



UNIVERSIDAD JOSÉ ANTONIO PÁEZ

**SISTEMA DE INFORMACIÓN PARA LA
PREDICCIÓN DE UN TARGET OBJETIVO
MEDIANTE EL CONSUMO DE UN
RENGLÓN CON BASE A UNA RED
NEURONAL.**

Autores:

Leopoldo J. Estrada P.

Alan O. Vásquez I.

Tutor: MSc. Jetro López

Urb. Yuma II, calle N° 3. Municipio San Diego

Teléfono: (0241) 8714240 (master) – Fax: (0241) 8712394



REPÚBLICA BOLIVARIANA DE VENEZUELA
UNIVERSIDAD JOSÉ ANTONIO PÁEZ
FACULTAD DE INGENIERÍA
ESCUELA DE COMPUTACIÓN

**SISTEMA DE INFORMACIÓN PARA LA
PREDICCIÓN DE UN TARGET OBJETIVO
MEDIANTE EL CONSUMO DE UN RENGLÓN
CON BASE A UNA RED NEURONAL.**

**Trabajo de grado presentado como requisito para optar al título de
INGENIERO EN COMPUTACIÓN**

Autores:
Leopoldo J. Estrada P.
C.I. V-25.754.287
Alan O. Vásquez I.
C.I. V-28.414.284
Tutor: MSc. Jetro López

San Diego, Junio de 2020



FI-C-013-2020-1CR (TG)

Valencia, 19 de junio de 2020

Ciudadanos:

Estrada P., Leopoldo J. C.I. V- 25.754.287

Vásquez I., Alan O. C.I. V-28.414.284

Presente-

Cumplo con informarle que la Comisión de Trabajo de Grado y Pasantías de la Facultad de Ingeniería en su reunión N° 04-2020 de fecha 13-02-2020 aprobó el proyecto de trabajo de grado titulado *SISTEMA DE INFORMACIÓN PARA LA PREDICCIÓN DE UN TARGET OBJETIVO MEDIANTE EL CONSUMO DE UN RENGLÓN CON BASE A UNA RED NEURONAL* presentado por usted (es) como requisito para optar al título de Ingeniero en Computación.

Se ratifica la designación del Ing. Jetro López C.I: 8.779.723 como Tutor Académico que los asesorara en el desarrollo de este proyecto.

Atentamente,



Prof. Luís Lira

Decano de la Facultad de Ingeniería

c.c. Coordinación de Pasantías y Trabajo de Grado (1).

Ll/a.a.



REPÚBLICA BOLIVARIANA DE VENEZUELA

UNIVERSIDAD JOSÉ ANTONIO PÁEZ

FACULTAD DE INGENIERÍA

ESCUELA DE COMPUTACIÓN

APROBACIÓN DEL TUTOR

Quien suscribe, Msc. Jetro López portador de la cédula de identidad N°, 8.779.723. En mi carácter de tutor del trabajo de grado presentado por el ciudadano Leopoldo José Estrada Pérez portador de la cédula de identidad N° 25.754.287 y el ciudadano Alan Omar Vásquez Izaguirre portador de la cedula de identidad N° 28.414.284, titulado **SISTEMA DE INFORMACIÓN PARA LA PREDICCIÓN DE UN TARGET OBJETIVO MEDIANTE EL CONSUMO DE UN RENGLÓN CON BASE A UNA RED NEURONAL**. Presentado como requisito parcial para optar al título de Ingeniero de Computación, considero que dicho trabajo reúne los requisitos y méritos suficientes para ser sometido a la presentación pública y evaluación por parte del jurado examinador que se designe.

En San Diego, a los veinticinco días del mes de junio del año dos mil veinte.

MSc. Jetro López

C.I: V- 8.779.723.

*A mi madre Dilia Pérez, quien me alienta todos los días para que no me
rinda y luche por mis sueños.*

*A mi abuela Gisela de Pérez, quien porta el corazón más noble del mundo y
quien me enseñó que para amar no se necesita más que un buen corazón.*

*A mi tía Betty Pérez, quien era mi segunda madre y quien hubiese dado todo
por estar a mi lado en todos mis triunfos.*

*A todos aquellos que creyeron y en los que no, en mí, porque me impulsaron
a superarme ante todas las adversidades.*

Leopoldo J. Estrada P.

Agradecimientos

A mis padres, Leopoldo y Dilia, quienes me han dado todo el apoyo posible para lograr mis objetivos. Agradecido eternamente por los valores y principios que me inculcaron y agradecido por todos sus sacrificios por asegurar mi futuro. Por tanto, espero mejorar cada día más, para que estén más orgullosos de mí.

A mis hermanos, Alejandro y Alejandra, por adoptar el rol de padres para cuidarme en todo momento, guiarme y estar siempre pendiente para cuando necesite su apoyo.

A mi abuela Gisela, quien sin importar las circunstancias y la distancia, siempre está ahí para brindar amor. A mi tía Betty P. porque sé que en ese hermoso lugar llamado cielo, me cuidas y me proteges.

A mi abuelo Leopoldo, quien me brindó su apoyo desde el inicio de mi carrera y quien me enseña el valor más importante, que es el respeto.

A mis Tíos, Miguel, Gisela, Janeth, Lucy, Braddy, a mis primos Andrea, Miguel, Diego, Sabrina, y a mis cuñados Glory y Carlos por ser parte de muchas de mis alegrías y que a pesar de la distancia, siempre estaremos unidos.

A mi compañero Alan, quien me soporto en todo este recorrido, discusiones, trasnochos y momentos únicos por el cual no me arrepiento hacer este proyecto con él, verdad tiene toda mi admiración.

A mi tutor Jetro López, agradecido por la hermosa experiencia de que fuese mi profesor, tutor, jefe, compañero y buen amigo, gracias por tantas buenas experiencias.

A mis hermanos de la vida, Daylibeth y José David, agradecido eternamente por su eterno apoyo, por tolerarme todos los días y por estar en todos esos momentos buenos, malos y regulares, me faltan las palabras para agradecerles por tanto.

A mis amigos y compañeros, Ely, Estefanía, Miguel G, Bryan, Rashel, Luis B, Nicol, Dimitri, Alan, Suljany, Miguel O., quienes me estuvieron siempre presentes.

A todos y cada uno de los profesores que fueron influyentes en mi carrera, Jetro López, Oneida Jiménez, José Luis, Mayerlin Maldonado, Belkys Araujo, José Saavedra, Oswaldo Rodríguez, Francisco Gelanze, Mary Zuluaga, María Silva, Franibe Rodríguez, Wilfredo y a todos aquellos a quienes ya no puedo recordar, Muchas gracias por haber impartido sus conocimientos y enseñanzas.

A mis amigos desde la primaria Loredanna, Melina y Jorge Luis, que sin importar el tiempo y la distancia, siempre están presentes en mi corazón.

Agradezco a la Universidad José Antonio Páez, por otorgarme la oportunidad de una beca, y agradezco a todas esas personas que he conocido en medio de este proceso, Mary Zuluaga, Michelle, Jose M. Maria Gabriela, La Sra. Luz, María, Diana y a las Autoridades de la universidad con quien tuve el placer de trabajar.

A mi sobrino perruno Zeus, quien llego en el momento preciso para brindarme todo su amor y cariño incondicional.

Leopoldo J. Estrada P

Agradecimientos

Quiero agradecer a mis padres por ser las personas que más me han apoyado en mi vida, por enseñarme que con dedicación y ganas de querer algo mejor en la vida puedes lograr lo que tanto sueñas solo con mucho esfuerzo, que en la vida no dependes de los demás sino de ti mismo y puedes lograr cosas por ti solo si tienes el talento necesario, y si no, te mejoras a ti mismo, gracias al reflejo de la vida de ambos me motivaron a ser igual a ellos y seguir el mismo camino.

A mis abuelos, al cuidarme me moldearon a los que soy actualmente, y me enseñaron los valores y respeto necesario hacia las otras personas y la tolerancia.

A mis mejores amigos y amigas, quienes posiblemente estén aquí o no, fueron ellos quienes en su mayoría me hicieron ver bastantes perspectivas de la vida, sin importar si fueron amigos de una semana o de toda la vida, si están aquí o simplemente no están, si seguimos siendo amigos o no, valoro todos los momentos buenos que me dieron en la vida y que realmente me hicieron sentir feliz, muchas gracias.

A mi compañero Leopoldo, sin importar los dolores de cabeza que tuvimos en este proyecto el cual nos esforzamos mucho, haciendo documentación y desarrollo, es una de las pocas personas que me han mostrado esa dedicación a la hora de trabajar, algo que no muchos tienen.

A mi tutor Jetro, no solo porque nos guio para poder hacer la tesis, sino porque a pesar de ser un profesor de alto temperamento, ha demostrado empatía hacia sus estudiantes, tanto en el aula de clases o en el área de la universidad, y como profesor, ingeniero y MSc, lo admiro, por algo tiene esos títulos.

A los profesores, que aunque no los nombrare para no crear favoritismo, hay profesores las cuales sus clases, personas que realmente en su momento no le prestaba tanta atención, gran practica en mi trabajo me ha demostrado que su conocimiento es muy valioso, persona que se dedican a enseñar, o cosas muy complejas o cosas muy esenciales, grandes profesores y grandes personas.

Alan O. Vásquez I.

ÍNDICE

CONTENIDO	Pág.
ÍNDICE DE FIGURAS	xi
ÍNDICE DE CUADROS	xi
RESUMEN	xii
INTRODUCCIÓN	1

CAPÍTULO

I EL PROBLEMA	3
1.1 Planteamiento del Problema.....	3
1.2 Formulación del Problema	5
1.3 Objetivos de la Investigación	6
1.3.1 Objetivo General	6
1.3.2 Objetivos específicos	6
1.4 Justificación del Problema	6
1.5 Alcance de la Investigación	7
II MARCO TEÓRICO	8
2.1 Antecedentes	8
2.2 Bases Teóricas.....	10
2.2.1 Redes Neuronales.....	10
2.2.2 Metodología XP	12
2.2.3 Python	15
2.2.4 Base de Datos No Relacionales (NOSQL)	15
2.2.5 CouchDB.....	16
2.2.6 Documentos de tipo JSON, BSON	18
2.2.7 Kivy.....	18
2.2.8 Numpy.....	18

2.2.9 Matplotlib.....	18
2.2.9 Pandas	18
2.3 Bases Legales	19
2.4 Definición de Términos.....	20
III MARCO METODOLÓGICO	22
3.1 Tipo de la Investigación	22
3.2 Diseño de la Investigación	22
3.3 Nivel de la Investigación.....	23
3.4 Población y Muestra.....	23
3.4.1 Población.....	23
3.4.2 Muestra.....	23
3.5 Técnicas de recolección de datos	23
3.5.1 Observación Directa.....	24
3.5.2 Entrevista no estructurada	24
3.6 Instrumento de recolección de datos	24
3.7 Validación del Instrumento	25
3.8 Confiabilidad del Instrumento	25
3.9 Fases de la Investigación.....	25
IV RESULTADOS	27
4.1 Fase I: Planificación.....	27
4.1.1 Requerimientos Funcionales:	32
4.1.2 Requerimientos No Funcionales	33
4.2 Fase II: Diseño.	33
4.2.1 Casos de Uso	35
4.2.2 Diagrama de Clases.....	43
4.2.3 Esquema de la Base de Datos	49
4.3 Fase III: Desarrollo o Diseño del sistema	50

4.4 Fase IV: Pruebas.	52
V CONCLUSIONES Y RECOMENDACIONES	55
5.1 Conclusiones.	55
5.2 Recomendaciones.....	55
REFERENCIAS BIBLIOGRÁFICAS	57

ÍNDICE DE FIGURAS

CONTENIDO	Pág.
FIGURA	
1. Escala de evaluación.....	28
2. Ítem N°1.....	28
3. Ítem N°2.....	29
4. Ítem N°3.....	30
5. Ítem N°4.....	30
6. Ítem N°5.....	31
7. Ítem N°6.....	32
8. Caso de uso; Crear conjunto de datos.....	36
9. Caso de uso; Crear modelos.....	38
10. Caso de uso; Entrenar Modelo.....	39
11. Caso de uso; Detección y Clasificación.....	41
12. Diagrama de Clase; DataManu.....	43
13. Diagrama de Clase; NeuralNetworkMenu.....	44
14. Diagrama de Clase; TrainningMenu.....	45
15. Diagrama de Clase; PredictionMenu.....	45
16. Diagrama de Clase; PantallaApp.....	46

17. Diagrama de Clase; neural_layer.....	46
18. Diagrama de Clase; neural_layer_json.....	47
19. Diagrama de Clase; Neural_Network.....	47
20. Diagrama de Clase; Database.....	48
21. Diagrama de Clase; Trainning.....	48
22. Diagrama de Clase; Graphic_Interface.....	48
23. Diagrama de Clase; Data_System.....	49
24. Esquema de la base de datos.....	49
25. Pantalla; Bases de Datos.....	50
26. Pantalla; Red Neuronal.....	50
27. Pantalla; Entrenamiento.....	51
28. Pantalla; Predicción.....	51

ÍNDICE DE CUADROS

CONTENIDO CUADRO	Pág.
1. Descripción de los actores: Usuario.....	34
2. Descripción de los actores: Administrativo.....	35
3. Descripción de los actores: Red Neuronal.....	35
4. Definición de caso de uso: Crear conjunto de datos.....	36
5. Definición de caso de uso: Crear modelos.....	38
6. Definición de caso de uso: Entrenar modelo.....	40
7. Definición de caso de uso: Detección y Clasificación.....	42
8. Caso de prueba 1.....	52
9. Caso de prueba 2.....	52
10. Caso de prueba 3.....	53
11. Caso de prueba 4.....	54



**REPÚBLICA BOLIVARIANA DE VENEZUELA
UNIVERSIDAD JOSÉ ANTONIO PÁEZ
FACULTAD DE INGENIERÍA
ESCUELA DE COMPUTACIÓN**

**SISTEMA DE INFORMACIÓN PARA LA PREDICCIÓN DE UN TARGET
OBJETIVO MEDIANTE EL CONSUMO DE UN RENGLÓN CON BASE A
UNA RED NEURONAL**

Autores: Estrada Pérez, Leopoldo José.
Vásquez Izaguirre, Alan Omar.
Tutor: MSc. Jetro López
Fecha: Junio, 2020

RESUMEN

El presente trabajo de grado plantea el desarrollo de un sistema avanzado para la predicción de un target objetivo mediante el consumo de un renglón con base a una red neuronal para minimizar el costo de las campañas y maximizar el retorno de inversión enviando anuncios publicitarios específicos de acuerdo con los intereses de los clientes, la misma se desarrolló con lenguaje de programación Phyton, se utilizó Jupites y VisualCode, además de un manejador de Base de datos CouchDB. En la actualidad se están incorporando de forma natural a la vida cotidiana de las personas y además de las operaciones de las empresas, a través de redes neuronales integradas por múltiples capas donde analiza una enorme cantidad de datos para el entrenamiento de los modelos de aprendizaje y con la experiencia obtenida se van creando y reforzando las conexiones para el aprendizaje. Se propuso el uso de la metodología de desarrollo de software XP, cuyo acrónimo significa eXtreme Programming (Programación Extrema), Metodológicamente es una investigación tipo proyecto factible, sustentada en una investigación de campo y documental de nivel descriptivo con una población de estudio de 60 empresas del BigLow, San Diego Edo. Carabobo dentro de los cuales se tomó como muestra 15, para lo cual se usó como técnica de recolección de datos la encuesta y la observación. La entrega del sistema se planifico para brindar una mejora a la toma de decisiones al momento de determinar los clientes potenciales para las empresas, negocios o personas mediante un análisis de la información existente dentro de su área en específico.

