicos y excluirlos del análisis.
3. detectar los grupos o clústeres con pocos elementos, para detectar que no se traten de casos atípicos y del ser el caso poderlos excluir del análisis.
4. tomar una muestra de cada uno de los grupos, tomando como tamaño mínimo de la muestra de acuerdo a la estimación de una proporción y conociendo el tamaño de la población de estudio.
5. con esta información tomar a la cantidad de usuarios definida en la estimación de la muestra y establecer un análisis de personalidad a cada uno de los usuarios de redes sociales incluidos en la muestra.
6. establecer el porcentaje de usuarios de cada una de las personalidades en cada uno de los clústeres.

- **Problema abordado.**

El propósito de este estudio monográfico es el plantear un enfoque personal de la aplicación de técnicas y tecnologías del big data en negocios de redes de mercadeo, que permitan agilizar el crecimiento de este tipo de compañías. se plantea la hipótesis de que es posible generar una mayor rentabilidad y crecimiento de la industria de las redes de mercado, a partir del uso del big data y de la minería de datos para clasificar a las personas con base en los patrones de conducta en las redes sociales y de esta manera generar estrategias de crecimiento para cada grupo. (Salgado Jaime, 2017)

- **Datos utilizados**

- Datos de la empresa.
- Número de personas relacionadas con su código.
- Datos básicos de las personas asociadas a su código (nombre, fecha de nacimiento, números de cedula, correo electrónico, ciudad de residencia).
- Datos de los consumos relacionados a cada código.

- **Herramientas de software.**

- 30D Tribus: permite a los usuarios acceder a audios e información sobre redes de mercado de manera gratuita.
- Node xl: Programa que permite hacer mapas a las redes sociales.
- Webometric analyst: permite extraer datos de redes sociales.

## **2) Sector de las comunicaciones.**

### **a. Creación de un método de analítica de noticias en medios de comunicación.**

A continuación, se muestran los pasos que se tuvieron en cuenta para el desarrollo de este proyecto:

1. Gestionar la llegada de los datos.

En esta etapa se identifican la llegada de los datos a partir de los clústeres o bases de datos en las cuales se realizan las consultas.

## 2. Procesamiento y clasificación de los datos.

En la etapa de procesamiento se escogen grandes cantidades de datos y se crean grupos o categorías definidas.

## 3. Transformación y análisis de los datos.

En esta etapa los datos toman nuevas formas, por medio de la calificación que se les da toman valor independiente.

- **Problema abordado.**

La problemática abordada en el proyecto de investigación tiene como objetivo crear un sistema que permita analizar todo el contenido de datos que se realiza en los diferentes sectores de los canales de televisión en un noticiero de comunicación, logrando por medio de técnicas de big data, agrupar y seleccionar un tipo de información que lleve a tener un análisis más detallado de las noticias y de las peticiones de los oyentes, y lectores de noticias de un canal informativo.

(Galvis Montero, 2020)

- **Datos obtenidos.**

- Información obtenida de los oyentes, noticias en tiempo real, solicitudes por parte los lectores.

- **Herramientas de software.**

- Advanced analytics herramienta de recopilación y gestión de grandes cantidades de datos.

### 3) Sector de la educación.

**a. Modelo de un sistema de software basado en las técnicas de learning analytics como herramienta de apoyo en la toma de decisiones académico-administrativas en las instituciones públicas de educación superior.**

Metodología del proceso unificado:

La metodología de proceso unificado con los objetivos planteados, que fueron utilizados los métodos básicos: inductivo, deductivo, analítico y sintético indistintamente, ajustándose a la necesidad requerida por cada ítem, los cuales ameritan a utilizar los métodos analítico y sintético, por cuanto proporciona profundidad a cada capítulo, manejando herramientas de la Ingeniería de Sistemas y Software que contribuyan al análisis y procesamiento de datos en la toma de decisiones. (Nieto Acevedo, 2016)

Fase 1: Inicio

En la fase de inicio se debe hacer el levantamiento de todos los requisitos del sistema para empezar con el modelamiento de datos.

Fase 2: Elaboración.

En esta fase se debe realizar una Recopilación y análisis detallado de toda la información

Fase 3: Construcción.

Diseño y construcción a partir del modelado de los datos.

Fase 4: Pruebas y validación.

Pruebas y validación del funcionamiento de la herramienta de apoyo.

- **Problema abordado.**

La problemática abordada en el proyecto es identificar un sistema de software basado en las técnicas de learning analytics suministre un sistema de apoyo en la toma de decisiones

académico-administrativas en las instituciones públicas de educación superior. (Nieto Acevedo, 2016)

- **Datos obtenidos**
  - Datos educativos.
- **Herramientas de software.**
  - Aplicación udlearn aplicación de interacción con los usuarios.
  - Oracle modelo cliente-servidor.

**b. Aplicación de big data e inteligencia de mercado en institución educativa tecnológica. (corporación tecnológica de Bogotá).**

- **Problema abordado.**

El problema abordado es crear una aplicaciones Big data urgente para crear sistemas que por medio de la información guardada por varios periodos permiten generar ganancias a través de modelos predictivos de ventas, y diferentes beneficios en áreas comerciales que ofrece manejar dentro del negocio herramientas de tecnología avanzada en este caso el big data como técnica del área de la inteligencia artificial, la cual incluye un modelo de análisis y clasificación de información, aplicando algoritmos basados en estudio de datos. (Novoa Urrego, 2017)

- **Datos utilizados.**
  - Base de datos de la institución.
- **Herramientas de software.**
  - No registra.

**c. Conociendo los beneficios de la relación de big data y cloud computing.**

- **Problema abordado.**

El propósito fundamental del presente documento es identificar la relación y los beneficios que existen en la utilización de cloud computing y big data en una organización, resaltando los atributos de cada una de ellas, partiendo del análisis detallado que se hace en cada una de las áreas evaluando aspectos más relevantes. (Martínez Beltrán, 2017)

- **Datos utilizados.**
  - Documentación de documentos relacionados con big data y cloud computing.
- **Herramientas de software.**
  - Sistemas de información.

#### **4) Sector de la agricultura.**

##### **a. Analítica de datos para el rendimiento en los cultivos de aguacate hass en Colombia.**

Metodología CRISP-DM:

Consiste en:

##### 1. Fase Comprensión del negocio.

Esta fase es una de las más importantes de la metodología en donde se debe tener la capacidad o habilidad de formular la pregunta correcta para aplicar el proceso de analítica de datos.

##### 2. Fase Comprensión del negocio – actividades

Como se observa en la ilustración anterior existen 4 actividades en donde se definen de la siguiente manera.

##### 3. Fase Comprensión de los datos.

En esta fase se comienzan a reunir los recursos para llevar a cabo el proceso de analítica de datos propuesto, es decir, basado en uno de los entregables de la fase anterior, como lo es la hipótesis inicial, se identifican aquellas posibles variables que tienen impacto en la pregunta formulada en

dicha fase, y una vez postuladas estas variables, se identifican los medios de recolección de la información necesaria en los periodos requeridos por cada una de ellas.

#### 4. Fase preparación de los datos.

Una vez recolectados los datos en la periodicidad necesitada, se procede a preparar los datos para el modelo, en donde se aplican los procesos de normalización propuestos para la fase anterior para lograr el nivel de calidad deseado, teniendo como premisa lo siguiente: “a mayor calidad de los datos recolectados, mayor es la confiabilidad del modelo de análisis predictivo”.

5. Fase evaluación en esta fase de debe evaluar los resultados del modelo ejecutado en la fase anterior, y una de las maneras más pragmáticas de hacerlo es tomando un grupo de registros históricos que no se tuvieron en cuenta en el desarrollo del modelo.

- **Problema abordado.**

La producción de aguacate hass en Colombia, se ha convertido en una gran oportunidad para el país, pero a su vez un gran reto para el agricultor, ya que requiere una gran inversión inicial, pero se ha tratado de materializar a través del conocimiento empírico, es decir, bajo la metodología del ensayo y el error, profundizando el riesgo de perder la siembra y por supuesto la inversión inyectada por los agricultores como consecuencia adversa, por lo tanto se planteó una solución por medio de analítica de datos para optimizar el rendimiento de los cultivos de aguacate hass en Colombia. (Lozano Vásquez & Suatena Cabrera, 2019)

- **Datos utilizados.**

- Datos meteorológicos actuales y futuros.

Estados climáticos:

- Humedad relativa.
- Temperatura.

- Precipitación.
- Brillo solar Producción.
- Agrícola.
- Área cosechada.
- Área Sembrada.
- Producción Rendimiento.
- **Herramientas de software.**
  - Microsoft Azure es un servicio de hospedaje web totalmente administrado que permite crear aplicaciones web.
  - Lenguaje R lenguaje de programación.

## **5) Sector publico**

### **a. Sistema inteligente para detectar y mapear eventos de violencia urbana en distintas zonas de Barranquilla, basado en información obtenida de Twitter.**

A continuación, se muestran los pasos que se tuvieron en cuenta para el desarrollo de este proyecto:

#### 1. Desarrolló del modelo.

En esta etapa se buscan los mejores procesos para implementar el desarrollo del modelo y de su funcionamiento, como está constituido y como va a ser sus funciones.

#### 2. Diseño de la solución.

En el diseño de la solución será utilizado los posibles resultados que se verán después cada prueba echa por el sistema.

#### 3. Desplegar el prototipo planeado.



En la etapa de despegar el prototipo planteado se busca dar a conocer cada proceso y cada solución que va a brindar el sistema según los datos necesarios.

#### 4. Validar el prototipo planeado.

En esta etapa se va a validar que el funcionamiento del sistema sea legible y correcto.

- **Problema abordado.**

Este proyecto surge de la necesidad de tener una forma rápida y gráfica de obtener información los distintos sucesos de violencia que están ocurriendo en los distintos barrios de la ciudad de Barranquilla diariamente. con base en esto se busca implementar un prototipo que reaccione diariamente, automatizado y que permita mostrar las distintas zonas donde están ocurriendo sucesos de violencia resaltados en un mapa, con base en información obtenida de tuits. se diseñó una arquitectura basada en microservicios cuya comunicación se realiza bajo el conjunto de protocolos tcp/ip y http. (Arrieta Villa, Estrada Vargas, & Ramírez Sánchez, 2020)

- **Datos utilizados.**

- Datos obtenidos api de Twitter.
- Reacciones y comentarios.

- **Herramientas de software.**

- IDE react entorno de desarrollo en la web.
- Firebase cloud firestore framework de desarrollo web.
- Api de Google maps herramienta de ubicación de Google.
- JavaScript lenguaje de programación.

## 6. Sector del gobierno.

### a) Informe final de práctica.

Metodología investigativa.

La metodología aplicada se enfoca en buscar información en diferentes fuentes, tales como plataformas reconocidas en el mercado de data lake, documentos, artículos y tesis enfocadas en los temas de interés, asesorías continuas con profesores de las áreas de Inteligencia artificial, gerencia y gestión de sistemas. (Gaviria Giraldo, 2019)

Consiste en:

1. Buscar información en diferentes fuentes.
2. Asesorías continuas con profesores de las áreas de inteligencia artificial, gerencia y gestión de sistemas.
3. Organización de los datos

En esta etapa se recolectan todos los datos, y se van ordenando de manera que sea posible obtener modelos de predicción.

4. Resultados finales.

En la etapa de resultados se obtienen a partir de las pruebas realizadas.<sup>7</sup>

- **Problema abordado.**

La alcaldía de Envigado en el área de las tic está enfocado a dar soluciones, por lo tanto se tomó la solución a través de formar profesionales sobre una base científica, tecnológica, administrativa, ética y humanística que les permita la utilización de las tecnologías de la comunicación y la información para brindar soluciones a los problemas que plantea el mundo moderno en los sectores productivo, financiero y de servicios a través de un análisis de grandes volúmenes de datos aplicando big data, y análisis de las funciones de los data lake en el trabajo de campo. (Gaviria Giraldo, 2019)

- **Datos utilizados.**

- Definiciones, ingesta, análisis, herramientas, gobernabilidad y arquitecturas para la transformación de datos en información

- **Herramientas de software.**

- Google cloud herramienta de almacenamiento de datos en la nube.

## **7) Sector del ambiente.**

### **a. Modelo de pronóstico de escasez y sobreoferta de agua en una cuenca altoandina basado en análisis big data.**

A continuación, se muestran los pasos que se tuvieron en cuenta para el desarrollo de este proyecto:

#### 1. Preguntas de investigación.

En esta etapa de la metodología se realizan unas preguntas detalladas que permitan recoger la cantidad información que se necesita para continuar con la investigación.

#### 2. Fuente de datos y estrategia de búsqueda.

En esta etapa de se define de que fuentes de información se van a recolectar los datos.

#### 3. Selección de estudios.

La selección de estudios se debe realizar de manera objetiva, centrando la investigación en la información precisa sobre la sobre oferta de agua.

#### 4. Clasificación de estudios.

La clasificación de los estudios permite obtener el modelo del pronóstico de la escasez de agua y la sobre oferta que se presentara según los datos planteados.

- **Problema abordado.**

Los cambios climáticos han sido fundamentales en la producción de ciertos sectores de la producción comercial en la economía , por lo que el agua es el recurso más caro, y se considera

muy importante crear modelos de detección de escasez a partir de la responsabilidad de la gestión y manejo del recurso hídrico, en la investigación se busca por medio de un modelado de grandes volúmenes de datos aplicando la técnica de inteligencia artificial big data, crear un modelo que permita obtener datos efectivos de un pronóstico de escasez de agua y la sobreoferta de la misma. (Valencia Payan, 2018)

- **Datos utilizados.**
  - Bases de datos, de investigación científica.
    - Scopus.
    - Ieee explore
    - Acm digital library (652 artículos relacionados).
- **Herramientas de software.**
  - SIG sistema de información geográfica.
  - Weka Plataforma de aprendizaje automático.
  - Software R, ide de programación.

## **8) Sector de la tecnología.**

### **a. Análisis comparativo del desempeño y costo computacional de una infraestructura de almacenamiento y procesamiento distribuido para el procesamiento de colecciones de texto.**

Metodología de investigación por fases secuenciales (bag of words).

La metodología establecida en el trabajo de grado es una investigación por fases secuenciales que estuvieron alineadas a los objetivos específicos representadas de forma general y que permitieron realizar todos los procesos en cada una de las etapas, en cada fase se tuvo

como propósito desarrollar uno o varios de los objetivos específicos de este proyecto. (Porras Garcia, 2018)

Consiste en 4 fases:

1. Selección e implementación de una Plataforma de procesamiento y almacenamiento distribuido con múltiples nodos para el almacenamiento del algoritmo distribuido mapreduce.

2. Implementación de la metodología bag of words para el procesamiento y representación de colecciones de documentos de textos utilizando el algoritmo MapReduce.

3. Implementación de un sistema de búsqueda de información por relevancia usando la representación Vector space model de los documentos y medida de similitud Coseno a partir de la representación bag of Words.

4. Evaluación de los resultados referentes a rendimiento computacional, tiempos de respuesta, costos de producción y ejecución en diferentes escenarios y numero de nodos vinculados al clúster.

