

UNIVERSIDAD NACIONAL DE TRUJILLO

Facultad de Ciencias Físicas y Matemáticas

Escuela Profesional de Informática



**Sistema Basado en Conocimiento Mediante RNA para
Pronósticar las Ventas del Restaurante Combiche en el Mall
Aventura**

TESIS

**PARA OPTAR EL TÍTULO PROFESIONAL DE INGENIERO
INFORMÁTICO**

AUTOR: Lopez Andrade Alex Michael Andre

ASESOR: Mg. Bravo Escalante Jorge David

TRUJILLO - PERÚ

2020

**SISTEMA BASADO EN CONOCIMIENTO MEDIANTE
RNA PARA PRONÓSTICAR LAS VENTAS DEL
RESTAURANTE COMBICHE EN EL MALL AVENTURA**

LOPEZ ANDRADE ALEX MICHAEL ANDRE

**SISTEMA BASADO EN CONOCIMIENTO MEDIANTE
RNA PARA PRONÓSTICAR LAS VENTAS DEL
RESTAURANTE COMBICHE EN EL MALL AVENTURA**

Tesis presentada a la Escuela Profesional de Informática en la Facultad de Ciencias Físicas y Matemáticas de la Universidad Nacional de Trujillo, como requisito parcial para la obtención del grado de Título profesional de Ingeniero Informático

ASESOR: Mg. BRAVO ESCALANTE JORGE DAVID

Trujillo - Perú

2020

HOJA DE APROBACIÓN

Sistema Basado en Conocimiento Mediante RNA para Pronósticar las Ventas del Restaurante Combiche en el Mall Aventura

Lopez Andrade Alex Michael Andre

Tesis defendida y aprobada por el jurado examinador:

Prof. Mg. Bravo Escalante Jorge David - Asesor
Departamento de Informática - UNT

Prof. Dr. Jorge Luis Gutierrez Gutierrez
Departamento de Informática - UNT

Prof. Mg. José Arturo Díaz Pulido
Departamento de Informática - UNT

Trujillo, 2020

Dedico esta tesis a:

Mis padres, por el gran apoyo recibido...

Lopez Andrade Alex Michael Andre

Agradecimientos

Agradezco a:

Mis padres, por el apoyo ...

Lopez Andrade Alex Michael Andre

Resumen

Actualmente, el área de ventas en las empresas se apoyan en tecnologías informáticas para el almacenamiento de información de manera que se puede determinar si van aumentando o disminuyendo en determinado tiempo. El presente trabajo de información brinda un sistema de conocimiento basando en Redes Neuronales que ha partir de la información sobre ventas que se tiene, pueda pronosticar el futuro de estas. Se concluyó que el sistema logra pronosticar eficientemente las ventas en el restaurante Combiche del mall Aventura Plaza.

Palabras claves: redes neuronales, ventas, pronostico, sistemas basado en conocimiento.

Abstract

Currently ...

Keywords: neural networks, , intelligent systems.

Índice de figuras

2.1. Red neuronal artificial.	12
2.2. Gráfico de la función sigmoide.	15
2.3. Representación de un red multicapa.	19
2.4. Ecuaciones del backpropagation.	21
3.1. Requerimiento funcional de investigación.	31
3.2. Modelo de diagrama de clases para la investigación de la tesis.	32
3.3. Modelo de estados para la investigación de la tesis.	33
3.4. Modelo de casos de uso para la investigación de la tesis.	33
3.5. Modelo de casos de uso para la investigación de la tesis.	34
3.6. Diagrama de componentes de nuestra investigación.	35
3.7. Modelo de secuencia para clasificar automáticamente documentos.	36
3.8. Ventana Principal del Sistema.	37
3.9. Diagrama de flujo para el desarrollo de la investigación.	39
3.10. Extracción de datos.	40

3.11. Descripción del patrón.	41
3.12. Arquitectura de la aplicación.	41
4.1. Contrastación por método de Chi-Cuadrado en Geogebra.	47

Índice de tablas

2.1. Componentes del perceptrón.	17
2.2. Indicadores para en nuestra investigación.	26
3.1. Requerimiento no funcional del sistema.	32

Índice general

Dedicatoria	I
Agradecimientos	II
Resumen	III
Abstract	IV
Índice de Figuras	VI
Índice de Tablas	VII
1. Introducción	1
1.1. Justificación de la investigación	2
1.2. Formulación del problema	2
1.3. Hipótesis	3
1.4. Objetivos	3

1.4.1. Generales	3
1.4.2. Específicos	3
1.5. Estructura de la tesis	4
2. Materiales y Métodos	6
2.1. Marco teórico	6
2.1.1. Inteligencia Artificial (IA)	7
2.1.2. Sistemas inteligentes	7
2.1.2.1. Capacidades requeridas	9
2.1.3. Red Neuronal Artificial (RNA)	10
2.1.4. Tipos de aprendizaje	11
2.1.4.1. Aprendizaje supervisado	12
2.1.5. Elementos básicos	12
2.1.5.1. Función de entrada (input function)	13
2.1.5.2. Función de activación (activation function)	13
2.1.5.3. Función de salida (output function)	16
2.1.6. Perceptrón	16
2.1.6.1. Perceptrón simple	17
2.1.6.2. Perceptrón Multicapa	18
2.1.7. Backpropagation	20

2.1.8.	Definición de regresión	22
2.1.8.1.	Fases del análisis de regresión múltiple	22
2.1.9.	Python	24
2.1.10.	TensorFlow	24
2.2.	Método de la investigación	25
2.2.1.	Tipo de investigación	25
2.2.2.	Variables de la Investigación	25
2.2.2.1.	Variable Dependiente	25
2.2.2.2.	Variable Independiente	26
2.2.3.	Operacionalización de la variable	26
2.3.	Recolección de datos para la elaboración del modelo	26
2.3.1.	Técnica	26
2.3.2.	Población	26
2.3.3.	Muestra	27
2.3.3.1.	Muestra por conveniencia	27
2.4.	Etapas de la investigación	27
3.	Sistema Basado en Conocimiento Mediante RNA para Pronósticar las Ventas del Restaurante Combiche en el Mall Aventura	29
3.1.	Análisis	30

3.1.1.	Requerimiento funcional	30
3.1.2.	Requerimiento no funcional	31
3.2.	Diseño	32
3.2.1.	Diagrama de clases	32
3.2.2.	Diagrama de modelos de estados	33
3.2.3.	Diagrama de casos de uso	33
3.2.4.	Diagrama de casos de uso del negocio	34
3.2.5.	Diagrama de componentes	35
3.2.6.	Diagrama de secuencia	36
3.2.7.	Diseño de interfaces	37
3.2.8.	Diseño de algoritmo	38
3.2.9.	Extracción de datos	40
3.2.10.	Arquitectura de la aplicación	41
3.3.	Implementación	42
3.4.	Funcionamiento de las redes neuronales	42
3.5.	Función activación	43
3.6.	Funciones del sistema:	44
4.	Resultados y discusión de la tesis	45
4.1.	Resultados computacionales	45

4.2. Discusión de Resultados	46
5. Consideraciones finales	48
5.1. Conclusiones	48
5.2. Trabajos futuros	49
Referencias bibliográficas	50

Capítulo 1

Introducción

Las empresas para elaborar sus planes de mercadeo y su planeación requieren la elaboración inicial de un pronóstico de ventas, el cual les permitirá proyectar las posibles ventas futuras basándose en datos históricos de la misma empresa, con el fin de planear, administrar y controlar los presupuestos necesarios para un buen uso de los recursos que se requerirán para cumplir con las metas propuestas.

El pronóstico de ventas es una herramienta comercial que permite estimar las ventas a futuro, con el fin de establecer metas en un determinado periodo, para su elaboración se tienen en cuenta los resultados históricos y las tendencias de ventas presentadas por el área comercial.