Descriptor: Predicción, Target, Clientes Potenciales, Red Neuronal, Sistema de Información (SI).

INTRODUCCIÓN

En el siguiente trabajo se plantea un sistema para la predicción de target objetivos con base a una red neuronal, para mejorar el proceso de la obtención del cliente potencial de una empresa o negocio, debido a que el uso de mercadotecnia ha disminuido en los últimos años. Una práctica muy común es la segmentación de la información para seleccionar y atraer de forma más efectiva a los clientes. En este sentido, se destacaran los problemas derivados del procesamiento actual de la información.

Este trabajo está orientado a resolver dichas problemáticas que ocurrieron en la actualidad, las cuales son abordadas de una manera idónea para su resolución siguiendo ciertos parámetros de investigación y donde se pretende satisfacer la necesidad que posee esta organización. Este proyecto se encuentra estructurado en cuatro capítulos de la manera siguiente:

Capítulo 1 El Problema: En este primer capítulo se describe el problema existente, el objetivo principal del proyecto: “Sistema de información para la predicción de un target objetivo mediante el consumo de un renglón con base a una red neuronal”, los pasos para lograrlo, es decir los objetivos específicos y la razón por la que este debe llevarse a cabo. De igual forma se dio a conocer de manera explícita el alcance tuvo el proyecto.

Capítulo II Marco Teórico: en este capítulo se establecen las teorías que sustentan la realización del proyecto al igual que los antecedentes existentes que puedan aportar algo al mismo.

Capítulo III Marco Metodológico: Capítulo en el cual se dio a conocer la metodología que se empleará para el desarrollo de este trabajo y se especificó los métodos utilizados para recolectar y analizar la información necesaria.

Capítulo IV Resultados: En el siguiente capítulo se presenta el desarrollo del proyecto en base a la metodología de desarrollo del software elegido, exponiendo así cada una de las fases de la metodología establecida.

Capítulo V Conclusiones y Recomendaciones: en este último capítulo se expondrán las conclusiones obtenidas en el desarrollo del proyecto, trayendo así recomendaciones para darle continuidad al proyecto o para futuros trabajos de investigación.

CAPÍTULO I

EL PROBLEMA

1.1 Planteamiento del Problema

Desde el inicio de la humanidad, el ser humano ha de realizar diferentes actividades y tareas que en su mayoría demandan conocimiento, donde dicho conocimiento es adquirido por la información que se obtiene por diversas fuentes y esta es debidamente procesada para luego así llevar a cabo sus actividades, desenvolverse satisfactoriamente, o para evitar posibles riesgos.

Por consecuencia de esta, se crea la necesidad de la búsqueda de información recurriendo así a diferentes medios incluyendo a las redes, que estas hacen aparición por primera vez en la década de los 60, esta red era única y exclusivamente textual y sin mucho contenido que ofrecer, poco a poco se ha mejorado esta red hasta a lo que hoy en día se le denomina Internet.

Sin duda en la actualidad, se aprecia en la evolución del Internet las nuevas necesidades del consumidor, siendo atraídos por la información y diversas actividades que esta le puede ofrecer, convirtiéndose en un extenso mercado que permite la facilidad y la rapidez de poder ofrecer servicio a clientes de manera virtual, haciendo que las empresas fueran adaptándose a esta área y permitiera un crecimiento económico mucho mayor.

Al mismo tiempo, es común que las empresas realicen pronósticos de ventas para poder planificar más acertadamente su mercado comparando los resultados obtenidos con los pronósticos de un método estadístico clásico y estableciendo los entornos más adecuados para su uso, utilizando métodos de captación de clientes para dar su target objetivo.

Sin embargo, no se podrá descuidar las estrategias de captación de nuevos compradores que garanticen la existencia y supervivencia de algún negocio o marca que, al fin y al cabo, dentro de una empresa se produce un flujo de entradas y salidas

de consumidores. Por esta razón, habrá de trabajar la captación para así garantizar el equilibrio entre el mantenimiento y esta; donde la captación se refiere a la búsqueda de clientes potenciales que tendrá que convertir en clientes finales.

El proceso de captación ha de ser un trabajo constante, aunque se pueda realizar campañas específicas en momentos puntuales o cuando las necesidades del negocio así lo requieran. Por ejemplo, resulta especialmente interesante realizar campañas de atracción de clientes en la inauguración de una tienda, en el lanzamiento de un nuevo producto, con negocios estacionales que tienen que multiplicar sus esfuerzos durante su temporada alta o para la captación de clientes bancarios, en cualquier caso, siempre que se tenga la oportunidad, habrá de recoger datos de nuestros leads o clientes potenciales, que son personas que han demostrado cierto interés por un producto o servicio. Sus datos son los cimientos con lo que se comenzará a gestionar y construir nuestra relación con él.

Para poder analizar los datos obtenidos de un flujo constante de información, se debe crear un conjunto de datos, de esta misma manera se pueden encontrar tanto como clasificación y selección del cliente, como las características y propiedades de dichos productos para así obtener patrones de conducta. Las grandes empresas o las empresas en crecimiento siempre tratan de conseguir nuevos consumidores para hacer que sus recursos siempre se enfoquen en los mayores clientes potenciales, para poder dirigir sus productos o servicios a ese público en concreto dedicándose a cautivarlos mediante estrategias de mercado específicas para que las ventas sean mayores.

Mientras que en algunas empresas se han presentado problemas de no conocer los clientes viables del mercado en todos los procesos, desde el registro y gestión de clientes, se realizan mediante herramientas muy básicas, esto significa un proceso considerado tedioso y tardío, además de propenso a errores de transcripción. Aunado a esto, se presenta un retraso en la difusión de los posibles clientes potenciales pasando muchos días, a veces, antes de que estos sean entregados al personal responsable de procesarlos y transmitirlos a los interesados.

Al mismo tiempo se ha notado la mejoría de los sistemas que han simplificado muchos procesos para la captación de clientes, sin embargo, estos mismos no son adaptables a las necesidades del usuario, teniendo así un sistema ineficiente e incapaz de evolucionar, es por ellos que se ha demostrado que un sistema con el uso de las redes neuronales puede ser el más óptimo y eficaz para el pronóstico de los posibles consumidores.

Por lo tanto, las redes neuronales son consideradas como un modelo computacional que se inspira en el comportamiento biológico de la mente humana, la cual está conformada por una red de neuronas artificiales conectadas entre sí con la finalidad de transferir datos de unas a otras, generando así diferentes valores dependiendo de la entrada, que posteriormente son estimadas en su salida.

Esta red neuronal aprende por medio de métodos de evaluación de aprendizaje genético creando en sus primeras etapas las redes neuronales aleatorias, que luego son evaluadas para obtener la red neuronal con más acierto, este se utiliza como modelo para las siguientes generaciones modificando de forma mínima cada una de estas redes individualmente e ir repitiendo este proceso constantemente hasta obtener un modelo con más precisión.

Al obtener un modelo capacitado, este se puede utilizar en pruebas reales y seguir evaluando los datos diarios e ir generando una nueva red neuronal constantemente que se adapte a los datos dinámicos que se reciben constantemente, y que haga predicciones sobre los posibles clientes potenciales que puedan consumir o contratar los servicios o productos de las empresas.

1.2 Formulación del Problema

¿De qué manera se puede determinar un cliente potencial por medio de herramientas de programación?

1.3 Objetivos de la Investigación

1.3.1 Objetivo General

Desarrollar un sistema de información para la predicción de un target objetivo que determine el consumo de un renglón con base a una red neuronal para la obtención del cliente potencial.

1.3.2 Objetivos específicos

-Describir los patrones de comportamiento de una red neuronal para el desarrollo de un sistema de información.

-Identificar los requerimientos funcionales y no funcionales en relación con el desarrollo de un sistema de información para la predicción de un target objetivo.

-Diseñar un sistema para la predicción de un target objetivo mediante el consumo de un renglón usando una red neuronal para la clasificación de la información.

-Construir un sistema, utilizando herramientas de programación, siguiendo la metodología XP (eXtreme Programming).

1.4 Justificación del Problema

El presente trabajo de investigación presentara beneficios para la sociedad referente al factor tiempo y sobre todo en recursos, ya que tiene la posibilidad de predecir con mayor asertividad el target objetivo por el consumo de un renglón mediante redes neuronales además de brindar una herramienta que permita a cualquier empresa, negocio o persona, almacenar y gestionar toda la información relacionada con el renglón, así como determinación del cliente potencial. De esta manera, se obtendrá un efecto positivo en los resultados, ya que ayudará a la planificación y elección de los diferentes clientes, aumentándose de esta forma el nivel de la captación.

Del mismo modo será de gran aprovechamiento para próximos trabajos de investigación a lo que respecta a la predicción de target objetivo tanto como en el desarrollo de una red neuronal. Siendo así de gran utilidad para los estudiantes de la

carrera de ingeniería en computación y para cualquier otra persona del que sea utilidad dicha información.

1.5 Alcance de la Investigación

Tomando en consideración lo antes expuesto, esta sección describe el alcance del proyecto y brevemente resalta las restricciones que limitan el enfoque de este trabajo de grado.

El conjunto de datos se tomará utilizando los conjuntos de datos proporcionados por las encuestas, este conjunto ha sido generado de forma sintética con la finalidad de simular un gran volumen de registros correspondientes a los clientes.

Así como los algoritmos de clasificación y de regresión, donde los algoritmos de clasificación serán seleccionados para mostrar un mejor desempeño frente a otros algoritmos, haciendo altamente eficiente, flexible y portable; En el caso de los algoritmos de regresión que serán experimentos limitados para mostrar un mejor desempeño.

Además, estará desarrollada bajo el concepto de determinación de clientes potenciales, y el lenguaje que se aspira emplear para la generación de la red neuronal será Python. Además la metodología que se empleará en este proyecto es la metodología XP (eXtreme Programming), la cual es un procedimiento rápido y ágil que permitirá aplicar las mejoras técnicas de desarrollo durante el ciclo de vida del sistema de una manera dinámica, obteniendo así mejores resultados y el tiempo del que se dispondrá para el desarrollo del proyecto es de seis (6) a ocho (8) meses, y el lugar en donde se puede desarrollar es en cualquier negocio o empresa con fines de captar el público o el cliente potencial que adquiera o consuma de su producto

CAPÍTULO II

MARCO TEÓRICO

2.1 Antecedentes

Para el desarrollo de esta investigación se tomaron en consideración diferentes referencias de proyectos anteriores con similitudes a la problemática ya planteada, de tal forma pueda complementar el proceso de alcanzar los objetivos planteados y de obtener los resultados esperados, de las cuales destacan las siguientes.

En Primer lugar, Castillo P. (2017) en su trabajo de grado titulado **“Aplicación de aprendizaje automático para la predicción de clientes potenciales en procesos de mercadotecnia”**, como requerimiento para la obtención de título de Maestro de Ingeniería de Software en Ingeniería en el Centro de Investigación en Matemáticas, A.C. (CIMAT), Guanajuato, México. Este trabajo tuvo como objetivo realizar un comparativo de técnicas de aprendizaje automático para predecir la probabilidad de compra de un producto por parte de un cliente para la empresa conocida en este estudio, definiendo un marco conceptual respecto el aprendizaje automático y que modelos de aprendizaje automático son necesarias para generar modelos que predigan la probabilidad de compra. Del presente trabajo se observó las distintas metodologías conocidas de los requerimientos básicos a tomar en cuenta para verificar cuál de los enfoques propuestos de aprendizaje automático es más preciso para la probabilidad de compra.

Así mismo, Macea y Matertelo (2016) en su trabajo de grado como modalidad de tesis titulado **“Predicción con redes neuronales de un proceso de secado del ñame (Dioscorea Rotundata) en hornos de microondas”**, presentado para la optar al título de Ingeniero en Alimentos en la Universidad de Córdoba en Montería, Colombia. En el cual se basa sobre redes neuronales, arquitecturas de las redes neuronales y el aprendizaje de las redes neuronales para el proceso de secado del

ñame (*Dioscorea Rotundata*) en hornos de microondas. Se obtuvo información respecto al entrenamiento de las redes neuronales tomando en consideración la variable dependiente en el proceso de secado, como también se todo en consideración la ecuación planteada para dicho aprendizaje con sus respectivos valores de entrada y su objetivo.

También, Lizarazo Y. (2018) en su trabajo de grado titulado “**Predicción de la dosis de coagulante para el tratamiento de agua potable utilizando redes neuronales artificiales**”, presentado para optar el título de Ingeniería Química en la Universidad Industrial de Santander en Bucaramanga, Colombia. En el cual propone el desarrollo de una arquitectura basada en una red neuronal artificial (RNA) como modelo de predicción para guiar la experimentación y ahorrar tiempo en la deducción de la dosis óptima de coagulante sulfato de aluminio ($Al_2(SO_4)_3 \cdot 18H_2O$) para el tratamiento de agua potable, utilizando para su entrenamiento y validación una de base de datos construida a partir de resultados experimentales. La arquitectura de red propuesta fue entrenada y validada mediante el algoritmo *backpropagation*, con el método de optimización Levenberg-Marquardt, además de aplicar técnicas como la repetición del entrenamiento y el método de regularización denominado *early stopping* para evitar sobreajustes, así como también la aplicación de funciones como *logsig* y *tansig* que fueron consideradas para la activación de las neuronas de la capa oculta.

Seguidamente, tenemos que Farias A. (2019) en su trabajo de grado titulado “**Aplicación predictiva de cuerpos, fenómenos y objetos astronómicos-espaciales mediante el uso de aprendizaje profundo (DNN, CNN) y visión artificial**”. Presentado para optar el título de Ingeniero en Computación en la Universidad José Antonio Páez en Carabobo, Venezuela. Farias presenta una propuesta de una herramienta de utilidad alternativa para procesar y clasificar los fenómenos, cuerpos y objetos astronómicos mediante el uso de aprendizaje profundo y visión artificial, ya

que, al reducir algunos pasos intermedios de un estudio clásico, el investigador podrá centrarse en las tareas de análisis e investigación.

