PRONÓSTICO DE DEMANDA UTILIZANDO INTELIGENCIA ARTIFICIAL

SERGIO DORADO GONZALEZ
VALENTINA VELÁSQUEZ ROBAYO

UNIVERSIDAD ICESI
FACULTAD DE INGENIERÍA
PROGRAMA DE INGENIERÍA INDUSTRIAL
CALI
JUNIO 2020

PRONÓSTICO DE DEMANDA UTILIZANDO INTELIGENCIA ARTIFICIAL

SERGIO DORADO GONZALEZ
VALENTINA VELÁSQUEZ ROBAYO

Proyecto de Grado para optar el título de Ingeniero Industrial

Director proyecto
ANDRÉS FELIPE OSORIO MURIEL

UNIVERSIDAD ICESI
FACULTAD DE INGENIERÍA
PROGRAMA DE INGENIERÍA INDUSTRIAL
CALI
JUNIO 2020

Tabla de Contenido

pág.

1. Contexto, Formulación y Justificación del Problema 9

1.1 Contexto 9

1.2 Formulación del problema 12

1.3 Justificación 12

2. Objetivos 13

2.1 Objetivo del Proyecto 13

2.2 Objetivos Específicos 13

3. Marco de Referencia 14

3.1 Antecedentes o Estudios Previos 14

3.2 Marco Teórico 17

3.3 Contribución Intelectual o Impacto del Proyecto 27

4. Metodología 28

5. Resultados 29

5.1 Selección de los modelos de pronóstico de demanda 29

5.2 Selección de las series de tiempo a utilizar en los pronósticos 30

5.2.1 Recursos utilizados 30

5.2.2 Procedimiento 31

5.3 Aplicación de los modelos 33

5.3.1 Pronósticos tradicionales 33

5.3.2 Pronósticos inteligencia artificial 34

5.4 Comparación de los métodos de pronostico 36

5.4.1 Interpretación de los resultados 38

5.5 Conclusiones 39

5.6 Recomendaciones 40

6 BIBLIOGRAFÍA 42

Lista de gráficas

Gráfica 1 Mercado global de IA.....10
Gráfica 2 Investigaciones sobre AM aplicada a series de tiempo.....14
Gráfica 3 Métodos de pronósticos. 15
Gráfica 4 Serie de tiempo costos de vivienda contra la cantidad de años.....18
Gráfica 5 Serie histórica con patrón cíclico.....19
Gráfica 6 Serie histórica del patrón estacional.....19
Gráfica 7 Serie histórica con variación irregular.....20
Gráfica 8 ARIMA/Dollar/RMSE 34
Gráfica 9 LSTM/Dollar/RMSE 35
Gráfica 10 Series de tiempo comparadas por medio del RMSE normalizado. 37

Lista de ilustraciones

Ilustración 1 De la industria 1.0 a la industria 4.0.....10
Ilustración 2 Bases tecnológicas de las fabricas inteligentes.....11
Ilustración 3 Clasificación de los pronósticos.....21
Ilustración 4 Splits de división de datos. 22
Ilustración 5 Ejemplo de una red neuronal interconectada 23
Ilustración 6 Proceso de pronóstico de la red neuronal. 23
Ilustración 7 Tipos de redes neuronales 24
Ilustración 8 Metodología.....28
Ilustración 9 Análisis grafico de la base de datos “DOLLAR”.....32
Ilustración 10 Análisis de tendencia de la base de datos “DOLLAR”.....32

Lista de ecuaciones

Ecuación 1 Formula Error MSE 26
Ecuación 2 Determinación de variabilidad.....31
Ecuación 3 Ecuación de la línea recta.....31

Lista de tablas

Tabla 1 Revisión de literatura 29
Tabla 2 Clasificación por naturalezas de las series de tiempo 33
Tabla 3 Comparación de los resultados.....36
Tabla 4 Método de pronostico recomendado.....38

Lista de Anexos

Anexo 1 Modelos de pronóstico.....45
Anexo 2 Redes neuronales artificiales.....46
Anexo 3 Análisis y clasificación bases de datos 47
Anexo 4 Resultados pronósticos.....47
Anexo 5 Códigos utilizados.....47

Glosario

AM: Aprendizaje de maquina

ARIMA: Modelo auto regresivo integrado de media móvil

BNN: Bayessian neural network. Traducción: Red neuronal bayesiana

CNN: Convolutional neural network. Traducción: Red neuronal convolucional.

DNN: Deep neural network. Traducción: Red neuronal profunda.

ELM: Extreme learning machine. Traducción: Máquina de aprendizaje extrema.

IA: Inteligencia artificial

LSSVM: Least square support vector machine. Traducción: Máquina de vector de soporte de mínimos cuadrados

LSTM: Long short-term memory. Traducción: Memoria de largo plazo.

MAE: Error absoluto medio

MAPE: Error porcentual absoluto medio

MLP: Multilayer perceptron. Traducción: Perceptrón multicapa.

MSE: Error cuadrático medio

RF*:* Random forest. Traducción: Bosque al azar*.*

RMLSE: Raíz del error logarítmico cuadrático medio

RMSE: Raíz del error cuadrático medio

SES: Suavización exponencial simple

SVM: Support vector machine. Traducción: Máquina vector de soporte.

Resumen

En este proyecto se evaluó la precisión de los métodos de pronóstico de demanda por medio de la comparación entre los modelos de inteligencia artificial frente a los modelos tradicionales, haciendo uso de series de tiempo de diferentes naturalezas (tendencia, estacionales, cíclicas e irregulares). Se examinó el resultado de la métrica de error RMSE para los pronósticos SES, Holt Winters y ARIMA, los cuales consideramos tradicionales contrastados con los modelos CNN, LSTM y MLP de aprendizaje de máquina. La comparación permitió sugerir la elección óptima del modelo según la naturaleza de la serie de tiempo que se desee pronosticar.

Para la realización de este proyecto se recurrió a la literatura de los últimos 3 años con la finalidad de reconocer los modelos más frecuentados, se utilizaron series de tiempo alojadas en repositorios de internet, estas se agruparon por fechas y volúmenes de demanda histórica. Se adaptaron modelos tradicionales y de aprendizaje de máquina por medio del lenguaje de programación Python en la plataforma de Google Collaborative. Finalmente, se logró demostrar que en los métodos de inteligencia artificial la precisión de la predicción es mayor a excepción de las series irregulares donde los métodos tradicionales resultaron más acertados.

Palabras clave: pronósticos de demanda, series de tiempo, inteligencia artificial, aprendizaje de máquina, redes neuronales artificiales, SES, Holt Winters, ARIMA, CNN, LSTM, MLP.

Introducción

Los pronósticos de demanda, se han constituido como una de las herramientas más útiles para las operaciones, permiten una estimación futura de la demanda en diferentes escenarios, por ejemplo, cantidad de materia prima, estimación de presupuestos en compras, contratación de mano de obra, manejo de inventarios, entre otros. (Anderson, Sweeney y Williams, 2004 como se cita en Botero y Álvarez, 2013).

El mercado presenta comportamientos muy variados y dinámicos, lo cual provoca que cada vez sea más complejo anticiparse a los cambios, por esta razón constantemente se generan nuevas herramientas que contribuyen a predecir los sucesos futuros, teniendo como finalidad reducir los errores en las predicciones, controlar los riesgos y obtener una disminución en los costos de operación.

Para obtener beneficios significativos en pronósticos de demanda es necesario utilizar las herramientas correctas. Esta investigación evalúa métodos de pronóstico tradicionales y de inteligencia artificial, situándolos en distintos comportamientos de series de tiempo o naturaleza de demanda, siendo su finalidad demostrar cuáles resultan más precisos, confiables y útiles para ejercicios académicos, comprendiendo su viabilidad y posible aplicabilidad a la industria.

1. Contexto, Formulación y Justificación del Problema

1.1 Contexto

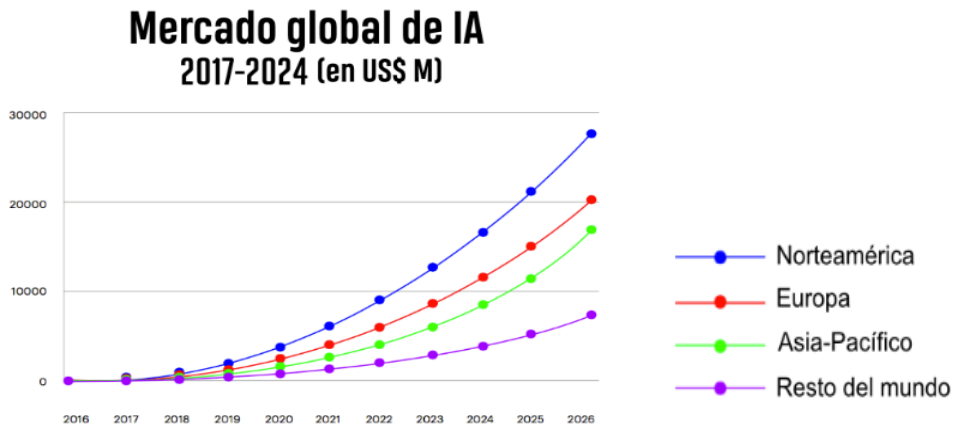
Los pronósticos son vitales para toda organización de negocios, así como para cualquier decisión importante de la gerencia (Chase, Jacobs y Aquilano, 2009). El pronóstico es la base de la planeación corporativa a largo plazo, es una herramienta vital para el correcto funcionamiento de una compañía, dado que ayuda a la planeación de presupuestos, control de costos, estimación de productos nuevos, toma de decisiones en selección de procesos, planeación de capacidades, distribución de instalaciones, planeación de la producción, programación e inventario.

Al largo de los años, se han ideado distintos métodos para pronosticar la demanda en las empresas, entre los más reconocidos y utilizados está la suavización, proyección de tendencia, análisis de regresión, entre otros. Sin embargo, cada vez es más frecuente la presencia de relaciones no lineales y la independencia entre variables que limitan estos modelos (Fallis, 2014).

Es indispensable que las herramientas se actualicen constantemente con los avances tecnológicos para lograr pronósticos más acertados, pues la disminución de la diferencia entre la demanda real y el resultado esperado en el pronóstico trae beneficios significativos para las empresas, por ejemplo, al lograr ajustar los efectos del precio, la actividad promocional y los gastos publicitarios a la demanda, aumentaría la rentabilidad en las estrategias de marketing (Hoptroff, 1993).

Los pronósticos por medio de aprendizaje de máquinas (Machine Learning), el cual es una rama de la inteligencia artificial, permite obtener resultados dinámicos, basado en la actualización del mercado y con una alta expectativa de mejora a lo largo de la práctica. A diferencia de los métodos tradicionales de pronóstico, estos no son lineales, se ajustan de forma matemática a eventualidades, según la revista Forbes (Columbus, 2019) la inteligencia artificial y el aprendizaje de máquinas tienen el potencial de crear un valor adicional de $ 2.6T US Marketing y Ventas, y hasta $ 2T US en fabricación y planificación de la cadena de suministro para el 2020.

Actualmente, los países desarrollados, como Alemania, Dinamarca, Estados Unidos, entre otros, están utilizando la inteligencia artificial (IA) para optimizar sus procesos. La *gráfica 1* ilustra el crecimiento esperado del mercado de la IA en millones de dólares para diferentes regiones geográficas proyectadas a una década, dejando claro la adopción acelerada que está teniendo esta tendencia tecnológica.



Gráfica 1 Mercado global de IA.

Fuente: Adaptada al español Imagen extraída de (Ward, T. 2015).

Las potencias mundiales están en la búsqueda de la globalización económica manifestándose en los rápido avances tecnológicos, los cuales obligan a las industrias a permanecer a la vanguardia para prevalecer. En palabras del Presidente Xi Jinping en la Ceremonia de Apertura de la Conferencia Anual de 2017 del Foro Económico Global: “La economía global se encuentra en un momento de cambio de ciclo, la maquinaria de crecimiento tradicional tiene un débil efecto sobre el impulso económico, la inteligencia artificial, la impresión en 3D y otras nuevas tecnologías emergen continuamente en gran cantidad”. Además, según Del Val Román (2016) la cuarta revolución industrial viene de la mano de la transformación digital, como se puede apreciar en la *ilustración 1* cada una de estas revoluciones se efectuaron con base en un cambio.

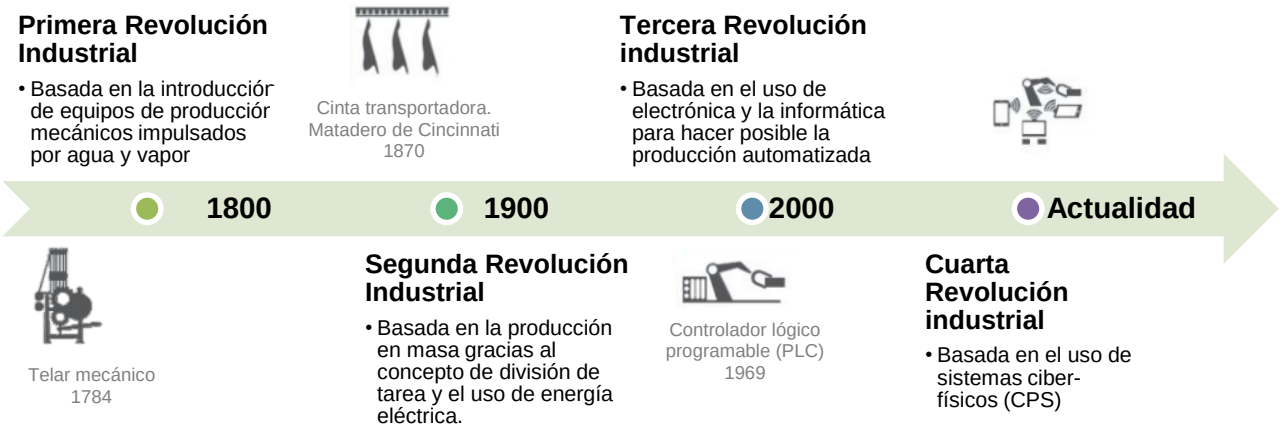


Ilustración 1 De la industria 1.0 a la industria 4.0.