En base a eso, la empresa (restaurant Combiche-Trujillo), mediante sus datos históricos de ventas, desea conocer cuáles serán las ventas que obtendrán en los siguientes periodos posterior a su análisis. Con el apoyo de las redes neuronales, se busca estimar las ganancias

que se obtendrán según el número de ventas pronosticadas , dado que en la actualidad, existen varios restaurantes en la ciudad de Trujillo, y que gracias a la aplicación de dicha herramienta computacional, contribuirá a que la empresa tome medidas de precaución cuando el mercado esté bajo.

1.1. Justificación de la investigación

La importancia de esta investigación desde un punto de vista informático se justifica académicamente, como un aporte para la investigación y una motivación para profundizar en inteligencia artificial, redes neuronales.

Este proyecto se justifica socialmente, ya que no solo beneficiará a la empresa, restaurant Combiche - Trujillo, sino a los comensales que desean degustar las variedades de platos que estan a la venta, y al ser realizado con sistemas basado en conocimiento mediante RNA la aplicación irá mejorando a medida que se vaya usando en el pronóstico de las ventas, ya que tiene la principal ventaja de ir aprendiendo.

1.2. Formulación del problema

El problema de la empresa es establecer medidas de precaución ante las bajas ventas que se puedan tener y para ello se necesita de una herramienta computacional que pronostique, en base a las ventas históricas de la empresa, si hay ganancias o pérdidas de las ventas para

tomar una decisión que beneficie a la empresa. Analizando nuestro problema, nos nace la siguiente pregunta:

¿Cómo pronosticar las ventas de un producto utilizando una red neuronal artificial?

1.3. Hipótesis

Mediante el desarrollo de un sistema basado en conocimiento utilizando R.N.A. se podrá pronosticar las ventas en el Mall Aventura Plaza.

1.4. Objetivos

1.4.1. Generales

Desarrollar un sistema basado en conocimiento mediante RNA para pronosticar las ventas de un producto.

1.4.2. Específicos

- Realizar una investigación bibliográfica para recolectar datos referentes al tema de investigación y analizar los resultados de una encuesta aplicada a un experto en ventas.
- Aplicar la metodología RUP para sistemas basados en conocimientos.
- Diseñar la arquitectura de nuestro sistema basado en conocimiento.
- Entrenar el software basándonos en el modelo computacional diseñado previamente.

- Analizar y comparar los resultados del software contra otros métodos de pronósticos de ventas (lineal, algorítmico).

1.5. Estructura de la tesis

En presente trabajo de investigación tiene los siguientes capítulos:

- En el primer capítulo se desarrolla los aspectos generales así como la justificación, formulación del problema, hipótesis de la investigación y objetivos del presente trabajo de investigación.
- En el segundo capítulo se realiza la recopilación de diferentes conceptos y definiciones que es de mucha importancia a lo largo de toda la investigación, como la metodología de investigación, inteligencia artificial, sistemas inteligentes, redes neuronales, perceptrón, perceptrón multicapa, backpropagation., también se define definiciones de clasificación automática de documentos, así como su representación vectorial y el centroide del conjunto de palabras, similitud aplicando métodos de clasificación automática de documentos, también describimos los tipos de documentos y por último como conceptos finales, consideramos a los gestores, plataformas y lenguajes de programación que utilizaremos así como Visual Studio Code, Tensorflow, Python.
- En el tercer capítulo se analiza, comprende, implementa y optimiza las soluciones de los sistemas inteligentes para la clasificación automática de documentos. Empezando

por algunos problemas de algoritmos actuales. También la prueba a seguir de redes neuronales, se diseñan los modelos propuestos y se explica la implementación del sistema.

- En el cuarto capítulo se contrasta los resultados con el prototipo de la aplicación de sistemas inteligentes para la clasificación automática de documentos.
- En el quinto capítulo se presenta las conclusiones de nuestro trabajo de investigación, demostrando los objetivos cumplidos.

Capítulo 2

Materiales y Métodos

2.1. Marco teórico

El siguiente capítulo presenta conceptos importantes que utilizamos como parte fundamental para el desarrollo de nuestra tesis, este material bibliográfico fue investigado por la relación que tiene con los temas de nuestra investigación. La investigación se tornó muy amplia, principalmente en los conceptos claves para nuestro conocimiento científico, el cual ayudó a desarrollar la tesis, inteligencia artificial, sistemas inteligentes, redes neuronales, perceptrón, backpropagation, conceptos importantes ya que sin ellos sería imposible el modelamiento y desarrollo.

2.1.1. Inteligencia Artificial (IA)

Empecemos esto con una pregunta clave: ¿Qué es la inteligencia artificial?. Aunque es un término de compleja definición, esta puede ser catalogada como una rama de la ciencia que se dedica al estudio de la forma en que se desarrolla el proceso del pensamiento humano para reproducirlo en un ente artificial. Es por ello que para poder lograrlo, la I.A. centra su estudio en dos áreas: el cuerpo humano (estudio de la inteligencia) y el ordenador electrónico. Teniendo en cuenta que para poder reproducir el pensamiento es necesario conocerlo, como primera meta para la investigación de los científicos fue entender los procesos cognoscitivos de la mente humana. De esta manera definir su modelo de inteligencia, para luego proceder a programarlo en ordenadores, simulando cada uno de los procesos y comprobando los resultados de sus teorías. (Prince, 2009).

Según McCarthy (1989), acuñó la expresión “inteligencia artificial”, y la definió como: “la ciencia e ingenio de hacer máquinas inteligentes, especialmente programas de cómputo inteligentes”.

2.1.2. Sistemas inteligentes

Los sistemas inteligentes radican en un conjunto de herramientas y aplicaciones que agrupados llevan a cabo la recopilación, extracción y formato de información obtenida de distintas fuentes con el único propósito de crear medios inteligentes y artificiales para múlti-

ples usos. Frecuentemente se usa para el soporte en la toma de decisiones, sin embargo, esta no es su única funcionalidad.

Son sistemas que, durante su existencia, aprenden y almacenan para luego actuar continuamente de forma interna o externa, de forma que pueda alcanzar su propósito superándose cada vez. Tiene su propio objetivo esencial y para alcanzarlo, escoge una acción tomada de las experiencias que ha almacenado en su memoria.

Para poder hablar de un sistema inteligente es necesario que exista un entorno con el cual el sistema pueda interactuar y, además, el sistema inteligente debe contener “sentidos” que le permitan recibir comunicaciones de dicho entorno y con ello poder transmitir información. El sistema actúa continuamente y cuenta con una memoria para archivar el resultado de sus acciones. Como se dijo anteriormente, este sistema tiene un objetivo, por lo cual, el conseguirlo significaría el hecho de seleccionar la respuesta adecuada ante cada estímulo. También se caracteriza porque a través de su memoria, durante su existencia, aprende de sus experiencias, logrando así, mejorar tanto su rendimiento como su eficiencia. Por último, es necesario indicar que este sistema consume energía, la cual utiliza en sus procesos internos y su actuar. (Cruz, 2012).

2.1.2.1. Capacidades requeridas

Según Cruz (2012), un sistema inteligente podrá ser considerado completo, si se incluye dentro de estas diversas funcionalidades, tales como :

- **Inteligencia:** Es la capacidad de alcanzar nuestros objetivos. La inteligencia incluye la capacidad de aprender a lograrlo.
- **Conceptualización:** Un concepto es el elemento primordial del pensamiento. Es el almacenamiento físico, material de información (en neuronas o electrones). Todos los conceptos de la memoria están interrelacionados en red. La capacidad de poder conceptualizar implica el desarrollo de distintos niveles de abstracción.
- **Reglas de actuación:** Una regla de actuación es la consecuencia de una experiencia o el resultado de interpretar la propia memoria, es decir, relaciona situaciones y efectos de la acción.
- **Memoria:** La memoria es un almacenamiento físico de conceptos y reglas de actuación, que incluye la experiencia del sistema.
- **Aprendizaje:** Posiblemente, es la capacidad más importante de un sistema inteligente, ya que este aprende conceptos a partir de la información obtenida a través de los sentidos. Aprende también reglas de actuación a base de su experiencia, la cual en oca-

siones, la actuación se almacena con su valor. Una regla de actuación aumenta en valor si permitió el logro de un objetivo. El aprendizaje incluye la inserción de conceptos abstractos, a base de ejemplos concretos y la creación de conceptos compuestos que contienen los conceptos de partes de un objeto. Entonces el aprendizaje también se define como la capacidad de detectar relaciones (patrones) entre la parte (“situación la parte”, “situación futura”) de una regla de actuación.