Por último, tenemos a Ochoa J. & Pereira S. (2020) en su trabajo de grado titulado “**Aplicación Web que diagnostique la inflación basada en inteligencia artificial**” Presentado para optar por el título de Ingeniero en Computación en la Universidad José Antonio Páez en Carabobo, Venezuela. En este trabajo los autores buscan que las empresas se prevengan de cualquier futura inflación y que de esta forma puedan mantenerse en el mercado nacional e internacional, también trae consigo, beneficios para el empleado u obrero para que los salarios de estos no sean afectados.

Por lo cual, el presente trabajo sirvió como aporte para la investigación actual en lo que compete al desarrollo de un sistema para predecir y así alcanzar todos los objetivos propuestos, además de proveer información relevante acerca del desarrollo de este sistema, desde el proceso de evaluación hasta los resultados y discusión del mismo, esto quiere decir que, el trabajo de grado previamente descrito contiene un alto significado hacia la presente investigación.

2.2 Bases Teóricas

2.2.1 Redes Neuronales

El cerebro humano contiene una inmensa cantidad de neuronas y de células nerviosas, por lo cual cada una de esas células estará siempre conectada a muchas más del mismo tipo. Esto crea una muy compleja red de transmisión de señales donde cada una de estas células neuronales recopila o recibe las entradas de las otras células neuronales a la cuales esté conectada y, en el caso de que alcance cierto umbral, enviará un impulso consecuente a todas las células a la cuales esté conectada (Borglin, 2011).

Esta breve descripción del cerebro humano permite en gran medida entender el concepto de las redes neuronales artificiales, pues, indica que la información se propaga en gran medida a través de una red de elementos interconectados que

reaccionan a un cierto impulso en específico, y que esto, además, determinará la consecuente propagación del impulso.

Así pues, a partir de la introducción y la analogía con el cerebro humano, Borglin (2011) describe que las redes neuronales artificiales son un clasificador modelado a partir del cómo trabaja el cerebro humano, es decir, imita el funcionamiento de una neurona mediante el uso del concepto de “perceptron”, que no es más que la unidad básica de referencia en este sistema. El perceptron, puede tomar múltiples entradas ponderadas y resumirlas, luego, determinará si la combinación de la entrada excede cierto umbral y de ser verdadero, el perceptron se activará y, consecuentemente, transmitirá una salida.

Un concepto más formal según Chen (1991), indica que una red neuronal artificial consiste de un conjunto de neuronas, denominadas como la unidad básica o elementos de procesado, una topología de conexión y un algoritmo de aprendizaje. Luego, describe que las señales de entrada se retransmiten a través de la red y que estas pueden venir del entorno externo de la red o, ser una salida de las otras unidades que conforman la red. Adicionalmente, señala que cada representación de una conexión de entrada entre las distintas unidades o neuronas no es más que un valor ajustable llamado peso o fuerza de la conexión.

Cabe destacar que esta red neuronal artificial se divide en capas, esto es: La primera capa conocida como la capa de entrada. La capa final, también conocida como la capa de salida. Luego, las $n-1$ capas intermedias a la capa de entrada y la capa de salidas, conocidas como las capas escondidas, pues su comportamiento es restringido de forma implícita solamente por los datos de entrenamiento, debido a que los datos no muestran la salida deseada para cada una de esas capas. Así bien, el número de estas n capas define la profundidad de la red (Borovkov, 2017).

De forma análoga, este concepto puede extenderse si se considera a la capa de entrada como las múltiples neuronas que reciben los impulsos de entrada. Luego, la capa escondida, este conjunto de capas se encargan de procesar y redirigir el impulso de entrada hacia la capa de salida y, puede ser desde una hasta n cantidad de capas

con m cantidad de neuronas en cada capa. Finalmente, la capa de salida, que no es más que el conjunto de neuronas que reaccionan y representan la respuesta al impulso inicial.

Finalmente, cabe destacar además el concepto de las redes neuronales profundas. Esto es pues, una derivación de las redes neuronales artificiales. Estas son pues, diferentes de los otros métodos populares del aprendizaje de máquina que permiten estructuras poco profundas (shallow). Típicamente, las redes neuronales profundas tienen más capas y parámetros, por lo tanto, tienen un mejor potencial para representar entradas más complejas. Estas redes complejas, permiten a las personas construir agentes inteligentes capaces de llevar a cabo complicadas tareas de percepción, e incluso, alcanzar el rendimiento o capacidad de los humanos en algunos casos (Wang, 2017).

2.2.2 Metodología XP

La metodología XP (eXtreme Programming), la cual es una metodología de desarrollo de la ingeniería de software formulada por Kent Beck, autor del primer libro sobre la materia, Extreme Programming Explained: Embrace Change (1999). Es el más destacado de los procesos ágiles de desarrollo de software. Al igual que éstos, la programación extrema se diferencia de las metodologías tradicionales principalmente en que pone más énfasis en la adaptabilidad que en la previsibilidad. La metodología XP tiene un conjunto importante de reglas y prácticas. En forma genérica, se pueden agrupar en:

1. Planificación.
2. Diseño.
3. Desarrollo.
4. Pruebas

Fase I: Planificación

La metodología XP plantea la planificación como un diálogo continuo entre las partes involucradas en el proyecto, incluyendo al cliente, a los programadores y a los

coordinadores o gerentes. El proyecto comienza recopilando “Historias de usuarios”, las que sustituyen a los tradicionales “casos de uso”. Una vez obtenidas las “historias de usuarios”, los programadores evalúan rápidamente el tiempo de desarrollo de cada una.

Si alguna de ellas tiene “riesgos” que no permiten establecer con certeza la complejidad del desarrollo, se realizan pequeños programas de prueba (“spikes”), para reducir estos riesgos. Una vez realizadas estas estimaciones, se organiza una reunión de planificación, con los diversos actores del proyecto (cliente, desarrolladores, gerentes), a los efectos de establecer un plan o cronograma de entregas (“Release Plan”) en los que todos estén de acuerdo. Una vez acordado este cronograma, comienza una fase de iteraciones, en dónde en cada una de ellas se desarrolla, prueba e instala unas pocas “historias de usuarios”.

Según Martín Fowler uno de los firmantes del “Agile Manifesto”, los planes en XP se diferencian de las metodologías tradicionales en tres aspectos:

1. Simplicidad del plan. No se espera que un plan requiera de un “gurú” con complicados sistemas de gerenciamiento de proyectos.
2. Los planes son realizados por las mismas personas que realizarán el trabajo.
3. Los planes no son predicciones del futuro, sino simplemente la mejor estimación de cómo saldrán las cosas. Los planes son útiles, pero necesitan ser 16 cambiados cuando las circunstancias lo requieren. De otra manera, se termina en situaciones en las que el plan y la realidad no coinciden, y en estos casos, el plan es totalmente inútil.

Fase II: Diseño

La metodología XP hace especial énfasis en los diseños simples y claros. Los conceptos más importantes de diseño en esta metodología son los siguientes:

1. Simplicidad: Un diseño simple se implementa más rápidamente que uno complejo. Por ello XP propone implementar el diseño más simple posible que funcione. Se sugiere nunca adelantar la implementación de funcionalidades que no correspondan a la iteración en la que se esté trabajando.

2. Soluciones “spike”: Cuando aparecen problemas técnicos, o cuando es difícil de estimar el tiempo para implementar una historia de usuario, pueden utilizarse pequeños programas de prueba (llamados “spike”), para explorar diferentes soluciones. Estos programas son únicamente para probar o evaluar una solución, y suelen ser desechados luego de su evaluación.

3. Recodificación: La recodificación (“refactoring”) consiste en escribir nuevamente parte del código de un programa, sin cambiar su funcionalidad, a los efectos de hacerlo más simple, conciso y/o entendible. Muchas veces, al terminar de escribir un código de programa, pensamos que, si lo comenzáramos de nuevo, lo hubiéramos hecho en forma diferente, más clara y eficientemente. Sin embargo, como ya está pronto y “funciona”, rara vez es rescrito.

4. Metáforas: Una “metáfora” es algo que todos entienden, sin necesidad de mayores explicaciones. La metodología XP sugiere utilizar este concepto como una manera sencilla de explicar el propósito del proyecto, y guiar la estructura y arquitectura del mismo. Por ejemplo, puede ser una guía para la nomenclatura de los métodos y las clases utilizadas en el diseño del código. Tener nombres claros, que no requieran de mayores explicaciones, redundan en un ahorro de tiempo.

Fase III: Desarrollo o Codificación

El cliente es una parte más del equipo de desarrollo; su presencia es indispensable en las distintas fases de XP. A la hora de codificar una historia de usuario su presencia es aún más necesaria. No olvidemos que los clientes son los que crean las historias de usuario y negocian los tiempos en los que serán implementadas. Antes del desarrollo de cada historia de usuario el cliente debe especificar detalladamente lo que ésta hará y también tendrá que estar presente cuando se realicen los test que verifiquen que la historia implementada cumple la funcionalidad especificada. La codificación debe hacerse ateniendo a estándares de codificación ya creados. Programar bajo estándares mantiene el código consistente y facilita su comprensión y escalabilidad.

Fase IV: Pruebas

Uno de los pilares de la metodología XP es el uso de test para comprobar el funcionamiento de los códigos que vayamos implementando.

El uso de los test en XP es el siguiente:

- Se deben crear las aplicaciones que realizarán los test con un entorno de desarrollo específico para test. Hay que someter a test las distintas clases del sistema omitiendo los métodos más triviales.
- Se deben crear los test que pasarán los códigos antes de implementarlos; en el apartado anterior se explicó la importancia de crear antes los test que el código.
- Un punto importante es crear test que no tengan ninguna dependencia del código que en un futuro evaluará.
- Como se comentó anteriormente los distintos test se deben subir al repositorio de código acompañados del código que verifican.
- Test de aceptación. Los test mencionados anteriormente sirven para evaluar las distintas tareas en las que ha sido dividida una historia de usuario.
- Al ser las distintas funcionalidades de nuestra aplicación no demasiado extensas, no se harán test que analicen partes de las mismas, sino que las pruebas se realizarán para las funcionalidades generales que debe cumplir el programa especificado en la descripción de requisitos.

2.2.3 Python

Para el autor Rossum, G. (2009). Define Python como un lenguaje de programación poderoso y fácil de aprender. Cuenta con estructuras de datos eficientes y de alto nivel y un enfoque simple pero efectivo a la programación orientada a objetos. La elegante sintaxis de Python y su tipado dinámico, junto con su naturaleza interpretada, hacen de éste un lenguaje ideal para scripting y desarrollo rápido de aplicaciones en diversas áreas y sobre la mayoría de las plataformas.

2.2.4 Base de Datos No Relacionales (NOSQL)

Son base de datos que no manejan un modelo relacional, los datos que se almacenan no requieren estructuras fijas como tablas, no soportan operaciones de

juntura, y tampoco garantizan la atomicidad, consistencia, aislamiento y durabilidad de datos (ACID), habitualmente escalan bien horizontalmente. (D. López, 2012)

Tipos de datos y funciones Los tipos de datos que manejan las bases de datos NoSQL y las funciones que cumplen son las siguientes.

Base de datos de documentos. Permite que una clave sea emparejada con un documento. Los documentos pueden contener una gran variedad de objetos como claves en arreglos y pares clave-valor, así como otros documentos.

Key-value stores. Almacenan información en forma de atributos y valores. En algunos casos, los valores pueden tener tipos como cadena y número entero.

Graph. Almacena datos sobre redes.

Wide-column stores. Son diseñadas el manejo de grandes cantidades de datos en forma de columnas.

2.2.5 CouchDB

Es un proyecto creado por Damien Katz en abril de 2005, lo definió como un “sistema de almacenamiento para una base de datos objetal de gran escala”. Sus objetivos eran que la base de datos se transformara en la base de datos de la Internet y que se diseñara desde un principio para prestar servicios a aplicaciones web. En febrero de 2008, pasó a ser un proyecto de Apache Incubator y la licencia se cambió a la Apache License, pocos meses más tarde, se graduó como proyecto top-level. Así se llegó a la publicación de la primera versión estable en julio de 2010. A principios de 2012, Damien Katz abandonó el proyecto a fin de concentrarse en Couchbase. El proyecto ha seguido adelante, liberando su versión 1.2 en abril de 2012. (CouchDB, 2013)

2.2.5.1 Definición CouchDB

Es un gestor de base de datos de código abierto, almacena sus datos con documentos JSON, accede a sus documentos con un navegador web a través de HTTP, consulta, combina y transforma los documentos con JavaScript. CouchDB no almacena los datos y sus relaciones en tablas, cada base de datos es una colección de

documentos independientes con sus propios datos y esquema. Una aplicación puede acceder a múltiples bases de datos. (CouchDB A., 2014)

2.2.5.2 Arquitectura del CouchDB

- HTTP Client: permite realizar al cliente peticiones al servidor de la base de datos.
- Erlang HTTP: permite el control y gestión de HTTP como gestión de peticiones, concurrencia, cabeceras HTTP, encriptado, autenticación.
- Mod_Couch: es el core de la máquina virtual de Erlang de CouchDB, este componente implementa la API y es el encargado de entender y gestionar las diferencias operacionales solicitadas por el cliente HTTP.
- ViewEngine: realiza este componente todo el proceso de las vistas, el diseño de los documentos, la indexación utilizando arboles B.
- TorageEngine: este componente se comunica activamente con ViewEngine y el Replicador ya que las vistas son incrementales y deben reflejar los cambios escritos a disco, gestiona las escrituras en disco.
- Replicador: permite a la replicación de los documentos, ayuda a la distribución entre diferentes dispositivos. Solo actúa si es activado manual o programando, pero no funciona de manera continua.

2.2.5.3 Características CouchDB

- Replicación fácil entre servidores a través de instancias.
- Soporte para resolución de conflictos y configuración del master.
- Actualizaciones de alimentación de datos por suscripción.
- Indexación, búsqueda y recuperación rápida.
- Interfaz web RESTful.
- Los documentos son accesibles a través de navegadores y los índices pueden ser consultados a través de HTTP.
- Operaciones de indexación, combinación y transformación realizadas con JavaScript.

- Operaciones para crear, leer, actualizar, borrar documentos simples (CRUD).
- MapReduce avanzado

2.2.6 Documentos de tipo JSON, BSON

JSON (JavaScript Object Notation) – Notación de objetos JavaScript) es un formato de documentos de formato ligero para el intercambio de datos, que utiliza objetos JavaScript, pero no requiere el uso de XML.

BSON acrónimo de Binary JSON, es una serialización binaria codificada de documentos JSON, apoya la incorporación de documentos y matrices dentro de otros documentos y matrices, también contiene extensiones que permiten la representación de los tipos de datos que no son parte de la especificación JSON. Por ejemplo, BSON tiene un tipo de fecha y un tipo BinData. BSON puede ser comparado con formatos de intercambio binarios, como Protocolos Buffers” (“BSON - Binary JSON,” 2013)

2.2.7 Kivy

Kivy es un framework multiplataforma desarrollado en Python y Cython, basado en OpenGL ES 2, junto con su propio lenguaje Kv que permite facilitar la estructura de la interfaz que posteriormente es mostrada en pantalla.

