

INSTITUTO TECNOLÓGICO Y DE ESTUDIOS SUPERIORES DE  
MONTERREY  
CAMPUS QUERÉTARO



**Tecnológico  
de Monterrey**

Uso de inteligencia artificial para crear un pronóstico de datos en  
una serie de tiempo

por

Emmanuel Enrique Bautista Palacios

Tutor: Dr. Pedro Oscar Pérez Murueta

Santiago de Querétaro, Querétaro, México

23/11/2023

## *Agradecimientos*

Ante todo, quisiera expresar mi sincera gratitud a mi tutor de tesis, el Dr. Pedro Oscar Pérez Murueta, por su inestimable apoyo y orientación a lo largo de todo el proceso de investigación. Su paciencia, dedicación y conocimientos han sido fundamentales para completar este trabajo.

También me gustaría dar las gracias a mis compañeros de clase por su apoyo y ánimo. Gracias a todos ustedes, este viaje ha sido mucho más agradable.

Un agradecimiento especial a mi familia y amigos por su constante apoyo y comprensión. Sin ustedes, no habría podido llegar hasta aquí.

Así mismo, agradezco a la empresa Hunai, por proveer los datos con los que fue posible llevar a cabo este proyecto.

Por último, me gustaría dar las gracias al Tecnológico de Monterrey por brindarme la oportunidad de realizar esta investigación.

**Abstract**— La variabilidad en la demanda tiene un impacto en la toma de decisiones de las empresas, por lo que es importante el uso de un pronóstico de tendencias para una planeación a largo plazo. Este documento aborda la integración de métodos de aprendizaje automático y técnicas de pronóstico de series de tiempo para abordar los desafíos enfrentados por las empresas que operan en mercados volátiles.

# Contenido

## Introducción

### 1.1 Antecedentes

### 1.2 Contexto

## Estado del arte

### 2.1 Trabajos previos

### 2.2 Conclusión

## Objetivo de la Investigación

### 3.1 Justificación de la propuesta

## Solución

### 4.1 Tecnología

### 4.2 Conjunto de datos

### 4.3 Desafíos

### 4.4 Desarrollo

### 4.5 Ajustes a los datos y al modelo

## Evaluación

### 5.1 Diseño de la evaluación

## Resultados y análisis

### 6.1 Resultados

### 6.2 Análisis

## Conclusión y trabajo futuro

### 7.2 Trabajo futuro

# Capítulo 1

## Introducción

### 1.1 Antecedentes

En el transcurso de la investigación de este proyecto, hablaremos de términos relacionados con la informática, así como con las matemáticas y la estadística. Para facilitar una mejor comprensión del proyecto, se dará un breve resumen de estos temas.

- Inteligencia artificial

La inteligencia artificial es la ciencia de crear máquinas inteligentes, especialmente programas informáticos inteligentes. Es un campo que utiliza las ciencias computacionales para resolver problemas. A lo largo del tiempo han existido diferentes iteraciones, sin embargo, es en años recientes que se encuentra en un punto donde puede ser utilizado de maneras imprevistas [1].

- Pronósticos en series de tiempo

En estadística inferencial, el pronóstico es un proceso matemático mediante el cual se hace una estimación del valor futuro de una o más variables, como puede ser la demanda. Un pronóstico es una predicción de uno o más eventos futuros a partir de indicios. El éxito de un pronóstico se encuentra en la exactitud con la que coincide con la realidad [2].

- Aprendizaje Automático

El aprendizaje automático es un campo de la inteligencia artificial que se centra en desarrollar algoritmos y modelos que permitan a los ordenadores aprender patrones y tomar decisiones basadas en datos sin intervención humana explícita. En lugar de seguir instrucciones programadas de forma convencional, los sistemas de aprendizaje automático utilizan los datos para aprender y mejorar su rendimiento con el tiempo. [2]

## 1.2 Contexto

Una empresa que se dedica a la compraventa opera en un mercado muy volátil. Las tendencias de los consumidores pueden cambiar rápidamente debido a factores como el cambio de modas, los avances tecnológicos o incluso acontecimientos imprevistos como las pandemias. En esta situación, las fluctuaciones de la demanda pueden ser imprevisibles y extremadamente variables [7].

Esta variabilidad de la demanda puede acarrear varios problemas a la empresa. Por un lado, si la empresa no es capaz de prever correctamente los cambios en la demanda, puede incurrir en exceso o escasez de existencias. El exceso conlleva costes de almacenamiento, obsolescencia y pérdida de oportunidades de invertir en otros productos con mayor demanda. Por otro lado, la escasez puede provocar pérdidas de ventas, clientes insatisfechos y daños a la reputación de la empresa.