Adaptado de: (Del Val Román, 2016)

Según Del Val Román (2016) para esta última revolución se necesita una trasformación de la industria, adaptarse a las necesidades del cliente, a productos personalizados, servicios innovadores, calidad en los procesos y una experiencia individualizada, para esto se requiere de soluciones inteligentes, innovación extendida, redes colaborativas agiles y conectadas, por último, se requiere de un control de la producción descentralizada en fábricas inteligentes. Para estas fábricas se demandan bases tecnológicas, como las que se muestran a continuación en la *ilustración 2*.

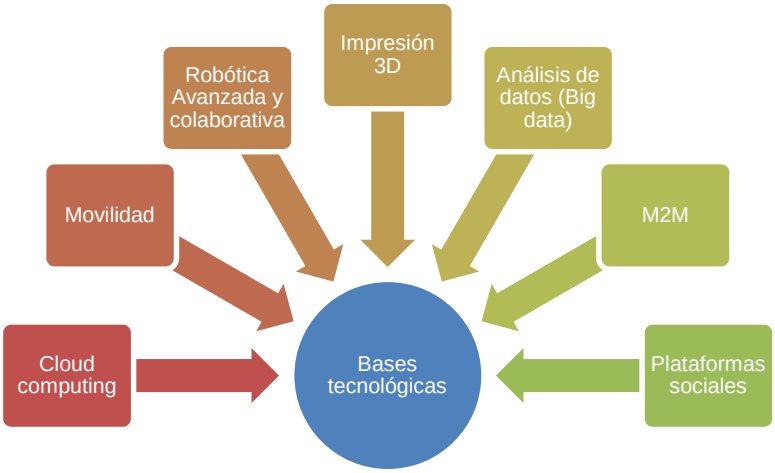


Ilustración 2 Bases tecnológicas de las fabricas inteligentes.

Adaptado de: (Del Van Román, 2016)

Para este proyecto las clasificaciones más relevantes resultan ser las relacionadas con el almacenamiento, el procesamiento, el análisis de datos y la robótica avanzada y colaborativa, pues la base de estos desarrollos la constituye la inteligencia artificial, la cual a medida que se desarrolla permite la automatización de procesos complejos y la interconexión de todos los entes participantes de las operaciones.

Por otro lado, el aprendizaje de máquina ha permitido que los desarrollos en materia de análisis de datos se efectúen, logrando realizar predicciones y pronósticos basados en la información que se tiene en el presente, esto ha permitido que empresas como Amazon Web Services (AWS), IBM, Google, entre otras, desarrollen algoritmos para realizar proyecciones de la demanda, lo que les permite crecer acorde a la demanda que exige el mercado actualmente.

Si estas empresas pueden utilizar esta tecnología para su beneficio, cualquier empresa en el mundo puede hacerlo, estos desarrollos fueron creados de forma abierta o como código abierto, lo que da acceso libre a todo el público para diseñar, crear, modificar o adaptar los métodos de pronósticos existentes a su demanda y participar de los inicios de la era digital aplicada a las industrias, pues esta les permitirá mejorar su competitividad y hacerse participes del mercado global.

1.2 Formulación del problema

¿Es posible mejorar la precisión en los pronósticos de demanda utilizando modelos de aprendizaje de máquina?

1.3 Justificación

Los métodos de pronóstico de demanda que se consideran tradicionales son usados desde los años 80. Existe evidencia que los cambios tecnológicos desde hace 20 años han llevado a una transformación sustancial de la población y sus necesidades. Las industrias con tecnología de pronóstico tienen una gran ventaja competitiva frente al mercado, debido a que, al tener un pronóstico de demanda más acertado, reducirán desperdicios en sobreproducción, inventarios, marketing y transporte.

Por medio de diversas fuentes se ha logrado demostrar que los métodos de pronóstico que se consideran tradicionales presentan linealidad en su metodología, lo que los hace rígidos e inflexibles ante los cambios en la demanda, es decir que no se ajustan de forma rápida a las variaciones del mercado, es por esto que por medio del uso de aprendizaje de máquina se pretende modelar una nueva forma de pronosticar la demanda bajo cualquier patrón que estas presenten, con la finalidad de comparar si realmente estas técnicas se adaptan mejor a las constantes fluctuaciones del mercado.

Al mismo tiempo, la preocupación por anticiparse a los movimientos del mercado es un factor determinante para las empresas, pues de esto depende su competitividad y estabilidad, es por esto que se buscan constantemente alternativa para mejorar los pronósticos, de acercarse a los clientes, entender la mente del consumidor y mejorar la percepción de este respecto a la compañía, este proyecto busca aportar a la toma de decisiones con base a resultados más precisos y confiables por medio de la predicción de la demanda acertada.

Se evidencia una gran oportunidad para facilitar la comprensión de la relación entre la naturaleza de las series de datos y los métodos de pronóstico de demanda tanto tradicionales como de aprendizaje de máquina. Por esto, el proyecto propuesto permitirá sentar bases sobre la correcta selección de los métodos de pronóstico, tomando en cuenta los avances tecnológicos que se presentan hoy día. Dicho análisis facilitará los futuros estudios e implementaciones sobre modelos de inteligencia artificial según el tipo de series de datos que utilice la organización, permitiendo un menor error de pronóstico.

2. Objetivos

2.1 Objetivo del Proyecto

Evaluar métodos de pronóstico tradicionales y de aprendizaje de máquina para estimar la demanda en series de tiempo de diferente naturaleza.

2.2 Objetivos Específicos

- Definir los modelos de pronóstico tradicionales y de aprendizaje de máquina a utilizar en el proyecto.
- Definir las series de tiempo que se usaran en los métodos de pronóstico de demanda seleccionados.
- Aplicar los métodos de pronóstico a las series de tiempo seleccionadas.
- Comparar la precisión entre los métodos de pronóstico tradicionales y de aprendizaje de máquina para las diferentes series de tiempo.

Entregables:

- Informe de los métodos de pronósticos tradicionales y de aprendizaje de máquina más utilizados.
- Informe de las series de tiempo clasificadas según la naturaleza que presenten.
- Informe de los métodos de pronóstico aplicados a las series de tiempo y la medición del error.
- Tabla comparativa de la precisión de los modelos tradicionales contra los de aprendizaje de máquina, con sus respectivas recomendaciones.

3. Marco de Referencia

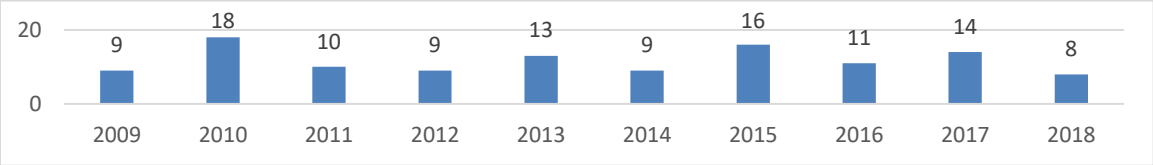
3.1 Antecedentes o Estudios Previos

Según Pérez (2017) la inteligencia artificial tiene sus bases con la creación de la máquina inteligente de Turing en el año 1950, bajo la propuesta de McCullong y Pitt en 1946 la cual simulaba el comportamiento de las neuronas biológicas por medio de un sistema lógico, posteriormente en la conferencia histórica de Dartmouth en 1956, Nawell, Simon y Minsky acuñan el nombre creado por John McCarthy de inteligencia artificial para este campo.

Dentro del naciente campo también aparecerían sus clasificaciones, el aprendizaje de máquina como una de ellas, donde se fue desarrollando de forma paralela con autores como Samuel (1967) que en 1952 creo el primer programa que aprendía a jugar checkers (Juego de mesa) por medio del análisis de muchas partidas, además, Rosenblatt (1961) quien diseño la primera red neuronal artificial aplicada a aprendizaje de máquina para 1958 con el fin de reconocer patrones y formas, finalmente en 1959 Widrow y Hoff (1960) lograron desarrollar una red neuronal artificial capaz de detectar patrones binarios y predecir la aparición del próximo bit en un flujo de bits, entre muchos otros desarrollos de este campo. Como se pudo observar, la idea central del desarrollo del aprendizaje de máquina se basa en analizar datos existentes y lograr transformarlos por medio de una función que le permita crear predicciones y suposiciones futuras basadas en estos.

Según (Tyralis et al., 2019) el fortalecimiento de la investigación sobre el aprendizaje de máquina aparece a principios de los años 90, afirmaron esto por medio de la recopilación de autores como Bates y Granger que en 1969 afirmaban que ningún método de aprendizaje de máquina era superior a los otros, décadas después autores como Granger y Ramanathan, en 1984, Wolpert en 1996 y muchos otros en conjunto, buscaban determinar si esta afirmación resultaba verdadera, logrando demostrar que la precisión de los métodos aumentaban si y solo si eran combinados entre sí, a esto le llamaron aprendizaje conjunto dentro del aprendizaje de máquina, de donde nacieron nuevos métodos.

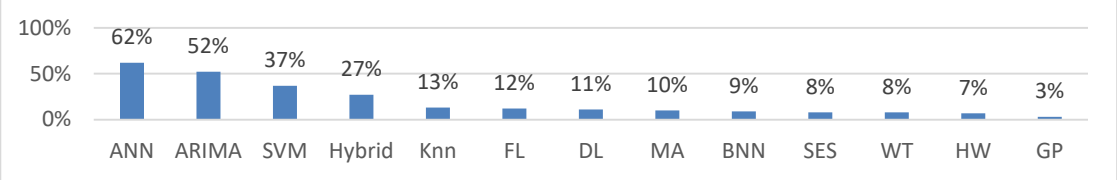
En la *gráfica 2* se muestra el análisis para 117 investigaciones de las cuales se describen los años de publicación y su enfoque, de manera que se puede observar como el estudio de los métodos de aprendizaje de máquina han tenido mayor acogida en las últimas décadas.



Gráfica 2 Investigaciones sobre AM aplicada a series de tiempo.

Adaptado de: (Parmezan, Souza y Batista, 2019)

Se puede observar que la mayor cantidad de estos artículos fueron publicados en periodos que comprenden el año 2009 y 2018 de manera tal que se está ante los inicios de las nuevas tecnologías de predicción. Por otro lado, de los 117 artículos solo se aplicaban métodos de pronósticos en 68 de ellos, de los cuales los más utilizados fueron: Redes neuronales artificiales (ANN), autor regresión integrada de media móvil (ARIMA), Soporte de máquinas de vectores (SVM), K-Vecino más cercano (kNN), Lógica difusa (FL), Aprendizaje profundo (DL), Redes neuronales bayesianas (BNN), Suavización exponencial simple (SES), Transformación de Wevelet (WT), Holt-Winter (HW), Proceso gaussiano (GP), como se ve en la *gráfica 3*.



Gráfica 3 Métodos de pronósticos.

Adaptado de: (Parmezan et al, 2019)

Demostrando que, de los métodos de aprendizaje de máquina los que resultan más estudiados son las redes neuronales artificiales, los cuales durante las últimas dos décadas, se han posicionado como métodos de pronóstico para series de tiempo en diversas áreas de la gestión empresarial, donde destacan: las finanzas, la generación de energía, los servicios, la medicina, las ciencias ambientales, las ciencias de los materiales, entre otras (Lao y Caridad, 2017)

Reconociendo el potencial de este campo aparecen investigaciones que permiten llevar a la práctica cada uno de los diseños que se han planteado a lo largo de los años, por ejemplo, por medio de una comparación empírica de algunos de los métodos de aprendizaje de máquina respecto a su capacidad de predicción para series de tiempo (Ahmed et al., 2010) logran establecer una escala que va desde los más precisos a los menos precisos, listando así: el perceptrón multicapa, procesos gaussianos, redes neuronales bayesianas, soporte de vectorregresión, redes neuronales de regresión generalizadas, entre otros, permitiendo obtener un marco de referencia para decisiones.

También para el mejoramiento de la una cadena logística, el proyecto de (Valencia Cárdenas et al., 2014) consistió en encontrar métodos eficientes para el pronóstico de la demanda de diferentes productos y que a su vez permitieran minimizar los costos del manejo de los inventarios. Por medio de la aplicación en repetidas ocasiones de diversas técnicas bayesianas y de aprendizaje de máquina, compararon la eficiencia de cada uno de estos mediante el indicador MAPE (por sus siglas en ingles *Median Absolute Percentage Error*) encontrando que el mejor método para pronosticar con pocos datos históricos es la metaheurística Tabú, con un mayor acierto en los pronósticos.

Por otro lado, (Suman y Kochak, 2015) que estudiaron la influencia de la previsión de la demanda por el método de pronóstico para series de tiempo usando redes neuronales mediante el programa MATLAB, en donde aplicaron un modelo de red multicapa, con el Algoritmo “Levenberg-Marquardt”. El proyecto fue investigado en una empresa manufacturera, arrojando como conclusión, que el método TrainLM[1] lleva a cabo de manera más eficaz que otros métodos cualitativos y la previsión es más fiable para este caso.

Posteriormente, Karakoyun y Çıbıkdiken (2018) Por medio de la comparación entre la autorregresión integrada de media móvil (ARIMA por sus siglas en inglés autoregressive integrated moving average) como modelo tradicional contra los modelos de inteligencia artificial (LSTM por sus siglas en inglés long Short-Term Memory) para evaluar el pronóstico en el precio de la moneda virtual Bitcoin, Obteniendo un error de predicción de 1.4% con el LSTM contra un 11,86% del ARIMA, llegando a la conclusión de que las técnicas de aprendizaje de máquina son convenientes para pronosticar el precio futuro de Bitcoin.

Un hallazgo importante surgió de la investigación de (Wolpert, 1996 citado (Cerqueira, Torgo, y Soares, 2019),p.8), quien haciendo uso de una **muestra pequeña**, comparó métodos de predicción, realizó la diferencia entre los métodos de aprendizaje de máquina y métodos tradicionales y llegó a la conclusión de que no todo modelo de inteligencia artificial es apto en todos los casos, es decir que en este escenario no se superó en precisión al modelo tradicional, se afirma que para muestras pequeñas se recomienda el uso de estos últimos.

Este campo está en auge y tiene el potencial de cambiar la dinámica de predicción como se conoce, las investigaciones aplicadas a demanda son escasas pero sus aplicaciones son bastante amplias, con este proyecto se apunta a contribuir al desarrollo de prácticas tecnológicas aplicadas al desarrollo intelectual y la mejora de las metodologías empresariales de pronósticos.