- **Problema abordado.**

La computación distribuida para el procesamiento en paralelo sobre un clúster de computadores ofrece un rendimiento computacional mayor con tiempos de respuesta mucho menores que si se realizara el mismo procesamiento en una sola máquina. Si se toma en cuenta que cada computador por sí solo implica una serie de recursos y costos para su funcionamiento, (económicos, humanos, etc.), Se realizo un análisis comparativo que se implementan a partir de técnicas y algoritmos de big data, para conocer detalladamente el costo y desempeño de una infraestructura de almacenamiento y procesamiento distribuido para el alojamiento y procesamiento de textos. (Porras Garcia, 2018)

- **Datos utilizados.**

- Clúster con 17 equipos conectados almacenando: (dataset) de 500 archivos de texto plano de tamaños variables con HDFS.
- **Herramientas de software.**
  - Apache hadoop distributed file system (HDFS).

Apache spark es una solución big data orientada al procesamiento de datos. en este artículo te enseñamos su funcionamiento básico.

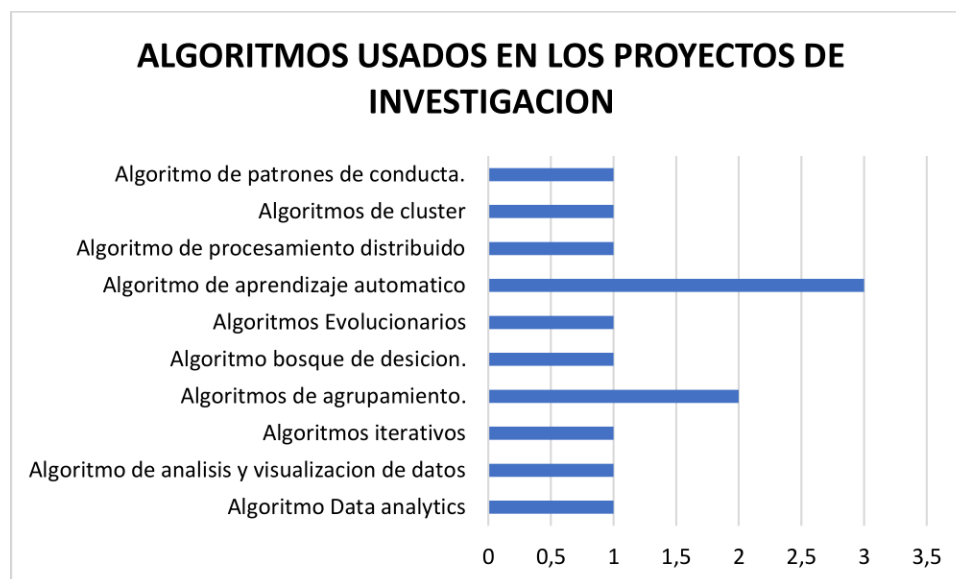


Figura 29. Grafica que representa el número de veces que se usaron determinadas metodologías dentro de la rama de la big data.

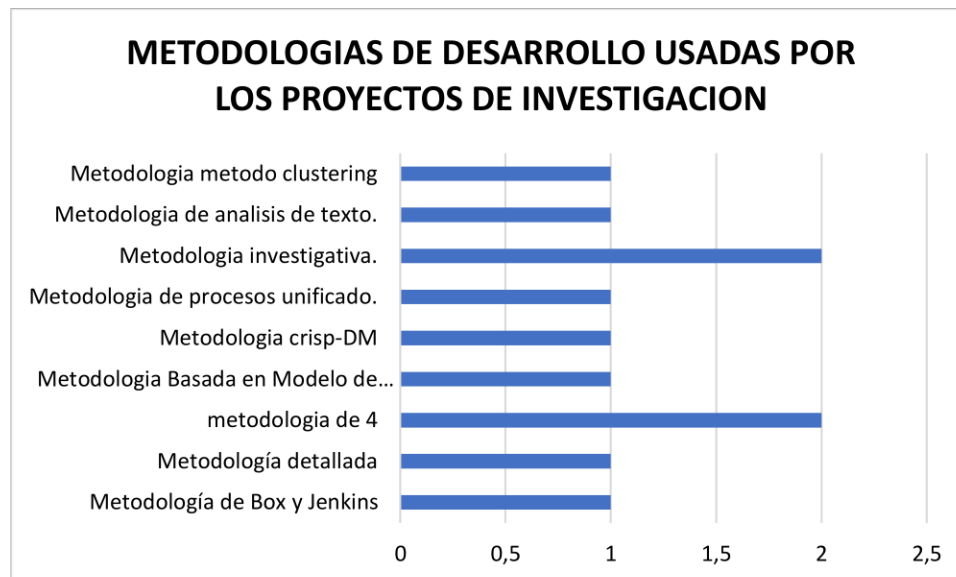


Figura 30. Grafica que representa el número de veces que se usaron determinadas metodologías dentro de la rama de la big data.

En los trabajos de grado estudiados donde se encuentra aplicada el área Big data, se obtuvo que los sectores más representantes en las tesis de grado son el sector del comercio, y de la educación, agricultura, y tecnología, para los cuales los algoritmos mas utilizados fueron, aprendizaje automático, data analytics, algoritmos clustering, donde el gran volumen de datos que son utilizados para el análisis y creación de modelos de predicción y de evaluación de rendimientos, son datos abiertos que se encuentran alojados en diferentes data sets, escogidos objetivamente para el desarrollo de los proyectos de investigación.

#### **4.1.11. Machine learning.**

El aprendizaje automático es una de las áreas de la informática de más rápido crecimiento con aplicaciones de gran alcance que permiten la integración de diferentes lenguajes de programación para lograr automatizar el funcionamiento de los sistemas, y así crear modelos de predicción que son inteligentes y cada vez más brindan soluciones a la vida cotidiana de las personas. (Shai Shalev & Shai Ben, 2014)

## **1) Sector de la educación.**

### **a. Machine learning applied to threat intelligence (aprendizaje de máquinas aplicado para amenazar la inteligencia).**

A continuación, se muestran los pasos que se tuvieron en cuenta para el desarrollo de este proyecto:

#### 1. Búsqueda y Selección de Información.

Dada la información de las familias de malware se investigó diferentes bases de datos que contuviera malware de forma segura para el análisis de forma abierta y recopilación accesible y segura.

#### 2. Extracción de Información.

Una vez obtenidos los ejecutables, se procede a analizarlos en el sandbox cuckoo. después de esto, cuckoo entrega un reporte con diferentes segmentos de los cuales nos enfocaremos en “summary” e “information on execution”.

#### 3. Normalización de datos.

Para que los datos sean más precisos y puedan usarse en la red neuronal artificial previamente escogida, se realizó un proceso de normalización y selección de datos. En el proceso de normalización, para los tres diccionarios con su severidad, se creó un vector de la misma longitud del vector de “signatures” o comportamientos llenos de ceros y unos, en el cual los “1” eran la posición exacta en donde la llave de los diccionarios coincide con el vector de las “signatures”.

#### 4. Aplicando aprendizaje automático (entrenamiento).



En el proceso de entrenamiento una vez establecido el conjunto de datos normalizados, los cuales fueron 150 malware distribuidos equitativamente entre las tres familias antes mencionadas.

#### 5. Aplicando aprendizaje automático (clasificador).

Para la prueba de la RNA se usaron 54 malwares diferentes a los usados en el entrenamiento de la red neural dando un 98% de clasificación acertada en las familias de malware antes expuestas; este porcentaje se mide basado en las salidas las cuales fueron 4, o de manera gráfica  $\{0,0,0,0\}$ .

#### 6. Resultados aplicación de algoritmo.

Una vez obtenidos los resultados de los 50 archivos maliciosos se logró observar un 98% de acierto en la clasificación de estos malware dentro de las categorías botnets, troyanos y crypters, logrando observar patrones y comportamientos en específicos en cada una de las familias, logrando crear una base de datos con esta información en la cual podemos ver que eventos son frecuentes en cada clasificación.

- **Problema abordado.**

El problema abordado en el proyecto de investigación fue crear un modelo basado en machine learning que demuestre que se puede superar el funcionamiento y almacenamiento de los datos en las personas. (Tatiana & Arboleda Diaz, 2017)

- **Datos utilizados.**

- Datos abiertos y analizados de la plataforma OSSIM de uso libre que clasifica la información de software maliciosos.

- **Herramientas de software.**

- Máquina virtual sistema operativo Ubuntu distribución de Linux.

- Cuckoo Análisis de Software malicioso.
- Weka es una plataforma de software para el análisis de datos con machine learning y minería de datos.

### **b) Framework para el aprendizaje de redes neuronales profundas.**

2 metodologías que son:

1. La metodología para el desarrollo del framework incluye las fases definidas en un (unified process) las cuales deben ser desarrolladas completamente haciendo un análisis de cada uno de los procesos, y se debe evaluar el cumplimiento de cada una de las tareas para el desarrollo de software.

#### a) Inicio.

Se define el alcance del proyecto.

#### b) Elaboración.

En esta fase se obtiene la visión refinada del proyecto a realizar, la implementación iterativa del núcleo de la aplicación, la resolución de riesgos altos, nuevos requisitos y se ajustan las estimaciones.

#### c) Construcción.

Esta abarca la evolución hasta convertirse en producto listo incluyendo requisitos mínimos. Aquí se afinan los detalles menores como los diferentes tipos de casos o los riesgos menores.

#### d) Transición.

En esta fase final, el programa debe estar listo para ser probado, instalado y utilizado por el cliente sin ningún problema. Una vez finalizada esta fase, se debe comenzar a pensar en futuras novedades para la misma.

1. Perceptrón multicapas.
2. Mecanismos de optimización. y regularización.
3. Red convolucional.
4. Red secuencial.
5. Problema abordado.

En el sector educativo existen diferentes problemáticas que hacen que los temas de la educación sean bastante complejos a la hora de practicar una pedagogía adaptable a los estudiantes, por lo tanto, se planteó el desarrollo de un software que sea capaz de por medio de actividades interactivas con los estudiantes, les permita tener un acercamiento mucho más efectivo en cuanto a la temática de redes neuronales y aprendizaje profundo, siendo así una herramienta que abarca tecnologías emergentes. (Anzola Avila & Moreno Silva, 2020)

- **Datos utilizados.**
  - Revisión de marco teórico de diferentes fuentes.
- **Herramientas de software.**
  - Keras biblioteca de código abierto.

### **c. Desarrollo de componente web parametrizable de clustering para análisis de datos.**

Metodología de desarrollo XP (Extreme Programming).

La metodología de desarrollo de proyectos XP está fundada en la agilidad y simplicidad, que, a diferencia de las metodologías clásicas o tradicionales, busca reducir el número de procesos reemplazándolos por el mínimo de procesos cuyo esfuerzo valga la pena, con el objetivo de tener un buen desarrollo de software, más allá de una buena documentación. (Rubio Tapias & Alba Rodriguez, 2018)

Consiste en 4 fases:

1. Fase de Exploración: Hace uso de las historias de usuario para definir el alcance general del proyecto, esta actividad se realiza junto con los clientes, se evalúan los tiempos iniciales de desarrollo, aunque no sean los finales, teniendo en cuenta que lo que se conoce del proyecto en esta instancia es algo muy general y los requerimientos se pueden modificar en cada iteración.

2. Fase de Planificación: Los requerimientos definidos por el usuario en la fase anterior son utilizados en la fase de planificación en la que se acuerda con los interesados del proyecto el orden en que serán implementados, esta fase es corta a comparación de las demás y el resultado de esta es un plan de entregas.

3. Fase de Iteraciones: Es quizá la fase más importante en el ciclo de la metodología XP. En esta los requerimientos son desarrollados y se hace la entrega, aquí también el usuario interactúa en la presente fase pues se debe hacer un análisis de la entrega y de lo que se requiere para así retroalimentar y solicitar cambios si es necesario, además en esta fase también es posible medir el avance del proyecto.

4. Fase de puesta en producción: al finalizar cada iteración se hace entrega de módulos funcionales y sin errores, el cliente puede solicitar no poner el sistema en producción hasta que se obtenga la funcionalidad completa del sistema. En esta fase no se realizan más desarrollos funcionales, pero es necesario estimar tareas de ajuste y soporte. para la presente fase ya se conocen los componentes definitivos revisados por los stakeholders, los cuales se ponen a disposición del usuario en una fase de despliegue.

- **Problema abordado.**

En general, en la mayoría de los procesos de investigación (de cualquier índole o tema) se realizan levantamientos o recolección de datos con los que se pretende extraer la información que ellos contienen, de manera que se puedan sacar conclusiones que generen valor a la investigación y que sean de utilidad para la misma, esto se realiza mediante técnicas de análisis de datos, por lo tanto la investigación se enfoca en desarrollar un componente web que le permita a los docentes tener una herramienta que les permita clasificar grandes cantidades de datos de acuerdo con sus características y similitudes, por medio de la técnica de análisis de clustering y que también puedan integrar estos resultados a diferentes software para sacar su valor generado. (Rubio Tapias & Alba Rodriguez, 2018)

- **Datos utilizados.**

- Datos alojados en clústeres a, b, c, datos utilizados paquete de datos de prueba “usarrests” que r ofrece.

- **Herramientas de software.**

- Etapa de desarrollo: frameworks como ‘Spring Boot’ y lenguajes de programación como java y r.
- Etapa de alojamiento: bases de datos: MongoDB, eclipse, STS, Spring Tools y rserve

## **2) Sector del gobierno.**

**a. Apoyo en la toma de decisiones en los procesos de fijación de cuota alimentaria, soportado en tecnologías cognitivas.**

- **Problema abordado.**

Debido a la facilidad con que se rompe el núcleo familiar en Colombia, y la ineffectividad por parte del sistema judicial colombiano para enfrentar el número de demandas por inasistencia

alimentaria se plantea que la tecnología cognitiva puede apoyar a los jueces y/o investigadores a mejorar la toma de decisiones en torno a los procesos de fijación de cuota alimentaria e incluso a los procesos jurídicos de inasistencia alimentaria y lograr establecer estrategias basadas en el uso de esta tecnología de aprendizaje automático. (Lozada Moreno & Martínez Mendoza, 2018)

- **Datos utilizados.**
  - Sistema de información de la registraduría nacional.
  - Dirección de impuestos y aduanas nacionales.
  - Migración Colombia.
  - Redes sociales.
- **Herramientas de software.**
  - No registra.

### **3) Sector de la agricultura.**

#### **a. Diseño e implementación de un sistema de reconocimiento de naranjas para el robot gio 1 usando visión asistida por computador.**

A continuación, se muestran los pasos que se tuvieron en cuenta para el desarrollo de este proyecto:

##### 1. Adquisición de la imagen.

Para la adquisición de la imagen es necesario de un dispositivo físico que sea sensible a una determinada banda del espectro electromagnético. Este dispositivo toma una señal análoga proporcional al nivel de energía detectado para luego ser digitalizado.

##### 2. Procesamiento de la imagen.

Para el procesamiento de imágenes obtenidas por el sensor kinect, fue necesario la instalación de la biblioteca libre openCV 2.4.11 en la plataforma Microsoft Visual Studio Code 2013.

### 3. Segmentación de la imagen.

La segmentación es la tarea fundamental más importante y desafiante de la visión por computador.