2.2.8 Numpy

Es una librería que permite implementar el uso de vectores y matrices multidimensionales de forma mucho más fácil, con el fin de poder aplicar funciones matemáticas de alto nivel.

2.2.9 Matplotlib

Es una librería basada en Matlab, que permite de gráficos 2D y 3D de alta calidad, utilizando los datos ingresados en forma de vectores o matrices.

2.2.9 Pandas

Es una librería que permite analizar y manipular datos de manera eficiente y flexible, permitiendo la lectura de múltiples formas como CSV, Microsoft Excel, bases de datos SQL.

2.3 Bases Legales

Las bases legales que dan soporte al proyecto en referencia, se encuentran plasmadas en la:

- Constitución de la República Bolivariana de Venezuela (1999)

Artículo 110: El Estado reconocerá el interés público de la ciencia, la tecnología, el conocimiento, la innovación y sus aplicaciones y los servicios de información necesarios por ser instrumentos fundamentales para el desarrollo económico social y político del país, así como para la seguridad y soberanía nacional.

Se infiere que todas las iniciativas en función de innovar los sistemas de información serán reconocidas como un instrumento para el desarrollo de las instituciones nacionales y por ende para el desarrollo nacional.

- Decreto Rango y Fuerza de Ley Orgánica de Ciencia, Tecnología e Innovación en Consejo de Ministros (2002)

Artículo 3.-Establece que forman parte del Sistema Nacional de Ciencia, Tecnología e Innovación, las instituciones públicas o privadas que generen y desarrollen conocimientos científicos y tecnológicos y procesos de innovación, y las personas que se dediquen a la planificación, administración, ejecución y aplicación de actividades que posibiliten la vinculación efectiva entre la ciencia, la tecnología y la sociedad. A tal efecto, forman parte del Sistema: las instituciones de educación superior y de formación técnica, academias nacionales, colegios profesionales, sociedades científicas, laboratorios y centros de investigación y desarrollo, tanto público como privado.

Artículo 22.-El Ministerio de Ciencia y Tecnología coordinará las actividades del Estado que, en el área de tecnologías de información, fueren programadas. Asumirá competencias que, en materia de

informática, ejercía la Oficina Central de Estadística e Informática (OCEI), así como las siguientes:

1. Actuar como organismo rector del Ejecutivo Nacional en materia de tecnologías de información.
2. Establecer políticas en torno a la generación de contenidos en la red, de los órganos y entes del Estado.
3. Establecer políticas orientadas a resguardar la inviolabilidad del carácter privado y confidencial de los datos electrónicos obtenidos en el ejercicio de las funciones de los organismos públicos.
4. Fomentar y desarrollar acciones conducentes a la adaptación y asimilación de las tecnologías de información por la sociedad.

En correspondencia con ambos artículos artículo el presente proyecto se ajusta a las aspiraciones de la ley por cuanto su desarrollado atiende a los lineamientos establecidos por la OCEI.

- Decreto N° 3.390 de la Presidencia de la República Bolivariana de Venezuela Gaceta 38.095 del 28/12/2004), sobre uso del Software Libre.

La Administración Pública Nacional empleará prioritariamente Software Libre desarrollado con Estándares Abiertos, en sus sistemas, proyectos y servicios informáticos. A tales fines, todos los órganos y entes de la Administración Pública Nacional iniciarán los procesos de migración gradual y progresiva de éstos hacia el Software Libre desarrollado con Estándares Abiertos.

2.4 Definición de Términos

A continuación, se definen los conceptos que dan inicio y hacen referencia a esta investigación.

Target: en función de marketing es el destinatario al que pretende llegar un servicio o un producto y sus correspondientes campañas de difusión.

Reglón: Serie de letras o palabras dispuestas horizontalmente en un escrito

Red neuronal: es un sistema compuesto de muchos elementos procesadores simples operando en paralelo, cuya función es determinada por la estructura de la red, fuerza en las conexiones y el procesamiento realizado por los elementos computacionales en los nodos.

Base de Datos: es una colección de información organizada de forma que un programa de ordenador pueda seleccionar rápidamente los fragmentos de datos que necesite.

Sistema de Información: es un conjunto de elementos que interactúan entre sí con un fin común; que permite que la información esté disponible para satisfacer las necesidades en una organización, aunque no siempre requiere contar con recurso computacional, aunque la disposición del mismo facilita el manejo e interpretación de la información por los usuarios.

Lenguaje de Programación: es un lenguaje diseñado para describir el conjunto de acciones consecutivas que un equipo debe ejecutar, es decir, es un modo práctico para que los seres humanos puedan dar instrucciones a un equipo.

Backpropagation: es un método de cálculo gradiente, que se utiliza en las redes neuronales para el cálculo de error en una red neuronal, comparando los valores de la salida con los deseados, evaluando cada salida con las capas anteriores.

CAPÍTULO III MARCO METODOLÓGICO

3.1 Tipo de la Investigación

La investigación realizada se vincula con la modalidad de proyecto especial, ya que por medio de esta se busca dar solución a los objetivos planteados. Según lo estipula las Normas de elaboración de Trabajo de Grado de la Universidad José Antonio Páez (UJAP, julio 2007), define Proyecto especial como:

“Consistirá en las creaciones tangibles, susceptibles de ser realizadas a problemas demostrados, o que respondan a necesidades o intereses de tipo cultural. Se incluyen en esta categoría los trabajos de elaboración de libros de texto y de materiales de apoyo educativo, el desarrollo de software y hardware, prototipos y productos tecnológicos en general (p. 5).”

3.2 Diseño de la Investigación

Según Arias, F. (2012), “El diseño de investigación es la estrategia general que adopta el investigador para responder al problema planteado. En atención al diseño, la investigación se clasifica en: documental, de campo y experimental” (p. 27). Esta investigación es clasificada como una investigación de campo, en la cual se realizan los criterios pautados por una investigación de campo. Según Arias, F. (2012)

La investigación de campo es aquella que consiste en la recolección de datos directamente de los sujetos investigados, o de la realidad donde ocurren los hechos (datos primarios), sin manipular o controlar variable alguna, es decir, el investigador obtiene la información, pero no altera las condiciones existentes. De allí su carácter de investigación no experimental (p.31).

De igual manera, el presente estudio presenta ser también una investigación del tipo documental, ya que la información adicional que respalda la presente investigación será extraída de autores pasados. De esta forma el autor ya mencionado, define la Investigación documental como “un proceso basado en la búsqueda, recuperación, análisis, crítica e interpretación de datos secundarios, es decir, los

obtenidos y registrados por otros investigadores en fuentes documentales: impresas, audiovisuales o electrónicas” (p.27).

3.3 Nivel de la Investigación

El trabajo de grado elaborado, presenta características de investigación descriptivo de acuerdo a Sabino (2004), que afirma “la investigación descriptiva busca especificar propiedades, características y rasgos importantes de cualquier fenómeno que se analice”. La investigación descriptiva suele utilizar técnicas como la observación como método descriptivo, permitiendo evaluar distintos enfoques de la problemática con mayor facilidad llegando a generar predicciones rudimentarias de los resultados que puede llegar a dar la solución planteada, a su vez se realiza una breve explicación de todos los pasos a ejecutar para el éxito del presente proyecto.

3.4 Población y Muestra

3.4.1 Población

Es necesario establecer una población a los cuales se les aplicará dichos instrumentos, los cuales son un grupo de individuos con características en común; dicha población es definida por Arias, F. (2012) como “un conjunto finito o infinito de elementos con características comunes para los cuales serán extensivas las conclusiones de la investigación. Ésta queda delimitada por el problema y por los objetivos del estudio.” (p. 81). Pasa esta investigación, la población está conformada por 60 empresas ubicadas en el BigLow, San Diego Edo. Carabobo, Venezuela.

3.4.2 Muestra

La muestra es la que permite evaluar la problemática, ya que esta genera datos por medio de los cuales se puede hacer inferencias o generalizar resultados de las fallas detectadas. Arias, F. (2012) define la muestra como “un subconjunto representativo y finito que se extrae de la población accesible” (p.82). Para fines de esta investigación la muestra será representada por 15 empresas.

3.5 Técnicas de recolección de datos

Con respecto a las técnicas e instrumentos de recolección de datos, Arias, F. (2012) expresa que:” se entenderá por técnica, el procedimiento o forma particular de obtener datos o información, y los instrumentos como medios materiales que se

emplean para recoger y almacenar información” (p.67). Para el desarrollo de la investigación se emplearán las siguientes técnicas e instrumentos de recolección de datos: observación directa y entrevista no estructurada.

3.5.1 Observación Directa

Según Tamayo y Tamayo (2012) la observación, “es aquella en el cual el investigador puede observar y recoger los datos mediante su propia observación”. (p.122). En base a esta definición, la técnica a aplicar para este proyecto será la observación de carácter directo por estar en contacto el fenómeno a investigar con el instrumento en sí, fue de observación simple o no participativa la cual Reyes T. (2012) define cómo: “Una observación con propósitos definidos. El investigador se vale de ella para obtener información y datos sin participar en los acontecimientos de la vida del grupo que estudia, permaneciendo ajeno al mismo”.

3.5.2 Entrevista no estructurada

Esta técnica se definió en base a lo mencionado por Arias, F. (2012) como “una técnica basada en un diálogo o conversación “cara a cara”, entre el entrevistador y el entrevistado acerca de un tema previamente determinado, de tal manera que el entrevistador pueda obtener la información requerida”. (p.73). En base a esto, se empleará como instrumento una entrevista de preguntas abiertas a las personas que se determinaron, con el objetivo de que las respuestas obtenidas sean más variadas y completa, cabe destacar que, cada una de esas interrogantes tendrá una ponderación.

3.6 Instrumento de recolección de datos

De acuerdo con Arias, F. (2012) define que “un instrumento de recolección de datos es cualquier recurso, dispositivo o formato (en papel o digital), que se utiliza para obtener, registrar o almacenar información. (p. 68).

En referencia a lo anterior, es necesario expresar que el instrumento seleccionado será la Lista de Cotejo que consiste en un listado de aspectos a evaluar (contenido, capacidades, habilidades, conductas, etc.), en el cual se puede calificar (“O” visto bueno, por ejemplo, una “X” si la conducta no es la lograda) un puntaje nota o algún concepto.

Además, se utilizará un Cuestionario del tipo Likert, debido a que el diseño de la investigación es de campo y la técnica es la encuesta no estructurada. Este instrumento hace uso de un conjunto de preguntas o ítems que se deben marcar o no, en función de la opinión de aquel que está siendo sometido al instrumento.

3.7 Validación del Instrumento

Según Sabino (2004), describen que “Al grado en que aparentemente un instrumento de medición mide la variable en cuestión, de acuerdo con expertos en el tema”. (p. 204). Es por esto que la validez del instrumento se obtendrá por medio de un juicio de experto en el área. Obteniendo apoyo de personas experta en el área y en el problema planteado, se espera que tales instrumentos sean sometidos a observaciones y corregidos de ser necesario.

3.8 Confiabilidad del Instrumento

Niebel y Freivalds (2009). Definen confiabilidad del instrumento como “El término confiabilidad define la probabilidad de éxito de un sistema, el cual necesariamente debe depender de la confiabilidad o el éxito de sus componentes. Un sistema podría ser ya sea un producto físico con componentes físicos o un procedimiento operativo con una secuencia de pasos o sub-operaciones que deben realizarse correctamente para que el procedimiento tenga éxito”.

3.9 Fases de la Investigación

La metodología que se utilizó para el desarrollo de este proyecto es la XP (eXtreme Programming), este método consta de cuatro fases para su correcto avance, el cual se divide en:

1. Planificación.
2. Diseño.
3. Desarrollo.
4. Pruebas.

Fase I: Análisis o Planificación.

Pressman, R. (2010) asegura sobre la planificación: “una actividad para recabar requerimientos que permite que los miembros técnicos del equipo XP entiendan el

contexto del negocio para el software y adquieran la sensibilidad de la salida y características principales y funcionalidad que se requieren” (P.61). Siguiendo los pasos de la planificación de la metodología XP se utilizará la entrevista no estructurada para obtener los requerimientos del sistema, y así describir la situación actual, en ella se describe un proceso semiautomático, en el cual requiere trabajo humano para la presentación del trabajo, como también las necesidades del nuevo sistema. Por lo tanto, se puede manifestar que esta es la fase más importante de la metodología para el usuario, ya que, en esta fase es donde se satisfacen sus necesidades.

Fase II: Diseño.

Pressman (2010), menciona sobre el diseño: “El diseño XP sigue rigurosamente el principio MS (mantenlo sencillo). Un diseño sencillo siempre se prefiere sobre una representación más compleja” (P.62). Para elaborar el sistema es necesario un diseño sencillo que el cliente irá moldeando a través de cambios, es decir se presentarán prototipos para así disminuir el riesgo de error cuando comience el proceso implementación verdadera y validar las estimaciones originales.

Fase III: Desarrollo o Codificación.

Posteriormente, luego de tener claro los requerimientos y el diseño del sistema de información se procede a la realización de pruebas unitarias para lograr tener una capacitación de lo que se quiere llegar. Cabe destacar que esta fase es en donde se desarrolla la funcionalidad del sistema, así como mejoras del diseño a través de la codificación siguiendo los estándares de codificación.

Fase IV: Pruebas.

Por último, en la fase de la metodología XP se realizarán las distintas pruebas al sistema para determinar el funcionamiento óptimo y planificando del mismo, en caso de haber fallar o errores, realizar respectivas correcciones.

CAPITULO IV

RESULTADOS

En el presente capítulo se exponen los resultados obtenidos en el desarrollo del sistema de información, para esto se requiere el análisis de los diferentes elementos de manera individual y en forma de conjunto. Se implementaron las fases metodológicas propuestas que se asocia con los objetivos planteados inicialmente, para el correcto funcionamiento del sistema.

Todos estos análisis son importantes para elaborar una estructura de trabajo, esto conlleva definir las metas y así tener un control eficiente de los recursos destinados para el proyecto. Con esta finalidad, surgen los distintos tipos de metodologías que permiten definir los parámetros específicos para cada tipo de proyecto existente.

Por esta razón se estudiaron distintas vías metodológicas a través de las cuales es más factible desarrollar una aplicación con estas características, llegando así a la conclusión de que la metodología que cumple con todos nuestros requisitos para el desarrollo del software es la XP, ya que esta es una metodología de desarrollo rápida o ágil.

Esta es basada en una serie de valores y de prácticas de forma secuencial, tal que esta persigue el objetivo de aumentar la productividad a la hora del desarrollo del software, logrando de esta manera dividir el desarrollo de este sistema de información en cuatro fases, tal y como dicta la metodología, siendo como objetivos específicos con los nombres de diagnóstico (planificación), diseño, desarrollo y pruebas, buscando lograr a través de estos el desarrollo óptimo del software.