[1] Redes neuronales con Levenberg-Marqueardt en MATLAB

3.2 Marco Teórico

Para mejorar el entendimiento de este proyecto se deben tener en cuenta diferentes conceptos como: demanda, pronósticos de demanda, series de tiempo, patrones, inteligencia artificial, aprendizaje de máquina, redes neuronales artificiales, lenguajes de programación, librerías y medidas de error, de forma que procedemos con el desarrollo de esta investigación.

La demanda se puede definir como el volumen total de un producto que será adquirido por un grupo de compradores determinado, por un periodo de tiempo establecido y contando con condiciones de entorno y esfuerzo comercial determinados (Ortiz, 2014 citado en (Chicaiza, 2019)). Con base en demandas de periodos pasados, se puede estimar la demanda que se tendrá en el futuro, pues estas pueden presentar patrones identificables que permiten que estimarlo por medio de pronósticos sea más factible.

Everet y Ebert (1991) definen el pronóstico como un proceso de estimación de un acontecimiento futuro, proyectando resultados a futuro con datos del pasado, los cuales se pueden combinar sistemáticamente en forma predeterminada para hacer una estimación de un proceso del futuro, con esto se logra determinar patrones que podrían replicarse posteriormente.

Según Heizer, Render y Munson (2017). Los pronósticos pueden ser clasificados por horizontes de tiempo futuros entre los cuales se encuentran:

- Corto plazo: De forma general va hasta 3 meses, su mayor utilidad se encuentra en la planificación de compras, fuerza de trabajo y asignaciones en producción.
- Mediano plazo: Desde 3 meses hasta 3 años, utilizado para la planificación de las ventas, la producción, el presupuesto y los planes operativos.
- Largo plazo: Generalmente posee un lapso de 3 años o más, son utilizados para la planificación de nuevos productos, planes de ubicación de instalaciones, expansiones e investigación y desarrollo.

Según Chopra y Meindl (2008) los pronósticos presentan 4 características esenciales, estas han sido propuestas a lo largo de los años por medio de la experimentación, comprenderlas permitiría entender la naturaleza de estos o abrir la posibilidad de redefinir estas características.

- Reconocer el valor del **error** de pronóstico, incluyendo el valor esperado de pronóstico con el fin de tomar mejores decisiones.
- Los pronósticos a **largo plazo** son menos precisos que los de corto plazo, esto se debe a una mayor desviación estándar.
- Se recomienda hacer pronósticos **agregados** pues estos son más exactos que los desagregados.
- Si la información es **distorsionada** serán menos efectivos los pronósticos.

Para realizar pronósticos se debe contar con valores que alimenten los modelos, estos serán los que permitirán predecir los acontecimientos futuros, a estos les llamamos, series de tiempo, estos son considerados un conjunto de datos en función de múltiples variables, donde una de las variables independientes es el tiempo y su transcurso natural.

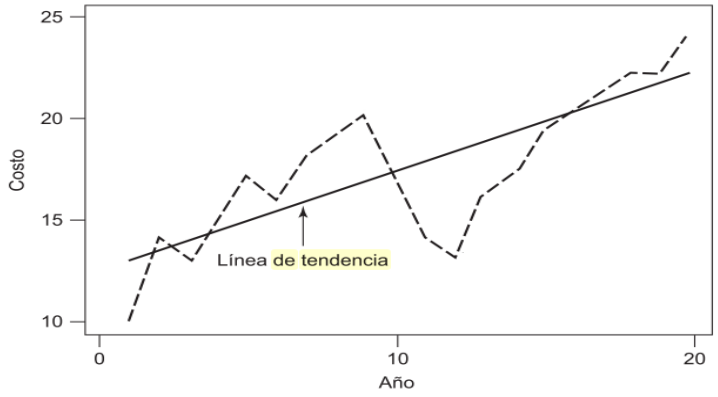
Para este caso particular se analizarán series de tiempo aplicadas a demandas las cuales son cantidades de productos o servicios en un tiempo determinado, demandas difiere de ventas pues estas últimas se pueden subestimar o sobreestimar para un tiempo en el futuro, por lo cual para lograr predecir el futuro de la demanda se deben reconocer patrones de comportamiento de los consumidores con datos de demandas pasadas.

Según Chapman (2006, p. 23).

> *La demanda pasada sigue cierto patrón, y si este patrón puede ser analizado puede utilizarse para desarrollar proyecciones para la demanda futura, suponiendo que el patrón continúa aproximadamente de la misma forma. Por último, esto implica el supuesto de que la única variable real independiente en el pronóstico de series de tiempo es, precisamente, el tiempo*

Las series de tiempo poseen patrones que pueden ser identificados, según Montemayor (2013). Los patrones que pueden presentar las series de tiempo son los siguientes:

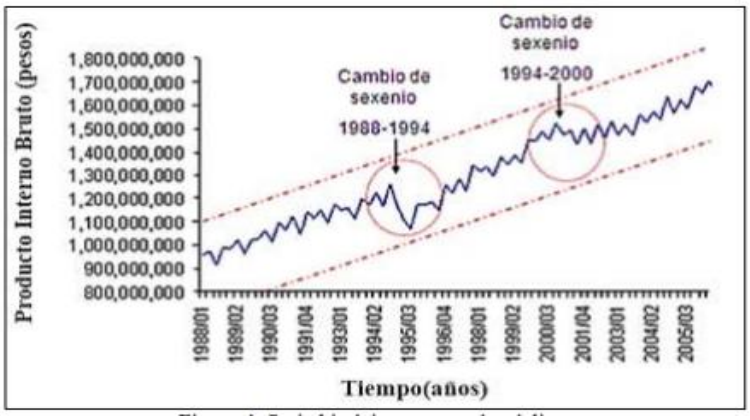
- **Tendencia:** Representa el incremento o decremento de forma consistente en la serie de tiempo, durante un extenso periodo.



Gráfica 4 Serie de tiempo costos de vivienda contra la cantidad de años.

Fuente: (Hanke y Wichern, 2010)

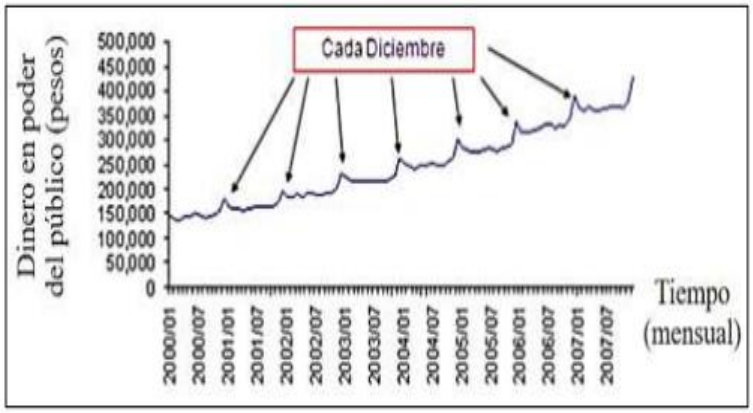
- **Cíclico:** existe cuando junto a la tendencia se observa un patrón adicional, un crecimiento o decrecimiento cada cierto tiempo. En este ejemplo se muestra el PBI mexicano en trimestres desde 1988 a 2005.



Gráfica 5 Serie histórica con patrón cíclico.

Fuente: (Montemayor, 2013)

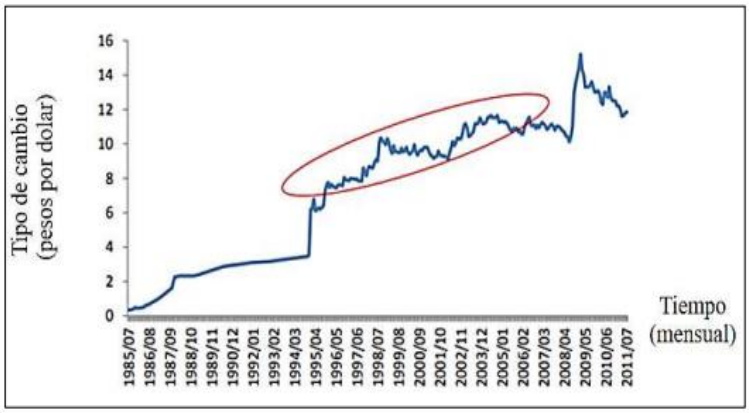
- **Estacionalidad:** Es el patrón de cambio que se repite año tras año, corresponde a una serie que es influenciada por factores que se repiten en la misma temporada del año, como por ejemplo la venta de los disfraces de Halloween o los adornos navideños.



Gráfica 6 Serie histórica del patrón estacional.

Fuente: (Montemayor, 2013)

- **Irregular o aleatoria:** Se encuentra en los patrones horizontales, de tendencia y estacionales, son difíciles de modelar matemáticamente debido a su alta variabilidad.



Gráfica 7 Serie histórica con variación irregular.

Fuente: (Montemayor, 2013)

Con el respectivo análisis de las series de tiempo, su naturaleza y reconociendo si presenta o no patrones se procede a seleccionar los métodos de pronósticos.

Según Chopra y Meindl (2008) existen diferentes tipos de pronósticos, como se ve en la *Ilustración 3,* los cuales abren la ventana de posibilidades a las empresas para elegir el que más se adapte a su demanda respecto a su mercado meta o realizar la combinación de diferentes métodos, estos son:

- Cualitativos: Carecen de estructura analítica definida, se usa generalmente en la inclusión de un nuevo producto en el mercado, son subjetivos y se apoyan en el juicio personal, genera resultados rápidos.
- Cuantitativos: Son los que utilizan herramientas matemáticas para determinar demandas futuras.
    - Series de tiempo: Utilizan demanda histórica para hacer pronósticos.
    - Causal: Correlación de la demanda con factores externos.

En la *ilustración 3*, se pueden ver cada una de las clasificaciones de los pronósticos con sus respectivas distinciones entre métodos:

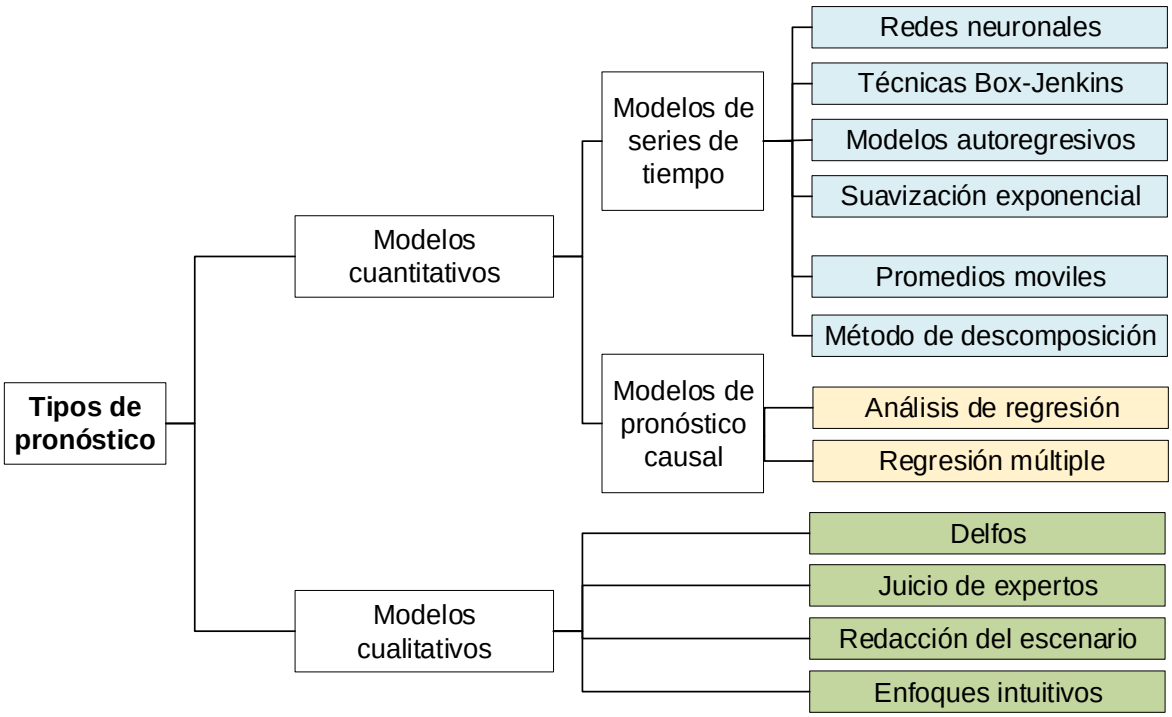


Ilustración 3 Clasificación de los pronósticos

Adaptado de: (Anderson, Senes y Williams, 2004, p. 206), (Montemayor, 2013).

Para ahondar en cada uno de los modelos que se describen en la *ilustración 3*, se puede dirigir al **Anexo 1**, en donde encontrara la descripción de cada uno de estos, es importante resaltar que en este proyecto nos centraremos únicamente en los modelos de series de tiempo, relacionando la demanda futura con datos históricos de demandas pasadas y considerando diferentes patrones de series.

Para el proyecto se realizó la división de los métodos que se encuentran dentro de los modelos de series de tiempo, estos fueron categorizados en dos clases, los cuales llamaremos, métodos tradicionales y métodos de aprendizaje de máquina.

Según (Chase et al., 2009) para pronosticar un resultado de una serie de tiempo existen diferentes técnicas y modelos, algunos ejemplos de estos son los siguientes, los cuales consideraremos como **modelos tradicionales**:

- **Promedio móvil simple**: Se calcula el promedio de un periodo que contiene varios puntos de datos dividiendo la suma de los valores de los puntos entre el número de éstos. Por lo tanto, cada uno tiene la misma influencia.
- **Promedio móvil ponderado:** Puede ser que algunos puntos específicos se ponderen más o menos que los otros, según la experiencia.
- **Suavización exponencial:** Los puntos de datos recientes se ponderan más y la ponderación sufre una reducción exponencial conforme los datos se vuelven más antiguos.

En contraposición a los modelos **tradicionales** de tratamiento de datos para series de tiempo en pronósticos, aparecen los modelos de **aprendizaje de máquina**.