2.1.3. Red Neuronal Artificial (RNA)

Una red neuronal es un sistema de procesadores paralelos conectados entre sí en forma de grafo dirigido. Esquemáticamente cada elemento de procesamiento (neuronas) de la red se representa como un nodo. Estas conexiones establecen una estructura jerárquica que tratando de emular la fisiología del cerebro busca nuevos modelos de procesamiento para solucionar problemas concretos del mundo real. Lo importante en el desarrollo de la técnica de las RNA es su útil comportamiento al aprender, reconocer y aplicar relaciones entre objetos y tramas de objetos propios del mundo real. (Freeman and Skapura, 1991)

Las ventajas de la red neuronal son:

1. Aprendizaje adaptativo: Son adaptables debido a la capacidad de autoajuste de los elementos procesales (neuronas) que componen el sistema.
2. Auto organización: La auto organización consiste en la modificación de la red neuronal

completa para llevar a cabo un objetivo específico. Cuando las redes neuronales se usan para reconocer ciertas clases de patrones, ellas auto organizan la información usada.

3. Tolerancia a fallos: Las redes neuronales pueden aprender a reconocer patrones con ruido, distorsionado o incompletos. Esta es una tolerancia a fallos respecto a los datos. Las redes pueden seguir realizando su función (con cierta degradación) aunque se destruya parte de la red.
4. Operación en tiempo real: Las redes neuronales se adaptan bien a esto debido a su implementación paralela.
5. Fácil inserción dentro de la tecnología existente: Con las herramientas computacionales existentes (no del tipo computadora), una red puede ser rápidamente entrenada, comprobada, verificada y trasladada a una implementación hardware de bajo coste.

2.1.4. Tipos de aprendizaje

Hay dos métodos de aprendizaje importantes, que son: aprendizaje supervisado y no supervisado, la diferencia entre ambos tipos depende en la existencia o no de un agente externo que controle todo el proceso. (Matich, 2001):

2.1.4.1. Aprendizaje supervisado

El aprendizaje supervisado se caracteriza porque el proceso de aprendizaje se realiza mediante un entrenamiento controlado por un agente externo (supervisor, maestro) que determina la respuesta que debería generar la red a partir de una entrada determinada.

En este tipo de aprendizaje se suelen considerar, a su vez, tres formas de llevarlo a cabo, que dan lugar a los siguientes aprendizajes supervisados: Aprendizaje por corrección de error, aprendizaje por refuerzo y aprendizaje estocástico. (Matich, 2001).

2.1.5. Elementos básicos

A continuación se puede ver, en la Figura 2.1, un esquema de una red neuronal:

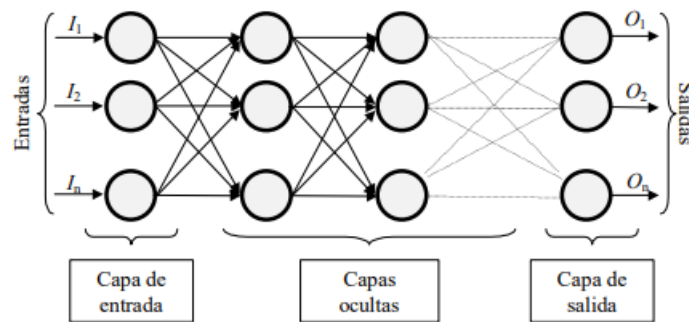


Figura 2.1: Red neuronal artificial.

Fuente: Erlinb (2017).

Esta red neuronal está constituida por neuronas interconectadas y arregladas en tres ca-

pas. Los datos ingresan por medio de la “capa de entrada”, pasan a través de la “capa oculta” y salen por la “capa de salida”. Cabe mencionar que la capa oculta puede estar constituida por varias capas.

2.1.5.1. Función de entrada (input function)

La neurona trata a muchos valores de entrada como si fueran uno solo; esto recibe el nombre de entrada global. Por lo tanto, ahora nos enfrentamos al problema de cómo se pueden combinar estas simples entradas $(in_{i1}, in_{i2}, \dots)$ dentro de la entrada global, gin_i . Esto se logra a través de la función de entrada, la cual se calcula a partir del vector entrada. La función de entrada puede describirse como sigue:

$$input_i = (in_{i1} \bullet w_{i1}) * (in_{i2} \bullet w_{i2} * \dots (in_{in} \bullet w_{in})) \quad (2.1)$$

donde: $*$ representa al operador apropiado (por ejemplo: máximo, sumatoria, productoria, etc.), n al número de entradas a la neurona n_i y w_i al peso. (Matich, 2001).

2.1.5.2. Función de activación (activation function)

Una neurona biológica puede estar activa (excitada) o inactiva (no excitada); es decir, que tiene un “estado de activación”. Las neuronas artificiales también tienen diferentes estados de activación; algunas de ellas solamente dos, al igual que las biológicas, pero otras pueden

tomar cualquier valor dentro de un conjunto determinado. (Matich, 2001).

La función activación calcula el estado de actividad de una neurona; transformando la entrada global (menos el umbral, σ_i) en un valor (estado) de activación, cuyo rango normalmente va de (0 a 1) o de (-1 a 1). Esto es así, porque una neurona puede estar totalmente inactiva (0 o -1) o activa (1). La función activación, es una función de la entrada global (gin_i) menos el umbral (σ_i). Las funciones de activación más comúnmente utilizadas se detallan a continuación:

- Función Lineal

con $x = gin_i - \sigma$, y $a > 0$.

Los valores de salida obtenidos por medio de esta función de activación serán: $a \cdot (gin_i - \sigma_i)$, cuando el argumento de $(gin_i - \sigma_i)$ esté comprendido dentro del rango $(-1/a, 1/a)$. Por encima o por debajo de esta zona se fija la salida en 1 o -1, respectivamente. Cuando $a = 1$ (siendo que la misma afecta la pendiente de la gráfica), la salida es igual a la entrada. (Matich, 2001).

- Función Sigmoidea

$$f(x) = \frac{1}{1 + e^{-gx}}, \text{ con } x = gin_i - \sigma, \quad (2.2)$$

Los valores de salida que proporciona esta función están comprendidos dentro de un rango que va de 0 a 1. Al modificar el valor de g se ve afectada la pendiente de la función de activación.

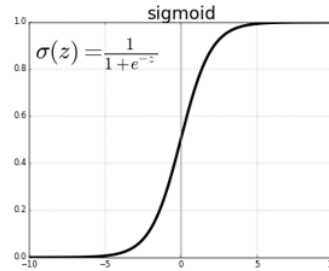


Figura 2.2: Gráfico de la función sigmoide.

Fuente: Elaboración propia.

■ Función Tangente hiperbólica

$$f(x) = \frac{e^{gx} - e^{-gx}}{e^{gx} + e^{-gx}}, \text{ con } x = g \cdot n_i - \sigma, \quad (2.3)$$

Los valores de salida de la función tangente hiperbólica están comprendidos dentro de un rango que va de -1 a 1. Al modificar el valor de g se ve afectada la pendiente de la función de activación.

2.1.5.3. Función de salida (output function)

El último componente que una neurona necesita es la función de salida. El valor resultante de esta función es la salida de la neurona i (out_i); por ende, la función de salida determina que valor se transfiere a las neuronas vinculadas. Si la función de activación está por debajo de un umbral determinado, ninguna salida se pasa a la neurona subsiguiente. Normalmente, no cualquier valor es permitido como una entrada para una neurona, por lo tanto, los valores de salida están comprendidos en el rango $[0, 1]$ o $[-1, 1]$. También pueden ser binarios 0, 1 o $-1, 1$. (Matich, 2001).