4.1 Fase I: Planificación.

Para el desarrollo de esta fase, inicialmente se evaluó la necesidad del desarrollo de un sistema para la predicción de targets objetivos, por lo cual se realizó

un cuestionario de 6 preguntas, para emplearla en una entrevista vía electrónica por la aplicación “ZOOM”, para así conocer la opinión de quienes serían los usuarios finales del sistema.

La entrevista constaba de 6 ítems y antes de iniciar se estableció una escala para la evaluación de cada una de estas, cuyos resultados fueron analizados y representados gráficamente.

En una escala del 1 al 10 por favor indique la opción que se ajuste a su punto de vista.

1: Totalmente en desacuerdo / Poco importante / No

10: Totalmente de acuerdo / Muy importante / Si

Figura 1: Escala de Evaluación
Fuente: Estrada & Vásquez (2020)

Ítem 1: ¿Se considera satisfecho con los métodos tradicionales para la determinación de clientes potenciales?

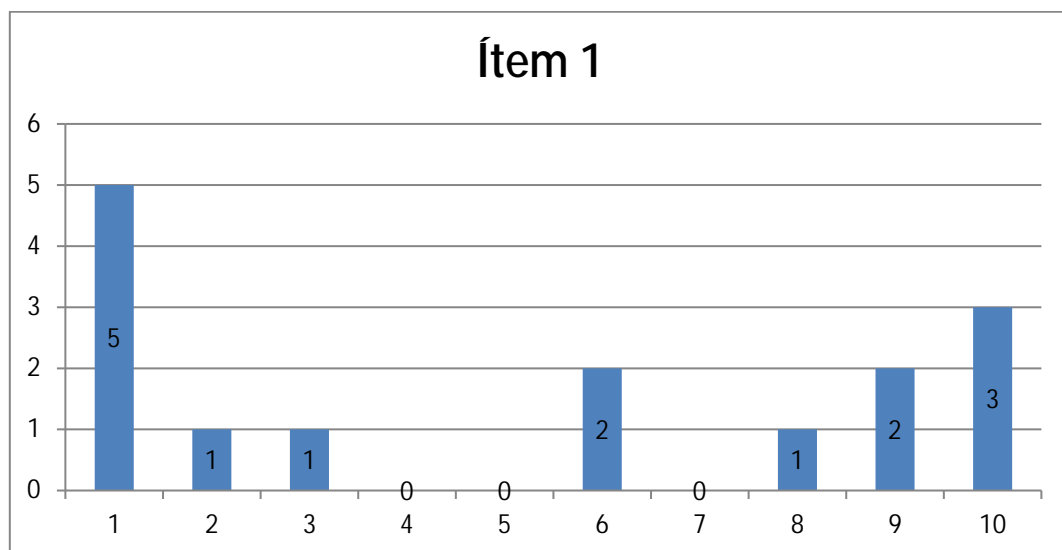


Figura 2: Ítem N° 1
Fuente: Estrada & Vásquez (2020)

De acuerdo con las respuestas emanadas por las personas que fueron entrevistadas se logra apreciar que la mayoría de estas se encuentran en grado de desacuerdo con respecto a los métodos tradicionales que se realizan para la obtención de clientes potenciales, en este caso se infiere que es debido una buena estructura del sistema, tal que, pueda ser amigable con el usuario y además pueda aportar resultados rápidos y precisos.

Ítem 2: De haber oído de ellas, o no de las Redes Neuronales ¿Consideras que una aplicación que determinar los clientes potenciales pueda ser útil?

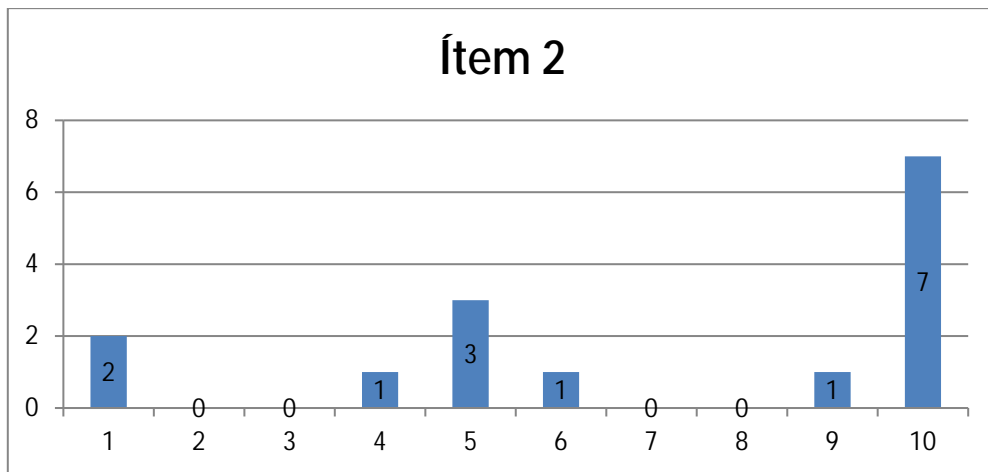


Figura 3: Ítem N° 2
Fuente: Estrada & Vásquez (2020)

Las respuestas obtenidas de este ítem, demuestra el interés que tienen los entrevistados en una aplicación que les permita obtener con mayor facilidad la determinación de un cliente potencial de forma rápida y precisa, de tal forma que esta le pueda ahorrar tiempo.

Ítem 3: De estar de acuerdo ¿Debería contar este sistema con un esquema de colores agradables, que facilite el uso de la aplicación durante largas horas de uso?

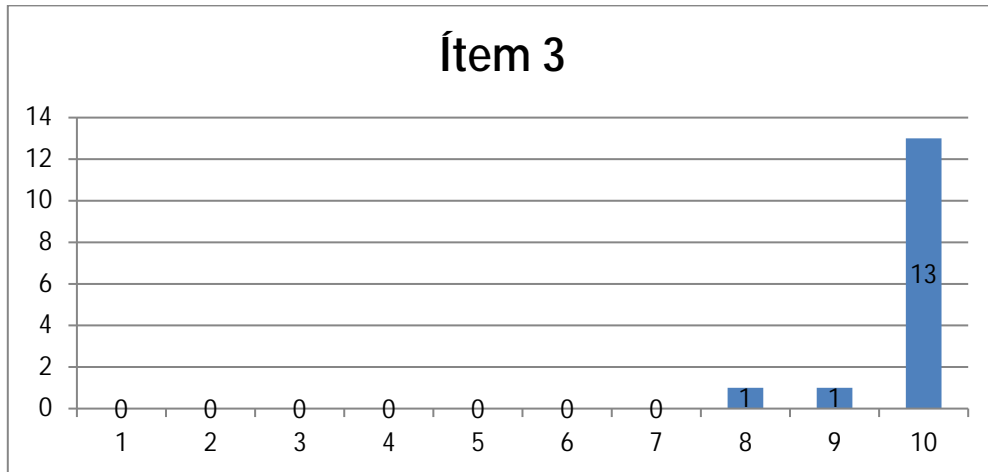


Figura 4: Ítem N° 3
Fuente: Estrada & Vásquez (2020)

Mediante el análisis del ítem, se ha demostrado que casi en su totalidad de los entrevistados se encuentran en total acuerdo, ya que manifiestan que los colores del esquema del sistema deben ser agradables para que facilite el uso de la aplicación por largas horas de uso, gran parte de los entrevistados expresaron una paleta de colores oscuros.

Ítem 4: ¿Le gustaría que los resultados obtenidos estadísticamente se muestren de manera gráfica?

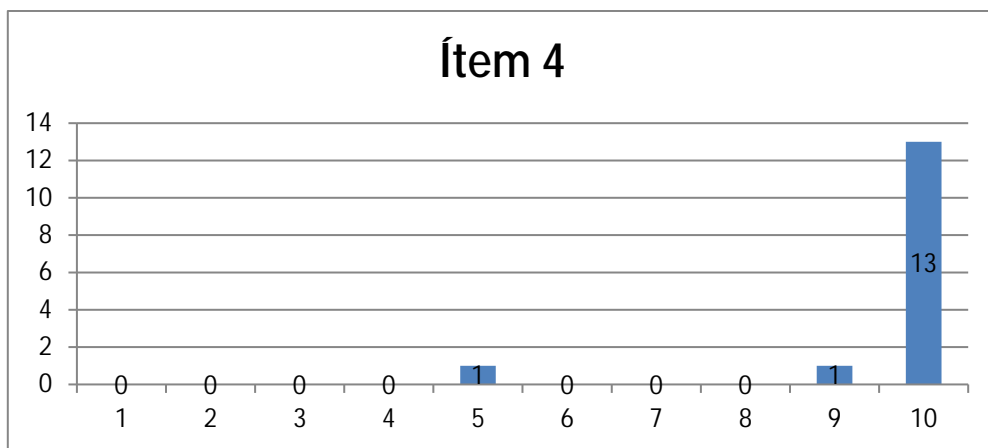


Figura 5: Ítem N° 4
Fuente: Estrada & Vásquez (2020)

Un análisis concienzudo de este ítem evidencia la necesidad de las estadísticas en representación gráfica de las predicciones obtenidas en la determinación del cliente potencial, ya que al poder tener acceso a estas se pueden utilizar para poder generar informes más presentables y veraces.

Ítem 5: ¿Debería contar la aplicación con módulos que separen las actividades de lectura, creación, entrenamiento y predicción de la red neuronal?

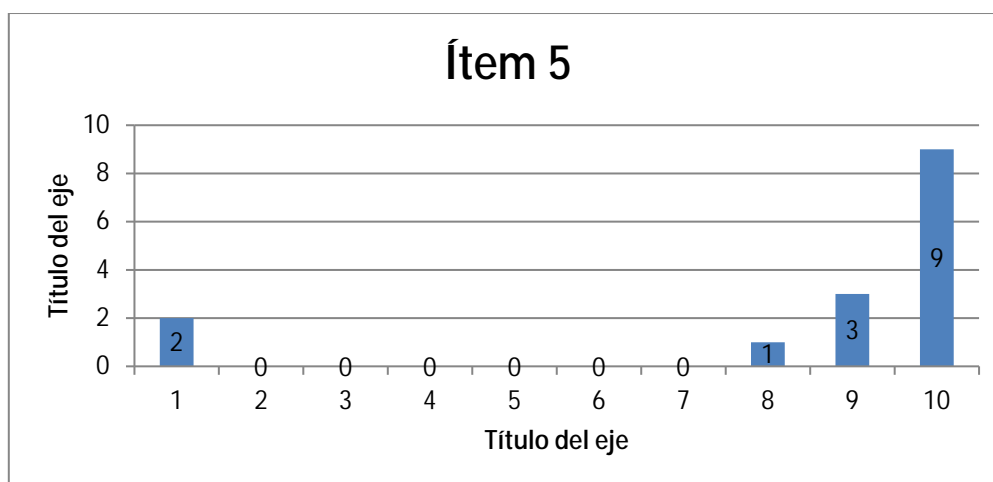


Figura 6: Ítem N° 5
Fuente: Estrada & Vásquez (2020)

Un estudio detallado de este ítem, refleja la necesidad de tener diferentes secciones que separen la lectura, creación, entrenamiento y predicción de la red neuronal para la obtención del target, ya que esto le permite manejar el sistema con mayor facilidad, también teniendo diferentes módulos se puede manejar el sistema de forma independiente de otras opciones.

Ítem 6: ¿Considera necesario el poder crear, guardar y leer conjuntos de datos mediante la aplicación, evitando así, tener que procesar individualmente los datos antes de la ejecución de lectura?

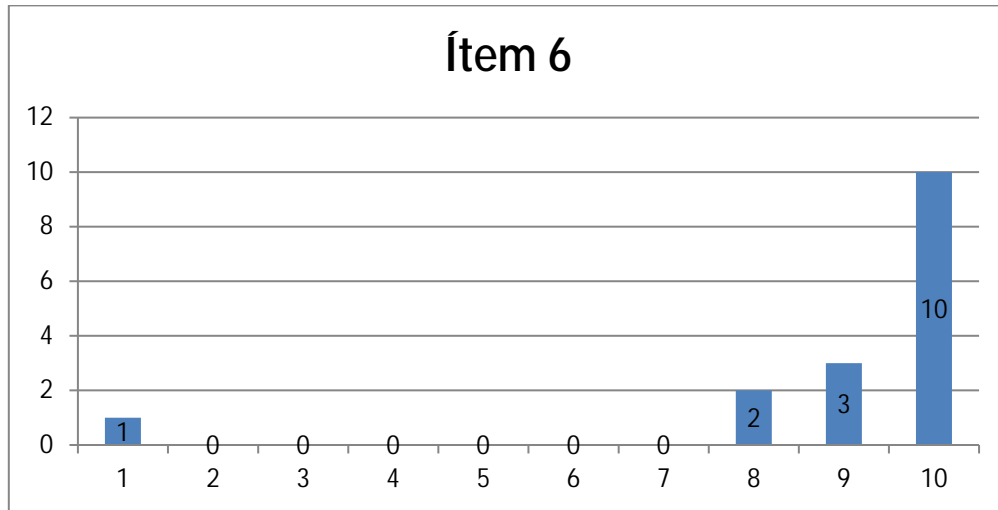


Figura 7: Ítem N° 6
Fuente: Estrada & Vásquez (2020)

Los resultados obtenidos a partir de este ítem, revela que existe un alto porcentaje de la población entrevistada que considera necesario poder suministrar, extraer y guardar conjuntos de datos dentro de la aplicación, ya que procesar individualmente los datos puede generar pérdida de tiempo, recurso valioso para las empresas y negocios.

De acuerdo con el estudio de cada una de las respuestas de la primera fase, se concluyó que existe la necesidad de desarrollar un sistema de información para la predicción de un target objetivo mediante el consumo de un reglón con base a una red neuronal para facilitarte a los ciudadanos una herramienta de apoyo y mejora en sus empresas y negocios, por medio de la cual se pueden utilizar sus recursos inmediatos para realizar predicciones de clientes con un mínimo de limitaciones. Debido a esto se han determinado los requerimientos funcionales y los requerimientos no funcionales del sistema.

4.1.1 Requerimientos Funcionales:

- Añadir archivos de historiales de ventas para actualizar el estado de ventas en el último periodo.

- Procesar los datos en un determinado periodo de tiempo ('desde los inicios', 'desde'- 'hasta').
- Mostrar estadísticas de cómo se comportan los clientes que compran más tus productos.
- Predicción de los clientes potenciales.
- Interfaz para visualizar toda la información de manera cómoda y sencilla.
- Capacidad de eliminar, modificar y añadir nuevos datos.