El aprendizaje de máquina es una subrama de la inteligencia artificial, con esta se busca desarrollar algoritmos que permitan a las máquinas aprender de forma similar a como lo hacen los seres humanos, esto por medio de la división de los datos en periodos de entrenamiento y periodos de pruebas, como se muestra en la *ilustración 4*.

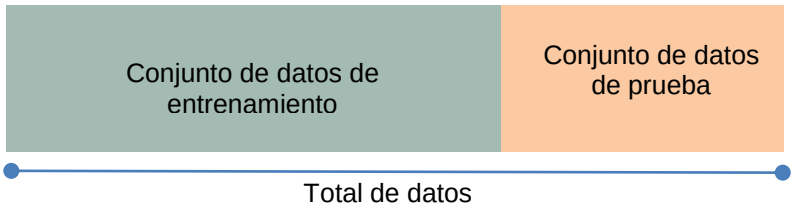


Ilustración 4 Splits de división de datos.

Fuente: elaboración propia

Se debe tener en cuenta que para programar la inteligencia artificial se requiere de una secuencia finita de instrucciones que especifique las diferentes acciones que ejecuta la computadora para resolver un determinado problema. Esta secuencia de instrucciones constituye la estructura algorítmica, (Benítez, Escudero, Kanaan, y Masip, 2014), siendo entonces la estructura algorítmica que programan los seres humanos lo que le da vida y funcionalidad a la inteligencia artificial.

El aprendizaje de máquina puede ser ampliamente definido como los métodos computacionales que usan la experiencia para mejorar el desempeño de las predicciones, logrando ser estas más precisas. Cuando nos referimos a experiencias hablamos específicamente de la información histórica recolectada que se utiliza para los procesos de entrenamiento. (Mohri, Rostamizadeh, y Talwalkar, 2012).

Aparecen entonces clasificaciones de aprendizaje de máquina aplicadas a los pronósticos de demanda existentes bajo series de tiempo, con la finalidad de alcanzar mejoras en los resultados de predicción, el modelo base que se utilizara para el desarrollo de este trabajo serán las redes neuronales artificiales.

Las redes neuronales son redes interconectadas en paralelo de elementos simples que usualmente son adaptativos y organizados jerárquicamente, estas interactúan con los objetos del mundo real de forma similar a como lo hace el sistema nervioso biológico (Ruiz y Matich, 2001).

Según (Ruiz y Matich, 2001) los elementos que componen una red neuronal son neuronas independientes interconectadas entre sí, arregladas en 3 capas, los datos de la capa de entrada pasan a una capa oculta y generan un resultado en la capa de salida, la capa oculta puede estar estructura por varias capas.

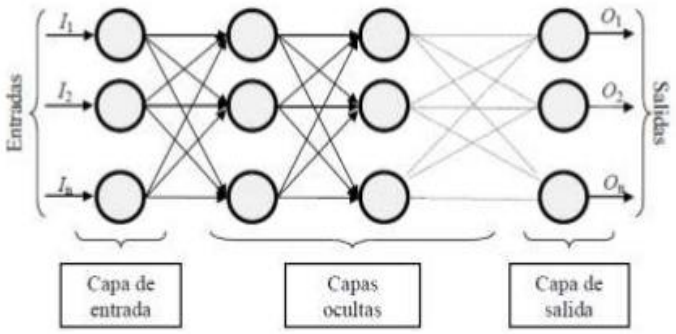


Ilustración 5 Ejemplo de una red neuronal interconectada

Según (Chicaiza, 2019) su uso en pronósticos se basa principalmente en modelos matemáticos simples, el modelo tendrá unas entradas influenciadas por los factores que afectan la demanda y la salida corresponderá a la demanda lo que permite llevar relaciones no lineales entre la variable de salida y la predicción final.

Para llevar realizar el pronóstico por medio de una red neuronal se deben llevar a cabo los siguientes pasos.

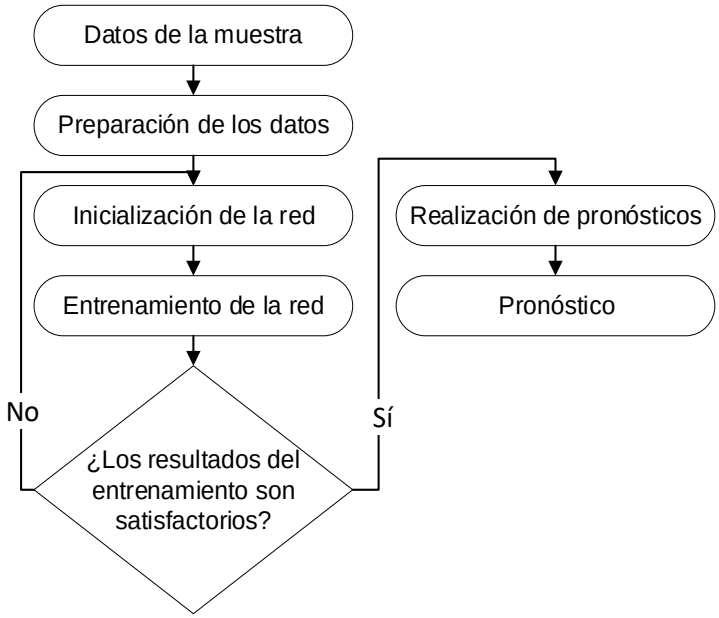


Ilustración 6 Proceso de pronóstico de la red neuronal.

Adaptada de: (Hu, sun y Wen, 2014 como se cita en Chicaiza, 2019)

Entre más capas contenga un modelo, más complejo resulta, lo que permitirá procesar la información que se registra y obtener pronósticos más aproximados a la realidad, a las redes que contienen varias capas se les conoce como multicapa, donde la entrada de una neurona es la salida de la anterior, que posteriormente se combinan por medio de una combinación lineal ponderada para pasar a la siguiente capa (Hyndman y Athnasopoulos, 2018).

Las redes neuronales se dividen en 3 clasificaciones las cuales contienen diferentes modelos funcionales para pronóstico, para ver la descripción de cada uno de ellos, ir al **Anexo 2** redes neuronales supervisadas, no supervisadas y redes hibridas, como se observa en la *ilustración 7*:

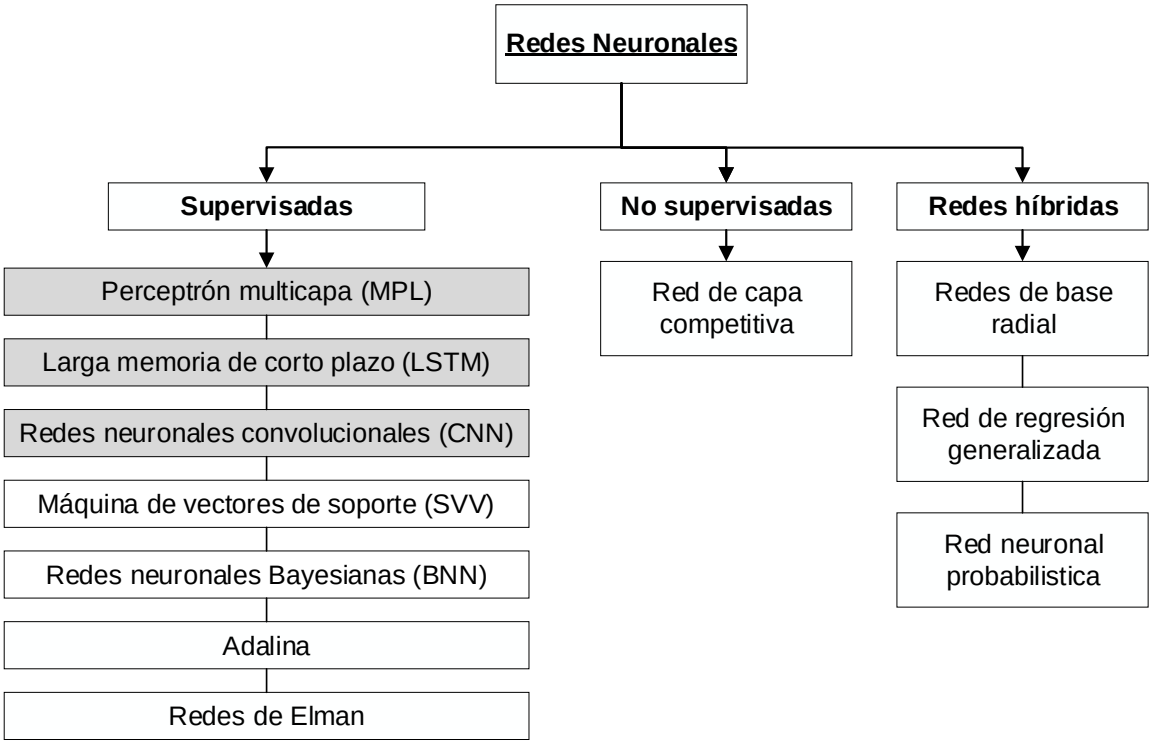


Ilustración 7 Tipos de redes neuronales

Adaptada de: (Valeru y Senabre, 2012 como se cita en Chicaiza, 2019)

**Nota: los recuadros sombreados son las redes neuronales que se usaron en este proyecto*

Las redes neuronales supervisadas son aquellas donde el algoritmo recibe datos de entrenamiento que contienen la respuesta correcta para cada ejemplo. El problema de estudio utiliza algoritmos de aprendizaje supervisado, donde el experto en compras da la respuesta correcta a cada ejemplo, es decir donde los datos que se están utilizando contiene etiquetas que permiten identificar la clasificación de cada uno de los datos (Garcete et al., 2017).

Las redes no supervisadas son aquellas donde el algoritmo busca estructuras en los datos de entrenamiento, como encontrar qué ejemplos son similares entre sí, y agruparlos en clúster, es decir que no existen etiquetas en los datos y el algoritmo debe aprender a identificarlos (Garcete et al., 2017).

Las redes hibridas son aquellas que no contienen etiquetas y no están clasificadas, los algoritmos deben aprender por medio del ensayo y error con los datos de entrenamiento a fin de aprender a identificarlos y procesarlos.

En este proyecto, se evaluaron los modelos de pronóstico de inteligencia artificial desde las redes neuronales supervisadas, específicamente MLP, LSTM y CNN.

-El perceptrón multicapa (MLP) es una estructura de red neuronal artificial y es un estimador no paramétrico que puede ser usado para la clasificación y regresión. Para utilizarlo se usa el algoritmo “Backpropagation” para diferentes aplicaciones (Alpaydin, 2014)

-Long short term memory (LSTM) es un modelo de aprendizaje profundo muy popular. Se usa particularmente para datos con correlaciones temporales, como textos, secuencias o series de tiempo (Huimei, Xingquan, & Ying, 2020)

-Las redes neuronales convolucionales (CNN) son un tipo de red neuronal profunda ampliamente usada en el reconocimiento, clasificación y análisis de información multimedia como imágenes, videos y sonidos (Fan et al., 2019). Su arquitectura trabaja con cada capa oculta como una convolución de sus entradas con el peso del vector. (Alpaydin, 2014).

Para la ejecución de los modelos de pronóstico de inteligencia artificial es necesario usar un lenguaje de programación, en este proyecto se optó por Python, ejecutado en el entono gratuito de Jupyter Notebook, “Google Colaboraty”. Esta elección, permite el uso de librerías preprogramadas, desarrolladas por expertos, lo cual facilita en gran medida la aplicación. A continuación, se hace una breve descripción de las librerías utilizadas en este proyecto:

Keras es una biblioteca a nivel de modelo, que proporciona bloques de construcción de alto nivel para el desarrollo de Modelos de aprendizaje profundo (Venables et al., 2008)

Tensorflow es una biblioteca de software de código abierto para máquinas inteligentes. permite trabajar de manera productiva utilizando un alto nivel de Keras y de las APIs de Estimator, y cuando necesita más control, proporciona acceso completo a la API principal de TensorFlow (TensorFlow y RStudio, 2018).

Matplotlib es una biblioteca completa para crear visualizaciones estáticas, animadas e interactivas en Python. (Hunter John et al, 2012)

Numpy es un paquete de Python que significa “Numerical Python”, es la librería principal para la informática científica, proporciona potentes estructuras de datos, implementando matrices y matrices multidimensionales. Estas estructuras de datos garantizan cálculos eficientes con matrices. (González, 2020)

Pandas es una librería de código abierto de Python que proporciona herramientas de análisis y manipulación de datos de alto rendimiento utilizando sus potentes estructuras de datos. El nombre de Pandas se deriva del término “Panel Data” y es la librería de análisis de datos de Python. (González, 2019)

Según Hanke (2010) la forma de medir la precisión de los pronósticos se da por medio de las medidas de error, con los cuales se realiza la evaluación del valor real contra el valor pronosticado y se mide la certeza de estos. En este proyecto se utilizó el error cuadrático medio (MSE), el cuál es muy usado en el aprendizaje de máquina supervisado. Posteriormente, se calculó el RMSE, teniendo como base que un valor de 0 o muy aproximado en RMSE indicaría un ajuste perfecto a los datos.

Error cuadrático medio mejor conocido como (*Mean Square Error* **MSE**), se calcula como se muestra en la *ecuación 1*, con el promedio de los cuadrados de las diferencias entre los valores pronosticados y observados. Al aplicar raíz cuadrada se obtiene el RMSE:

Ecuación 1 Formula Error MSE

$$\text{MSE} = \frac{\sum(\text{errores de pronóstico})}{n}$$

En el análisis realizado para la selección del medidor existían cinco posibilidades, entre estas se encuentran el MAE, MAPE, MSE, RMSE y el RMSLE. Las dos últimas derivadas del error MSE, por un lado, el RMSE como la reducción del MSE por medio de la aplicación de la raíz cuadrada y el RMSLE con la misma función por medio de la diferencia logarítmica.

El error MAPE como una medida porcentual, que en caso de obtener resultados superiores a 1 dificultaría su interpretación. Por otro lado, el error MAE el cual resulta insensible ante las altas tasas de frecuencia y variabilidad de los datos que para este proyecto son comunes, dejando de lado valores que podrían ser significativos para el análisis de los resultados.

Siendo elegido finalmente el MSE, y como medida de reducción el RMSE el cual viene predeterminado por las librerías de Python, facilitando así su obtención automatizada y su análisis ya que un resultado cercano cero demostraría un ajuste perfecto de los datos pronosticados a los reales.