Dos de las funciones de salida más comunes son cuando: la salida es la misma que la entrada (función identidad), o binaria, que devuelve uno cuando el act_i es mayor o igual que el umbral ϵ_i , caso contrario es cero.

2.1.6. Perceptrón

El perceptrón lee los valores de entrada, después suma todas las entradas teniendo en cuenta los pesos y por último el resultado lo introduce en una función de activación que nos genera el resultado final. (Espitia Corredor, Daniel Eduardo, 2017).

2.1.6.1. Perceptrón simple

Se determina los pesos sinápticos y el umbral que proporcione el óptimo ajuste de la entrada con la salida, estas variables, para determinar estas variables se sigue un proceso adaptativo, el cual comienza con valores aleatorios y se van modificando según la diferencia entre los valores deseados y los calculados por la red.

Recordar que el perceptrón sólo es capaz de representar funciones lineales, dado que no dispone de capas ocultas, para esto existen los perceptrones multicapa.

Tabla 2.1: Componentes del perceptrón.

Componentes del perceptrón	Definición
Entradas	Es la información que recibe el perceptrón.
Pesos	Son los valores numéricos que se encargan de establecer la influencia de una entrada en la salida deseada.
Bias	Es un parámetro que tienen algunos modelos de redes neuronales el cual permite encontrar fácilmente la separación entre posibilidades de salida de una red neuronal.
Función de activación	Se encarga de determinar un valor de salida una vez se han procesado cada una de las entradas.

Fuente: Elaboración propia según Espitia Corredor, Daniel Eduardo (2017).

En la Tabla 2.1 se representan los componente del perceptrón simple.

El entrenamiento del perceptrón es un proceso iterativo y sigue los siguientes pasos, hasta lograr reducir el error:

- Paso 1: Inicializar los pesos y el bias.
- Paso 2: Calcular las salidas (net) con los pesos y el bias.
- Paso 3: Obtener la salida utilizando la función de activación y calcular cada valor del error.
- Paso 4: Corregir el Bias y los pesos.

2.1.6.2. Perceptrón Multicapa

El Perceptrón multicapa es una red de alimentación hacia delante compuesta por una capa de N neuronas de entrada (sensores), otra capa formada por M neuronas de salida y un número determinado de capas ocultas. (Ver Figura 2.3). El tamaño de éstas dependerán de la dificultad de la correspondencia a implementar. (Munt, 2018).

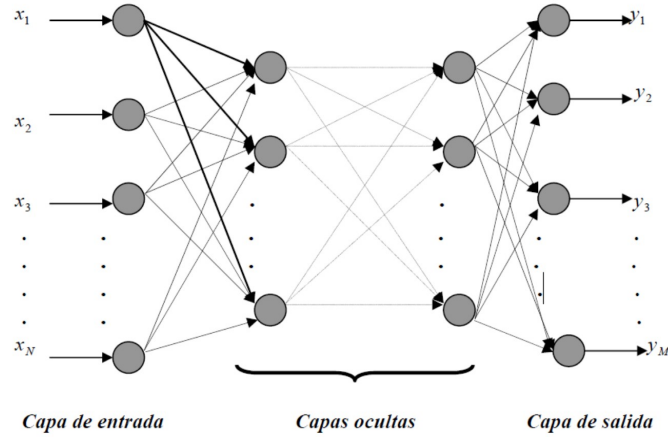


Figura 2.3: Representación de un red multicapa.

Fuente: Alba Morera (2018).

El objetivo que se busca con este tipo de red es el mismo, establecer una correspondencia entre un conjunto de entrada y un conjunto de salidas deseadas, de manera que:

$$(x_1, \dots, x_n) \in \mathbb{R}_N \rightarrow (y_1, \dots, y_n) \in \mathbb{R}_M \quad (2.4)$$

Para ello se dispone de un conjunto de p pares de entrenamiento de manera que sabemos perfectamente que al patrón de entrada (x_1^k, \dots, x_N^k) le corresponde la salida (y_1^k, \dots, y_M^k) , $k = 1, \dots, p$. Así, nuestro conjunto de entrenamiento es:

$$(x_1^k, \dots, x_N^k) \rightarrow (y_1^k, \dots, y_M^k), k=1, \dots, p \quad (2.5)$$

2.1.7. Backpropagation

La red neuronal propaga la señal de los datos de entrada hacia adelante a través de sus parámetros hacia el momento de la decisión, y luego propaga hacia atrás la información sobre el error, a la inversa a través de la red, para que pueda alterar los parámetros. Esto sucede paso a paso:

- La red adivina los datos, utilizando sus parámetros.
- La red se mide con una función de pérdida.
- El error se propaga hacia atrás para ajustar los parámetros equivocados.

La propagación hacia atrás de errores o retropropagación (del inglés backpropagation) es un método de cálculo del gradiente utilizado en algoritmos de aprendizaje supervisado utilizados para entrenar redes neuronales artificiales. El método emplea un ciclo propagación (adaptación de dos fases). Una vez que se ha aplicado un patrón a la entrada de la red como estímulo, este se propaga desde la primera capa a través de las capas siguientes de la red, hasta generar una salida. La señal de salida se compara con la salida deseada y se calcula una señal de error para cada una de las salidas.

Las salidas de error se propagan hacia atrás, partiendo de la capa de salida, hacia todas las neuronas de la capa oculta que contribuyen directamente a la salida. Sin embargo, las neuronas de la capa oculta sólo reciben una fracción de la señal total del error, basándose

aproximadamente en la contribución relativa que haya aportado cada neurona a la salida original. Este proceso se repite, capa por capa, hasta que todas las neuronas de la red hayan recibido una señal de error que describa su contribución relativa al error total.

La importancia de este proceso consiste en que, a medida que se entrena la red, las neuronas de las capas intermedias se organizan a sí mismas de tal modo que las distintas neuronas aprenden a reconocer distintas características del espacio total de entrada. Después del entrenamiento, cuando se les presente un patrón arbitrario de entrada que contenga ruido o que esté incompleto, las neuronas de la capa oculta de la red responderán con una salida activa si la nueva entrada contiene un patrón que se asemeje a aquella característica que las neuronas individuales hayan aprendido a reconocer durante su entrenamiento.

$$\delta^L = \nabla_a C \odot \sigma'(z^L) \quad (\text{BP1})$$

$$\delta^l = ((w^{l+1})^T \delta^{l+1}) \odot \sigma'(z^l) \quad (\text{BP2})$$

$$\frac{\partial C}{\partial b_j^l} = \delta_j^l \quad (\text{BP3})$$

$$\frac{\partial C}{\partial w_{jk}^l} = a_k^{l-1} \delta_j^l \quad (\text{BP4})$$

Figura 2.4: Ecuaciones del backpropagation.

Fuente: Michael A. Nielsen (2015).

Donde: BP1 es para el error en la capa de salida, BP2 es para el error en términos del error en la siguiente capa, BP3 es para la tasa de cambio del costo con respecto a cualquier

sesgo en la red y BP4 es para la tasa de cambio del costo con respecto a cualquier peso en la red.

2.1.8. Definición de regresión

Según Pat Fernández et al. (2013), el término de regresión es uno de los pilares estadísticos más modernos el cual hace referencia al análisis simultaneo de dos o más variables relacionadas entre sí.

Una de las variables se le conoce como variable dependiente (y) y la otra como variable independiente (x).

$$Y = B_0 + B_1X_1 + B_2X_2 + \dots + B_kX_k \quad (2.6)$$

Donde: Y : es la variable dependiente, la cual también es denominada variable respuesta X_i : es la variable independiente i, la cual también se llama exploratoria B_i : es el coeficiente del modelo para la variable X_i

Tanto la variable independiente como las independientes deben ser métricas, aunque las independientes también pueden tener valores cualitativos

2.1.8.1. Fases del análisis de regresión múltiple

Según Pat Fernández et al. (2013), las fases son:

- Identificar problema o área de oportunidad.
- Seleccionar las variables dependientes e independientes.
- Recolectar variables.
- Realizar análisis descriptivo del tipo de relación entre variables.
- Seleccionar método.
- Calcular coeficientes del modelo de regresión lineal múltiple para construir la función.
- Identificar problemas de colinealidad o multicolinealidad.
- Realizar prueba global de la ecuación.
- Efectuar pruebas individuales de los coeficientes.
- Probar cumplimiento de los supuestos del análisis.
- .Interpretar coeficientes de determinación, correlación, determinación ajustado y error estándar.
- Analizar los coeficientes de la ecuación de regresión.
- Elaborar pronósticos puntuales y por intervalo.