- **Problema abordado.**

La agricultura de precisión ha sido en los últimos años uno de los más importantes objetos de investigación en muchas universidades del mundo. En el encontramos que los principales trabajos son dirigidos a sistemas de cosecha autónomos para vegetales/frutas, sistemas de fertilización autónomos o sistemas de irrigación autónomos, entre otros, A partir de esto se busca diseñar un sistema de reconocimiento de naranjas para el robot GIO 1 usando visión asistida por computador con el fin de investigar la agricultura de precisión. (Gonzales Galvis & Alexis, 2015)

- **Datos utilizados.**

- 100 muestras de naranjas, y para distinguir que el algoritmo solo clasifica naranjas se utilizó: limón pera mango mandarina manzana: haciendo un estudio. Estas tres categorías que son:

1. pruebas
2. aciertos
3. fallos

- **Herramientas de software.**

- Sensor kinect de Microsoft.

- Microsoft Visual Studio Code 2013.
- Software Arduino.
- OpenCV algoritmos para el procesamiento de imágenes en lenguaje C++,  
Openni.

#### **4) Sector público.**

##### **a. Aplicación del aprendizaje automático en la clasificación de textos cortos: un caso de estudio en el conflicto armado colombiano.**

A continuación, se muestran los pasos que se tuvieron en cuenta para el desarrollo de este proyecto:

- 1) Determinar las plataformas a usar para la realización del proyecto.
- 2) Determinar los entornos a utilizar para el desarrollo de los algoritmos.
- 3) Determinar los algoritmos a usar para su desarrollo.
- 4) Adicionar los tweets para su análisis.
- 5) Clasificar las bolsas de palabra obtenidas de los tweets seleccionados.

- **Problema abordado.**

Se quiere determinar cómo se puede aplicar el aprendizaje automático para identificar la orientación pacifista, guerrerista o neutra de un comunicado corto emitido por un actor del conflicto armado colombiano. (Paramo Lozada & Espitia Betancourt, 2018)

- **Datos utilizados.**

- Los Tweets cortos realizados por los dos bandos de la guerra colombiana  
FARC-EP y el gobierno colombiano.

- **Herramientas de software.**

- Weka plataforma de aprendizaje automático.



- Rapidminer estudio plataforma de consultas de grandes volúmenes de datos.
- Knime plataforma de minería.
- Matlab software para cálculos matemáticos.

**b) Metodología para la evaluación de la integridad de tuberías corroídas mediante la predicción del factor estimado de reparación a través de la aplicación de algoritmos de machine learning.**

A continuación, se muestran los pasos que se tuvieron en cuenta para el desarrollo de este proyecto:

1. Bases de datos ILI.

En el primer reporte, se especifica cada uno de los componentes (como válvulas, puntos de soldaduras, accesorios tipo te, entre otros) y cada una de las anomalías identificadas (como corrosión, grietas, abolladuras, desprendimientos, ovalidad, entre otras).

2. Segmentación clustering.

Una vez realizada la exploración de los datos se propone establecer una segmentación de la tubería de tal forma, que se establezcan zonas críticas afectadas por la corrosión y con el objetivo de poder evaluar y analizar la tubería en dos instantes de tiempo diferentes.

3. Procesamiento de datos.

Después de realizar el proceso de segmentación de la tubería y la respectiva caracterización es indispensable realizar un tratamiento de la base de datos que se va a utilizar para entrenar los modelos.

4. Entrenamiento de los modelos.

Una vez pre-procesados los datos se realiza el entrenamiento de los modelos de Machine Learning. Para realizar el ajuste de los modelos propuestos se utiliza la parte de la base de datos destinada para el entrenamiento.

- **Problema abordado.**

El problema abordado en esta investigación es la reparación de las tuberías de aguas residuales y limpias de la ciudad, esta reparación se lleva a cabo o su objeto de estudio parte de crear una metodología que permita evaluar la integridad de las tuberías, realizando un estudio de la predicción del factor estimado de reparación aplicando Machine Learning para crear la solución. (Barrera Martinez, 2019)

- **Datos utilizados.**

- Datos de inspecciones a las tuberías.
- Dislong Distancia longitudinal de registro del defecto de corrosión desde el punto de referencia.
- Poshoraria Posición horaria en la tubería del defecto.
- Superficie es la superficie en la que se localiza el defecto. interna o externa.
- Largo l es longitud del defecto  $a$ .
- Ancho w ancho del defecto.

- **Herramientas de software.**

- Software r: hclust () para realizar el clustering jerárquico.
- Lm () para realizar las regresiones lineales, PCR ()).

## 5) Sector de la salud.

**a. Sistema e-health para capturar, transmitir y almacenar el estado de la presión arterial y otros datos relevantes, con el fin de calcular la probabilidad de padecer hipertensión arterial.**

A continuación, se muestran los pasos que se tuvieron en cuenta para el desarrollo de este proyecto:

1. Análisis de información

Esta fase consiste en la búsqueda y análisis de información sobre los temas más relevantes para el desarrollo de este proyecto de grado de grado, tales como: soluciones basadas en técnicas de inteligencia artificial.

2. Requerimientos para la captación de variables:

Requerimientos generales del sistema En esta primera parte se definen los requerimientos en general del sistema.

3. Almacenamiento y visualización.

En la fase de almacenamiento es donde se definen donde alojar los grandes volúmenes de información de donde luego será extraída y clasificada.

4. Implementación y validación.

La implementación de los datos permite que al realizarle un análisis detallado a todos los que se usaron se logre obtener la validación del funcionamiento efectivo del sistema.

- **Problema abordado.**

El problema abordado en el proyecto de investigación es implementar un sistema health que permita capturar, transmitir y almacenar y el estado en tiempo real en el cual se encuentran los pacientes que tienden a sufrir de esta enfermedad, y emitir una especie de alarma a través de los datos registrados. (Baracaldo Perez, 2020)

- **Datos utilizados.**

- Variables.

1. Edad 2. Genero 3. Raza 4. Diabetes 5. Antecedentes 6. Practica deporte 7. Fumador 8. Consume alcohol 9. Sobrepeso.

- Categoría de la presión arterial.

Normal, elevada, presión arterial alta (hipertensión nivel 1) Presión arterial alta (hipertensión nivel 2) Crisis de hipertensión rangos definidas por edades: sistólica mm hg diastólica mm hg.

- **Herramientas de software.**

- Genie es una herramienta que permite establecer las probabilidades de las variables dependientes e independientes. archivos CSV.

**b) Modelo de machine learning para la clasificación de pacientes en términos del nivel asistencial requerido en una urgencia pediátrica con área de cuidados mínimos.**

A continuación, se muestran los pasos que se tuvieron en cuenta para el desarrollo de este proyecto:

1. Investigación de las técnicas de Machine Learning Existentes.
2. Se procede a la recolección y análisis de datos con los que se entrenara la herramienta de machine learning seleccionada.

3. Se realizar el entrenamiento y las pruebas a la herramienta de machine learning.

Problema abordado.

Se busca evaluar el aporte que pueden hacer las herramientas basadas en técnicas de machine learning para clasificar (Triage) los pacientes de una urgencia pediátrica en términos de la atención requerida, ya sea en el área de cuidados mínimos o en el área de alta complejidad,

como una herramienta de apoyo al profesional del área para tomar una decisión. (García Gazabon, 2014)

- **Datos utilizados.**
  - Los datos utilizados para entrenar la técnica seleccionada fueron un total de 17.132 historias clínicas con las que contaba el sistema de información del hospital.
- **Herramientas de software.**
  - Sistemas de información

## 6) Sector de la tecnología.

**a. Estudio para modelar matemáticamente el comportamiento de un robot lego ev3 con python.**

- **Problema abordado.**

En los últimos años, el uso de tecnología en el sector educativo ha cobrado fuerza logrando, de esta manera, crear e implementar diversas herramientas que permiten realizar ciertas actividades de forma más ágil y atractiva para los estudiantes, el problema abordado en la investigación es crear un estudio que permita modelar matemáticamente el comportamiento de un robot lego. (Jaimes Pardo, 2020)

- **Datos utilizados.**
  - Pruebas de rendimiento de los motores.
- **Herramientas de software.**
  - Librería de sklearn.
  - IDE desarrollo micropython.
  - Libretas de júpiter.

7. Sector del comercio.

**a) Proyecto vanir - pattern analysis using interactive visualization.**

A continuación, se muestran los pasos que se tuvieron en cuenta para el desarrollo de este proyecto:

1. Fuentes de datos.

Los datos por visualizar mediante Vanir pueden llegar en distintos formatos, lo cual dificulta la estandarización y adaptabilidad de cada una de las visualizaciones debido que estas son definidas a partir de un conjunto de datos en donde se definen los parámetros y datos a visualizar establecidos previamente.

2. Parsing y unificación de fuentes.

Para adaptar las fuentes de datos es necesario formalizar el formato que será procesado por Vanir para el despliegue de visualizaciones.

3. Uso de vectores y d3.

Para el desarrollo de las visualizaciones dinámicas se trabaja sobre una plataforma web, alimentada por un archivo JSON.

4. Decisiones de diseño

Las visualizaciones VANIR fueron realizadas usando la librería de javascript d3, la cual hace uso de gráficos vectoriales redimensionarles (scalable vector graphics), los cuales mediante el uso del formato XML se definen figuras como vectores, lo cual permite su escalabilidad y adaptación sencilla a nuevas fuentes de alimentación de información.

5. Desarrollo, implementación y pruebas.

funcionamiento del producto final, especificando el ambiente de desarrollo, escenarios de calidad, interactividad, uso y despliegue, manejo de colores, y desarrollo de pruebas de usuario para cada una de las visualizaciones.

### **Problema abordado.**

El proyecto de investigación busca implementar una plataforma o sistema que permita interactuar las imágenes tomadas por un usuario y que estas mismas tengan reconocimiento automático por parte del sistema, y de esta manera se clasifiquen en datos de mayor valor.

(Cifuentes de la Portilla, 2016)

### **Datos utilizados.**

- Data sets:
  - Datos: Actividades: acción realizada por una persona en un lugar y hora en específico.
  - Tiempo: Intervalos de tiempo en las que una actividad es realizada.
  - Lugares: Esta variable representa el lugar en la que se realizan las actividades, en estos lugares.
- **Herramientas de software.**
  - Json es un formato de texto sencillo para el intercambio de datos.
  - Html5 y javascript lenguajes de programación orientados a la web.

### **b) Aplicación de deep learning en robótica móvil para exploración y reconocimiento de objetos basados en imágenes.**

A continuación, se muestran los pasos que se tuvieron en cuenta para el desarrollo de este proyecto:

1. Desarrollo de red neuronal convolucional.

La parte de obtención de características de la imagen, llamada convolución, intenta resaltar bordes, esquinas de la imagen, con el objetivo de resaltar el objeto en esencia, para ello se aplican una serie de filtros que son vistas como capas.

## 2. Aplicación de deep learning.

El objetivo del robot es explorar el terreno para identificar tipos de estructuras físicas (corredores, ascensores, escaleras, etc.) que lo componen. Por medio de reconocimiento de imágenes y sensores láser, analizar su navegabilidad en el ambiente.

## 3. Diseño de arquitectura del robot móvil neuronal.

Teniendo en cuenta el diseño de la red neuronal se plantea tener dos redes neuronales, una convolucional para el reconocimiento y aprendizaje de las estructuras, escaleras, ascensores, etc.

## 4. Descripción de la implementación.

El sistema cuenta con una instancia controladora de todos los componentes “neural driver”, la cual a su vez nos menciona las etapas de implementación del sistema.

## 5. Resultados esperados.

Se espera que, al implementar el entrenamiento respectivo de los elementos, en tiempo real el sistema reconozca de forma correcta con al menos un 90%.

- **Problema abordado.**

Con el fin de mejorar y hacer más eficientes la reestructuración o reconstrucción de un edificio se pretende diseñar e implementar un robot que realice reconocimiento de estructuras físicas y navegación al interior de espacios cerrados por medios de una red neuronal, partiendo del análisis de objetos basado en imágenes (Contreras Rojas, 2015).

- **Datos utilizados.**



- 20500 datos de las imágenes de entrenamiento a la red neuronal, Datos de proximidad del sensor láser.
- **Herramientas de software.**
  - Targetmanager y librería de desarrollo cpickle especializada en persistir arreglos y matrices, “imagedriver capta las imágenes.

**c) Diseño de un Modelo matemático para la predicción de fuga de clientes en el área de telefonía.**

- **Problema abordado.**

Hoy en día, debido a la necesidad de estar comunicados telefónicamente y a la gran cantidad de empresas que ofrecen este servicio en Colombia, la telefonía móvil enfrenta una feroz competencia por retener usuarios, buscando evitar la fuga de sus clientes a otras entidades y no incurrir en costos adicionales al tratar de atraer a nuevos clientes. Por lo que es de suma importancia para las empresas de este sector, el predecir si un usuario abandonará sus servicios. Como solución a esta problemática, se propone el diseño de modelos matemáticos que predigan intención de fuga de clientes mediante el uso de técnicas clasificadoras de minería de datos.

(Camargo Ramos, Morales Cotes, & Orozco Polo, 2020)

- **Datos utilizados**
  - Base de datos históricos perteneciente a una empresa de telefonía, (7000 clientes).
- **Herramientas de software.**
  - Software R. análisis roc (receiver operating characteristic).

**d) Análisis y desarrollo de una herramienta tipo Marketplace utilizando técnicas de aprendizaje automático bajo algoritmos no supervisados.**

### Metodología scrum:

La metodología scrum es un framework adaptable que fue diseñado para mostrarle un valor a los clientes durante todas las etapas del desarrollo del proyecto, y su principal objetivo es lograr que el cliente quede satisfecho a partir de todas sus necesidades y un entorno de comunicación asertiva, una responsabilidad colectiva y un progreso que se vea reflejado de manera continua. (Bonilla Mariota & Guerreiro Campo, 2018)

#### 1. Planificación del sprint.

Si entendemos el significado del sprint como un mini proyecto dentro del proyecto principal, cada uno de ellos tiene un objetivo en particular. Por ejemplo, el primer intervalo puede ser plantear cuál será el presupuesto general a utilizar, por lo que se necesitará de un equipo de profesionales expertos en el tema económico.

2. Etapa de desarrollo cuando el trabajo del sprint está en curso, los encargados deben garantizar que no se generen cambios de último momento que puedan afectar los objetivos del mismo. Además, se asegura el cumplimiento de los plazos establecidos para su término.

3. Revisión del sprint al final del desarrollo del intervalo, es posible analizar y evaluar los resultados. si es necesario, todo el equipo colaborará para saber qué aspectos necesitan ser cambiados. en esta fase se fomenta la colaboración y retroalimentación entre todos. se incluyen los siguientes puntos: colaboración entre equipos, supervisores, jefes y dueños de productos. Se admiten análisis externos como forma de complementación.

4. Retroalimentación Los resultados pueden ser entregados para recibir un feedback no solo por parte de los profesionales dentro del proyecto, sino también de las personas que utilizarán directamente lo que se desea lograr; es decir, los clientes potenciales. Las lecciones

aprendidas durante esta etapa permitirán que el siguiente sprint pueda ser mucho más efectivo y ágil.