Por último, algunos de los factores que pueden influir en la elección de un método de pronóstico se encuentran:

- Costo y precisión
- Datos disponibles
- Tiempo
- Naturaleza de productos y servicios
- Respuesta de impulso y amortiguación de ruido

3.3 Contribución Intelectual o Impacto del Proyecto

Este proyecto tiene como finalidad incentivar la investigación en el uso de tecnologías para el desarrollo de habilidades en el área de pronósticos y mejoras en las capacidades de los ingenieros industriales y afines, ahondado entorno a la programación sencilla y al reconocimiento de nuevos métodos de pronósticos que existen hoy día, apostándole a una generación digital que se encamine a la toma de decisiones basadas en datos confiables y resultados oportunos.

Con este proyecto se invita a aprender a discernir entre los métodos de pronóstico, debido a que, no todos los métodos que se ofrecen en el mercado son adecuados. La razón es que, si se poseen modelos de pronóstico de vanguardia, pero no se tienen datos confiables, se automatizarán practicas erróneas. Es por esto que, el proyecto invita a mejorar la cultura de recolección de datos e invertir en estudios de información propia, con el fin de determinar cuáles son los patrones que afectan la demanda de cada uno de los productos o servicios y mejorar la selección de los métodos que más se ajusten a estas.

Además, concientizar a los profesionales en que la competitividad futura viene de la mano de mejores prácticas tecnológicas y el uso de pronósticos acertados puede mejorar la posición en el mercado de cualquier empresa, y más aún, utilizar la información que no es utilizada, donde se pierde contenido valioso por carencias tecnológicas, falta de inversión e interés, por ejemplo: cantidad de pedidos, planeación de producción, unidades demandadas, número de clientes, materia prima, nuevos proveedores, entre otras.

Por último, se espera contribuir a la ingeniería industrial en mayor medida respecto al uso de los métodos de pronóstico de demanda existentes más allá de los convencionales, pues en la actualidad las empresas tienen reprocesos y perdidas debido a la inadecuada estimación de su demanda. También, contribuir a futuros estudios en donde se desarrollen implementaciones de herramientas de aprendizaje de máquina a la industria, modelos de selección automatizada e investigaciones sobre industria 4.0, entre otros.

4. Metodología

Este proyecto se desarrolló en cuatro etapas, como se puede observar en la *Ilustración 8*. Cada una de las etapas establecidas corresponden a cada uno de los objetivos específicos, de forma tal que están correlacionados de forma secuencial, logrando que con la implementación de cada fase se logre dar cumplimiento a los objetivos planteados al inicio de este proyecto.

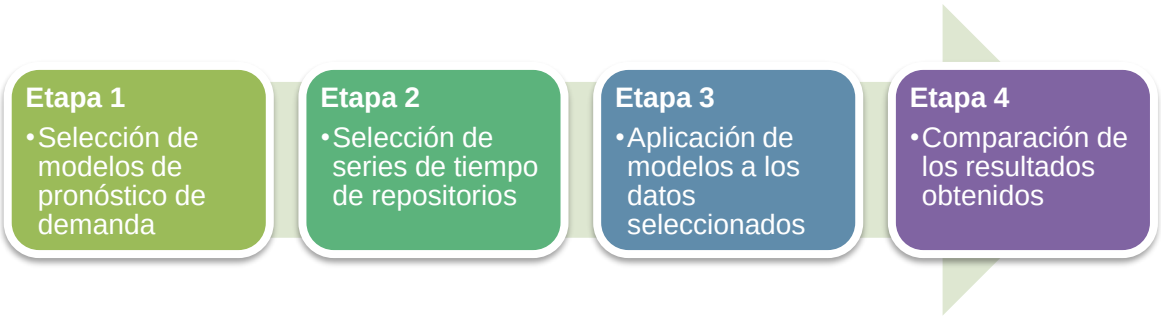


Ilustración 8 Metodología
Fuente: Elaboración propia

Para el desarrollo de la **primera etapa** se realizó una *revisión de la literatura* sobre los modelos de aprendizaje de máquina y métodos tradicionales de pronóstico de demanda registrados, los cuales se evidenciaron en una tabla de relación cuyo contenido es: autor, modelos utilizados, métrica (error) y número de registro de datos. Con esta tabla se logró seleccionar los métodos y modelos más utilizados y, el número de datos aproximados con los cuáles se debía realizar la búsqueda de series de datos a aplicar a cada modelo.

Posteriormente, en la **segunda etapa,** se realizó una búsqueda de series de tiempo en repositorios encontrados en la web, con la herramienta Power Query de Excel se limpiaron, filtraron y clasificaron las bases de datos según el tipo de variación (tendencia, variación estacional, variación cíclica, variación irregular).

Una vez seleccionados los *métodos tradicionales, modelos de aprendizaje de máquinas y series de tiempo,* se procedió a la **tercera etapa***,* en la cual se realizó la aplicación de todas las series de tiempo seleccionadas a cada uno de los métodos tradicionales y de aprendizaje de máquina, usando el lenguaje de programación *Python*, implementando librerías como *Tensor Flow, Keras, Sklearn, Matplotlib, Numpy y Pandas*. En cada corrida, se registró el error de pronóstico RMSE; teniendo en cuenta el método y la variación de la serie de datos aplicada.

Para concluir la **cuarta etapa**, se realizó una comparación de errores de pronóstico teniendo en cuenta la información recopilada de los resultados obtenidos. Con base a estos, se desarrollaron sugerencias sobre el correcto uso de los métodos y modelos de pronósticos estudiados, teniendo en cuenta el menor error según el tipo de serie de tiempo pronosticado.

5. Resultados

5.1 Selección de los modelos de pronóstico de demanda

Para la selección de los modelos a revisar se exploraron diferentes bases de datos confiables como Science direct, Oxford Academy, EBSCOhost, entre otras. La elección de los artículos de pronóstico con inteligencia artificial a revisar estuvo basada en tres criterios. El primero, que en el trabajo usara modelos de inteligencia artificial para pronosticar demanda o series de tiempo. El segundo, que se usaran los modelos de inteligencia artificial conocidos con pocas modificaciones en el código. El tercero, que estuviera publicado como máximo hace 2 años (2018), teniendo en cuenta el rápido avance que se ha tenido en esta área en los últimos años.

En la *tabla 1*, se puede observar la revisión de literatura realizada bajo la temática "pronóstico de demanda con redes neuronales".

Tabla 1 Revisión de literatura
Fuente: elaboración propia

Cita	Pronóstico de	Modelo[2]
(Vijai y Bagavathi Sivakumar, 2018)	Demanda de agua en el Reino Unido	LSSVM
		RF
		ELM
		MLP
(Bon y Tieng, 2019)	Demanda turística desde China a Sabah	MLP
		LSTM
(Siami-Namini, Tavakoli, y Namin, 2019)	Proyecciones financieras de series temporales	LSTM
		LSTM+CNN
(Kim, Lee, & Sohn, 2019)	Demanda de bicicletas compartidas por hora	CNN
(Hamedmoghadam, Joorabloo, y Jalili, 2018)	Demanda de electricidad a largo plazo de Australia	MLP
(Cheng, Xu, Mashima, Thing, y Wu, 2019)	Demanda de electricidad neuronal	LSTM
(Khalid et al., 2020)	Demanda de electricidad y pronóstico de precios	LSTM
(Hu, Szerlip, Karaletsos, y Singh, 2019)	Demanda horaria de pasajeros de Uber en 8 ciudades de EE. UU.	BNN
(Poorzaker Arabani y Ebrahimpour Komleh, 2019)	Demanda de efectivo en cajeros automáticos de la banca iraní	MLP
		CNN
(Poorzaker Arabani y Ebrahimpour Komleh, 2019)	Demanda de dinero en cajeros ATM	SVM
		CNN

[2] Las definiciones de estas abreviaturas se encuentran en el glosario

En la revisión de literatura se evidenció una frecuencia en el uso de modelos específicos para pronóstico de demanda en series de tiempo. Estos fueron los métodos: MLP, LSTM y CNN.

Los métodos seleccionados poseen diferentes ventajas y limitaciones, entre estas encontramos que el modelo MLP tiene la capacidad de aproximar funciones no lineales, poca sensibilidad a la variabilidad de los datos, capaz de trabajar con entradas multivariables y finalmente con la capacidad para pronosticar múltiples pasos, su limitación es la necesidad de un mapeo significativo entre las entradas y las salidas. El CNN posee todos los beneficios del MLP, además de poseer capacidad de memorización de patrones, su limitación es la dificultad para almacenar dependencias temporales, finalmente el LSTM el cual posee los beneficios del CNN, siendo el modelo más robusto, posee la capacidad de reconocer dependencias temporales. Los modelos tradicionales son simples en estructura y de fácil deducción, su principal limitación es la necesidad de datos completos, el trabajo con relaciones lineales y la dificultad para pronósticos a largo plazo.

Por otra parte, los métodos de pronóstico de demanda tradicionales que se decidió usar fueron: suavización exponencial simple, Holt-Winter y ARIMA, los cuales son explicados con mayor rigurosidad en el Marco teórico y se muestran en el **Anexo 1**.

El motivo de la elección de estos métodos de pronóstico tradicionales se debe a la complejidad en su sistema de predicciones que respecto a otros modelos de pronóstico permite obtener mejores resultados, pues consideran una amplia gama de parámetros que logran aproximar mejor los cálculos y de esta manera lograr una mejor comparación respecto a los avances en tecnología. Además, son frecuentemente comparados con los modelos IA.

5.2 Selección de las series de tiempo a utilizar en los pronósticos

5.2.1 Recursos utilizados

Por medio de la búsqueda de bases de datos en la web se obtuvieron fuentes que se ajustan a los patrones requeridos para el proyecto. Entre las series de tiempo que se analizaron se encontraron diferentes naturalezas, categorías y variedad entre la cantidad de registros, esto permite enriquecer el análisis, pues se inicia con el supuesto de su posible aplicación a todos los campos existentes en la industria y en general.

Las bases de datos fueron depuradas, limpiadas y organizadas por medio de la utilización de Power Query de Excel, herramienta que permite el trabajo flexible con grandes volúmenes de datos.

En esta oportunidad se obtuvieron 12 series de tiempo, entre las clasificaciones por naturaleza se encuentran estacionales, con tendencia, cíclicas y erráticas. Algunas

de las categorías a las que estas pertenecen son el sector salud, la movilidad, economía entre otros.

Para realizar la clasificación de las series de tiempo se inició un proceso de análisis grafico que permitiera determinar de forma cualitativa su naturaleza, por otro lado, se realizaron cálculos que facilitan establecer de forma cuantitativa y con mayor certeza los patrones de estas series.

El análisis grafico se llevó a cabo por medio de los paquetes Pandas y Matplotlib en Python, estos paquetes facilitaron el proceso de análisis estadístico y gráfico, por otro lado, los cálculos se realizaron en Excel por medio de las fórmulas mostradas en la *ecuación 2 y 3*.

Ecuación 2 Determinación de variabilidad

$$A = \frac{Desviacion\ estandar}{Media}$$

Ecuación 3 Ecuación de la línea recta

$$B = (\mathrm{y} = \mathrm{mx} + \mathrm{b})$$

Para la clasificación de patrones erráticos y de tendencia se establecen las siguientes convenciones:

- Si A ≥ 1 entonces la naturaleza de la serie temporal es “Errático”
- Si en B, m < 0 entonces corresponde a una “Tendencia” negativa
- Si en B, m > 0 entonces corresponde a una “Tendencia” positiva

Se espera establecer la forma de cálculo para determinar patrones “estacionales”, pero una medida confiable resulta del análisis visual, logrando determinar patrones en periodos determinados de tiempo que muestren que el comportamiento de la demanda se mantiene por encima o por debajo de un promedio preestablecido y este patrón continua a lo largo de la escala temporal.

5.2.2 Procedimiento

Se procedió con la búsqueda de las bases de datos en la web, se utilizaron fuentes de datos gratuitas tales como: data.world, kaggle.com, github.com entre otras.

Para explicar el procedimiento, se ilustra un ejemplo que muestra de forma específica cómo se trabajó con la base de datos **3**, “DOLLAR” está se asocia con el precio del dólar histórico, las otras 11 bases de datos se encuentran en el **Anexo 3**.

La forma matemática de reconocimiento para el patrón cíclico es escaso o no son replicables, por esto por medio de la observación y teniendo en cuenta que estos presentan estacionalidad y tendencia, de forma gráfica decidimos realizar el análisis correspondiente.

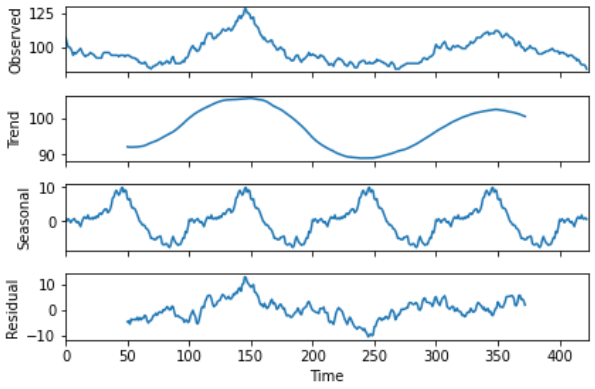


Ilustración 9 Análisis grafico de la base de datos "DOLLAR"

Se observa en la descomposición que se presentan tres momentos de tendencia, dos momentos positivos y tres negativos, por otro lado, es evidente la estacionalidad que se presenta en las mediciones, con un patrón estacional cada 50 datos, esto nos lleva a pensar que es un patrón cíclico, pues cumple con la tendencia y la estacionalidad, sus estadísticas son:

El valor de la desviación es: 9.069735774085139
El valor de la media es: 95.47695035460993
El valor de A es: 0.09499398274032979 <1, No es errática

La tendencia analizada con Excel da como resultado: -0,0091x siendo una tendencia ligeramente negativa. B<0

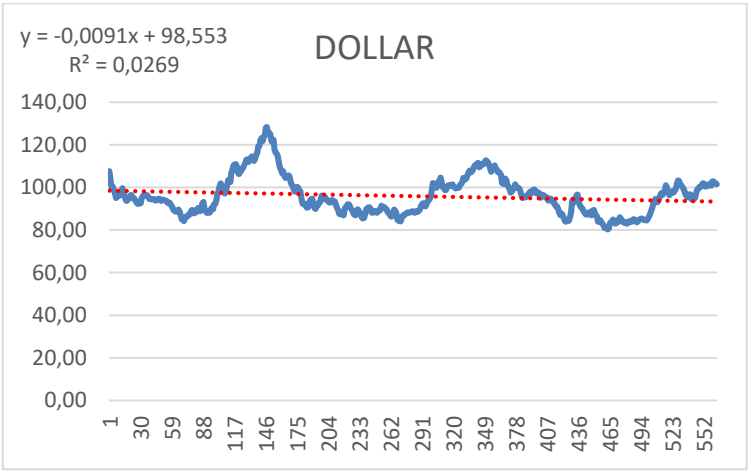


Ilustración 10 Análisis de tendencia de la base de datos "DOLLAR"

Se obtiene así un análisis concluyente de la naturaleza cíclica de los datos correspondientes al precio del dólar. Este procedimiento se repite para todas las series de tiempo seleccionadas, hasta obtener la clasificación de cada una de las fuentes evidenciadas en la *tabla 2*.