2.1.9. Python

Python es un lenguaje de programación interpretado simple pero poderoso que cierra la brecha entre la programación de C y la de shell y, por lo tanto, es ideal para la "programación desechable" y la creación rápida de prototipos. Su sintaxis se construye a partir de construcciones tomadas de una variedad de otros lenguajes; las más destacadas son las influencias de ABC, C, Modula-3 e Icon. El intérprete de Python se puede ampliar fácilmente con nuevas funciones y tipos de datos implementados en C. Python también es adecuado como lenguaje de extensión para aplicaciones C altamente personalizables, como editores o administradores de ventanas. Python está disponible para varios sistemas operativos, entre los que se encuentran varios tipos de UNIX (incluido Linux), el sistema operativo Apple Macintosh, MS-DOS, MS-Windows 3.1, Windows NT y OS / 2. Es conciso, pero intenta ser exacto y completo. La semántica de los tipos de objetos incorporados no esenciales y de las funciones y módulos incorporados se describe en la Referencia de la biblioteca de Python. (Rossum, 1995).

2.1.10. TensorFlow

TensorFlow implementa una visión de la teoría de la probabilidad adaptada al paradigma moderno de aprendizaje profundo de computación diferenciable de extremo a extremo. Basado en dos abstracciones básicas, ofrece bloques de construcción flexibles para el cálculo probabilístico. Las distribuciones proporcionan métodos rápidos y numéricamente estables

para generar muestras y calcular estadísticas, por ejemplo, densidad logarítmica. Los bijectors proporcionan transformaciones de seguimiento de volumen componibles con almacenamiento en caché automático. Juntos, permiten la construcción modular de distribuciones y transformaciones de alta dimensión que no son posibles con bibliotecas anteriores (por ejemplo, pixelCNN, flujos autorregresivos y redes residuales reversibles). Son el caballo de batalla detrás de los sistemas de programación probabilísticos profundos como Edward y potencian la inferencia rápida de caja negra en modelos probabilísticos contruidos sobre componentes de redes profundas.. (Dillon et al., 2017).

2.2. Método de la investigación

2.2.1. Tipo de investigación

La presente tesis es del tipo experimental.

2.2.2. Variables de la Investigación

2.2.2.1. Variable Dependiente

Pronosticar las ventas en el Mall Aventura Plaza.

2.2.2.2. Variable Independiente

Sistema basado en conocimiento mediante RNA.

2.2.3. Operacionalización de la variable

Nos permiten realizar mediciones y determinar la validez de la hipótesis que fue planteada en la investigación.

Tabla 2.2: Indicadores para en nuestra investigación.

Variable dependiente	Descripción	Indicador	
Pronosticar las ventas en el Mall Aventura Plaza	Técnica inteligente para pronosticar las ventas	Sensibilidad:	$\frac{\text{Verdaderos positivos}}{\text{Verdaderos positivos} + \text{falsos negativos}}$
		Precisión:	$\frac{\text{Verdaderos positivos}}{\text{Verdaderos positivos} + \text{falsos positivos}}$
		Exactitud:	$\frac{\text{Verdaderos positivos} + \text{falsos positivos} + \text{falsos negativos}}{\text{Verdaderos positivos} + \text{Verdaderos negativos} + \text{falsos positivos} + \text{falsos negativos}}$
		Tiempo:	$\frac{\text{Total de documentos}}{\text{Representado en segundos}}$

Fuente: Elaboración propia.

2.3. Recolección de datos para la elaboración del modelo

2.3.1. Técnica

2.3.2. Población

La población para nuestra investigación son las ventas del 2018 hasta el 2020 en el Restaurante Combiche en el Mall Aventura Plaza - Trujillo.

2.3.3. Muestra

Este conjunto de datos está conformado por un dataset de 968 registros, que representan las ventas durante el período enero del 2018 hasta agosto del 2020.

2.3.3.1. Muestra por conveniencia

Permite seleccionar aquellos casos accesibles que acepten ser incluidos. Esto, fundamentado en la conveniente accesibilidad y proximidad de los sujetos para el investigador. (Otzen and Manterola, 2017).

2.4. Etapas de la investigación

Nuestra investigación comprenderá las siguientes etapas:

- La investigación bibliográfica a través de búsqueda de artículos en la web, libros y casos de estudio en relación con el tema de investigación (pronósticos de ventas).
- La aplicación de la metodología RUP.
- Recolección de nuestra población, ventas, los cuales se utilizó en el desarrollo de la investigación.
- Teniendo en cuenta que se necesita entrenar a las redes neuronales, puesto que son que

las redes neuronales son heurísticas se requieren un constante ajuste para obtener un óptimo resultado.

- Comparar resultados del software con la información del dataset.

Capítulo 3

Sistema Basado en Conocimiento

Mediante RNA para Pronósticar las

Ventas del Restaurante Combiche en el

Mall Aventura

3.1. Análisis

3.1.1. Requerimiento funcional

La siguiente imagen (ver Figura: 3.1) hace referencia y describen las actividades del sistema.

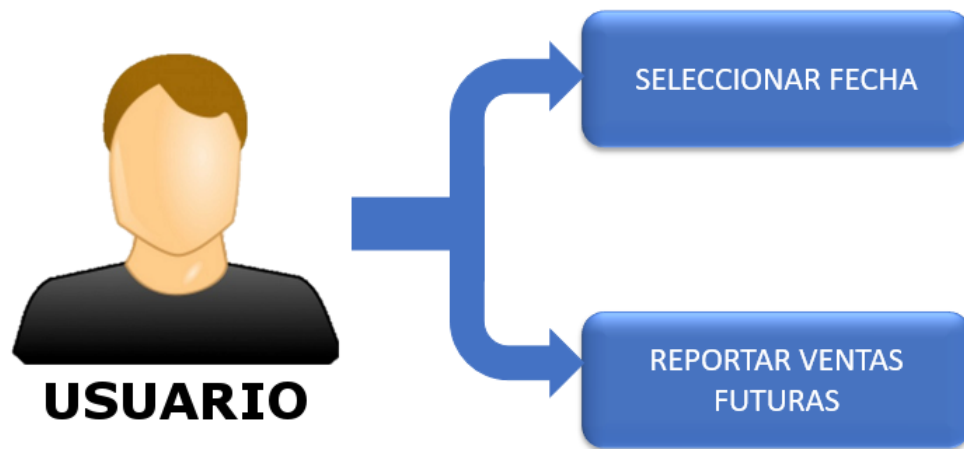


Figura 3.1: Requerimiento funcional de investigación.

Fuente: Elaboración propia.

Usuario:

- Seleccionar la fecha a pronósticar.
- Reportar futuras ventas gráficamente.

3.1.2. Requerimiento no funcional

Describe otras prestaciones, características y/o limitaciones del sistema.

Tabla 3.1: Requerimiento no funcional del sistema.

Requerimiento no funcional	Descripción
Usabilidad	Aplicación de fácil manejo, con interfaz gráfica sencilla.
Eficiencia	Aplicación eficiente, ligera y con respuesta rápida.
Dependibilidad	Aplicación disponible. Aplicación confiable. Aplicación íntegro, sin alteraciones. Fácil mantenimiento de la aplicación.

Fuente: Elaboración propia.

3.2. Diseño

3.2.1. Diagrama de clases

Modelo el cual seguiremos en el desarrollo del sistema.