- **Problema abordado.**

En la actualidad, las plataformas marketplace para la venta de servicios mediante cupones no cuenta con un canal dedicado a las empresas que les permita crear sus propios con un diseño igual al que verían los usuarios potencialmente compradores y en el cual la empresa esté al pendiente de cuanto está activo y cuando se ha comenzado a vender el cupón, además cuentan con un alto porcentaje en el descuento de cada de cada servicio, por lo tanto el problema abordado en la investigación es enfocado a la mejora de un sitio de ventas en línea aplicando nuevas tecnologías, para mejorar las herramientas de búsqueda de y atención al cliente en cuanto al tema de hotelería, ubicaciones, reservas accesos. (Bonilla Mariota & Guerreo Campo, 2018)

- **Datos utilizados.**

- Procesos de desarrollo de software, en el sitio.

- **Herramientas de software.**

- Microsoft Visual Studio Code IDE de desarrollo.
- Node.js s un entorno en tiempo de ejecución multiplataforma, de código abierto, para la capa del servidor basado en el lenguaje de programación.
- MongoDB, elasticsearch motores de bases de datos.

**e) Desarrollo de un sistema recomendador web para la toma de decisiones durante el proceso de adquisición de equipos de cómputo utilizando árboles de decisión.**

A continuación, se muestran los pasos que se tuvieron en cuenta para el desarrollo de este proyecto:

1) Se ubico los establecimientos y el lugar de los mismos que participarían facilitando su lista de precios de componentes y equipos de cómputo.

2) Se establecieron los perfiles a tener en cuenta, en este caso seria, el cliente y el vendedor.

3) Se realiza un diagnóstico inicial de la situación, en este caso es que los vendedores quieren ofrecer sus equipos de cómputo y componentes de pc al mejor precio posible en un horario establecido y hay clientes que quieren acceder a estos pero su falta de conocimiento en el ámbito tecnológico o falta de tiempo para acercarse a los establecimientos en los horarios establecidos se les dificulta realizar una mejor evaluación de las diferentes opciones del mercado.

4) Para el desarrollo de lo mencionado anteriormente se utilizó la metodología Kanban.

- **Problema abordado.**

El problema abordado en el proyecto de investigación es diseñar un sistema recomendador web que integre una tienda virtual de equipos de cómputo con un espacio en el cual los usuarios puedan obtener asesoría del sistema, y la presencia de distribuidores nacionales donde puedan realizar sus compras. (Diaz Garcia & Lozano Garzon, 2017)

- **Datos utilizados.**

- La lista de precios ofrecidas por los establecimientos de tecnología con los cuales los árboles de decisión realizaran su análisis.

- **Herramientas de software.**

- Php myAdmin motor de Bases de datos.
- Anaconda continuum plataforma de búsqueda de minería de datos.
- Xampp servidor local.

- Django es un framework de desarrollo web de código abierto, escrito en Python.
- Scikit learn es una biblioteca de software de aprendizaje automático para el lenguaje de programación Python.

**f) Estudio comparativo de algoritmos de inteligencia artificial y minería de datos enfocados a la toma de decisiones empresariales de selección de personal.**

Metodología scrum:

La metodología scrum es un framework adaptable que fue diseñado para mostrarle un valor a los clientes durante todas las etapas del desarrollo del proyecto, y su principal objetivo es lograr que el cliente quede satisfecho a partir de todas sus necesidades y un entorno de comunicación asertiva, una responsabilidad colectiva y un progreso que se vea reflejado de manera continua. (Romero Romero, 2018)

Consiste en:

1. Planificación del sprint.

Si entendemos el significado del sprint como un mini proyecto dentro del proyecto principal, cada uno de ellos tiene un objetivo en particular. Por ejemplo, el primer intervalo puede ser plantear cuál será el presupuesto general a utilizar, por lo que se necesitará de un equipo de profesionales expertos en el tema económico.

2. Etapa de desarrollo cuando el trabajo del sprint está en curso, los encargados deben garantizar que no se generen cambios de último momento que puedan afectar los objetivos del mismo. Además, se asegura el cumplimiento de los plazos establecidos para su término.

3. Revisión del sprint al final del desarrollo del intervalo, es posible analizar y evaluar los resultados. Si es necesario, todo el equipo colaborará para saber qué aspectos necesitan ser

cambiados. En esta fase se fomenta la colaboración y retroalimentación entre todos. Se incluyen los siguientes puntos: colaboración entre equipos, supervisores, jefes y dueños de productos.

4. Retroalimentación los resultados pueden ser entregados para recibir un feedback no solo por parte de los profesionales dentro del proyecto, sino también de las personas que utilizarán directamente lo que se desea lograr; es decir, los clientes potenciales. Las lecciones aprendidas durante esta etapa permitirán que el siguiente sprint pueda ser mucho más efectivo y ágil.

- **Problema abordado.**

El presente trabajo de tesis cuyo objetivo es la realizar un estudio de todos los algoritmos de inteligencia artificial y minería de datos que son utilizados y programados para obtener decisiones empresariales mediante técnicas emergentes y así hacer la selección de personal en una empresa. (Romero Romero, 2018)

- **Datos utilizados.**

- Estudio de archivos de investigación de algoritmos de inteligencia artificial y minería de datos.

- **Herramientas de software.**

- Php lenguaje de programación orientada a la web.
- Base de datos mysql.

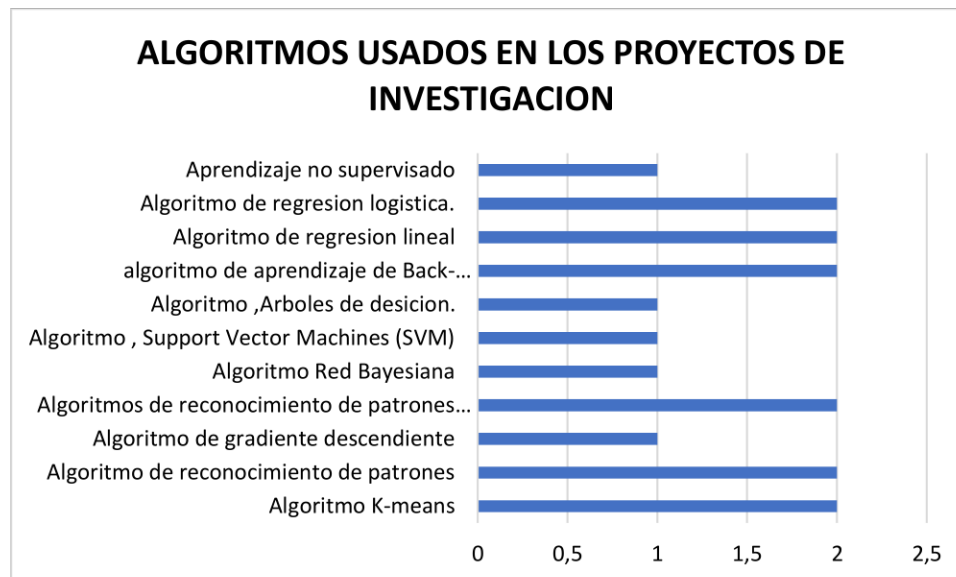


Figura 31. Grafica que representa el número de veces que se usaron determinadas metodologías dentro de la rama de machine learning.

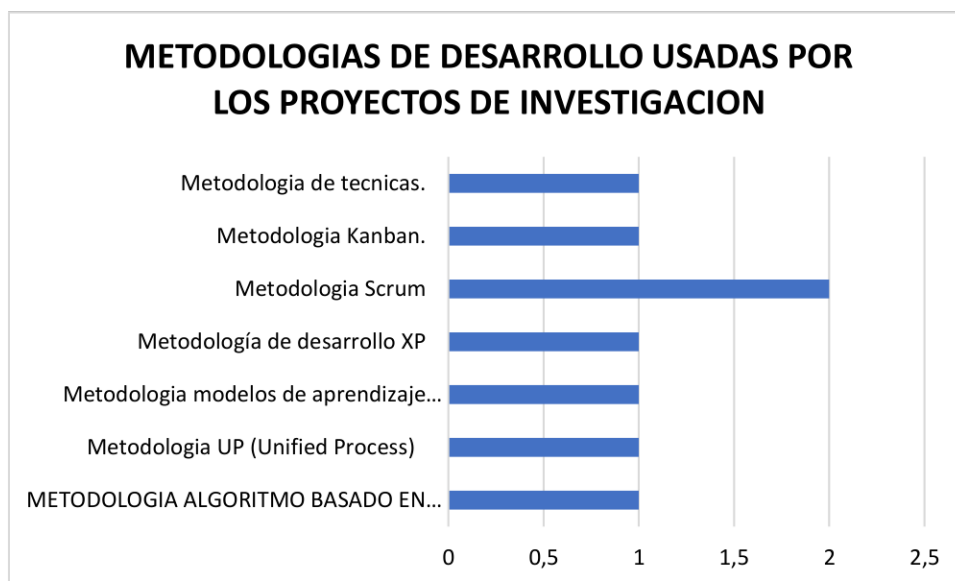


Figura 32. Grafica que representa el número de veces que se usaron determinados algoritmos en las tesis relacionadas con del machine learning.

En los trabajos de grado estudiados donde se encuentra aplicada el área Machine learning, se obtuvo que los sectores más representantes en las tesis de grado son el sector del comercio, educación, público y de la tecnología, para estos proyectos de investigación los algoritmos que

más se utilizaron fueron el aprendizaje supervisado y no supervisado, arboles de decisiones, algoritmo knn, clustering , regresión lineal, regresión logística, para la clasificación de la información estos algoritmos obtienen el conjunto de datos en bases de datos escogidas objetivamente ya que su contenido es abierto, permiten el acceso a la información, en muchos casos la información esta organizada en categorías, y en otros estas categorías deben ser definidas.

### **Conclusiones.**

1. A partir de la búsqueda realizada en la página de la gobernación SNIES, se obtuvieron datos abiertos que determinaron resultados muy importantes, los cuales mostraron la organización que tienen todas las universidades de Colombia, en cuanto a los repositorios de proyectos de grado donde se alojan todas las investigaciones que son realizadas por parte de los estudiantes que finalizan sus estudios en programas que conforman y son categorizados en el sector de las tic, logrando encontrar que existen repositorios que no permiten el acceso, en otros casos el acceso a los estudios de trabajo de grado es delimitado por un usuario y contraseña asignado a los estudiantes activos de las instituciones, en muy pocos casos los repositorios no se encuentran activos, y en un menor porcentaje de universidades se logró acceder, logrando así seleccionar los repositorios y analizar las tesis, que son el pilar del estudio que nos permitieron conocer cómo se está abriendo campo al conocimiento en nuevas tecnologías, en este caso la inteligencia artificial e IoT y de esta manera conocimos diferentes aspectos acerca de cómo se está preparando el país para hacerle frente a la industria 4.0 o más conocida como revolución industrial.



2. Para concluir que los trabajos de grados realizados en las universidades de Colombia sobre inteligencia artificial e Iot, son caracterizados por diferentes campos que definen la orientación y el objetivo planteado al realizar los proyectos de investigación, concluimos así que las áreas más representantes en las tesis presentadas en las diferentes instituciones de educación superior son la minería de datos, machine learning, redes neuronales artificiales e internet de las cosas, ya que sus tesis de investigación conforman el 74 , 35% de los estudios realizados por profesionales en las carreras ofertadas por las universidades colombianas que hacen parte de las tecnologías de la informacion y comunicación Tic.

3. Finalmente logramos concluir que en Colombia de los 32 departamentos a los cuales se le realizo la búsqueda a fondo, en la cual el análisis expuesto nos indicó que los que más se encuentran actualizados y generando propuestas tecnológicas basadas en inteligencia artificial e IoT son, Bogotá, Valle del Cauca, Atlántico, Bolívar y Norte de Santander ya que los temas de investigaciones hacen parte del 75,86% de las tesis que se realizaron entre los años 2015 y 2020, y que su enfoque fue realizar propuestas de desarrollo tecnológico que permitió optimizar tareas en diferentes sectores como es el de la salud, ventas, automatización de procesos industriales, y mejoramiento de la circulación vehicular en diferentes ciudades, logrando así ultimar que las áreas de la IA e IoT más utilizadas fueron la minería de datos aplicando el análisis de grandes volúmenes de información obtenidos de datos abiertos que permitían crear modelos de predicción que permitían en el sector de la salud obtener diagnósticos más efectivos, el aprendizaje automático (machine learning) aplicado en sectores industriales y empresas de

producción que requerían de modelos de funcionamiento rápido y eficiente aplicado por medio la técnica de aprendizaje supervisado, las redes neuronales en la solución de modelos de predicción de comportamiento de vías y soluciones de flujo vehicular, y para concluir con la investigación el internet de las cosas (iot) aplicado en diferentes modelos de predicción y organización de funcionamiento de sensores, y de más soluciones aplicados a procesos de automatización.

## **Recomendaciones**

Recomendamos a las universidades de Colombia, que al diseñar e implementar el funcionamiento de los repositorios institucionales se debe tener muy en cuenta incluir al desarrollo de las plataformas el protocolo de generacion y obtencion de meta datos OAI-PMH que genera y promueve los más destacados estándares de interoperabilidad que facilitan a los usuarios la difusión, intercambio y accesibilidad a documentos de diferentes naturaleza, ya que este protocolo tiene la capacidad de almacenar todos los meta datos en un solo lugar, para así lograr que las consultas a la información sean más eficaces y el tiempo de respuesta de los motores de búsqueda sean mucho más rápidos, a través del uso del protocolo de búsqueda en la web HTTP.

Sugerimos impulsar a las universidades de educación superior a la implementación de repositorios institucionales de acceso abierto que permitan a los investigadores gestionar, preservar y visibilizar las producciones científicas institucionales de tal manera que se disponga un marco de estándares para la conformación de grandes bases de datos abiertas a investigadores para el desarrollo de servicios de información virtual y así conformar un gran personal de formación de talento humano en documentación científica que atiendan a las necesidades del desarrollo y la apropiación de la ciencia, la tecnología y la innovación de los estudiantes de las instituciones de educación superior.

Proponemos a las instituciones de educación superior aplicar estos estándares que permitirán consolidar una oferta de la producción científica del país para lograr dar visibilidad y acceso a la información científica nacional, y de esta manera facilitar la inclusión en redes internacionales a través de la estandarización e interoperabilidad de los diferentes repositorios de las instituciones

miembros, y consolidar grandes bases de datos de información abierta a los diferentes investigadores, que a partir de estos grandes volúmenes de datos, crean modelos influyentes del comportamiento que tienen cada uno de los proyectos de investigación desarrollados por los estudiantes de instituciones de educación superior.

## Referencias

(s.f.). Obtenido de

<http://repositorio.ucundinamarca.edu.co/bitstream/handle/20.500.12558/2497/Trabajo%20de%20grado.pdf?sequence=3&isAllowed=y>

Accenture; Fundacion de la innovacion Bankinter. (2011). *El internet de las cosas. En un mundo conectado de objetos inteligentes*. Obtenido de

<https://www.fundacionbankinter.org/documents/20183/42758/PDF+Internet+de+las+cosas>

Acuña Ocampo, R. D. (2017). *Elaborar un prototipo de escritorio para reconocer imágenes de vista frontal por medio de inteligencia artificial con Python OpenCV [Trabajo de grado de ingeniería, Universidad Tecnológica del Pereira]*. Repositorio institucional. Obtenido de

<http://repositorio.utp.edu.co/dspace/bitstream/handle/11059/9216/T005.3%20A189.pdf?sequence=1&isAllowed=y>

Albrieu, R., Rapetti, M., López Brest, C., Larroulet, P., & Sorrentino, A. (2018). *Inteligencia artificial y crecimiento económico. Oportunidades y desafíos para Colombia*. Buenos Aires: CIPPEC.