Tabla 2 Clasificación por naturalezas de las series de tiempo
Fuente: elaboración propia

Naturaleza de la demanda	Cantidad de datos	Periodicidad	Nombre de la Base de datos	Sector
Tendencia	631	Mensual	Construcción	Construcción
	523	Diario	Gasolina	Transporte
	364	Diario	Nacimientos	Salud
Cíclico	4017	Diario	Contaminación	Medio ambiente
	565	Mensual	Dollar	Economía
	469	Mensual	Pasajeros	Transporte
Estacional	851	Mensual	Crimen	Seguridad
	547	Mensual	Halloween	Temporada
	3649	Diario	Temperaturas	Medio ambiente
Errática	468	Semanal	Producto0001	Producción
	8759	Diario	Solar	Medio ambiente
	543	Diario	Tv	Entretenimiento

De forma tal que el siguiente paso procedimental corresponde con la aplicación de los algoritmos de predicción a las series de tiempo y su posterior evaluación.

5.3 Aplicación de los modelos

Para la correcta ejecución de los modelos se usó *Google collaborative*, el cual permitió correr los modelos en la nube, sin ocupar grandes espacios de memoria en un computador y, con la facilidad de tener, sin instalación previa, todas las librerías del lenguaje de programación Phyton.

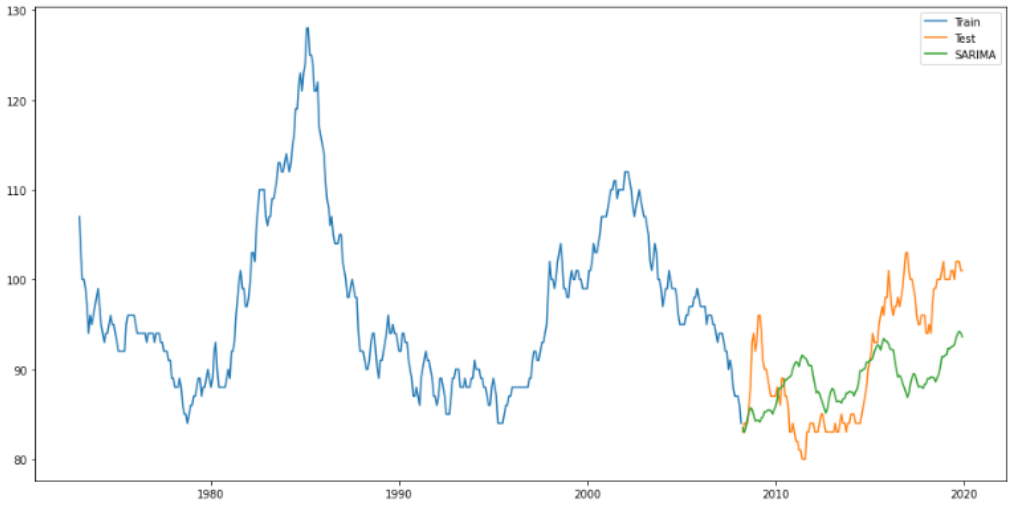
5.3.1 Pronósticos tradicionales

Los pronósticos tradicionales que se utilizaron para la implementación y desarrollo de las predicciones fueron:

- Suavización exponencial simple (SES)
- Holt-Winter
- ARIMA

Las bases de datos fueron clasificadas en dos rangos, un rango de entrenamiento y un rango de pruebas, 70% de los registros de cada base de datos corresponden al primer rango y el otro 30% al rango siguiente, por medio de este establecimiento de rangos se logra establecer un periodo de tiempo en el que el modelo se entrena y “aprende” de los datos para finalmente arrojar una predicción.

La predicción que se obtiene como resultado es comparada respecto al rango de pruebas que se establece y de esta manera se obtiene el valor de los errores, este procedimiento se repitió 36 veces para los pronósticos tradicionales, en el **Anexo 4** se observan los resultados para todas las bases de datos, a manera de ejemplo usaremos la base de datos **3**, llamada “DOLLAR”, para el modelo ARIMA.



Gráfica 8 ARIMA/Dollar/RMSE

RMSE = 6,684769266467668

5.3.2 Pronósticos inteligencia artificial

Para el caso de los modelos de inteligencia artificial se eligieron tres modelos:

- Red neuronal convolucional (CNN)
- Memoria de corto-largo plazo (LSTM)
- Perceptron multicapa (MLP).

El procedimiento de ejecución de los códigos CNN, LSTM y MLP, desglosados en el **Anexo 4**, fue el siguiente:

-Importar librerías: Se importaron las librerías *pandas*, para codificar las series de tiempo; *keras,* desde donde se importan los modelos de pronóstico; *sklearn* desde donde se importa el error (MSE, RMSE); *matplotlib* y *numpy,* que permiten graficar la diferencia entre lo datos pronosticados y los testeados.

-Definir series de supervisión*:* Se convierten las series de tiempo en aprendizaje supervisado.

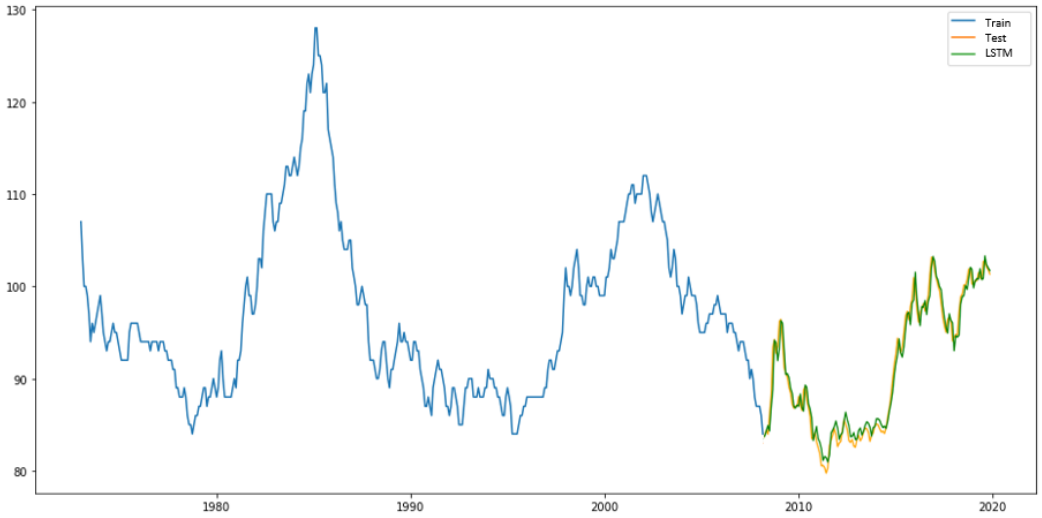
-Importar datos desde cada archivo cvs: Los datos son importados, codificados en el alfabeto latino, y separados por “,”

-Dividir los datos de entrenamiento: Para este proyecto, todas las series de tiempo se dividieron en 70% datos de entrenamiento y, 30% datos de testeo.

-Diseño y entrenamiento del modelo: Para esto, es necesario hacer un redimensionamiento de datos con la función “reshape”, según lo requiera el modelo e, importar el modelo correspondiente. En esta misma línea de código se define el error a calcular.

-Entrenar los modelos y predecir: Al entrenar los modelos, se definió el número de episodios como “500” ya que, entre más episodios resulta más acertado el cálculo. Una vez finalizado el modelo lanzará una predicción.

-Graficar pronosticado vs datos reales: Con ayuda de pyplot, se graficaron los datos arrojados en la predicción (ŷ) contrastados con los valores reales de la serie de tiempo (y). Al correr el modelo, se obtuvo la siguiente información. En la *gráfica 9*, se utiliza como referente la serie de tiempo “DOLLAR”, el resultado obtenido corresponde al modelo LSTM con el error RMSE.



Gráfica 9 LSTM/Dollar/RMSE

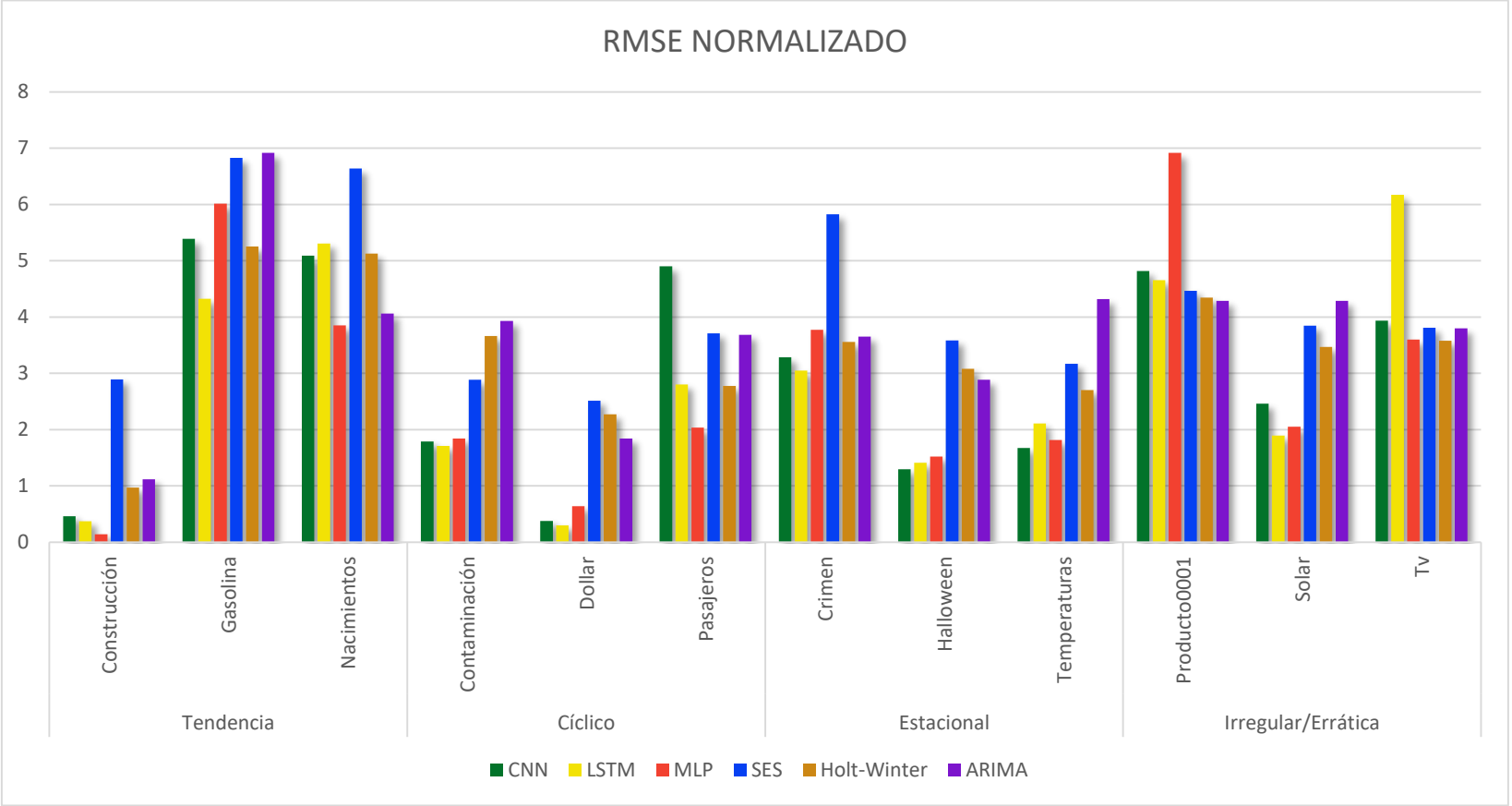
RMSE = 1,08

5.4 Comparación de los métodos de pronostico

Tabla 3 Comparación de los resultados

Fuente: Elaboración propia

Naturaleza de la demanda	Cantidad de datos	Periodos pronosticados	Nombre de la Base de datos	IA			Tradicionales		
				CNN	LSTM	MLP	SES	Holt Winter	ARIMA
				RMSE	RMSE	RMSE	RMSE	RMSE	RMSE
Tendencia	631	189	Construcción	3,02	2,43	0,92	18,87	6,32	7,28
	523	157	Gasolina	432,01	346,35	481,96	547,31	420,8	554,15
	364	109	Nacimientos	8,61	8,97	6,51	11,22	8,67	6,87
Cíclico	4017	1205	Contaminación	2,33	2,23	2,4	3,76	4,78	5,13
	565	166	Dollar	1,36	1,08	2,31	9,11	8,24	6,68
	469	141	Pasajeros	1950,14	1115,07	808,93	1475,75	1104,88	1467,19
Estacional	851	156	Crimen	24,13	22,39	27,71	42,8	26,13	26,81
	547	164	Halloween	5,53	6,03	6,49	15,3	13,15	12,31
	3649	1095	Temperaturas	2,42	3,05	2,62	4,58	3,9	6,24
Errática	468	140	Producto0001	1618	1563,28	2323,24	1499,67	1459,99	1440,32
	8759	2628	Solar	1352,64	1042,38	1128,71	2116,94	1907,31	2357,79
	543	163	Tv	125783,67	197178,93	114925,77	121739,89	114370,24	121410,92



Gráfica 10 Series de tiempo comparadas por medio del RMSE normalizado.