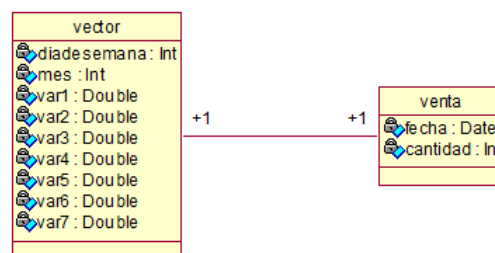


Figura 3.2: Modelo de diagrama de clases para la investigación de la tesis.

Fuente: Elaboración propia.

3.2.2. Diagrama de modelos de estados

Modelo de estados el cual seguiremos en el desarrollo del sistema.

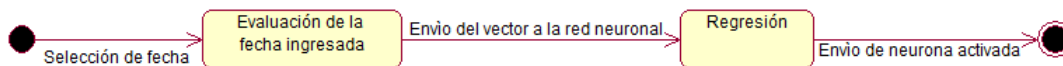


Figura 3.3: Modelo de estados para la investigación de la tesis.

Fuente: Elaboración propia.

3.2.3. Diagrama de casos de uso

Modelo de caso de uso.

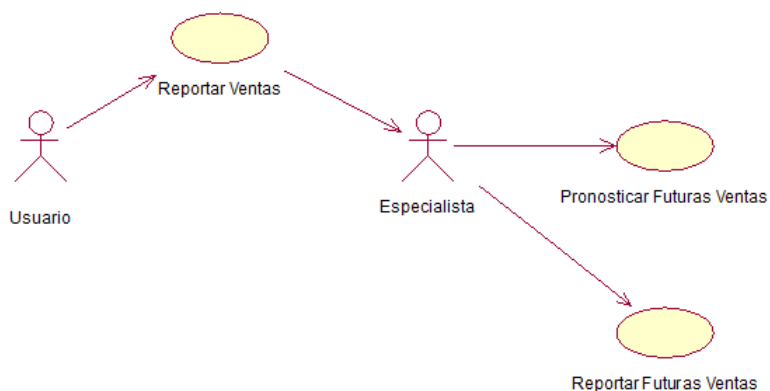


Figura 3.4: Modelo de casos de uso para la investigación de la tesis.

Fuente: Elaboración propia.

- Reportar Ventas, el usuario reporta las ventas a un especialista.
- Pronosticar Futuras Ventas, el especialista pronostica las futuras ventas a partir de los reportes.
- Reportar Futuras ventas, el especialista brinda el pronóstico de las futuras ventas en un reporte.

3.2.4. Diagrama de casos de uso del negocio

Modelo de caso de uso de negocio.

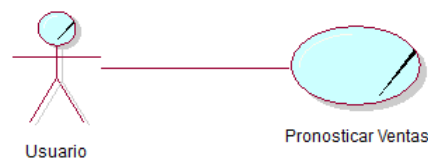


Figura 3.5: Modelo de casos de uso para la investigación de la tesis.

Fuente: Elaboración propia.

- Pronosticar Ventas, El usuario pronostica directamente las futuras ventas mediante un sistema basado en conocimiento.

3.2.5. Diagrama de componentes

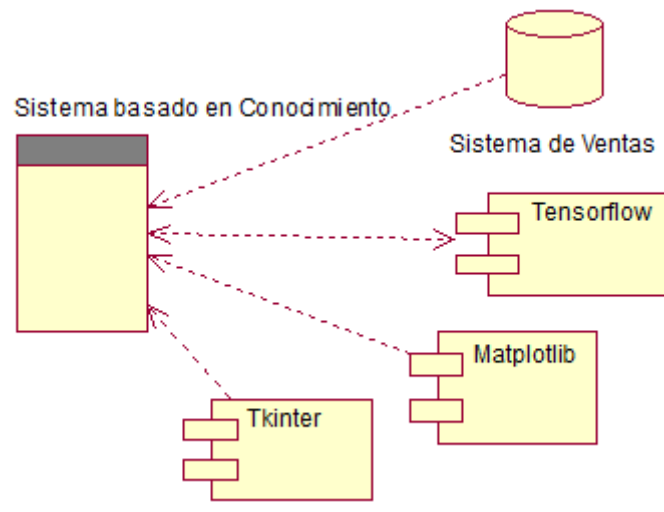


Figura 3.6: Diagrama de componentes de nuestra investigación.

Fuente: Elaboración propia.

3.2.6. Diagrama de secuencia

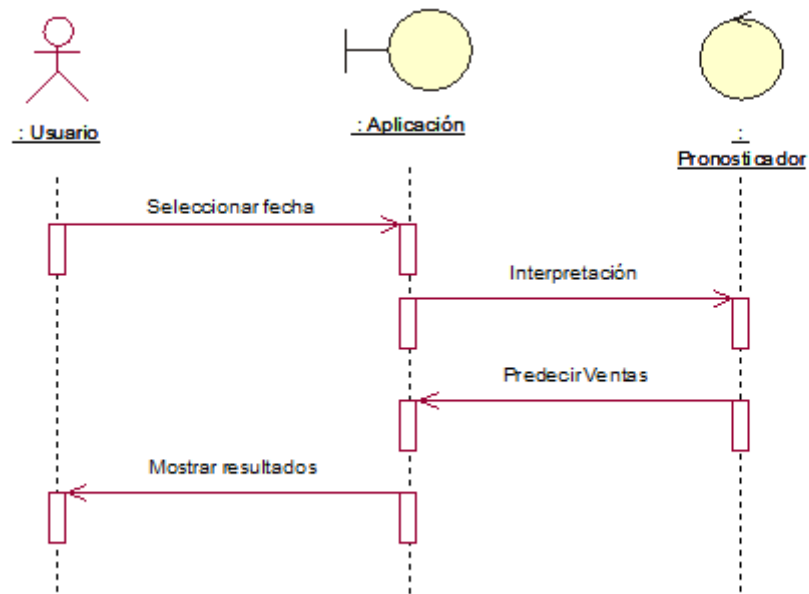


Figura 3.7: Modelo de secuencia para clasificar automáticamente documentos.

Fuente: Elaboración propia.

3.2.7. Diseño de interfaces

A continuación se muestran la interfaz del sistema:

En la Figura 3.8, se muestra la ventana principal.

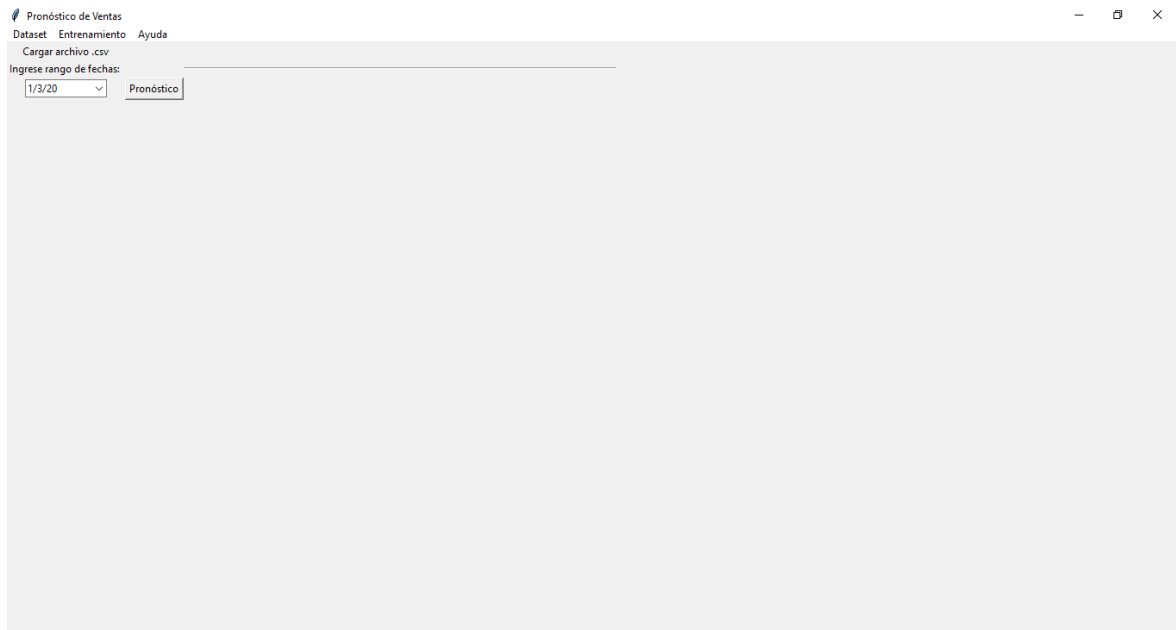


Figura 3.8: Ventana Principal del Sistema.