Alvarez V., J. A., Hurtado G., S. V., & Trujillo M., H. (1995). *Algoritmos Genéticos [Proyecto de investigación, Universidad ICESI]*. Repositorio institucional. Obtenido de [https://repository.icesi.edu.co/biblioteca\\_digital/bitstream/10906/2529/1/Algoritmos\\_geneticos.pdf](https://repository.icesi.edu.co/biblioteca_digital/bitstream/10906/2529/1/Algoritmos_geneticos.pdf)

Álvarez, V., Padilla, J., & Sepúlveda, L. (2020). *Aplicación Software para la clasificación de pacientes con Alzheimer basado en técnicas de Data Mining [Trabajo de grado de*

- ingeniería, Universidad del Norte*]. Repositorio institucional. Obtenido de <http://manglar.uninorte.edu.co/handle/10584/8873#page=1>
- Angarita García, D. A., & Muñoz Franco, J. J. (2016). *Búsqueda de patrones en el comportamiento de los visitantes de la plataforma “Oferto” de la cámara de comercio de Armenia y del Quindío, a través de la aplicación de minería web [Tesis de maestría, Universidad Autónoma de Manizales]*. Repositorio Institucional. Obtenido de <http://repositorio.autonoma.edu.co/jspui/handle/11182/637>
- Anzola Avila, A., & Moreno Silva, J. A. (2020). *Framework para el aprendizaje de redes neuronales profundas. [J]*. Repositorio institucional. Obtenido de <https://repositorio.escuelaing.edu.co/handle/001/1236>
- Aravena Garrido, C., Bahamondes Acevedo, V., Diaz Muñoz, H., Galvez Mella, C., Hadweh Briceño, M., Veliz Veliz, C., & Maureira Cid, F. (2018). *¿Que es la inteligencia?* (F. Maureira Cid, Ed.) España: Bubok Publishing S.L. Obtenido de *Que es la inteligencia*: <https://sibdigital.ufps.edu.co:2303/es/ereader/ufps/51359>
- Arias Roa, E. F. (2019). *Modelamiento predictivo del ganador de un partido de futbol de la categoria A del futbol profesional colombiano usando aprendizaje de maquina [Trabajo de grado de ingeniería, Universidad Católica de Colombia]*. Repositorio institucional. Obtenido de [https://repository.ucatolica.edu.co/bitstream/10983/23643/1/PrediccionFutbolColombiano\\_625504.pdf](https://repository.ucatolica.edu.co/bitstream/10983/23643/1/PrediccionFutbolColombiano_625504.pdf)
- Arrieta Villa, O., Estrada Vargas, J. S., & Ramírez Sánchez, B. E. (2020). *Sistema inteligente para detectar y mapear eventos de violencia urbana en distintas zonas de Barranquilla, basado en información obtenida de Twitter. [Trabajo de grado de ingeniería,*

- Universidad del Norte*]. Repositorio institucional. Obtenido de <http://manglar.uninorte.edu.co/handle/10584/8854?show=full>
- Arroyave Zuluaga, L. A., & Mosquera Renteria, A. d. (2015). *Diseño e implementacion de un kit para robotica educativa*. Trabajo de grado, Corporacion Universitaria Minuto de Dios, Facultad de educacion virtual y a distancia, Bello. Obtenido de <https://core.ac.uk/download/pdf/160118576.pdf>
- Avellaneda Díaz, A. F. (2015). *Sistema de reconocimiento optico para mapeo conversion Barille - Español en el INCI [Trabajo de grado de ingenieria, Universidad Libre]*. Repositorio institucional. Obtenido de <https://repository.unilibre.edu.co/bitstream/handle/10901/8916/SISTEMA%20DE%20RECONOCIMIENTO%20OPTICO%20PARA%20MAPEO%20CONVERSI%c3%93N%20%20BRAILLE-%20ESPA%c3%91OL%20EN%20EL%20INCI.pdf?sequence=1&isAllowed=y>
- Baracaldo Perez, R. A. (2020). *Sistema e-health para capturar, transmitir y almacenar el estado de la presión arterial y otros datos relevantes, con el fin de calcular la probabilidad de padecer hipertensión arterial. [Trabajo de grado de ingenieria, Universidad de Colombia]*. Repositorio institucional. Obtenido de <https://repository.ucatolica.edu.co/bitstream/10983/24504/2/Trabajo%20de%20Grado.pdf>
- Barrera Martinez, B. (2019). *Metodología para la evaluación de la integridad de tuberías corroídas mediante la predicción del factor estimado de reparación a través de la aplicación de algoritmos de Machine learning. [Tesis de maestria, Unviersidad de los Andes]*. Repositorio institucional. Obtenido de

<https://repositorio.uniandes.edu.co/bitstream/handle/1992/43861/u830339.pdf?sequence=1&isAllowed=y>

Barriga Mariño, J. C. (2017). *Desarrollo y aplicacion de una herramienta de extraccion y almacenamiento de datos de twitter a un contexto social de violencia politica [Trabajo de grado de ingenieria, Universidad Catolica de Colombia]*. Repositorio institucional.

Obtenido de

[https://repository.ucatolica.edu.co/bitstream/10983/14597/1/DocumentoProyectoDeGrado\\_Camilo\\_Barriga\\_625140.pdf](https://repository.ucatolica.edu.co/bitstream/10983/14597/1/DocumentoProyectoDeGrado_Camilo_Barriga_625140.pdf)

Barrios Moises, A. (2018). *Internet de las cosas* (Primera ed.). Madrid, España: REUS S.A.

Recuperado el Agosto de 2020, de

[https://sibdigital.ufpso.edu.co:2303/es/ereader/ufpso/121519?as\\_all=Robotica&as\\_all\\_op=unaccent\\_\\_icontains&prev=as](https://sibdigital.ufpso.edu.co:2303/es/ereader/ufpso/121519?as_all=Robotica&as_all_op=unaccent__icontains&prev=as)

Becerra , R. (2011). *Sistemas expertos para la realización de diagnóstico de trastornos neuromusculares con electromiografía*. Trabajo de grado, Universidad del Aconcagua, Facultad de ciencias sociales y administrativas , Mendoza. Obtenido de

[http://bibliotecadigital.uda.edu.ar/objetos\\_digitales/265/tesis-3970-sistemas.pdf](http://bibliotecadigital.uda.edu.ar/objetos_digitales/265/tesis-3970-sistemas.pdf)

Bembibre , C. (07 de 2009). *DefinicionABC*. Obtenido de

<https://www.definicionabc.com/general/contexto.php>

Bernal Ariza, C. C. (2018). *Modelo inteligente de decisión de espectro que mejora el desempeño en redes de radio cognitiva. [Tesis de maestria, Universidad Distrital Francisco Jose de Caldas]*. Repositorio institucional. Obtenido de

<http://repository.udistrital.edu.co/handle/11349/13505>



- Bonilla Mariota, G. L., & Guerreo Campo, F. (2018). *Análisis y desarrollo de una herramienta tipo Marketplace utilizando técnicas de aprendizaje automático bajo algoritmos no supervisados. [Trabajo de grado de ingeniería, ].* Repositorio institucional. Obtenido de <http://repository.lasallista.edu.co/dspace/handle/10567/2564>
- Buendia Robinson, J. C., & Covilla Yarce, L. C. (2015). *Sistema para la captura y análisis de requerimientos de los proyectos basados en inteligencia computacional para la Universidad de Cartagena. [Trabajo de grado para de ingeniería, Universidad de Cartagena].* Repositorio institucional. Obtenido de [https://repositorio.unicartagena.edu.co/bitstream/handle/11227/2932/Bee-%20Tesis\\_LCovilla\\_JBuendia.pdf?sequence=1&isAllowed=y](https://repositorio.unicartagena.edu.co/bitstream/handle/11227/2932/Bee-%20Tesis_LCovilla_JBuendia.pdf?sequence=1&isAllowed=y)
- Cajamarca Montoya, Y. J. (2017). *Evaluación de una metodología para la generación de modelos digitales del terreno a partir de estereofotogrametría digital aplicando el modelo de correlación por mínimos cuadrados mediante una red neuronal artificial. [Tesis de maestría, Universidad Distrital Francisco Jose de Caldas]* Repositorio institucional. Obtenido de <http://repository.udistrital.edu.co/handle/11349/5141>
- Camacho Barragán, A. F., & Carranza Rivera, D. A. (2019). *Construcción de un prototipo de software para el procesamiento y reconocimiento de las marchas básicas de un caballo en los ejes (x,y,z), con el fin de generar modelos de movimientos aproximados. [Trabajo de grado de ingeniería, Universida de Cundinamarca].* Repositorio institucional. Obtenido de <http://repositorio.ucundinamarca.edu.co/handle/20.500.12558/2414>
- Camargo Ramos, J. J., Morales Cotes, E. A., & Orozco Polo, I. A. (2020). *Diseño de un Modelo matemático para la predicción de fuga de clientes en el área de telefonía. [Trabajo de*

- grado de ingeniería, Universidad del Norte*]. Repositorio institucional. Obtenido de <http://manglar.uninorte.edu.co/handle/10584/8881?show=full>
- Cañon Miranda, D. A. (2016). *Seguridad de la información en la internet de las cosas. [Tesis de especialización, Universidad Piloto de Colombia]*. Repositorio institucional. Obtenido de Seguridad de la información en la internet de las cosas
- Carrillo Álvarez, F. (2016). *Plan de implementación de tecnologías Big Data para la optimización de estrategias comerciales y de segmentación [Tesis de maestría, Universidad Autónoma de Bucaramanga]*. Repositorio institucional. Obtenido de <https://repository.unab.edu.co/handle/20.500.12749/3369>
- Carrillo Lindado, H. C., & Llanos Terán, D. J. (2017). *Reconocimiento y conteo de monedas colombianas mediante procesamiento digital de imágenes y redes neuronales convolucionales. [Trabajo de grado de ingeniería, Universidad del Norte]*. Repositorio institucional. Obtenido de <http://manglar.uninorte.edu.co/handle/10584/7795?show=full>
- Chagui Romero, J. I. (2018). *Prototipo de sistema de monitoreo no invasivo de pacientes crónicos que padecen hipertensión, asma y / o epoc. [Trabajo de grado de ingeniería, Universidad de Cartagena]*. Repositorio institucional. Obtenido de <https://repositorio.unicartagena.edu.co/handle/11227/8505>
- Chagui Romero, J. I., & Martínez Franco, E. d. (2018). *Prototipo de sistema de monitoreo no invasivo de pacientes crónicos que padecen hipertensión, asma y/o EPOC. [Trabajo de grado de ingeniería, Universidad de Cartagena]*. Repositorio institucional. Obtenido de <https://repositorio.unicartagena.edu.co/bitstream/handle/11227/8505/PROYECTO%20D E%20GRADO%20CHAG%c3%9cI- MARTINEZ%20%28Corregido%29.pdf?sequence=1&isAllowed=y>

- Cifuentes Cortes, J. D., & Martinez Naizaque, J. M. (2018). *Revision sistematica de las tecnicas de mineria web de contenido*[Trabajo de grado de ingenieria, Universidad Catolica de Colombia]. Repositorio institucional. Obtenido de <https://repository.ucatolica.edu.co/bitstream/10983/22602/1/Trabajo%20de%20Grado%20-%20Final.pdf>
- Cifuentes de la Portilla, M. A. (2016). *PROYECTO VANIR - pattern analysis using interactive visualization*. [Trabajo de grado de ingenieria, Universida de los Andes]. Repositorio institucional. Obtenido de <https://repositorio.uniandes.edu.co/bitstream/handle/1992/15278/u728634.pdf?sequence=1&isAllowed=y>
- Codari, M., Schiaffino, S., Sardanelli, F., & Trimboli, R. M. (Febrero de 2019). Artificial Intelligence for Breast MRI in 2008-2018: A Systematic Mapping Review. *American Journal of Roentgenology*, 212(2), 292. Obtenido de <https://pubmed.ncbi.nlm.nih.gov/30601029/>
- Congreso de Colombia. (04 de Enero de 2009). Ley N° 1273 de 2009. Colombia. Obtenido de <https://www.mintic.gov.co/portal/inicio/3705:Ley-1273-de-2009>
- Congreso de Colombia. (30 de Julio de 2019). Ley N° 1341 del 2009. Colombia. Obtenido de [https://www.mintic.gov.co/portal/604/articles-3707\\_documento.pdf](https://www.mintic.gov.co/portal/604/articles-3707_documento.pdf)
- Consejo Nacional de Politica Economica y Social. (08 de Noviembre de 2019). Politica nacional para la transformacion digital e inteligencia artificial. *Documento CONPES*, 63. Obtenido de Política Nacional para la Transformación Digital e Inteligencia Artificial, Documento CONPES 3975: [https://www.mintic.gov.co/portal/604/articles-107147\\_recurso\\_1.pdf](https://www.mintic.gov.co/portal/604/articles-107147_recurso_1.pdf)

- Contreras Rojas, J. S. (2015). *Aplicación de Deep Learning en robótica móvil para exploración y reconocimiento de objetos basados en imágenes*. [Trabajo de grado de ingeniería, Universidad de los Andes]. Repositorio institucional. Obtenido de <https://repositorio.uniandes.edu.co/bitstream/handle/1992/18454/u721833.pdf?sequence=1&isAllowed=y>
- Corrales, C. J. (2015). *Predicción adaptativa de la calidad del agua mediante técnicas de inteligencia computacional*. [Tesis de maestría, Unviersidad del Cauca]. Repositorio institucional. Obtenido de <http://repositorio.unicauca.edu.co:8080/bitstream/handle/123456789/1362/PREDICCI%c3%93N%20ADAPTATIVA%20DE%20LA%20CALIDAD%20DEL%20AGUA%20MEDIANTE%20T%c3%89CNICAS%20DE%20INTELIGENCIA%20COMPUTACIONAL.pdf?sequence=1&isAllowed=y>
- Corzo Sosa, D. E. (2018). *Modelo conceptual para determinar el proximo presidente de estado a traves de DATA MINING y twitter* [Trabajo de grado de ingeniería, Universidad Catolica de Colombia]. Repositorio Institucional. Obtenido de <https://repository.ucatolica.edu.co/bitstream/10983/22428/1/TwitterDatamining.pdf>
- Cuellar Chacon, C. M. (2018). *Generacion automatica de resúmenes de multiples documentos mediante la higratacion de la metaheuristica de la mejor busqueda armonica global y el algoritmo basado en grafos LexRank*. [Tesis de maestría, Unviersidad del Cauca]. Repositorio institucional. Obtenido de <http://repositorio.unicauca.edu.co:8080/bitstream/handle/123456789/1300/Generaci%c3%b3n%20Autom%c3%a1tica%20de%20Res%c3%bamenes%20de%20M%c3%baltiples%20Documentos.pdf?sequence=1&isAllowed=y>

- Diaz Garcia, E. W., & Lozano Garzon, A. (2017). *Desarrollo de un sistema recomendador web para la toma de decisiones durante el proceso de adquisición de equipos de cómputo utilizando árboles de decisión. [Trabajo de grado de ingeniería, Universidad de Cundinamarca]*. Repositorio institucional. Obtenido de <http://repositorio.ucundinamarca.edu.co/bitstream/handle/20.500.12558/738/Sistema%20de%20Recomendaci%c3%b3n%20Basado%20en%20Arboles%20de%20Decision.pdf?sequence=6&isAllowed=y>
- Diaz Pedroza, K. Y., & Chindoy Chasoy, B. Y. (2019). *Desarrollo de un software que permita la detección de estudiantes desertores del programa de ingeniería de sistemas utilizando procesos y elementos de minería de datos [Trabajo de grado de ingeniería, Universidad Francisco de Paula Santander Ocaña]*. Repositorio institucional. Obtenido de <http://repositorio.ufpso.edu.co/bitstream/123456789/490/1/32790.pdf>
- Dinero. (20 de Agosto de 2020). *¿Cuáles son las empresas más innovadoras del país?* Obtenido de Revista Dinero: <https://www.dinero.com/empresas/articulo/estas-son-las-empresas-mas-innovadoras-del-pais/296258>
- Dorronsoro Díaz , B. (12 de 2006). *Diseño e implementación de algoritmos genéticos celulares para problemas complejos. . Tesis Doctoral, Universidad de Málaga, Lenguajes y Ciencias de la Computación , Málaga*. Obtenido de <https://neo.lcc.uma.es/tesis/PhD-Bernabe.pdf>
- Durán Gómez, N. (2016). *Esquematización de un sistema para la ayuda en la toma de decisiones médicas al momento de identificar dolor en pacientes dentro de la unidad de cuidados intensivos neonatal [Trabajo de grado de ingeniería, Universidad de los Andes]*.