5.4.1 Interpretación de los resultados

En la *tabla 3* se resaltaron en color verde los valores mínimos de RMSE para cada una de las series de tiempo, logrando evidenciar que en mayor medida los resultados más próximos a cero corresponden a los modelos de pronóstico de aprendizaje de máquina. Se puede observar que para las series de tiempo erráticas existe una particularidad, esta consiste en que los métodos tradicionales son menores respecto a la métrica de error que los de aprendizaje de máquina, esto puede deberse a la dificultad en la predicción de este tipo de datos y la necesidad de considerar variables externas para lograr mejores resultados en la parametrización de los mismo que permitan una mayor precisión.

Para obtener una gráfica capaz de evidenciar la diferencia entre los resultados fue necesario la normalización de los valores obtenidos de RMSE ya que cada serie de tiempo tiene una escala diferente. La normalización logra contar el número de desviaciones estándar respecto a la media de cada una de las series de tiempo, así el eje **y** corresponde a este número, en la *gráfica 10* se puede observar que la barra amarilla (LSTM) y la naranja (MLP) son menores en la mayoría de los casos, es decir que estos modelos son los más precisos respecto a la métrica de error RMSE.

Finalmente se obtiene la *tabla 4* en la cual se muestran cuales resultan ser los modelos recomendados en este proyecto para cada naturaleza de serie de tiempo, esto debido a la frecuencia en la aparición del modelo como más preciso respecto a los otros.

Tabla 4 Método de pronostico recomendado

Fuente: elaboración propia

Naturaleza de la demanda	**Método recomendado**
Tendencia	MLP
Cíclico	LSTM
Estacional	CNN
Errática	Tradicionales

5.5 Conclusiones

El presente proyecto tuvo como objetivo comprobar la superioridad asertiva de los métodos de pronóstico de demanda utilizando inteligencia artificial frente a los métodos tradicionales. Lo expuesto a lo largo de este trabajo permite arribar a las siguientes conclusiones:

En cuanto a la revisión del estado del arte, se muestra claramente la dirección que han tomado los avances tecnológicos en cuanto a la utilización de la inteligencia artificial para los últimos 3 años, es importante rescatar que nos encontramos aun en la superficie de su capacidad, pero resulta sorprendente el poder que posee para simplificar la solución de problemas complejos, por ejemplo, para el pronóstico de demanda logrando resultados cada vez más precisos. Los repetidos estudios sobre estas facilitaron la elección de los modelos a evaluar en este proyecto.

La utilización de recursos gratuitos tales como los repositorios de bases de datos contribuyen al desarrollo colectivo, potencian la comunidad educativa y mejoran el desempeño de los estudiantes e investigadores interesados por el desarrollo de tecnologías relacionadas con la inteligencia artificial. Actualmente empresas de todo el mundo comparten su información con la finalidad de premiar a quienes desarrollen soluciones creativas haciendo uso del aprendizaje de máquina para solucionar problemas reales.

La herramienta Google Collaborative está disponible de forma gratuita y posee gran cantidad de librerías de gran utilidad, esta permitió la aplicabilidad tanto de los modelos tradicionales como los de inteligencia artificial de manera automatizada por medio de la programación en lenguaje Python, siendo este recurso una pieza esencial para responder a la pregunta inicial, pues por este medio se logró mejorar significativamente la precisión de los pronósticos acercándolos cada vez más al valor real por medio del aprendizaje de máquina.

Como se pudo evidenciar en este proyecto resultan más precisos los modelos de inteligencia artificial, esto debido a que pueden manipular gran cantidad de datos en poco tiempo y entrenarse por medio del reconocimiento sucesivo de patrones, esto los hace infalibles ante otros métodos de pronóstico, de manera que podamos reconocer estos modelos, aplicarlos y evaluarlos con la finalidad de obtener comparaciones significativas que permitan obtener mejores resultados en la realización de predicciones.

El alcance de este proyecto solo nos permitió analizar 12 series de tiempo clasificadas en 4 naturalezas y la técnica de pronostico elegida como recomendada se obtuvo debido a la frecuencia de aparición con menor RMSE en el desarrollo del mismo, Este proyecto surge como la parte inicial de una investigación más robusta, una que permita determinar con exactitud qué modelo se ajusta a cada naturaleza y de esta manera estandarizar los métodos desarrollados, esto aún no puede ser

posible ya que a pesar de que cada serie de tiempo pertenezca a una clasificación, su variabilidad es única y esto provoca que el método de pronostico a utilizar sea determinado por el comportamiento de la serie mas no por la naturaleza de la misma.

Sin embargo, este proyecto busca recomendar de forma intuitiva un modelo que sea utilizado inicialmente para la evaluación de una naturaleza de serie de tiempo determinada, esto por medio de la observación obtenida de los resultados de este, pero en caso de que no presenten un ajuste congruente, la elección dependen totalmente del usuario que utilice los pronósticos, de esta manera sin el desarrollo de un análisis más robusto, es decir, con un mayor número de series de tiempo analizadas no se puede afirmar en su totalidad que los modelos aquí recomendados sean los óptimos para cada naturaleza de serie de tiempo.

5.6 Recomendaciones

Para futuros proyectos de aplicación de modelos de pronóstico de demanda, se recomienda contar con las series de tiempo como primera medida, especialmente a proyectos que trabajen con datos de organizaciones específicas, ya que esto permitirá tener una directriz frente al desarrollo de los modelos.

Es necesario realizar una limpieza minuciosa de los datos que se usarán para el desarrollo de pronósticos, ya que esto facilitará la aplicación de los modelos, se debe realizar un riguroso análisis que permita descartar información menos importante y eliminar cada uno de los errores que puedan detener la ejecución de los modelos, tales como valores nulos o celdas en blanco.

Se recomienda a las facultades relacionadas con temas de planeación de demanda o realización de pronóstico que se planteen la posibilidad de enseñar estas herramientas que han surgido en los últimos años, las cuales tienen una gran utilidad y aplicabilidad.

Es importante resaltar la sinergia que debe existir entre múltiples disciplinas, la conexión entre la programación y la solución de casos reales es cada vez más necesaria, por esto es preciso que se motive a los estudiantes a aprender a trabajar en equipos compuestos por diferentes profesiones con la finalidad de lograr combinar conocimientos que permitan el desarrollo de soluciones creativas haciendo uso de tecnologías de información.

Aprender sobre avances tecnológicos, estar enterado de nuevos desarrollos y hacer uso de las herramientas existentes es un deber de los estudiantes del siglo XXI, la inteligencia artificial se encuentra aún en su fase de desarrollo, es por esto que constituye una gran oportunidad para incursionar en un mundo que está en constante transformación y el cual probablemente dentro de pocos años logre controlar todo a su alrededor por medio de algoritmos desarrollados en esta modalidad.

Para este tipo de proyectos es posible utilizar dos medidores de error, que para nuestro caso corresponderían al MAE y el RMSE, estos no pueden ser comparados entre sí, pero si pueden ser comparados con sus semejantes en otras series de tiempo, lo que construiría una recomendación más sólida para la elección de los métodos de pronóstico.

Finalmente, la invitación a arriesgarse a manipular herramientas tecnológicas, Python como un lenguaje de programación interpretado permite desde la creación de aplicaciones hasta ser una de las herramientas más utilizadas en ciencia de datos, el mercado actual se moviliza en la dirección de la administración de información y quienes dominen estas herramientas tendrán mayores oportunidades en el futuro.

6 BIBLIOGRAFÍA

Ahmed, N. K., Atiya, A. F., El Gayar, N., & El-Shishiny, H. (2010). An empirical comparison of machine learning models for time series forecasting. *Econometric Reviews*, *29*(5), 594–621. https://doi.org/10.1080/07474938.2010.481556

Alpaydin, E. (2014). *Introduction to machine learning*. (M. I. of Tecnology, Ed.) (Third edit). London, England.

Amat, J. (2017). Máquinas de Vector Soporte (Support Vector Machines, SVMs). Retrieved from https://rpubs.com/Joaquin_AR/267926

Bon, A. T., & Tieng, C. S. (2019). Develop tourism demand forecasting model using Artificial Neural Network. *Proceedings of the International Conference on Industrial Engineering and Operations Management*, *2019*(MAR), 3321–3326.

Cerqueira, V., Torgo, L., & Soares, C. (2019). Machine Learning vs Statistical Methods for Time Series Forecasting: Size Matters. Retrieved from http://arxiv.org/abs/1909.13316

Chase, R. B., Jacobs, F. R., & Aquilano, N. J. (2009). *Administración de Operaciones*. (McGRAW-HILL, Ed.). México, D.F.

Cheng, Y., Xu, C., Mashima, D., Thing, V. L. L., & Wu, Y. (2019). PowerNet: Neural Power Demand Forecasting in Smart Grid. Retrieved from http://arxiv.org/abs/1904.11979

Chicaiza, J. A. (2019). *Diseño de un modelo para el pronóstico de la demanda para una empresa embotelladora de agua purificada mediante redes neuronales*. Ecuador.

Columbus, L. (2019). Roundup Of Machine Learning Forecasts And Market Estimates For 2019. Retrieved October 16, 2019, from https://www.forbes.com/sites/louiscolumbus/2019/03/27/roundup-of-machine-learning-forecasts-and-market-estimates-2019/#3c8fd6697695

Fallis, A. . (2014). *Diseño de un sistema de pronostico de la demanda de nitrogeno y oxigeno para la compañía cryogas, s.a. basado en redes neuronales. Journal of Chemical Information and Modeling*. https://doi.org/10.1017/CBO9781107415324.004

Fan, Y., Zhang, Y., Hua, Y., Wang, Y., Zhu, L., Zhao, J., … Liu, H. (2019).

Investigation of Machine Intelligence in Compound Cell Activity Classification. *Molecular Pharmaceutics*, *16*(11), 4472–4484. https://doi.org/10.1021/acs.molpharmaceut.9b00558

Garcete, A., Benítez, R., Pinto-Roa, D., & Vázquez, A. (2017). Técnica de pronóstico de la demanda basada en Business Intelligence y Machine Leaming. *Simposio Argentino Sobre Tecnología y Sociedad*, 193–202. Retrieved from http://sedici.unlp.edu.ar/bitstream/handle/10915/64728/Documento_completo.pdf-PDFA.pdf?sequence=1

González, L. (2019). Introducción a la Librería Pandas de Python . Retrieved April 18, 2020, from https://ligdigonzalez.com/introduccion-a-la-libreria-pandas-de-python-parte-1/

González Ligdi. (2020). Introducción a la librería NumPy de Python . Retrieved April 18, 2020, from https://ligdigonzalez.com/introduccion-a-numpy-python-1/

Hamedmoghadam, H., Joorabloo, N., & Jalili, M. (2018). Australia's long-term electricity demand forecasting using deep neural networks. Retrieved from http://arxiv.org/abs/1801.02148

Hoptroff, R. G. (1993). The principles and practice of time series forecasting and business modelling using neural nets. *Neural Computing & Applications*, *1*(1), 59–66. https://doi.org/10.1007/BF01411375

Hu, X., Szerlip, P., Karaletsos, T., & Singh, R. (2019). Applying SVGD to Bayesian Neural Networks for Cyclical Time-Series Prediction and Inference, (1), 1–5. Retrieved from http://arxiv.org/abs/1901.05906

Huimei, H. A. N., Xingquan, Z. H. U., & Ying, L. I. (2020). Generalizing long short-term memory network for deep learning from generic data. *ACM Transactions on Knowledge Discovery from Data*, *14*(2). https://doi.org/10.1145/3366022

Hunter John; Dale Darren, F. E. D. M. (2012). Matplotlib: Python plotting — Matplotlib 3.2.1 documentation. Retrieved April 18, 2020, from https://matplotlib.org/

Khalid, R., Javaid, N., Al-zahrani, F. A., Aurangzeb, K., Qazi, E. U. H., & Ashfaq, T. (2020). Electricity load and price forecasting using jaya-long short term memory (JLSTM) in smart grids. *Entropy*, *22*(1), 10. https://doi.org/10.3390/e22010010

Kim, T. S., Lee, W. K., & Sohn, S. Y. (2019). Graph convolutional network approach applied to predict hourly bike-sharing demands considering spatial, temporal, and global effects. *Plos One*, *14*(9), 1–17. https://doi.org/10.1371/journal.pone.0220782

Lao, O., & Caridad, M. (2017). Redalyc.Procedimiento para el pronóstico de la demanda mediante redesneuronales artificiales.

Poorzaker Arabani, S., & Ebrahimpour Komleh, H. (2019). The Improvement of Forecasting ATMs Cash Demand of Iran Banking Network Using Convolutional Neural Network. *Arabian Journal for Science and Engineering*, *44*(4), 3733–3743. https://doi.org/10.1007/s13369-018-3647-7

Ruiz, C. A., & Matich, D. J. (2001). Redes Neuronales: Conceptos Básicos y Aplicaciones.

Sahoo, B. B., Jha, R., Singh, A., & Kumar, D. (2019). Long short-term memory (LSTM) recurrent neural network for low-flow hydrological time series forecasting. *Acta Geophysica*, *67*(5), 1471–1481. https://doi.org/10.1007/s11600-019-00330-1

Siami-Namini, S., Tavakoli, N., & Namin, A. S. (2019). A Comparative Analysis of Forecasting Financial Time Series Using ARIMA, LSTM, and BiLSTM. Retrieved from http://arxiv.org/abs/1911.09512

Suman, S., & Kochak, A. (2015). Demand Forecasting Using Neural Network for Supply Chain Managment, *4*(1).

Tyralis, H., Papacharalampous, G., Langousis, A., Parmezan, A. R. S., Souza, V. M. A., Batista, G. E. A. P. A., … Smith, D. M. (2019). Super learning for daily streamflow forecasting: Large-scale demonstration and comparison with multiple machine learning algorithms. *ArXiv Preprint ArXiv:1909.04131*, *1*, 36.

Valencia Cárdenas, M., Ramírez Agudelo, S., Fernando Tabares Muñoz, J., & Andrés Velásquez Galvis, C. (2014). *MÉTODOS DE PRONÓSTICOS-CLÁSICOS Y BAYESIANOS CON APLICACIONES*.