Fuente: Elaboración propia.

3.2.8. Diseño de algoritmo

El siguiente diagrama de flujo en el cual mostramos el proceso de implementación computacional que seguimos para el desarrollo.

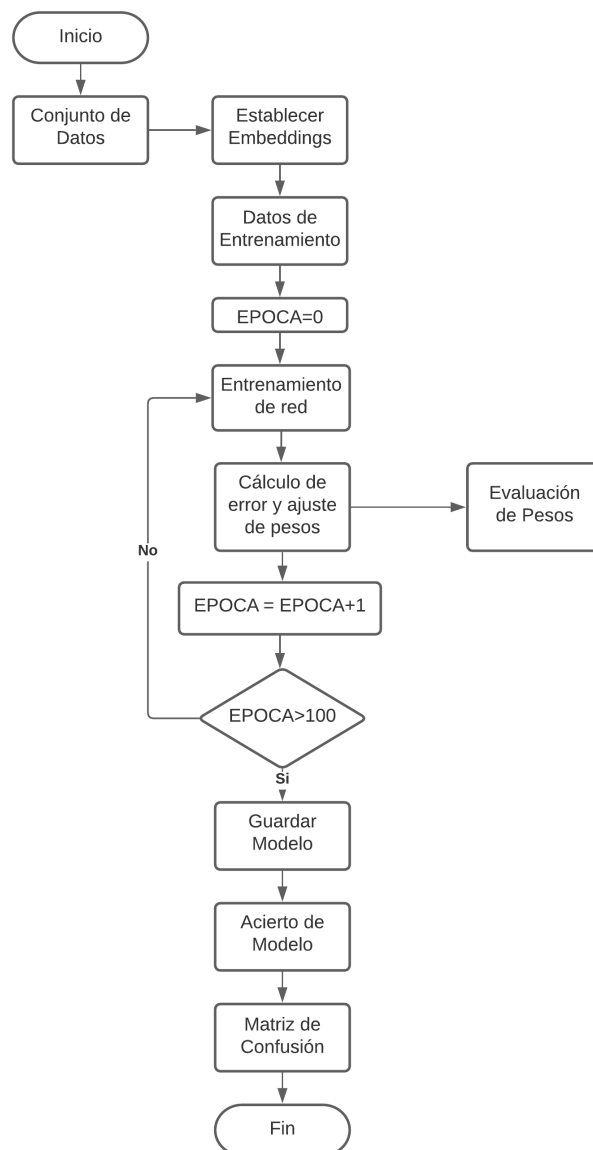


Figura 3.9: Diagrama de flujo para el desarrollo de la investigación.

Fuente: Elaboración propia.

3.2.9. Extracción de datos

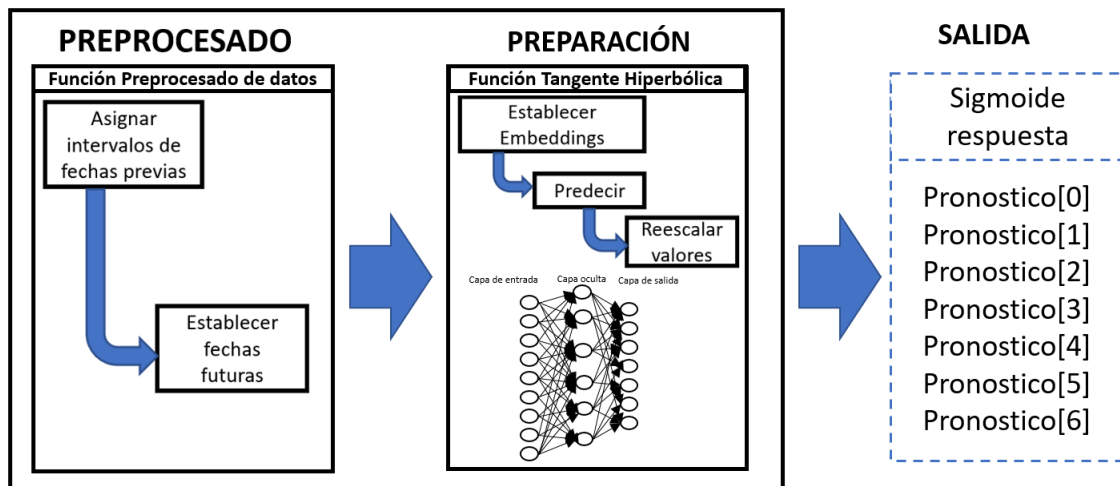


Figura 3.10: Extracción de datos.

Fuente: Elaboración propia.

Preprocesado: Selección de las fechas previas(15 días) a la fecha que se quiere pronosticar, de esta manera se tiene los valores de ventas previas a dicha semana, se prepara las fechas futuras(1 semana) a partir de la fecha seleccionada.

Preparación: Usando los valores de ventas y fechas se le establece sus respectivos valores de Embeddings, se re-escalan los valores para luego predecir en la red neuronal pre-entrenada y se re-escala los valores obtenidos.

Salida: La función tangente hiperbólica nos da como respuesta un vector, el cual indica los valores de las ventas en la próxima semana iniciando en la fecha indicada.

5.0	1.0	-0.533333	-0.202604	-0.507507	-0.103704	-0.965075	-0.103704
-----	-----	-----------	-----------	-----------	-----------	-----------	-----------

Figura 3.11: Descripción del patrón.

Fuente: Elaboración propia.

En este vector se puede visualizar que los primeros dos valores son extraídos del día de la semana y el mes de la fecha, mientras que los otros siete indican los valores de las fechas siguiente pero re-escalado.

3.2.10. Arquitectura de la aplicación

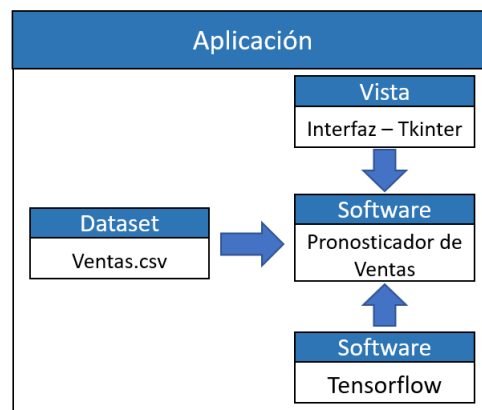


Figura 3.12: Arquitectura de la aplicación.

Fuente: Elaboración propia.

3.3. Implementación

Nos enfocamos en la resolución de la tesis y el desarrollo del sistema basado en conocimiento, un software realizado en python independiente de la IDE, usando tensorflow para el entrenamiento y predicción de la red neuronal feedforward, el cual nos mostrará los resultados óptimos de nuestra pronóstico a través de nuestro sistema.

En el proceso de entrenamiento de la red neuronal se hizo un total de 100 épocas. Los datos utilizados fueron extraídos de la base de datos del restaurante Combiche en el Mall Aventura y almacenados en un dataset (en formato .csv).

3.4. Funcionamiento de las redes neuronales

Al seleccionar la fecha, se establece vectores de fechas previas para predecir la semana siguiente, se establecen embeddings y se interpreta en la red neuronal. Los vectores que se obtienen en la prediccion son re-escalados para poder entenderlos.

3.5. Función activación

Las redes neuronales feedforward tienen la característica de ir hacia adelante, no presentan bucles. La función de activación nos indica el valor de venta de la fecha siguiente a las fechas ingresadas. Considerando que dicho valor aún necesita re-escalarsse.