- Repositorio institucional. Obtenido de  
<https://repositorio.uniandes.edu.co/handle/1992/13614>
- ENDEAVOR. (2018). El impacto de la inteligencia artificial en el emprendimiento. Mexico.
- Escobar Ortiz, A. F. (2018). *Diseño de un clasificador de señales EEG provenientes de una interfaz BCI utilizando redes neuronales para controlar un robot hexápodo. [Trabajo de grado de ingeniería, Universidad Autónoma de Occidente]*. Repositorio institucional. Obtenido de <https://red.uao.edu.co/bitstream/10614/10354/4/T08014.pdf>
- Estevez Blanco, J. A. (2017). *Ciudades inteligentes y datos abiertos: Un dashboard basado en minería de datos [Trabajo de grado de ingeniería, Universidad Católica de Colombia]*. Repositorio Institucional. Obtenido de <https://repository.ucatolica.edu.co/bitstream/10983/15407/1/Jose%20Arnulfo%20Estevez%20Blanco%20625314.pdf>
- (s.f.). *Framework para el aprendizaje de redes neuronales profundas. []*.
- Galindo, F., Ruiz, S., & Ruiz, F. (2017). Competencias digitales ante la irrupción de la Cuarta Revolución Industrial. *Estudios en Comunicación, 1*, 11. Obtenido de <http://ojs.labcom-ifp.ubi.pt/index.php/ec/article/view/277/144>
- Gallego Gallego, M. (2015). *Descubrimiento de conocimiento en una empresa de outsourcing de TI de la ciudad de Medellín aplicando técnicas de minería de datos que permita identificar potencialidades en el éxito de los proyectos de desarrollo de software. [Tesis de maestría, Universidad Autónoma de Manizales]*. Obtenido de <http://repositorio.autonoma.edu.co/handle/11182/604>
- Galvis Montero, J. (2020). *Creación de un método de analítica de noticias en medios de comunicación. [Tesis de maestría, Escuela Colombiana de Ingeniería Julio Garavito]*.

Repositorio institucional. Obtenido de

<https://repositorio.escuelaing.edu.co/handle/001/1261>

Gantiva Henao, L. A. (2020). *Gestión de Riesgos en el Internet de las Cosas (IoT) [Tesis de especialización, Universidad Piloto de Colombia]*. Repositorio institucional. Obtenido de <http://repository.unipiloto.edu.co/handle/20.500.12277/6868>

García Gazabon, G. I. (2014). *Modelo de Machine Learning para la Clasificación de pacientes en términos del nivel asistencial requerido en una urgencia pediátrica con Área de Cuidados Mínimos. [Tesis de maestría, Universidad Tecnológica de Bolívar]*.

Repositorio institucional. Obtenido de

<https://repositorio.utb.edu.co/bitstream/handle/20.500.12585/1200/0068210.pdf?sequence=1&isAllowed=y>

García Montero, Y. (2018). *Red neuronal para diagnóstico de hipertensión arterial. [Trabajo de grado de ingeniería, La Universidad en Internet]*. Repositorio institucional. Obtenido de <https://reunir.unir.net/handle/123456789/6937>

García Reina, L. F. (2018). *Asistente virtual tipo chatbot*. Trabajo de grado, Universidad Católica de Colombia, Facultad de Ingeniería, Bogotá. Obtenido de [https://repository.ucatolica.edu.co/bitstream/10983/17726/1/ASISTENTE%20VIRTUAL%20TIPO%20CHATBOT\\_final.pdf](https://repository.ucatolica.edu.co/bitstream/10983/17726/1/ASISTENTE%20VIRTUAL%20TIPO%20CHATBOT_final.pdf)

Garzón Castillo, G. L. (2017). *Watson analytics una solución cloud para procesos de big data. [Trabajo de grado de ingeniería, Universidad Libre]*. Repositorio institucional. Obtenido de <https://repository.unilibre.edu.co/bitstream/handle/10901/11219/Watson%20Analytics%2>

0una%20Soluci%c3%b3n%20Cloud%20para%20procesos%20de%20BigData.pdf?seque  
nce=1&isAllowed=y

Gaviria Giraldo, G. D. (2019). *Investigación de data lake Alcaldía de Envigado [Trabajo de grado de ingeniería, Institucion Universitaria de Envigado]*. Repositorio institucional. Obtenido de <http://bibliotecadigital.iue.edu.co/xmlui/handle/20.500.12717/491>

Giraldo Giraldo, F., Guzmán Luna, J., & Ovalle Carranza, D. (2005). Sistema multiagente para la gestion de material bibliografico. *Revista Colombiana de Tecnologia de Avanzada*, 2(6). Obtenido de [http://www.unipamplona.edu.co/unipamplona/portalIG/home\\_40/recursos/01\\_general/revista\\_6/13102011/17.pdf](http://www.unipamplona.edu.co/unipamplona/portalIG/home_40/recursos/01_general/revista_6/13102011/17.pdf)

Giraldo Zuluaga, W. A. (2019). *La minería de de texto como herramienta para la identificación del perfil investigativo: Caso gestion del conocimiento en America Latina de 2012 a 2018 [Trabajo de grado de ingeniería, Universidad Cooperativa de Colombia]*. Repositorio Institucional. Obtenido de [https://repository.ucc.edu.co/bitstream/20.500.12494/7902/1/2019\\_investigacion\\_gestion\\_conocimiento.pdf](https://repository.ucc.edu.co/bitstream/20.500.12494/7902/1/2019_investigacion_gestion_conocimiento.pdf)

Gonzales Galvis, J. L., & Alexis, P. A. (2015). *Diseño e implementación de un sistema de reconocimiento de naranjas para el robot gio 1 usando visión asistida por computador. [Trabajo de grado de ingeniería, Universidad Catolica de Colombia]*. Repositorio institucional. Obtenido de <https://repository.ucatolica.edu.co/bitstream/10983/2461/1/DISE%c3%91O%20E%20IMPLEMENTACI%c3%93N%20DE%20UN%20SISTEMA%20DE%20RECONOCIMIEN>



TO%20DE%20NARANJAS%20PARA%20EL%20ROBOT%20GIO%201%20USAND  
O%20VISI%c3%93N%20ASISTIDA%20POR%20COMPUTADOR.pdf

González Larin, Y. G. (2017). *El internet de las cosas y sus riesgos para la privacidad. [Tesis de especialización, Universidad Piloto de Colombia]*. Repositorio institucional. Obtenido de <http://repository.unipiloto.edu.co/handle/20.500.12277/2681>

Grandez Marquez, M. A. (2017). *Aplicacion de mineria de datos para determinar patrones de consumo en futuros clientes de una distribuidora de suplementos nutricionales*. Trabajo de grado, Universidad San Ignacio de Loyola, Facultad de Ingenieria., Lima. Obtenido de [http://repositorio.usil.edu.pe/bitstream/USIL/2763/1/2017\\_Granda\\_Aplicacion-de-mineria-datos.pdf](http://repositorio.usil.edu.pe/bitstream/USIL/2763/1/2017_Granda_Aplicacion-de-mineria-datos.pdf)

Han, J., & Kamber, P. J. (s.f.). *Data Mining. Concepts and Techniques* (Vol. 3). Obtenido de <http://myweb.sabanciuniv.edu/rdehkharghani/files/2016/02/The-Morgan-Kaufmann-Series-in-Data-Management-Systems-Jiawei-Han-Micheline-Kamber-Jian-Pei-Data-Mining.-Concepts-and-Techniques-3rd-Edition-Morgan-Kaufmann-2011.pdf>

Hoyos Monroy, J. D. (2019). *Prototipo de software para el reconocimiento facial por medio de visión artificial usando la metodología de análisis EigenFaces [Trabajo de grado de ingeniería, Universidad Tecnológica del Pereira]*. Repositorio institucional. Obtenido de <http://repositorio.utp.edu.co/dspace/bitstream/handle/11059/10960/T006.2483995%20H868.pdf?sequence=1&isAllowed=y>

Hueso, A., & Cascant, J. (2012). *Metodologías y técnicas cuantitativas de la investigación* (Primera ed.). Valencia, España: Unviersitat Politecnica de Valencia. Recuperado el Agosto de 2020, de <https://riunet.upv.es/bitstream/handle/10251/17004/Metodolog%C3%ADa%20y%20t%C>

3%A9cnicas%20cuantitativas%20de%20investigaci%C3%B3n\_6060.pdf?sequence=3&isAllowed=y

Iglesias Martínez, S., & Iliás del Río, J. D. (2019). *Plataforma adaptativa para refuerzo de sumas utilizando redes neuronales y sistemas de reputación bayesiano*. [Trabajo de grado de ingeniería, Universidad del Norte]. Repositorio institucional. Obtenido de <http://manglar.uninorte.edu.co/handle/10584/8474>

Jaimes Pardo, L. F. (2020). *Estudio Para Modelar Matemáticamente El Comportamiento De Un Robot Lego Ev3 Con Python*. [Trabajo de grado de ingeniería, Universidad de Cundinamarca]. Repositorio institucional. Obtenido de <http://repositorio.ucundinamarca.edu.co/bitstream/handle/20.500.12558/3012/ESTUDIO%20PARA%20MODELAR.pdf?sequence=1&isAllowed=y>

Jiménez García, E. E., & Montoya Campo, J. D. (2017). *Sistema eléctrico de filtrado de información basado en inteligencia computacional para recomendación de atractivos turísticos del caribe colombiano*. [Trabajo de grado de ingeniería, Universidad de Cartagena]. Repositorio institucional. Obtenido de <https://repositorio.unicartagena.edu.co/handle/11227/5941>

Kaartemo, V., & Helkkula, A. (17 de Octubre de 2018). A Systematic Review of Artificial Intelligence and Robots in Value Co-creation: Current Status and Future Research Avenues. *4*(2), 228. Obtenido de <https://journals.sagepub.com/doi/10.1177/2394964318805625>

Larco Ampudia, E. A. (2015). *Uso del PSP (Personal Software Process) en el desarrollo de software*. [Trabajo de grado de ingeniería, Escuela Politécnica Nacional]. Repositorio

- institucional. Obtenido de <https://bibdigital.epn.edu.ec/bitstream/15000/345/1/CD-0760.pdf>
- Liu, R., Zio, E., Chen, X., & Yang, B. (s.f. de s.f. de 2018). Artificial intelligence for fault diagnosis of rotating machinery: A review. *Mechanical Systems and Signal Processing*, 108, 47. Obtenido de [https://www.researchgate.net/publication/326742898\\_Artificial\\_intelligence\\_for\\_fault\\_diagnosis\\_of\\_rotating\\_machinery\\_A\\_review](https://www.researchgate.net/publication/326742898_Artificial_intelligence_for_fault_diagnosis_of_rotating_machinery_A_review)
- Lozada Moreno, W. A., & Martínez Mendoza, H. F. (2018). *Apoyo en la toma de decisiones en los procesos de fijación de cuota alimentaria, soportado en tecnologías Cognitivas [Tesis de maestría, Escuela Colombiana de Ingeniería Julio Garavito]*. Repositorio institucional.
- Lozano Vásquez, C. A., & Suatena Cabrera, J. E. (2019). *Analítica de datos para el rendimiento en los cultivos de aguacate Hass en Colombia. [Tesis de maestría, Universidad Externado de Colombia]*. Repositorio institucional. Obtenido de <https://bdigital.uexternado.edu.co/handle/001/1892>
- Manterola, C., Astudillo, P., Arias, E., Claros, N., & MINCIR, G. (27 de Octubre de 2011). Revisiones sistematicas de la literatura. Que se debe saber acerca de ellas. *ELSEVIER*, 7. Recuperado el Agosto de 2020, de [https://www.researchgate.net/publication/251553765\\_Revisiones\\_sistematicas\\_de\\_la\\_literatura\\_Que\\_se\\_debe\\_saber\\_acerca\\_de\\_ellas](https://www.researchgate.net/publication/251553765_Revisiones_sistematicas_de_la_literatura_Que_se_debe_saber_acerca_de_ellas)
- Martínez Beltrán, J. P. (2017). *Conociendo los beneficios de la relación de Big Data y Cloud. [Trabajo de grado de ingeniería, Universidad Libre]*. Repositorio institucional.