4.1.2 Requerimientos No Funcionales

- Instalación correcta de sistema eléctrico (los CPU a su máximo funcionamiento consumen mucho voltaje).
- El sistema se basará en uso de lenguaje Python, utilizando Numpy para aplicar funciones matemáticas que permiten el proceso de información en las redes neuronales
- Se mostrará las gráficas del sistema utilizando Matplotlib, para poder visualizar la manera como el sistema determina los datos (depende de la cantidad de datos suministrados), y graficas que muestren el porcentaje de predicción a lo largo del tiempo
- Los datos suministrados se cargarán con Pandas, para poder leerlos y luego analizarlos en el sistema
- Se utilizará el framework de Kivy para tener un control grafica de la información a mostrar
- El sistema tiene que aplicar métodos de aprendizaje para conseguir la red con mejor predicción

4.2 Fase II: Diseño.

Siguiendo la metodología XP, dentro de la fase de diseño se deben realizar distintos prototipos simples, a través de los cuales se utilice el menos tiempo y esfuerzo posible a la hora de ser maquetados y anexados al sistema. Aunado a esto, deben estar ligados al fácil entendimiento del usuario.

Para poder dar inicio a esta fase, se procedió a estudiar a través de distintas estrategias las características del sistema para así realizar un diseño completamente adaptado a los requerimientos determinados anteriormente, iniciando así con la descripción de los actores y consecuente, los diagramas de casos de uso, el cual, ofrece al desarrollador una idea concreta y simplificada de cómo debe comportarse el sistema desde el punto de vista de los usuarios, facilitando de esta manera la planificación del desarrollo, el modelado de los datos y además dejando clara las principales funciones que el sistema debe cumplir.

Cuadro 1

Descripción de los actores: Usuario

Actor	Descripción
Usuario	Este actor será como el usuario básico de la aplicación. Podrá crear, entrenar, visualizar y predecir los resultados provenientes de la red neuronal. Adicional a esto, podrá crear conjuntos de datos necesarios para entrenar la red neuronal.

Cuadro 2

Descripción de los actores: Administrador

Actor	Descripción
--------------	--------------------

Administrador	<p>El actor “Administrador” se encargará de alterar el estado visual de la aplicación. Pudiendo cambiar la forma en el cual se presenta el Menú, botones, widgets y otros componentes de la interfaz gráfica.</p> <p>Funciona como un puente entre la interfaz del usuario y los demás actores, obteniendo la información que generan y adaptándola para mostrar en pantalla.</p>
---------------	---

Cuadro 3

Descripción de los actores: Red Neuronal

Actor	Descripción
Red Neuronal	<p>La función de este actor es ser el intermediario entre la API de bajo nivel y el actor “Administrador”. Es decir, se encarga de obtener el flujo de datos necesarios para crear la estructura de la red neuronal, guardarla y restaurarla.</p>

4.2.1 Casos de Uso

Según Sommerville (2011), los casos de uso identifican las interacciones individuales entre el sistema y sus usuarios y otros sistemas. En su forma más sencilla, un caso de uso identifica a los actores implicados en una interacción y nombra el tipo de interacción. Cada caso de uso debe documentarse con una descripción textual (p. 107).

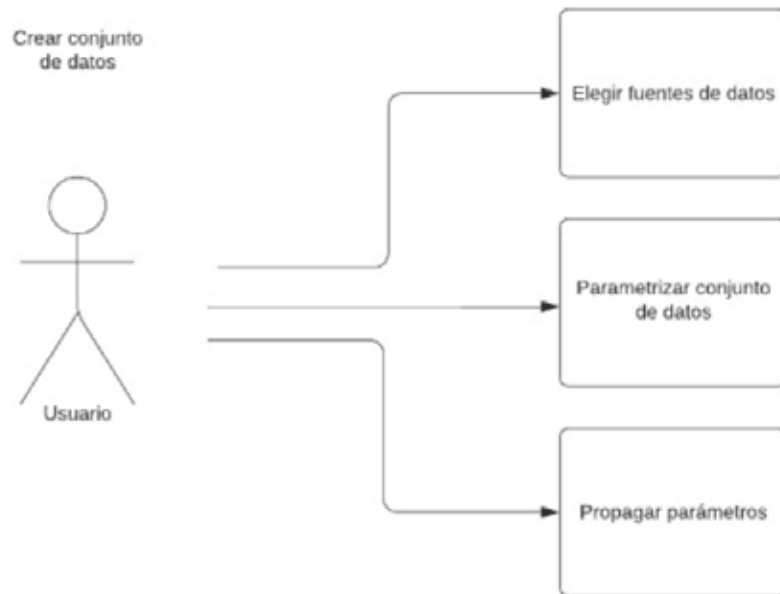


Figura 8: Caso de uso; Crear conjunto de datos

Fuente: Estrada & Vásquez (2020)

Cuadro 4

Definición de caso de uso: Crear conjunto de datos.

Definición del Caso de Uso	
Nombre del Caso de Uso	Crear Conjunto de datos
Actor Principal	Administrador
Actor Secundario	Usuario

<p align="center">Objetivo en Contexto</p>	<p>Crear un conjunto de datos jerárquico, para eliminar la necesidad de tener que procesar y normalizar los datos cada vez que se cargan para ser utilizados en los procesos de entrenamiento y validación</p>
<p align="center">Precondiciones</p>	<p>Deben existir datos en un directorio, ordenado en forma tal que, cada clase identificada sea un directorio único.</p>
<p align="center">Disparador</p>	<p>El administrador decide crear un conjunto de datos con características específicas.</p>
<p align="center">Condición de Término</p>	<p>El conjunto de datos se guarda en el disco de forma exitosa.</p>
<p align="center">Condición de Término Fallida</p>	<p>El espacio de almacenamiento es insuficiente, o el administrador cancela la operación.</p>
<p align="center">Prioridad</p>	<p>Baja. Puede implementarse para extender las capacidades de la aplicación y facilitar la creación de una red neuronal.</p>

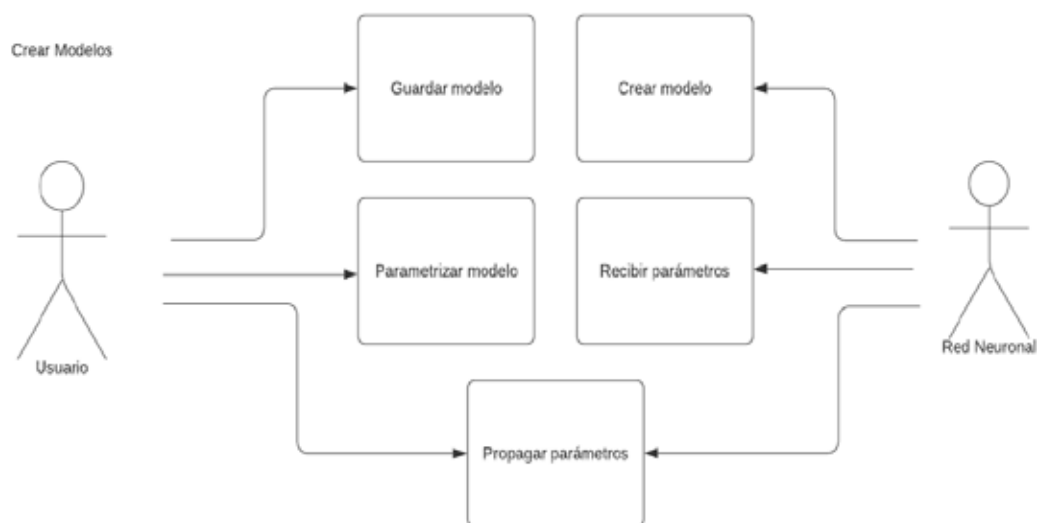


Figura 9: Caso de uso; Crear modelos
Fuente: Estrada & Vásquez (2020)

Cuadro 5

Definición de caso de uso: Crear modelos.

Definición del Caso de Uso	
Nombre del Caso de Uso	Crear modelos
Actor Principal	Red Neuronal
Actor Secundario	Usuario, Administrador
Objetivo en Contexto	Crear un modelo de red neuronal a partir de los parámetros introducidos por el usuario. Esto es para que cada capa pueda ser parametrizada en función de su rol que cumple dentro del modelo.
Precondiciones	Debe seleccionarse un conjunto de datos con la forma definida en los parámetros. Este conjunto es necesario para inicializar el modelo. Se recomienda tengan al menos

	el 10% del total de los datos.
Disparador	El usuario decide crear un modelo de red neuronal.
Condición de Término	El modelo se guarda en disco, o se crea e inicializar y se guarda la inicialización en el disco.
Condición de Término Fallida	El conjunto de datos seleccionado para inicializar no cumple con los factores de forma especificados en la inicialización.
Prioridad	Alta. Debe implementarse obligatoriamente.

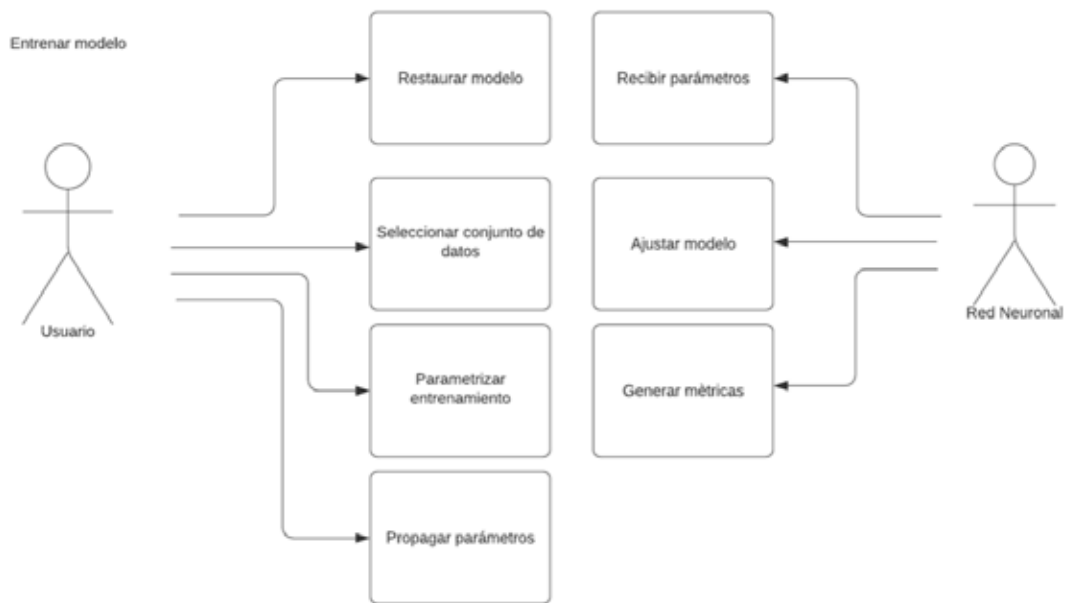


Figura 10: Caso de uso; Entrenar Modelo

Fuente: Estrada & Vásquez (2020)

Cuadro 6

Definición de caso de uso: Entrenar modelo.

Definición del Caso de Uso	
Nombre del Caso de Uso	Entrenar modelo
Actor Principal	Red Neuronal
Actor Secundario	Usuario, Administrador
Objetivo en Contexto	Entrenar el modelo creado, a partir de la información contenida en los conjuntos de datos de entrenamiento
Precondiciones	Debe seleccionarse el modelo a entrenar, el conjunto de datos con la forma correcta (definida en el modelo guardado) y parametrizar el entrenamiento
Disparador	El usuario decide entrenar un modelo de red neuronal con un conjunto de datos, con forma de entrada igual a la definida en el modelo
Condición de Término	Se completa el entrenamiento según los parámetros especificados y se guarda la información del modelo para poder restaurarlos y reentrenar, en caso de que se requiera
Condición de Término Fallida	El conjunto de datos seleccionado para el entrenamiento no cumple con los factores de forma especificados en la inicialización. O los parámetros de entrenamiento no son los adecuados para el equipo de cómputo, lo que produciría un consumo anormal de los recursos del sistema y detendría

	completamente la ejecución del mismo
Prioridad	Alta. Debe implementarse obligatoriamente.

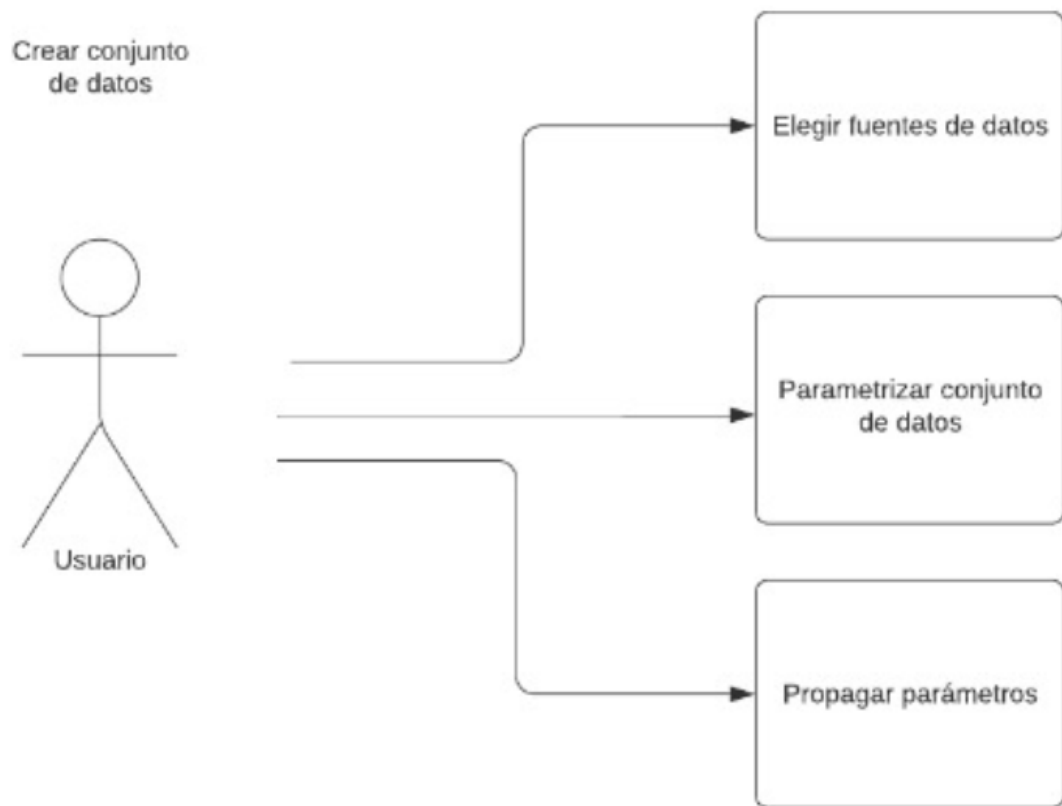


Figura 11: Caso de uso; Detección y Clasificación.
 Fuente: Estrada & Vásquez (2020)

Cuadro 7

Definición de caso de uso: Detección y Clasificación.