3.6. Funciones del sistema:

Una de las funciones es mostrar gráficamente la pérdida(valores de la función de coste con los datos de entrenamiento).

Además, reporta gráficamente los valores de testing y el pronóstico para las fechas a probar. Principalmente, muestra gráficamente los valores del pronóstico de la semana de la fecha seleccionada.

Capítulo 4

Resultados y discusión de la tesis

En la aplicación del sistema basado en conocimiento conseguimos resultados esperados que van de acuerdo con la hipótesis planteada.

4.1. Resultados computacionales

La presente investigación tiene como finalidad pronósticar las ventas de la siguiente semana utilizando un set de datos que entrena a una red neuronal fast forward en un sistema basado en conocimiento. Se consultó el pronóstico para la semana del 23 de febrero del 2020, tomando en cuenta que es semanas antes de la cuarentena.

En la tabla 4.1 se muestra el dataset con los valores de venta observados y la fecha, además

de una columna con los valores de venta pronosticados por el sistema basado en conocimiento.

AÑO	MES	DIA	VENTAS OBSERVADAS	PREDICCIÓN
2020	2	23	5	1
2020	2	24	0	5
2020	2	25	16	7
2020	2	26	5	5
2020	2	27	3	3
2020	2	28	2	4
2020	2	29	5	5

4.2. Discusión de Resultados

Para la contrastación se utilizó el método Chi-Cuadrado. Utilizando la Tabla 4.1, podemos decir:

$$df = 6$$

$$Chi - Square = 11,4039$$

$$p - value = 0,0766682$$

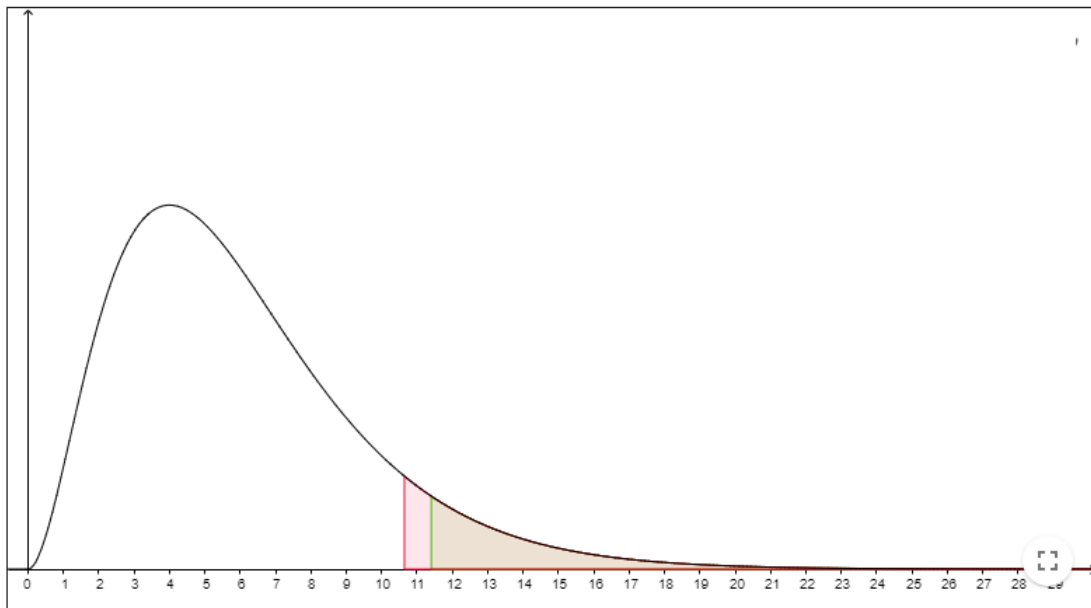


Figura 4.1: Contrastación por método de Chi-Cuadrado en Geogebra.

Fuente: Elaboración propia.

Con los resultados obtenidos por la contrastación, se obtiene el $p\text{-value} = 0.0766682$, valor inferior a 0.1 por lo tanto, llegamos a la conclusión que la hipótesis es estadísticamente significativa.

Capítulo 5

Consideraciones finales

5.1. Conclusiones

Al finalizar la investigación se pudo cumplir con los objetivos específicos propuestos, los cuales se detallan a continuación:

1. Durante esta investigación se pudo explicar la problemática de la pronosticación de ventas así como su solución aplicando un sistema basado en conocimiento, para este caso se determinó usar redes neuronales fast forward, la cual nos permitió pronosticar las ventas de la próxima semana.
2. Se cumplió con la implementación de un prototipo del sistema basado en conocimiento para pronósticos, la cual es muy útil para la toma decisiones.

3. En la investigación se logró descartar la hipótesis nula con un nivel de significancia del 90 %.

5.2. Trabajos futuros

Si bien esta investigación cumple con la hipótesis y objetivos específicos planteados, durante la presente tesis, se presentaron algunas ideas que podrían mejorar la aplicación, por ejemplo:

- Utilizar datasets con una mayor cantidad de registros, para un mejor entrenamiento de la red neuronal y así una obtener una predicción más eficaz.
- Proponer automatización del sistema o módulo de compras al conectarlo con el sistema basado en conocimiento para establecer las compras que se realizaran para esa semana de acuerdo a las ventas pronosticadas.
- Crear un app móvil para acceder al sistema basado en conocimiento de manera remota.

Referencias bibliográficas

Cruz, A. (2012). Definiciones de sistemas inteligentes y 5 ejemplos de sistemas inteligentes. url=<http://www.academia.edu/10376257/Definiciones-de-sistemas-inteligentes-y-5-ejemplos-de-sistemas-inteligentes> . [Online; accedado Octubre 05, 2019].

Dillon, J. V., Langmore, I., Tran, D., Brevdo, E., Vasudevan, S., Moore, D., Patton, B., Alemi, A., Hoffman, M., and Saurous, R. A. (2017). Tensorflow distributions. *arXiv preprint arXiv:1711.10604*.

Erlinb (2017). Redes neuronales: Conceptos básicos y aplicaciones. url=www.perfectcodex.com/redneuronal/2017/12/27/conceptos-basicos/. [Online; accedado Octubre 01, 2019].

Espitia Corredor, Daniel Eduardo (2017). Entrenamiento del perceptron. url=www.platzi.com/tutoriales/1352-ia/2619-entrenamiento-del-perceptron/. [Online; accedado Octubre 06, 2019].

- Freeman, J. A. and Skapura, D. M. (1991). *Neural networks: algorithms, applications, and programming techniques*. Addison Wesley Longman Publishing Co., Inc.
- Matich, D. J. (2001). *Redes neuronales: Conceptos básicos y aplicaciones*. Universidad Tecnológica Nacional, México.
- McCarthy, J. (1989). Artificial intelligence, logic and formalizing common sense. In *Philosophical logic and artificial intelligence*, pages 161–190. Springer.
- Michael A. Nielsen (2015). *Redes neuronales y aprendizaje profundo: Cómo funciona el algoritmo de backpropagation*. url=www.neuralnetworksanddeeplearning.com/chap2.html. [Online; accedado Octubre 01, 2019].
- Munt, A. M. (2018). *Introducción a los modelos de redes neuronales artificiales el perceptrón simple y multicapa*. Universidad de Zaragoza, España.
- MySQL Workbench (2019). *Mysql workbench*. url=www.mysql.com/products/workbench/. [Online; accedado Octubre 07, 2019].
- Otzen, T. and Manterola, C. (2017). Técnicas de muestreo sobre una población a estudio. *International Journal of Morphology*, 35(1):227–232.
- Pat Fernández, L. A., Menchaca, M., Fernández, P., Martínez Luis, D., et al. (2013). *Introducción a los modelos de regresión*. Number EE/519.536 I5.

Prince, B. D. (2009). Inteligencia artificial. ¿qué es? url=<http://inteligencia-artificial-delabarrita.blogspot.pe/2009/10/en-que-consiste-y-en-que-se-fundamenta.html>. [Online; accedado Octubre 04, 2019].

Rossum, G. (1995). Python reference manual.