Vijai, P., & Bagavathi Sivakumar, P. (2018). Performance comparison of techniques for water demand forecasting. *Procedia Computer Science*, *143*, 258–266. https://doi.org/10.1016/j.procs.2018.10.394

Anexos

Anexo 1 Modelos de pronóstico

Fuente: (Anderson, Sweeney y Williams, 2004, p. 206), (Montemayor, 2013).
Elaboración propia

Modelos de pronostico			
Modelo	**Método**	**Descripción**	**Computo**
Cualitativo	**Delfos**	Consenso en grupo, un panel de expertos es consultado de manera independiente para dar respuesta a varios cuestionarios, la meta es producir un despliegue de opiniones relativamente reducido dentro del cual coincida la mayoría de los expertos.	No
	Juicio de expertos	Cada uno de los participantes hace una serie de consideraciones de los factores que posiblemente pueden influir en el pronóstico que se está realizando y, posteriormente, realizan un consenso de sus conclusiones dando como resultado el pronóstico.	No
	Redacción del escenario	Elaboración de un escenario conceptual de futuro teniendo en cuenta un conjunto de suposiciones muy bien definidas, la persona encargada debe contar con un juicio muy razonable para determinar la probabilidad de la presentación de cada escenario antes de llevar a cabo cualquier toma de decisión.	No
	Enfoques intuitivos	Comúnmente se desarrolla en grupo en donde los participantes realizan una lluvia de ideas.	No
Series de tiempo	**Redes neurales**	Usa un programa complejo de computadora para asimilar datos relevantes y reconocer patrones mediante "aprendizaje" como lo hacen los humanos	Si
	Técnicas Box-Jenkins	No supone ningún patrón particular en los datos históricos de la serie que va a ser pronosticada; usa un método iterativo de identificación y ajuste de un modelo posiblemente útil tomado de un grupo general de modelos	Si
	Modelos autorregresivos	Se emplean con variables económicas para explicar las relaciones entre observaciones adyacentes en una serie de tiempo	Si
	Suavización exponencial	Similar a los promedios móviles, pero los valores son ponderados exponencialmente, otorgando mayor peso a los datos más recientes	Si
	Promedios móviles	Se usan para eliminar la aleatoriedad en una serie de tiempo; el pronóstico se basa en los datos de la serie de tiempo suavizados por un promedio móvil	No
	Método de descomposición	Pronóstico explicativo: supone una relación de causa-efecto entre el tiempo y la información resultante del sistema; el sistema se descompone en sus elementos	Si
Modelos causales de pronósticos	**Regresión múltiple**	Pronóstico explicativo: supone una relación de causa-efecto entre la información introducida en el sistema y la información que resulta de éste	Si
	Análisis de regresión	Pronóstico explicativo; supone una relación de causa-efecto entre la información de entrada y la información resultante	Si

Anexo 2 Redes neuronales artificiales

Adaptada de: (Chicaiza, 2019).
Fuente: Elaboración propia

Redes neuronales artificiales		
Modelo	**Método**	**Descripción**
Supervisadas	**Máquina de vectores de soporte (SVV)**	Es uno de los mejores clasificadores para un amplio abanico de situaciones, por lo que se considera uno de los referentes dentro del ámbito de aprendizaje estadístico y machine learning. Las Máquinas de Vector Soporte se fundamentan en el Maximal Margin Classifier, que, a su vez, se basa en el concepto de hiperplano. (Amat, 2017)
	Perceptrón multicapa (MLP)	La red Neuronal como perceptrón nace a partir de añadir capas ocultas a un perceptrón simple. Esta arquitectura suele entrenarse mediante el algoritmo de retro propagación de errores o Back Propagation (BP)
	Larga memoria de corto plazo (LSTM)	Es una extensión de las redes neuronales recurrentes. "El LSTM consiste en una célula de memoria que consta de cuatro elementos básicos: una puerta de entrada, una neurona con una conexión autorreparadora (una conexión a sí misma), una puerta de olvidar y una puerta de salida. Esta multiplicación tiene lugar en cada celda de entrada y salida por sus respectivas puertas, mientras que la puerta multiplica el estado anterior (la conexión auto-corriente de la célula de memoria) y permite a la célula olvidar o recordar su estado anterior usando la función de activación sigmoide." Traducido de (Sahoo, Jha, Singh, & Kumar, 2019)
	Adalina	Es un tipo de red considerada como elástica, ya que utiliza una neurona similar a la del perceptrón, pero da una respuesta lineal cuyas entradas pueden ser continuas. Esto permite que las salidas puedan tomar cualquier valor, en cambio las del perceptrón debe ser de 0 o 1
	Redes de Elman	Es un tipo de Red BP de dos capas, la primera de tipo sigmoidea y la segunda lineal, con realimentación desde la salida de la primera capa hasta la entrada de esta
	Red de capa competitiva	Este tipo de red es no supervisada y unidireccional. Esto quiere decir que las neuronas de la capa competitiva se distribuyen por sí solas para reconocer los vectores de entrada presentados de forma frecuente
No supervisadas	**Redes de base radial (RBF)**	El modelo de redes de base radial, RBF, aunque de reciente introducción, cada vez cuenta con más aplicaciones prácticas gracias a su simplicidad, generalidad y rapidez de aprendizaje
Redes hibridas	**Red de regresión generalizada (GRNN)**	Las redes GRNN se usan normalmente para aproximación de funciones. Poseen una capa de neuronas de base radial y una capa lineal especial
	Red neuronal probabilística (PNN)	Las redes PNN pueden ser usadas en problemas de clasificación. Cuando se presenta una entrada, la primera capa computa las distancias desde el vector de entrada a los vectores de entrada del entrenamiento y produce un vector cuyos elementos indican lo próxima que está la entrada respecto a la entrada de entrenamiento

Anexo 3 Análisis y clasificación bases de datos

Inicialmente por medio de Excel y su herramienta de clasificación y organización de bases de datos, Power Query se realizó la filtración, clasificación y separación de los elementos necesarios para el desarrollo del proyecto, así para todas las bases de datos se homogenizo la cantidad de columnas a 2, las cuales llamamos “Fecha” y “Cantidad”, cada base de datos con su respectiva cantidad de filas e información.

Haciendo uso de los siguientes códigos adaptados para la revisión de bases de datos, se clasificaron las bases de datos según la naturaleza de sus patrones:

El 75% de los datos fueron tomados como entrenamientos y los restantes como sets de pruebas y comparación, el estudio de estacionalidad y tendencia se realizó para el set de prueba, por medio de este código que fue adaptado para la visualización grafica de las bases de datos, al igual que la limpieza, y eliminación de errores.

Dirigirse al anexo 3 en el archivo Word adjunto

Anexo 4 Resultados pronósticos

Para la visualizacion de cada una de las series de tiempo, su respectivo codigo y resultado **dirigirse al anexo 4 en el archivo word adjunto**

Anexo 5 Códigos utilizados

Modelo LSTM

Modelo completo en el Link Google collaborative:
https://colab.research.google.com/drive/1-6nyO0lEk4z1w36kYmEvtjpR6MuRzdvo

```
#Importar librerías
from numpy import concatenate
import numpy as np
import pandas as pd
from matplotlib import pyplot
from pandas import read_csv
from pandas import DataFrame
from pandas import concat
from sklearn.preprocessing import MinMaxScaler
from sklearn.preprocessing import LabelEncoder
from sklearn.metrics import mean_squared_error
```

```
from sklearn.metrics import mean_absolute_error
from sklearn.metrics import mean_absolute_error
from keras.models import Sequential
from keras.layers import Dense
from keras.layers import LSTM


#Definir series de supervisión (Series de tiempo a aprendizaje supervisado)
def series_to_supervised(data, n_in=1, n_out=1, dropnan=True):
n_vars = 1 if type(data) is list else data.shape[1]
df = pd.DataFrame(data)
cols, names = list(), list()
for i in range(n_in, 0, -1):
cols.append(df.shift(i))
names += [('var%d(t-%d)' % (j+1, i)) for j in range(n_vars)]
for i in range(0, n_out):
cols.append(df.shift(-i))
if i == 0:
names += [('var%d(t)' % (j+1)) for j in range(n_vars)]
else:
names += [('var%d(t+%d)' % (j+1, i)) for j in range(n_vars)]
agg = concat(cols, axis=1)
agg.columns = names
if dropnan:
```

Modelo CNN

Modelo completo en el Link Google collaborative:
https://colab.research.google.com/drive/1BDOOa5wDkjFQevcRILCtPXgLUX6GpabV

```
# Importar librerías
import pandas as pd
import numpy as np
from matplotlib import pyplot
from numpy import array
from keras.models import Sequential
from keras.layers import Dense
from keras.layers import Flatten
from pandas import concat
from keras.layers.convolutional import Conv1D
from keras.layers.convolutional import MaxPooling1D
from sklearn.preprocessing import MinMaxScaler
from sklearn.metrics import mean_squared_error
from sklearn.metrics import mean_absolute_error
```

```
# Dividir la secuencia de datos en conjuntos
def split_sequence(sequence, n_steps):
X, y = list(), list()
for i in range(len(sequence)):
end_ix = i + n_steps
if end_ix > len(sequence)-1:
break
seq_x, seq_y = sequence[i:end_ix], sequence[end_ix]
X.append(seq_x)
y.append(seq_y)
return array(X), array(y)

#Definir series de supervisión
def series_to_supervised(data, n_in=1, n_out=1, dropnan=True):
n_vars = 1 if type(data) is list else data.shape[1]
df = pd.DataFrame(data)
cols, names = list(), list()
for i in range(n_in, 0, -1):
cols.append(df.shift(i))
names += [('var%d(t-%d)' % (j+1, i)) for j in range(n_vars)]
for i in range(0, n_out):
cols.append(df.shift(-i))
if i == 0:
```

Modelo MPL

Modelo completo en el Link Google Collaborative: https://colab.research.google.com/drive/1zAZHWXgBlhGfGH9Y3NftuDGeiP9m94m1

```
def TSF_B_using_TF():
  import warnings
  warnings.filterwarnings("ignore")
import numpy as np
import matplotlib
import matplotlib.pyplot as plt
import pandas as pd
import statsmodels.api as sm
from pandas.plotting import autocorrelation_plot
df=pd.read_csv('NACIMIENTOS.csv')#Definir la cantidad para entrenar
from sklearn.model_selection import train_test_split
train, test = train_test_split(df, shuffle=False)
```

```
df.Timestamp=pd.to_datetime(df.Fecha,format='%d-%m-%Y')
df.index=df.Timestamp

train.Timestamp=pd.to_datetime(train.Fecha,format='%d-%m-%Y')
train.index=train.Timestamp

test.Timestamp=pd.to_datetime(test.Fecha,format='%d-%m-%Y')
test.index=test.Timestamp

print(train.shape)
print(test.shape)
print(train.head())
train.dropna(how='any', inplace=True)
test.dropna(how='any', inplace=True)
print(train.shape)
print(test.shape)
print(train.head())
train['Cantidad']=train.Cantidad.astype(int)
test['Cantidad']=test.Cantidad.astype(int)
```

Modelo SES

Modelo completo en el Link Google Collaborative:
https://colab.research.google.com/drive/1Im-GpUQRzxHutFsPa0gOfz1XDKIfl--w

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
df=pd.read_csv('NACIMIENTOS.csv')
df.head()
df.tail()
df=pd.read_csv('NACIMIENTOS.csv')#Definir la cantidad para entrenar
from sklearn.model_selection import train_test_split
train, test = train_test_split(df, shuffle=False)
df.Timestamp=pd.to_datetime(df.Fecha,format='%d-%m-%Y')
df.index=df.Timestamp

train.Timestamp=pd.to_datetime(train.Fecha,format='%d-%m-%Y')
train.index=train.Timestamp

test.Timestamp=pd.to_datetime(test.Fecha,format='%d-%m-%Y')
test.index=test.Timestamp
```

```
print(train.shape)
print(test.shape)
print(train.head())
train.dropna(how='any', inplace=True)
test.dropna(how='any', inplace=True)
print(train.shape)
print(test.shape)
print(train.head())
train['Cantidad']=train.Cantidad.astype(int)
test['Cantidad']=test.Cantidad.astype(int)
```

Modelo Holt Winters

Modelo completo en el Link Google Collaborative:
https://colab.research.google.com/drive/1AOWtaHZIFEhttdsd-8XD07lZPAifQbqx

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
df=pd.read_csv('CONTAMINACION.csv')
df.head()
df.tail()
df=pd.read_csv('CONTAMINACION.csv') #Definir la cantidad para entrenar
from sklearn.model_selection import train_test_split
train, test = train_test_split(df, shuffle=False)
df.Timestamp=pd.to_datetime(df.Fecha,format='%d-%m-%Y')
df.index=df.Timestamp

train.Timestamp=pd.to_datetime(train.Fecha,format='%d-%m-%Y')
train.index=train.Timestamp

test.Timestamp=pd.to_datetime(test.Fecha,format='%d-%m-%Y')
test.index=test.Timestamp

print(train.shape)
print(test.shape)
print(train.head())
train.dropna(how='any', inplace=True)
test.dropna(how='any', inplace=True)
print(train.shape)
print(test.shape)
```

```
print(train.head())
train['Cantidad']=train.Cantidad.astype(int)
test['Cantidad']=test.Cantidad.astype(int)
```

Modelo ARIMA

Modelo completo en el Link Google Collaborative:
https://colab.research.google.com/drive/1s_JZy3b8Gv5bw9MvigHPHk3gAP3Rd3Lv

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
df=pd.read_csv('PASAJEROS.csv')
df.head()
df.tail()
df=pd.read_csv('PASAJEROS.csv') #Definir la cantidad para entrenar
from sklearn.model_selection import train_test_split
train, test = train_test_split(df, shuffle=False)
df.Timestamp=pd.to_datetime(df.Fecha,format='%d-%m-%Y')
df.index=df.Timestamp

train.Timestamp=pd.to_datetime(train.Fecha,format='%d-%m-%Y')
train.index=train.Timestamp

test.Timestamp=pd.to_datetime(test.Fecha,format='%d-%m-%Y')
test.index=test.Timestamp

print(train.shape)
print(test.shape)
print(train.head())
train.dropna(how='any', inplace=True)
test.dropna(how='any', inplace=True)
print(train.shape)
print(test.shape)
print(train.head())
train['Cantidad']=train.Cantidad.astype(int)
test['Cantidad']=test.Cantidad.astype(int)
```