Obtenido de <https://repository.unilibre.edu.co/bitstream/handle/10901/11216/IFM%20-%20Big%20Data.pdf?sequence=1&isAllowed=y>

Martinez Bielostotzky, D. A., & Rueda Iglesias, A. F. (2019). *Análisis y predicción de accidentes de tránsito en Barranquilla mediante técnicas de minería de datos [Trabajo de grado de ingeniería, Universidad del Norte]*. Repositorio institucional. Obtenido de <http://manglar.uninorte.edu.co/handle/10584/8445#page=1>

Martinez, A. M. (2011). *Conocimiento y base de datos: Una propuesta de integración inteligente*. Tesis Doctoral, Universidad de Cantabria, Departamento de organización de empresas. Obtenido de CONOCIMIENTO Y BASES DE DATOS: [https://www.tdx.cat/bitstream/handle/10803/31767/0de3.MAMprevio\\_cap1.pdf;jsessionid=7EE2D82882158E25BEE7C76CAE753CFE?sequence=1](https://www.tdx.cat/bitstream/handle/10803/31767/0de3.MAMprevio_cap1.pdf;jsessionid=7EE2D82882158E25BEE7C76CAE753CFE?sequence=1)

Mayorga García, G. (2017). *Uso de analíticas para predecir los computadores afectados por Malware, en una institución financiera en Colombia, 2017 [Trabajo de grado de ingeniería, Escuela Colombiana de Ingeniería Julio Garavito]*. Repositorio institucional. Obtenido de <https://repositorio.escuelaing.edu.co/handle/001/565>

Maza Figueroa, N. P., & Orozco Acevedo, M. A. (2017). *Modelo de gestión estratégica para el desarrollo de un campus inteligente basado en conceptos de Smart City en la Universidad de Cartagena - Campus Piedra de Bolívar. [Trabajo de grado de ingeniería, Universidad de Cartagena]*. Repositorio institucional. Obtenido de <https://repositorio.unicartagena.edu.co/handle/11227/5931>

Mendoza Ibarra, J. F. (2018). *Arquitectura de aplicaciones de software embebido en microcontroladores para tarjetas de captura de datos de la IOT. [Tesis de maestría, Universidad de San Buenaventura Colombia]*. Repositorio institucional. Obtenido de

[http://bibliotecadigital.usb.edu.co/bitstream/10819/5794/3/Arquitectura\\_software\\_embebido\\_Mendoza\\_2018.pdf](http://bibliotecadigital.usb.edu.co/bitstream/10819/5794/3/Arquitectura_software_embebido_Mendoza_2018.pdf)

Mendoza Ibarra, J. F. (2018). *Arquitectura de aplicaciones de software embebido en microcontroladores para tarjetas de captura de datos de la IOT. [Tesis de maestría, Universidad San Buenaventura Colombia]*. Repositorio institucional. Obtenido de <http://bibliotecadigital.usb.edu.co/handle/10819/5794>

Mina Trujillo, J. C. (2017). *Análisis de sentimientos en entornos de mercadeo móvil [Trabajo de grado de ingeniería, Unviersida de San Buenaventura Colombia]*. Repositorio institucional. Obtenido de <http://bibliotecadigital.usb.edu.co/handle/10819/5568>

Ministerio de comercio, industria y turismo. (27 de Junio de 2013). Decreto N° 1377 de 2013. Colombia. Obtenido de [https://www.mintic.gov.co/portal/604/articles-4274\\_documento.pdf](https://www.mintic.gov.co/portal/604/articles-4274_documento.pdf)

Miquelarena, R., Cerrolazaa, M., Febresa, J., & Priettoa, C. (2010). Sistema de vision artificial para el control de movimiento de un asistente robotico medico. *Mecanica Computacional, XXIX*, 6619-6629. Obtenido de <http://venus.ceride.gov.ar/ojs/index.php/mc/article/view/3473/3390>

Moine, J. M., Gordillo, S., & Haedo, A. S. (2011). Análisis comparativo de metodologías para la gestión de. *XVII CONGRESO ARGENTINO DE CIENCIAS DE LA COMPUTACIÓN*. Obtenido de [https://digital.cic.gba.gob.ar/bitstream/handle/11746/3516/11746\\_3516.pdf-PDFA.pdf?sequence=1&isAllowed=y](https://digital.cic.gba.gob.ar/bitstream/handle/11746/3516/11746_3516.pdf-PDFA.pdf?sequence=1&isAllowed=y)

Molina García, J. A. (2019). *La importancia de la gestión de riesgos y seguridad en el internet de las cosas (IOT). [Tesis de especialización, Universidad Piloto de Colombia]*.

- Repositorio institucional. Obtenido de  
<http://repository.unipiloto.edu.co/handle/20.500.12277/6754>
- Montano, J. (s.f. de s.f. de s.f.). *Lifeder*. Obtenido de <https://www.lifeder.com/investigacion-no-experimental/>
- Montenegro Bermudez, A. F., & Parada Rojas, C. D. (2015). *Diseño e implementación de un sistemas de detección de malezas en cultivos cundiboyacenses [Trabajo de grado de ingeniería, Universidad Católica de Colombia]*. Repositorio institucional. Obtenido de <https://repository.ucatolica.edu.co/bitstream/10983/3202/4/DISE%c3%91O%20E%20IMPLEMENTACI%c3%93N%20DE%20UN%20SISTEMA%20DE%20DETECCI%c3%93N%20DE%20MALEZAS%20EN%20CULTIVOS%20CUNDIBOYACENSES.pdf>
- Morales Guerrero, F. (2015). *Caracterización y segmentación de público objetivo para Scouting universitario a partir de minería de datos [Trabajo de grado de ingeniería, Universidad de los Andes]*. Repositorio institucional. Obtenido de <https://repositorio.uniandes.edu.co/flexpaper/handle/1992/17346/u713441.pdf?sequence=1&isAllowed=y#page=1>
- Moreno, R. (2019). La llegada de la inteligencia artificial a la educación. *Revista de Investigación en Tecnologías de la Información (RITI)*, 7(14), 270. Obtenido de <https://dialnet.unirioja.es/descarga/articulo/7242777.pdf>
- Mosquera Mosquera, A. S., & Martínez Rendon, J. P. (2018). *Reconocimiento óptico de caracteres en placas vehiculares haciendo uso de redes neuronales convolucionales. [Trabajo de grado de ingeniería, Unviersidad Tecnológica de Pereira]*. Repositorio institucional. Obtenido de

<http://repositorio.utp.edu.co/dspace/bitstream/handle/11059/10027/T005.12%20M912.pdf?sequence=1&isAllowed=y>

Muñoz Garcia, G. I. (2017). *Implementación de un sistema de captura de gestos usando un leap motion y redes neuronales para su clasificación. [Trabajo de grado de ingeniería, Universidad Tecnológica de Pereira]*. Repositorio institucional. Obtenido de <http://repositorio.utp.edu.co/dspace/bitstream/handle/11059/8948/T006.32%20M971.pdf?sequence=1&isAllowed=y>

Murcia Jimenez, H. D. (2020). *Sistema de visión artificial para detección y ubicación del golpe de un balón en superficie. [Trabajo de grado de ingeniería, Univesidad de Cundinamarca]*. Repositorio institucional. Obtenido de <http://repositorio.ucundinamarca.edu.co/bitstream/handle/20.500.12558/2921/MurciaHernan.pdf?sequence=1&isAllowed=y>

Nieto Acevedo, Y. V. (2016). *Modelo de un sistema de software basado en las técnicas de Learning Analytics como herramienta de apoyo en la toma de decisiones académico-administrativas en las instituciones públicas de educación superior. [Tesis de maestría, Universidad Distrital Francisco Jose de Caldas]* Repositorio institucional. Obtenido de <http://repository.udistrital.edu.co/handle/11349/2772>

Novoa Urrego, G. A. (2017). *Aplicación de Big Data e inteligencia de mercado en institución educativa tecnológica. (Corporación Tecnológica de Bogotá). [Trabajo de grado de ingeniería, Universidad Libre]*. Repositorio institucional. Obtenido de <https://repository.unilibre.edu.co/bitstream/handle/10901/11210/APLICACION%20DE%20BIG%20DATA%20E%20INTELIGENCIA%20DE%20MERCADERO%20EN>

%20INSTITUCI%c3%93N%20EDUCATIVA%20TECNOL%c3%93GICA%20%28CORPORAC2.pdf?sequence=1&isAllowed=y

- Ohlhorst, F. (2013). *BIG DATA Turning big data into big money*. New Jersey: John Wiley & Sons, Inc. Obtenido de [http://www.bdbanalytics.ir/media/1121/big-data-analytics\\_turning-big-data-into-big-money.pdf](http://www.bdbanalytics.ir/media/1121/big-data-analytics_turning-big-data-into-big-money.pdf)
- Ordoñez Erazo, H. A. (2018). *Un middleware para la gestión labores agrícolas en un ambiente de internet de las cosas. [Tesis de maestría, Universidad de San Buenaventura Colombia]*. Repositorio institucional. Obtenido de <http://bibliotecadigital.usb.edu.co/handle/10819/6650>
- Osma-Molina, J. J., & Villamizar-Monar, D. G. (2020). *Diseño de un sistema IoT para la medición de temperatura, humedad y luz en huertas urbanas. [Trabajo de grado de ingeniería, Universitaria Uniagustiniana]*. Repositorio institucional. Obtenido de <http://repositorio.uniagustiniana.edu.co/handle/123456789/1349>
- Osorio Salcedo, K. P. (2016). *Design and Implementation of a Data Analysis Platform in the Educational Sector [Trabajo de grado de ingeniería, Universidad del Norte]*. Repositorio institucional. Obtenido de <http://manglar.uninorte.edu.co/handle/10584/5860?show=full>
- Ospino Pinedo, M. E. (2017). *Modelo de gestión estratégica para el desarrollo de un campus inteligente basado en conceptos de smart city en la universidad de Cartagena - campus piedra de bolívar. [Trabajo de grado de ingeniería, Universidad de Cartagena]*. Repositorio institucional. Obtenido de [https://repositorio.unicartagena.edu.co/bitstream/handle/11227/5931/INFORME%20FIN](https://repositorio.unicartagena.edu.co/bitstream/handle/11227/5931/INFORME%20FINAL.pdf?sequence=1&isAllowed=y)  
AL.pdf?sequence=1&isAllowed=y



- Otzen, T., & Manterola, C. (s.f. de s.f. de 2017). Técnicas de Muestreo sobre una Población a Estudio. 35(1), 232. Recuperado el Agosto de 2020, de <https://scielo.conicyt.cl/pdf/ijmorphol/v35n1/art37.pdf>
- Oxford Insights. (s.f. de s.f. de 2019). *Oxford Insights*. Obtenido de <https://www.oxfordinsights.com/ai-readiness2019>
- Pachon Espinel, D. L. (2018). *Prototipo de sistema automatizado con vision artificial para la seleccion de empaques de plastico, vidrio y lata en el proceso de reciclaje [Trabajo de grado de ingenieria, Universidad de Cundinamarca Exension Chia]*. Repositorio institucional. Obtenido de <http://repositorio.ucundinamarca.edu.co/bitstream/handle/20.500.12558/2497/Trabajo%20de%20grado.pdf?sequence=3&isAllowed=y>
- Paramo Lozada, J. P., & Espitia Betancourt, C. A. (2018). *Aplicación del aprendizaje automático en la clasificación de textos cortos: un caso de estudio en el conflicto armado colombiano. [J]*. Repositorio Institucional. Obtenido de <https://repository.ucatolica.edu.co/bitstream/10983/22546/1/APLICACION-DEL-APRENDIZAJE-AUTOMATICO-EN-LA-CLASIFICACION-DE-TEXTOS-CORTOS-UN-CASO-DE-ESTUDIO-EN-.pdf>
- Pignani, J. M. (s.f.). *Sistemas expertos*. Obtenido de [https://www.frro.utn.edu.ar/repositorio/catedras/quimica/5\\_ano/orientadora1/monograias/pignani-sistemasexpertos.pdf](https://www.frro.utn.edu.ar/repositorio/catedras/quimica/5_ano/orientadora1/monograias/pignani-sistemasexpertos.pdf)
- Pinea Cortes, S., & Pinzon Rivera, R. (2015). *Optimización en la distribución de rutas de colegios privados en Pereira con ayuda de algoritmos de inteligencia artificial. [Trabajo de grado de ingenieria, Universidad Tecnologica del Pereira]*. Repositorio institucional.

Obtenido de Optimización en la distribución de rutas de colegios privados en Pereira con ayuda de algoritmos de inteligencia artificial.

Pinzón Ortiz, J. E. (2019). *Modelo de interacción inteligente del reglamento del Congreso de Colombia para dispositivos móviles "MIRC"* [Tesis de maestría, Universidad Distrital Francisco Jose de Caldas]. Repositorio institucional. Obtenido de <http://repository.udistrital.edu.co/handle/11349/15362>

Porras Garcia, Y. F. (2018). *Análisis comparativo del desempeño y costo computacional de una infraestructura de almacenamiento y procesamiento distribuido para el procesamiento de colecciones de texto.* [Trabajo de grado de ingeniería, Universidad de los Llanos]. Repositorio institucional. Obtenido de

[https://repositorio.unillanos.edu.co/bitstream/001/1139/1/RUNILLANOS%20SIS%200301%20ANALISIS%20COMPARATIVO%20DEL%20DESEMPE%  
c3%91O%20Y%20COSTO%20COMPUTACIONAL%20DE%20UNA%20INFRAESTRUCTURA%20DE%20ALMACENAMIENTO%20Y%20PROCESAMIENTO%20DISTRIBUIDO%20PARA%20EL%](https://repositorio.unillanos.edu.co/bitstream/001/1139/1/RUNILLANOS%20SIS%200301%20ANALISIS%20COMPARATIVO%20DEL%20DESEMPE%c3%91O%20Y%20COSTO%20COMPUTACIONAL%20DE%20UNA%20INFRAESTRUCTURA%20DE%20ALMACENAMIENTO%20Y%20PROCESAMIENTO%20DISTRIBUIDO%20PARA%20EL%20)

Prieto Alvarez, S. N. (2018). *Modelo de evaluacion de impacto de los programas de pregrado y postgrado de la facultad de ingenieria de la Universidad Catolica de Colombia,* [Trabajo de grado de ingenieria, Universidad Catolica de Colombia]. Repositorio Institucional.