Definición del Caso de Uso	
Nombre del Caso de Uso	Detección y clasificación
Actor Principal	Red Neuronal
Actor Secundario	Usuario, Administrador
Objetivo en Contexto	Predecir el cliente potencial, mediante un modelo de red neuronal entrenado en función de los parámetros establecidos por el usuario en la fase de creación y entrenamiento
Precondiciones	Debe existir un modelo entrenado antes de poder predecir la clase de datos o conjunto de datos. Este modelo se podrá cargar desde el directorio en que se haya guardado luego de entrenar el modelo.
Disparador	El usuario decide probar la efectividad de los parámetros de entrenamiento al evaluar la precisión de la predicción del modelo de red neuronal al evaluar la clase de una o varios datos sin necesidad de que hayan sido procesadas.
Condición de Término	Se predice la clase de datos o conjunto de datos. Si es un conjunto de datos puede cancelarse la operación en medio de la operación para que se detenga la evaluación de los elementos.

Condición de Término Fallida	No se encuentra el modelo a cargar en la ruta especificada. Los datos de validación no son los indicados para el modelo cargado.
Prioridad	Alta. Debe implementarse obligatoriamente.

4.2.2 Diagrama de Clases

Los diagramas de clase nos permiten modelar las clases de un sistema en una visión estática en cual se mostraran los atributos, operaciones y relaciones que ejecutan las mismas.

La estructura de esta, es en forma de cuadro o cajas que son la forma representativa más utilizada para las clases e interfaces. Cada cuadro se divide en partes horizontales donde la parte superior se define el nombre de la clase; la sección media hace mención de sus atributos y por último, en la sección más baja, contiene los comportamientos o métodos de la clase.

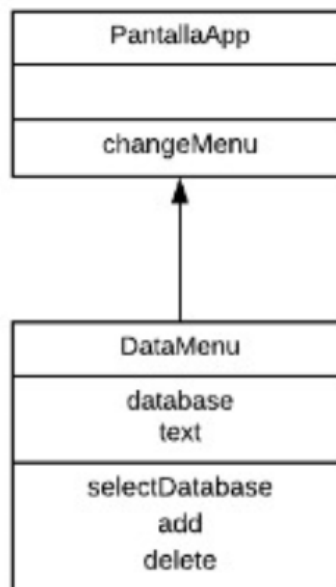


Figura 12: Diagrama de Clase; DataManu.
Fuente: Estrada & Vásquez (2020)

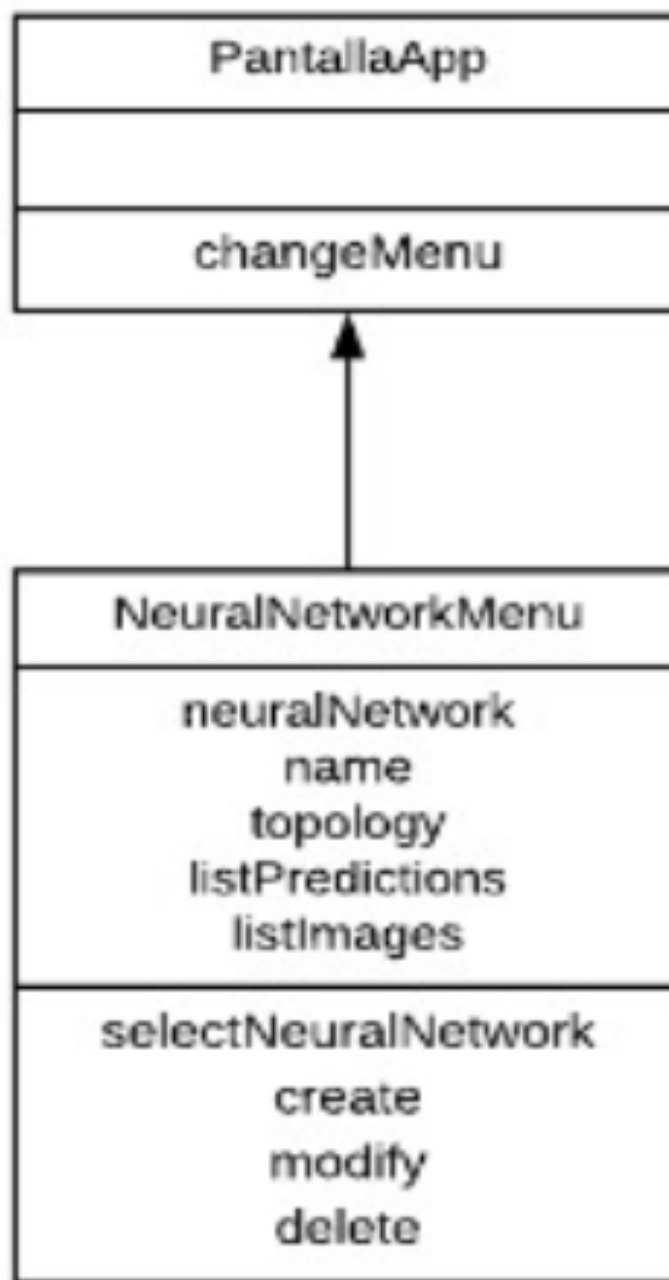


Figura 13: Diagrama de Clase; NeuralNetworkMenu.
Fuente: Estrada & Vásquez (2020)

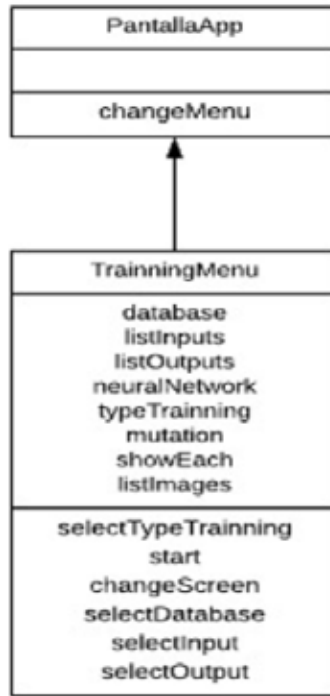


Figura 14: Diagrama de Clase; TrainingMenu.
Fuente: Estrada & Vásquez (2020)

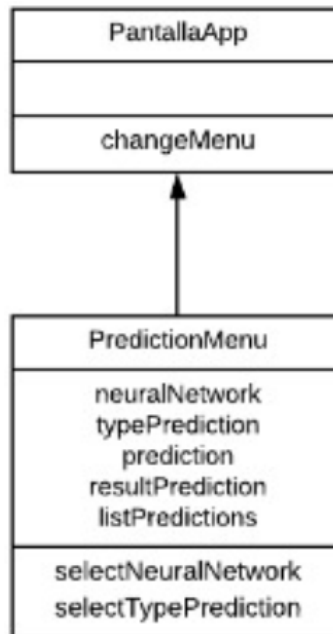


Figura 15: Diagrama de Clase; PredictionMenu.
Fuente: Estrada & Vásquez (2020)

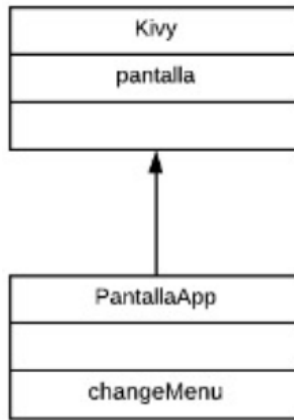


Figura 16: Diagrama de Clase; PantallaApp.
Fuente: Estrada & Vásquez (2020)

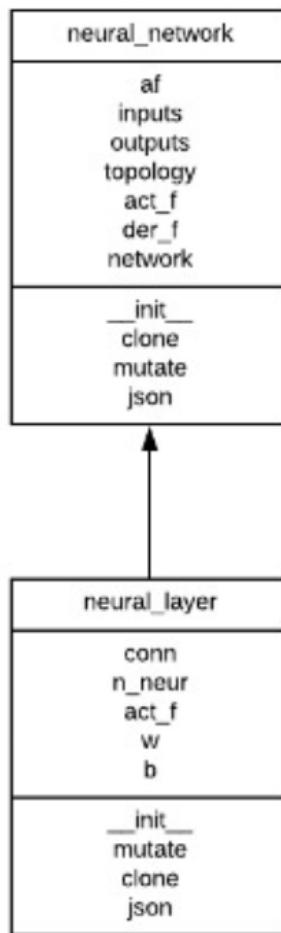


Figura 17: Diagrama de Clase; neural_layer.
Fuente: Estrada & Vásquez (2020)

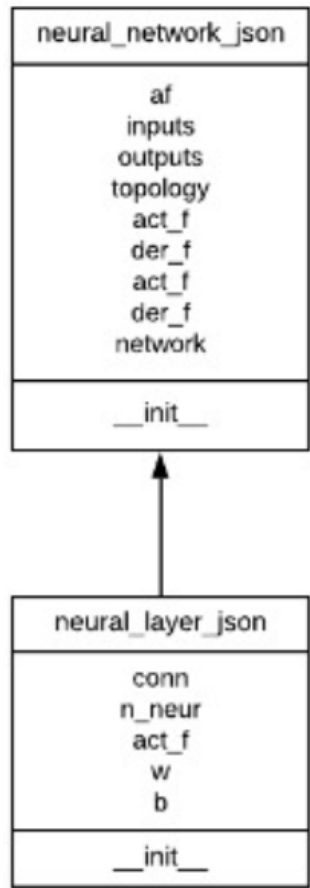


Figura 18: Diagrama de Clase; neural_layer_json.
Fuente: Estrada & Vásquez (2020)

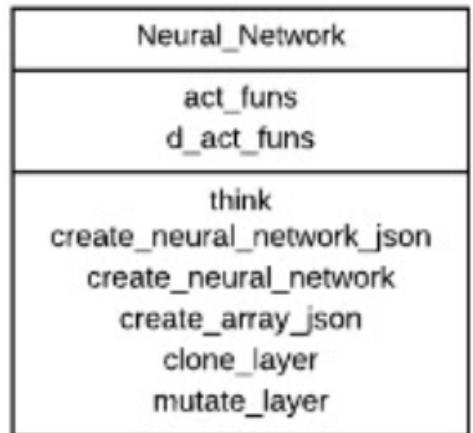


Figura 19: Diagrama de Clase; Neural_Network.
Fuente: Estrada & Vásquez (2020)

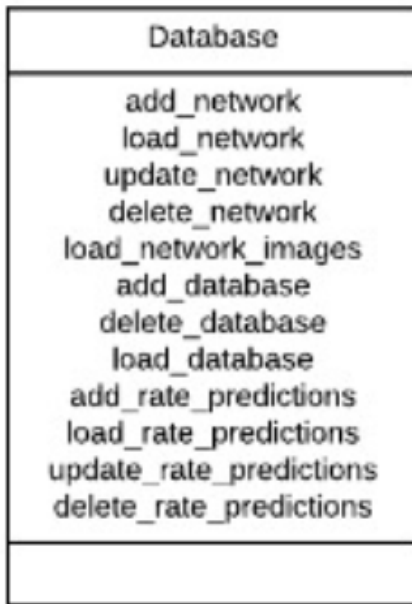


Figura 20: Diagrama de Clase; Database.
 Fuente: Estrada & Vásquez (2020)

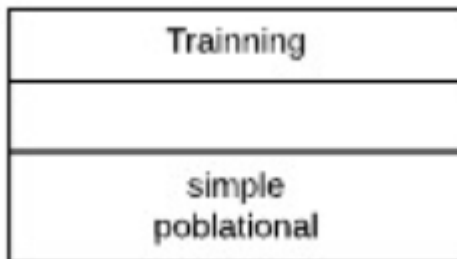


Figura 21: Diagrama de Clase; Trainning.
 Fuente: Estrada & Vásquez (2020)

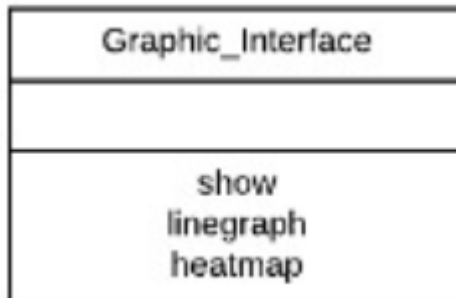


Figura 22: Diagrama de Clase; Graphic_Interface.
 Fuente: Estrada & Vásquez (2020)

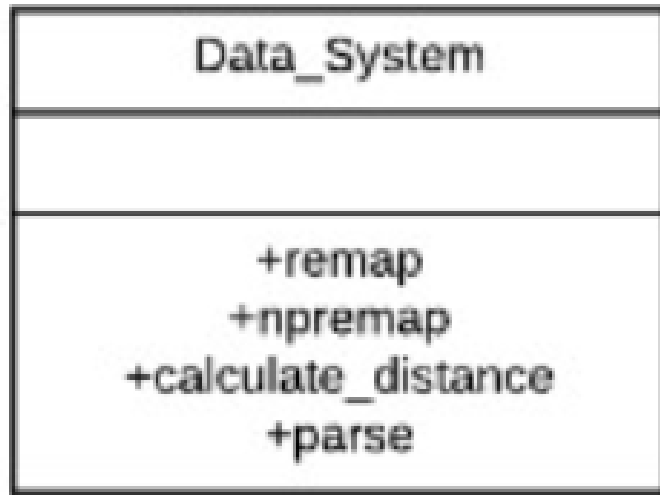


Figura 23: Diagrama de Clase; Data_System.
Fuente: Estrada & Vásquez (2020)

4.2.3 Esquema de la Base de Datos

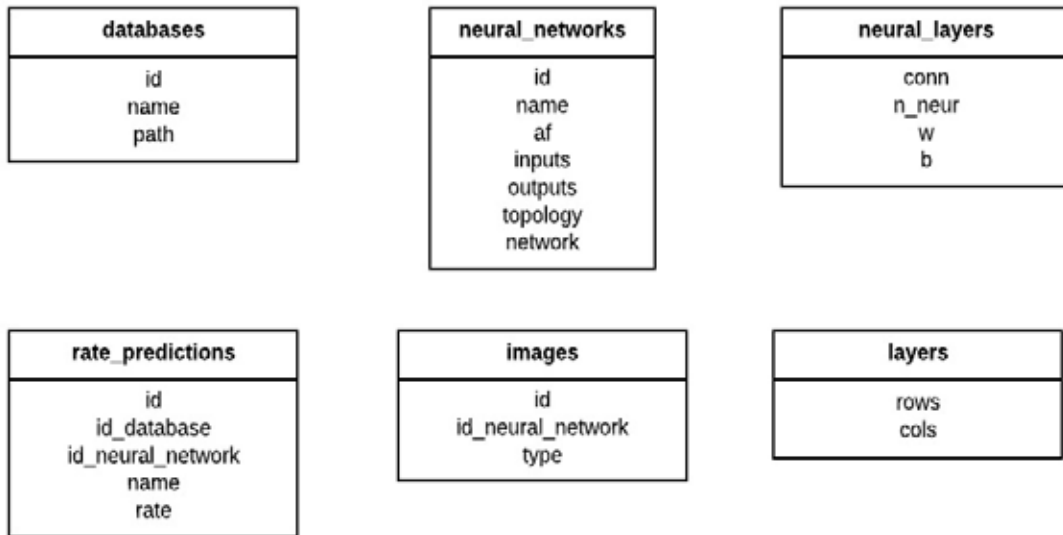


Figura 24: Esquema de la base de datos.
Fuente: Estrada & Vásquez (2020)

4.3 Fase III: Desarrollo o Diseño del sistema



Figura 25: Pantalla; Bases de Datos.
Fuente: Estrada & Vásquez (2020)

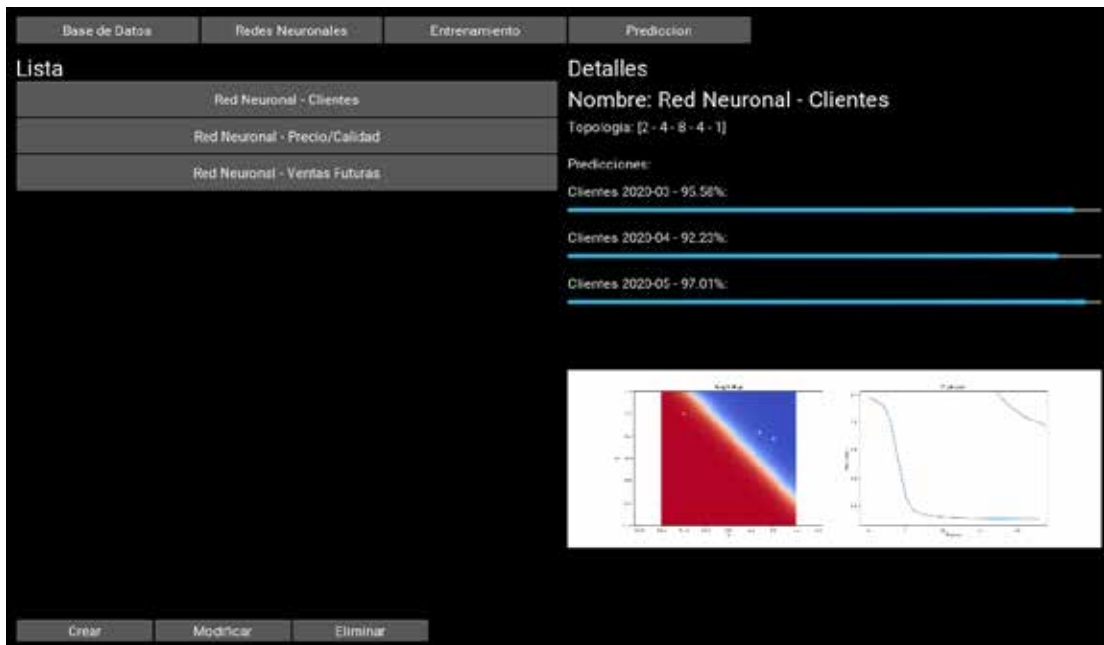


Figura 26: Pantalla; Red Neuronal.
Fuente: Estrada & Vásquez (2020)

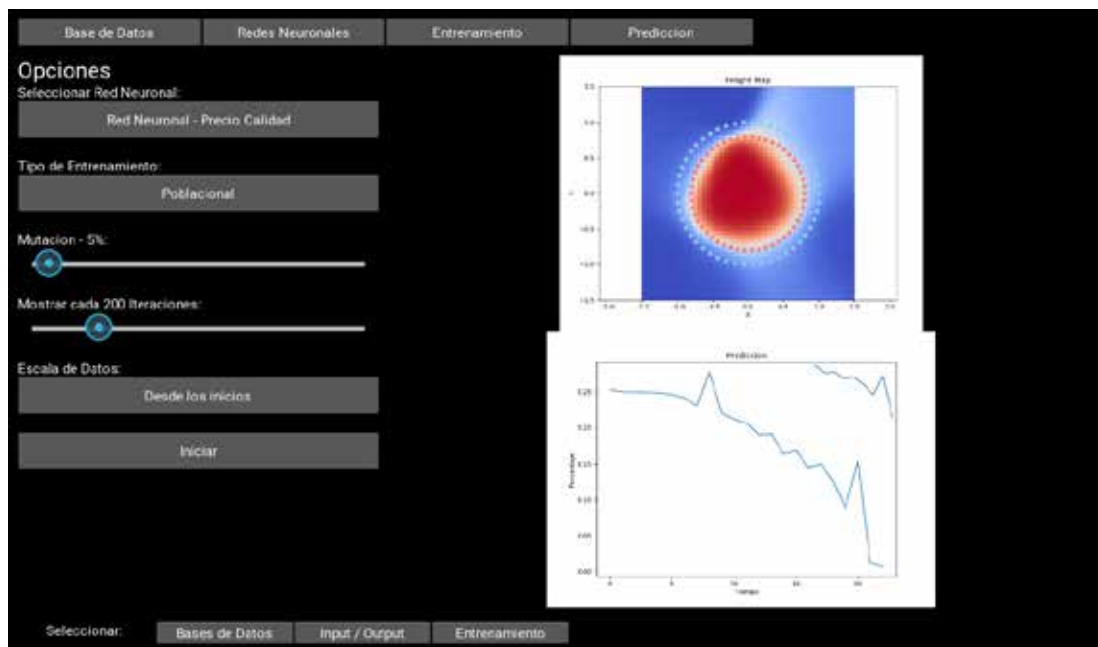


Figura 27: Pantalla; Entrenamiento.
Fuente: Estrada & Vásquez (2020)

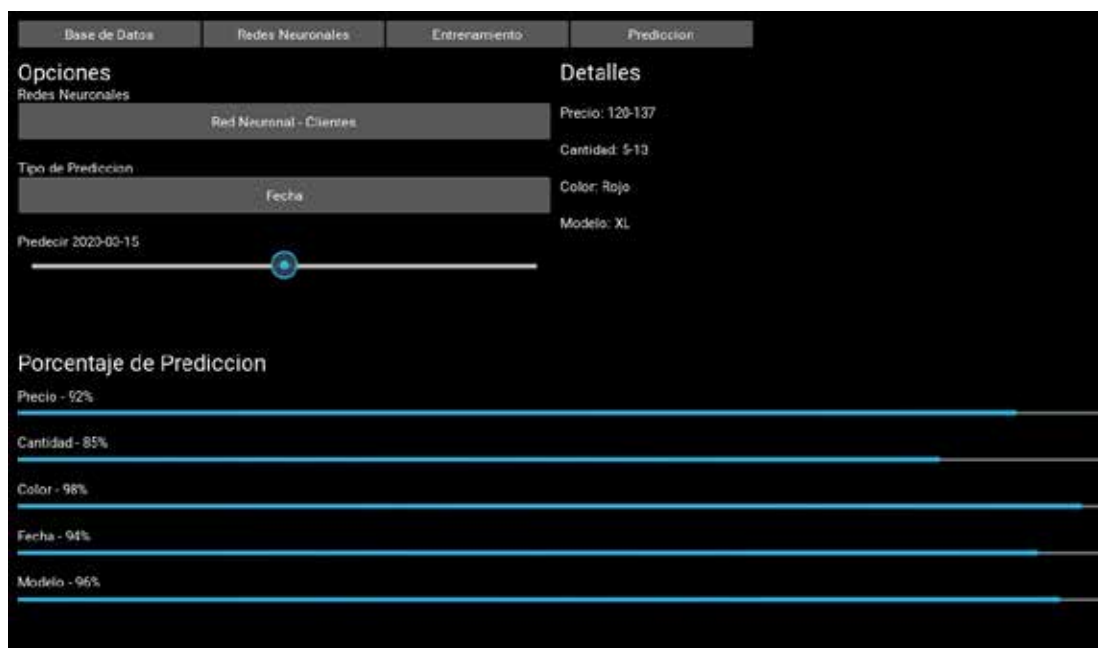


Figura 28: Pantalla; Predicción.
Fuente: Estrada & Vásquez (2020)

4.4 Fase IV: Pruebas.

Esta es la última fase completada de la metodología XP en la misma se llevaron a cabo las pruebas pertinentes para garantizar la efectividad del sistema, estas pruebas se aplicaron a los diferentes módulos y acciones del sistema.

Cuadro 8: Caso de Prueba 1

Añadir una base de datos	
Número 1	Nombre: Caja negra
Descripción: Se comprobó que el usuario puede seleccionar el archivo que desea utilizar	
Condición de ejecución: Formato correcto de la base de datos.	
Entrada: Base de datos.	
Salida: El archivo no pertenece a los formatos admitidos y no clasifica como base de datos.	
Evaluación de Prueba: la base de datos no se cargó.	
Decisión: Adaptar el módulo de carga de datos para admitir ese tipo de formato.	

Cuadro 9: Caso de Prueba 2

Crear Red Neuronal	
Número 2	Nombre: Caja blanca

Descripción: Establecer los parámetros de la nueva red neuronal
Condición de ejecución: Parámetros validos
Entrada: Parámetros
Salida: Creación de la red neuronal
Evaluación de Prueba: Ninguna.
Decisión: Establecer un mínimo aceptable para la topología

Cuadro 10: Caso de Prueba 3

Visualización de las estadísticas	
Número 3	Nombre: Caja negra
Descripción: Comprobación de la gráfica arrojada por la red neuronal	
Condición de ejecución: haber creado la base de datos y la red neuronal	
Entrada: Parámetros de entrada.	
Salida: Estadísticas predictivas	

Evaluación de Prueba: Se visualizan las estadísticas y la gráfica correspondiente de la red neuronal.
Decisión: Tener una base de datos fija para evitar iniciar sin datos.

Cuadro 11: Caso de prueba 4.

Visualizar el comportamiento de la red neuronal	
Número 4	Nombre: Caja negra
Descripción: Comprobación del correcto funcionamiento de la red neuronal	
Condición de ejecución: haber creado la base de datos y la red neuronal	
Entrada: valores estadísticos	
Salida: Comportamiento de la red neuronal	
Evaluación de Prueba: Se muestran correctamente los valores.	
Decisión: Modificar la fuente para mejor visualización.	

CAPITULO V

CONCLUSIONES Y RECOMENDACIONES

5.1 Conclusiones.

Una vez finalizado el desarrollo del Sistema de información para la predicción de un target objetivo mediante el consumo de un renglón con base a una red neuronal, y tomando como referencia los resultados obtenidos en la presente investigación, surgen las siguientes conclusiones:

Mediante la entrevista y la recolección de datos, se logró el correcto uso de las herramientas permitiendo así, realizar una planificación completa de las empresas, pudiendo así hacer todo el levantamiento de los requerimientos funcionales y no funcionales necesarios para formar la mejor estructura a su modelo de implementación sistemática.

Con el desarrollo de una interfaz practica y cómoda, busca la mayor ergonomía para los usuarios que podrán utilizar la aplicación de forma fluida y agradable. Aunado a esto, se logró comprender la necesidad del desarrollo de una interfaz más intuitiva para el correcto funcionamiento de cualquier sistema siendo parte del cuerpo de cualquier investigación.

Por medio del uso de las pruebas, se logró obtener un grado de respuesta adecuada, verificando así, los resultados obtenidos de las predicciones a través de estas, buscando que sean las más precisas y óptimas para la realización de este sistema.

5.2 Recomendaciones.

Actualmente el sistema trabaja de manera local, el cual se tiene que instalar en un computador y posteriormente ingresas las bases de datos conjuntos a la red neuronal. Por ello, se recomienda hacer una implementación es un sistema web, de tal manera que no sea necesario almacenar el sistema en un computador, esto traería

como beneficio que el entrenamiento de la red no utilice los recursos propios del computador y que además los datos sean almacenados en un servidor dedicado, haciendo que sea mucho más rápido y eficiente.

Con una implementación de sistema web, se recomienda crear un sistema de usuario para que estos puedan crear sus propias bases de datos y conjunto a la red neuronal de forma remota y así poder observar el entrenamiento en tiempo real, posteriormente tener acceso a la predicción.

Se recomienda la optimización para la utilización del GPU, ya que esto permitirá que múltiples redes puedan trabajar en simultáneo, por tanto actualmente el sistema funciona en base al procesador.

Por último, se recomienda hacer una expansión para que el sistema pueda aceptar mayor cantidad de formatos de bases de datos, ya que el sistema actual posee la capacidad de utilizar documentos básicos de Excel, sería beneficioso para el sistema poder aceptar cualquier tipo de base de datos que tengan estructuras diferentes.

REFERENCIAS BIBLIOGRÁFICAS

- Acuña, B. (2008). **Como se Elabora un Proyecto de Investigación.**
- Arias, F. (2012). **El proyecto de Investigación: Introducción a la Metodología científica.** Edición N°6. Caracas: Editorial Episteme.
- Lennon J. (2009) **Beginning CouchDB.** Editorial Apress.
- BSON – Binary JSON. (2013). Extraído de: <http://bsonspec.org/>
- Castillo P. (2017). **Aplicación de aprendizaje automático para la predicción de clientes potenciales en procesos de mercadotecnia.** [Documento en línea] Disponible en: <https://cimat.repositorioinstitucional.mx/jspui/handle/1008/755>
- Farias A. (2019). **Aplicación predictiva de cuerpos, fenómenos y objetos astronómicos-espaciales mediante el uso de aprendizaje profundo (DNN, CNN) y visión artificial.**
- Lizarazo Y. (2018). **Predicción de la dosis de coagulante para el tratamiento de agua potable utilizando redes neuronales artificiales.** [Documento en línea] Disponible en: <http://noesis.uis.edu.co/bitstream/123456789/4564/1/172752.pdf>
- López, D. (2012). **Bases de Datos No Relacionales (NoSQL): Cassandra, CouchDB, MongoDB y Neo4j.** [Documento en línea] Disponible en: <http://www.slideshare.net/dipina/nosql-cassandra-couchdb-mongodb-y-neo4j>
- Macea y Matertelo (2016). **Predicción con redes neuronales de un proceso de secado del ñame (Dioscorea Rotundata) en hornos de microondas.** [Documento en línea] Disponible en: <https://repositorio.unicordoba.edu.co/handle/123456789/1030>
- Niebel y Freivalds (2009). **Métodos, estándares y diseño de trabajo.**
- Pressman, R. (2010) **Ingeniería de software un enfoque práctico.** [Documento en línea] Disponible en: <http://cotana.informatica.edu.bo/downloads/Id-Ingenieria.de.software.enfoque.practico.7ed.Pressman.PDF>
- Rossum, G. (2009). **Lenguaje de programación Python**

- Sabino, C. (2004) **El proceso de la investigación**. [Documento en línea] Disponible en: http://paginas.ufm.edu/sabino/ingles/book/proceso_investigacion.pdf
- Sommerville, I. (2011) **Ingeriría de software**. (9º, ed.). México: Pearson Educación de México, S.A. de C.V.
- Tamayo, M. (2012) **Instrumento de recolección de datos**. [Documento en línea] Disponible en: <https://es.slideshare.net/sarathrusta/el-proceso-de-investigacion-cientifica-mario-tamayo-y-tamayo1>
- UJAP. (2007). **Normas para la elaboración y presentación de los anteproyectos, proyectos y trabajos de grado**. Valencia, Venezuela.