Obtenido de

[https://repository.ucatolica.edu.co/bitstream/10983/16214/1/Modelo%20de%20evaluaci%  
c3%b3n%20de%20impacto-%20Nicol%20Sebasti%  
c3%a1n%20Prieto%20%  
c3%81lvarez-538868.pdf](https://repository.ucatolica.edu.co/bitstream/10983/16214/1/Modelo%20de%20evaluaci%c3%b3n%20de%20impacto-%20Nicol%20Sebasti%c3%a1n%20Prieto%20%c3%81lvarez-538868.pdf)

- Prieto Arciniegas, C. E. (2015). *Uso de regresión logística para predecir deserción estudiantil temprana [Trabajo de grado de ingeniería, Universidad de los Andes]*. Repositorio institucional. Obtenido de <https://repositorio.uniandes.edu.co/handle/1992/18302>
- Quiroga Saavedra, L. F. (2016). *Prototipo de tutor inteligente para el aprendizaje de la programación de computadores. [Trabajo de grado de ingeniería, Universidad Católica de Colombia]*. Repositorio institucional. Obtenido de <https://repository.ucatolica.edu.co/bitstream/10983/14968/1/PROTOTIPO%20DE%20UN%20SISTEMA%20TUTOR%20INTELIGENTE.pdf>
- Ramírez Rincón, C. A. (2015). *Perfil de identificación de dolor en neonatos a partir de variables fisiológicas [Trabajo de grado de ingeniería, Universidad de los Andes]*. Repositorio institucional. Obtenido de <https://repositorio.uniandes.edu.co/flexpaper/handle/1992/18359/u722062.pdf?sequence=1&isAllowed=y#page=1>
- Rieder, R., & Patricio, D. I. (Octubre de 2018). Computer vision and artificial intelligence in precision agriculture for grain crops: A systematic review. *Computers and Electronics in Agriculture*, 153, 81. Obtenido de <https://www.sciencedirect.com/science/article/abs/pii/S0168169918305829>
- Rodriguez Saldaña, O. E. (2017). *La ventaja competitiva del Big Data en el marketing empresarial y la innovación de productos. [Monografía de ingeniería, Universidad Libre]*. Repositorio institucional. Obtenido de <https://repository.unilibre.edu.co/bitstream/handle/10901/11209/Monografia%20Big%20Data%20->

%20Oscar%20Eduardo%20Rodriguez%20Salda%c3%b1a.pdf?sequence=1&isAllowed=y

- Rojas Gutierrez, E. A., & Aguilar, J. S. (2017). *Mineria de datos para el descubrimiento de patrones en enfermedades respiratorias en Bogota, Colombia [Trabajo de grado de ingenieria, Unviersidad Catolica de Colombia]*. Repositorio institucional. Obtenido de <https://repository.ucatolica.edu.co/bitstream/10983/15329/1/Trabajo%20de%20grado.pdf>
- Romero Castillo, J. a. (2018). *BIG DATA IFCT12890* (Primera ed.). Malaga, Antequera, España: IC Editorial. Recuperado el Agosto de 2020, de BIG DATA IFCT12890: [https://sibdigital.ufpso.edu.co:2303/es/ereader/ufpso/124254?as\\_all=Robotica&as\\_all\\_op=unaccent\\_\\_icontains&prev=as](https://sibdigital.ufpso.edu.co:2303/es/ereader/ufpso/124254?as_all=Robotica&as_all_op=unaccent__icontains&prev=as)
- Romero Guzman, J. E., Gonazalez Castro, C. A., & Jaime Fuentes, D. M. (2017). *Mecanismo de monitoreo para bicicletas en movimiento, usando enfoque de IOT. [Trabajo de grado de ingenieria, Universitaria Agustiniana]*. Repositorio institucional. Obtenido de <http://repositorio.uniagustiniana.edu.co/handle/123456789/220>
- Romero Moreno, F. Y. (2019). *Las redes sociales como factor de predicción de resultados electorales [Tesis de maestria, Universidad Distrital Francisco José de Caldas]*. Repositorio institucional. Obtenido de <http://repositorio.udistrital.edu.co/handle/11349/25142>
- Romero Romero, C. A. (2018). *Estudio comparativo de algoritmos de inteligencia artificial y mineria de datos enfocados a la toma de decisiones empresariales de seleccion de personal. [Trabajo de grado de ingenieria, Universidad de Cundinamarca]*. Repositorio institucional. Obtenido de <http://repositorio.ucundinamarca.edu.co/handle/20.500.12558/1086>

- Romero Torres, S. (02 de 07 de 2020). *IMPACTOTIC*. Obtenido de <https://impactotic.co/los-5-paises-lideres-en-la-cuarta-revolucion-industrial-colombia-es-el-56/>
- Romero, J., Dafonte, C., Gomez, A., & Penousal, F. (2007). *Inteligencia artificial y computacion avanzada*. (F. A. Brañas, Ed.) Fundacion Alfredo Brañas. Obtenido de INTELIGENCIA ARTIFICIAL Y: <https://cdv.dei.uc.pt/wp-content/uploads/2014/03/ms07.pdf>
- Rory J, P. (s.f. de s.f. de 2013). How to write a systematic literature review: a guide for medical students. Edimburgo, Escocia. Recuperado el Agosto de 2020, de How to write a systematic literature review:: <https://sites.cardiff.ac.uk/curesmed/files/2014/10/NSAMR-Systematic-Review.pdf>
- Rubio Tapias, J. A., & Alba Rodriguez, C. A. (2018). *Desarrollo de componente web parametrizable de clustering para análisis de datos*. [Trabajo de grado de ingeniería, Universidad Catolica de Colombia]. Repositorio institucional. Obtenido de [https://repository.ucatolica.edu.co/bitstream/10983/23778/1/DESARROLLO%20DE%20COMPONENTE%20WEB%20PARAMETRIZABLE%20DE%20CLUSTERING%20PARA%20AN%cc3%81LISIS%20DE%20DATOS\\_625178\\_625167\\_20.pdf](https://repository.ucatolica.edu.co/bitstream/10983/23778/1/DESARROLLO%20DE%20COMPONENTE%20WEB%20PARAMETRIZABLE%20DE%20CLUSTERING%20PARA%20AN%cc3%81LISIS%20DE%20DATOS_625178_625167_20.pdf)
- Rueda, F. (1993). *La inteligencia artificial: Su identidad, sus conflictos, sus potencialidades y limitaciones*.
- Salazar Cardenas, D., & Vargas Trujillo, I. A. (2017). *Diseño e implementación de un sistema creador de arte digital basado en el concepto de cadáver exquisito utilizando inteligencia artificial*. [Trabajo de grado de ingeniería, Universidad Autonoma de Occidente]. Repositorio institucional. Obtenido de <https://red.uao.edu.co/bitstream/10614/9866/1/T07534.pdf>

- Salcedo Toro, D. F. (2015). *Tendencias y características de los viajeros que visitan la ciudad de Pereira por medio de técnicas de minería de datos [Tesis de maestría, Universidad Autónoma de Manizales]*. Repositorio institucional. Obtenido de <http://repositorio.autonoma.edu.co/jspui/handle/11182/630>
- Salgado Jaime, C. A. (2017). *Consideraciones generales para la aplicación del Big Data en negocios de redes de mercado. [Monografía de ingeniería, Universidad Libre]*. Repositorio institucional. Obtenido de <https://repository.unilibre.edu.co/bitstream/handle/10901/11212/CONSIDERACIONES%20GENERALES%20PARA%20LA%20APLICACION%20DEL%20BIG%20DATA%20EN%20NEGOCIOS%20DE%20REDES%20DE%20MERCADO.pdf?sequence=1&isAllowed=y>
- Sánchez Pérez, A. P. (2020). *Mapa de la situación académica colombiana a través del análisis de las bases de datos del ICFES [Tesis de maestría, Universidad Autónoma de Bucaramanga]*. Repositorio institucional. Obtenido de <https://repository.unab.edu.co/handle/20.500.12749/11937>
- Sánchez Prieto, K. S., & Alvarado Pérez, K. J. (s.f.). *ECIBOT: Agente inteligente para la interacción en lenguaje natural. [Trabajo de grado de ingeniería, Escuela Colombiana de Ingeniería Julio Garavito]*. Repositorio institucional. Obtenido de <https://repositorio.escuelaing.edu.co/bitstream/001/780/1/Alvarado%20Kevin%20Julian%202018.pdf>
- Sánchez, M. F., Millán Rodríguez, F., Salvador Bayarri, J., Palou Redorta, J., Rodríguez Escovar, F., Esquena Fernández, S., & Villavicencio Mavrich, H. (2007). *Historia de la*

- robótica: de Arquitas de Tarento al robot (Parte I). *Actas de Urología Españolas*.  
Obtenido de <https://scielo.isciii.es/pdf/aue/v31n2/original1.pdf>
- Sarmiento Martínez, C. A. (2015). *Motores de riesgo transaccionales ¿solución eficaz para reducir agresivamente el fraude electrónico? [Tesis de especialización, Universidad Piloto de Colombia]*. Repositorio institucional. Obtenido de <http://repository.unipiloto.edu.co/handle/20.500.12277/2921>
- Segura Gazcón, M. Á., & Velásquez Hernández, R. R. (2017). *Arquitectura de un sistema de telemonitorización para hospitalización domiciliaria de adultos mayores apoyada en tecnologías de Internet de las cosas (IOT)*. [Trabajo de grado de ingeniería, Universidad de Cartagena]. Repositorio institucional. Obtenido de <https://repositorio.unicartagena.edu.co/handle/11227/5943>
- Serrano Gonzales, D. (2015). *Predicción de compra sobre consultas de información en comercio electrónico [Trabajo de grado de ingeniería, Universidad de los Andes]*. Repositorio institucional. Obtenido de <https://repositorio.uniandes.edu.co/flexpaper/handle/1992/18301/u721905.pdf?sequence=1&isAllowed=y#page=1>
- Shai Shalev, S., & Shai Ben, D. (2014). *Understanding Machine Learning: From theory to algorithms*. Canada. Obtenido de <https://www.cs.huji.ac.il/~shais/UnderstandingMachineLearning/understanding-machine-learning-theory-algorithms.pdf>
- Silva Monsalve, A. E. (2018). *Determinación y post-procesamiento de perfiles con técnicas no supervisadas de minería de datos como soporte a la toma de decisiones en la Estrategia Unidos para la Superación de la Pobreza en Colombia*. [Tesis de maestría, Escuela

- Colombiana de Ingenieria Julio Gravito]. Obtenido de  
<https://repositorio.escuelaing.edu.co/handle/001/783>
- SM, R., AP, B., J, T., & AD., O. (20 de Agosto de 2009). Systematic review of dermoscopy and digital dermoscopy/ artificial intelligence for the diagnosis of melanoma. *British Journal of Dermatology*, 161(3), 604. Obtenido de <https://pubmed.ncbi.nlm.nih.gov/19302072/>
- Tatiana, M. R., & Arboleda Diaz, J. N. (2017). *Machine learning applied to threat intelligence (aprendizaje de máquinas aplicado para amenazar la inteligencia)*. [Trabajo de grado de ingeniería, Escuela Colombiana de Ingeniería Julio Garavito]. Repositorio institucional. Obtenido de  
<https://repositorio.escuelaing.edu.co/bitstream/001/702/2/Arboleda%20D%c3%adaz%2c%20Joseph%20Nikolay%20-%202017.pdf>
- Tavera Sanchez, D. (2015). *Diseño e implementacion de agentes para la plataforma juego del agua*. [Trabajo de grado de ingeniería, Universidad de los Andes]. Repositorio institucional. Obtenido de  
<https://repositorio.uniandes.edu.co/bitstream/handle/1992/17617/u713934.pdf?sequence=1&isAllowed=y>
- Tello Castañeda, M. L. (2016). *Modelo predictivo para la determinación del nivel de riesgo de deserción estudiantil en tecnología en electrónica de La Universidad Distrital empleando redes neuronales artificiales* [Tesis de maestría, Universidad Distrital Francisco Jose de Caldas]. Repositorio institucional. Obtenido de  
<https://repository.udistrital.edu.co/handle/11349/2769?show=full>
- Toledo Cubillos, E. (2016). *Diseño de un sistema de soporte de decisiones para control de inundaciones en áreas urbanas mediante algoritmos genéticos NSGA-II. Evolutivos*



- basado en preferencias y redes neuronales. [Tesis de maestría, Universidad Santiago de Cali]. Repositorio institucional. Obtenido de*  
<https://repository.usc.edu.co/handle/20.500.12421/2219>
- Torres Robles, E. A., & Rincon Saavedra, W. A. (2017). *Implementación de una herramienta de autogestión y autoconfiguración para la implementación de servicios en proyectos de big data. [Trabajo de grado de ingeniería, Universidad Católica de Colombia]. Repositorio institucional. Obtenido de*  
<https://repository.ucatolica.edu.co/bitstream/10983/18238/1/HERRAMIENTA%20DE%20AUTOGESTION%20PARA%20BIG%20DATA.pdf>
- Torres Samboni, L. A. (2015). *Análisis de sentimientos sobre el posconflicto colombiano utilizando herramientas de minería de texto, [Trabajo de maestría, Escuela Colombiana de Ingeniería Julio Garavito]. Repositorio institucional. Obtenido de*  
<https://repositorio.escuelaing.edu.co/bitstream/001/403/1/Torres%20Samboni%2c%20Lina%20Andrea%20-%202016.pdf>
- Valencia Payan, C. H. (2018). *Modelo de pronóstico de escasez y sobreoferta de agua en una cuenca altoandina basado en análisis Big Data. [Tesis de maestría, Universidad del Cauca]. Repositorio institucional. Obtenido de*  
<http://repositorio.unicauca.edu.co:8080/bitstream/handle/123456789/1370/MODELO%20DE%20PRON%26%93STICO%20DE%20ESCASEZ%20Y%20SOBREFERTA%20DE%20AGUA%20EN%20UNA%20CUENCA%20ALTOANDIDA%20BASADO%20EN%20AN%26%81LISIS%20BIG%20DATA.pdf?sequence=1&isAllowed=y>
- Valiente Cristancho, A. (2017). *Integración de internet de las cosas en sistema embebido system on chip, con aplicación a domótica. [Tesis de maestría, Universidad Distrital Francisco*

- Jose de Caldas*]. Repositorio institucional. Obtenido de  
<http://repository.udistrital.edu.co/handle/11349/7281>
- Vargas Antolínez, E. A. (2018). *Modelado basado en datos para la clasificación semiautomática de correspondencia electrónica: caso de estudio para la Administración Pública Colombiana [Tesis de maestría, Escuela Colombiana de Ingeniería Julio Garavito]*. Repositorio institucional. Obtenido de  
<https://repositorio.escuelaing.edu.co/handle/001/829>
- Vargas Orozco, C. A. (2015). *Un modelo de credit scoring utilizando inteligencia artificial para el fondo de empleados de salud en Risaralda – fesser*. Repositorio institucional. Obtenido de  
<http://repositorio.utp.edu.co/dspace/bitstream/handle/11059/6111/0063V297.pdf?sequence=1&isAllowed=y>
- Viviescas Sandoval., J. N. (2019). *Sistema para el reclutamiento y selección de personal. [Trabajo de grado de ingeniería, Institución universitaria de Envigado]*. Repositorio institucional. Obtenido de  
<http://bibliotecadigital.iue.edu.co/xmlui/handle/20.500.12717/535>
- Zabala Jaramillo, L. A. (2016). *Gestión de la seguridad en el internet de las cosas. [Tesis de especialización, Universidad Piloto de Colombia]*. Repositorio institucional. Obtenido de  
<http://repository.unipiloto.edu.co/handle/20.500.12277/2721>
- Zapata Garrido, L., & Diaz Mojica, H. (s.f. de s.f. de 2007). *Predicción del tipo de cambio peso-dolar utilizando redes neuronales artificiales*. Tesis de Maestría, Universidad del Norte, Facultad de Ciencias Administrativas, Barranquilla. Obtenido de  
<https://pdfs.semanticscholar.org/e2bd/c3a810fbae004d4b5b072e60caf413ea39f0.pdf>

Zuluaga Agudelo, L. G. (2019). *Propuesta pedagógica para incentivar el pensamiento científico y creativo a través de la robótica en niños y niñas de 3 a 6 años del colegio Carlos Castro Saavedra. [Trabajo de grado de ingeniería, Universidad Libre].* Repositorio institucional. Obtenido de <https://repository.unilibre.edu.co/bitstream/handle/10901/17292/PROPUESTA%20PEDAG%c3%93GICA%20PARA%20INCENTIVAR%20EL%20PENSAMIENTO%20CIENT%c3%8dFICO.pdf?sequence=1&isAllowed=y>

Zuluaga Valencia, C. D. (2017). *Predicción del tiempo de la estancia hospitalaria de pacientes con trastorno afectivo bipolar en la clínica San Juan de Dios de Manizales por medio de técnicas de minería de datos [Tesis de maestría, Universidad Autónoma de Manizales].* Repositorio institucional. Obtenido de <http://repositorio.autonoma.edu.co/jspui/handle/11182/696>

## **Apéndices.**