Además, el planteamiento de estos pronósticos puede resultar complicado para las pequeñas y medianas empresas, por lo que necesitan aún más formas de facilitar la toma de decisiones sin que sea un factor operativo importante. [8]

# Capítulo 2

## Estado del arte

En este capítulo, analizaremos las investigaciones relevantes que se han realizado en estas áreas, destacando el uso del aprendizaje automático y sus aplicaciones dentro del pronóstico de datos.

### 2.1 Trabajos previos

Los trabajos realizados por Gilliland (2020)[[10](#)], Pavlyshenko (2019)[[11](#)], Ptotic et al. (2022) [[12](#)] y Vera et al (2023) [[6](#)] son los más relevantes para esta investigación.

Examinemos primero los modelos de aprendizaje. El primer trabajo compara el rendimiento de los modelos tradicionales de series temporales, como ARIMA y GARCH, con los métodos de aprendizaje automático más utilizados, como la regresión logística, la red neuronal artificial y la máquina de vectores de apoyo. El estudio también analiza el rendimiento de un modelo de aprendizaje profundo, el autoencoder "Denoising".

La investigación utilizó datos históricos de tres índices bursátiles: Dow 30, S&P 500 y Nasdaq. Los resultados mostraron que los métodos de aprendizaje automático superaron significativamente a los modelos tradicionales de series temporales en términos de precisión de la predicción. El modelo de aprendizaje profundo, el autoencoder Denoising, obtuvo el mejor rendimiento global.[[11](#)]

Dicho esto, existen diferentes modelos que pueden ser usados para fines de pronóstico de datos. Los modelos de aprendizaje automático más utilizados para pronósticos de series temporales son:

- **Regresión lineal:** La regresión lineal es un modelo estadístico que predice una variable continua (ventas) en función de una o más variables independientes (por ejemplo, precio, gastos de marketing, etc.). La regresión lineal es un modelo sencillo de implementar y comprender, pero puede no ser adecuado para series temporales con patrones complejos.

- Árboles de decisión: Los árboles de decisión son un modelo de aprendizaje automático que predice una variable (ventas) en función de una o más variables independientes (por ejemplo, precio, gastos de marketing, etc.). Los árboles de decisión son capaces de aprender patrones complejos en los datos, pero pueden ser difíciles de interpretar.
- Bosques aleatorios: Los bosques aleatorios son un conjunto de árboles de decisión que se entrenan en diferentes subconjuntos de datos. Los bosques aleatorios son más resistentes al sobreajuste que los árboles de decisión individuales, pero pueden ser más computacionalmente intensivos de entrenar.
- Redes neuronales artificiales: Las redes neuronales artificiales son un tipo de deep learning que pueden aprender patrones complejos en los datos. Las redes neuronales artificiales pueden ser muy precisas para los pronósticos de series temporales de ventas, pero pueden ser difíciles de entrenar y requieren una gran cantidad de datos.

Los modelos de aprendizaje automático son una herramienta poderosa para los pronósticos de series temporales y demostraron tener un mejor desempeño en precisión cuando se compara con métodos tradicionales.[11]. Además, los modelos de aprendizaje automático son más complejos de implementar y requieren más datos de entrenamiento que un modelo tradicional.[12]

Author	sMAPE	MASE	OWA	Runtime (mins)	Runtime vs. Comb
Smyl	11.374	1.536	0.821	8056.0	242×
Montero-Manso, etal.	11.720	1.551	0.838	46108.3	1388×
Legaki & Koutsouri	11.986	1.601	0.861	25.0	0.75×
Theta	12.309	1.696	0.897	12.7	0.38×
Comb	12.555	1.663	0.898	33.2	–
ARIMA	12.669	1.666	0.903	3030.9	91×
Naïve2	13.564	1.912	1.000	2.9	0.09×
Naïve1	14.208	2.044	1.058	0.2	0.01×

Figura 2.1: Resultados entre métodos convencionales y de aprendizaje automático [10]

Las aportaciones clave son que los métodos de aprendizaje automático tienen el potencial de mejorar el pronóstico de datos siempre y cuando se elija el método adecuado para las características de los datos de la serie temporal [12].

Por último, se presentó el uso de estas herramientas para proveer un dashboard que ofreciera información importante para la empresa. Este dashboard se utiliza para tomar decisiones sobre el marketing y las ventas.

La propuesta se basa en los siguientes pasos:

- Recopilación de datos: Los datos necesarios para el dashboard se recopilan de diversas fuentes, como sistemas de información empresarial, bases de datos y redes sociales.
- Preparación de datos: Los datos se preparan para el análisis.
- Análisis de datos: Los algoritmos de aprendizaje automático se utilizan para analizar los datos y extraer información.
- Visualización de datos: La información extraída de los datos se presenta de forma visual en el dashboard.



Figura 2.2: Dashboard de alto impacto [6]

El estudio de caso muestra que el dashboard es capaz de proporcionar información valiosa a los gerentes, lo que les permite tomar decisiones informadas.[6]

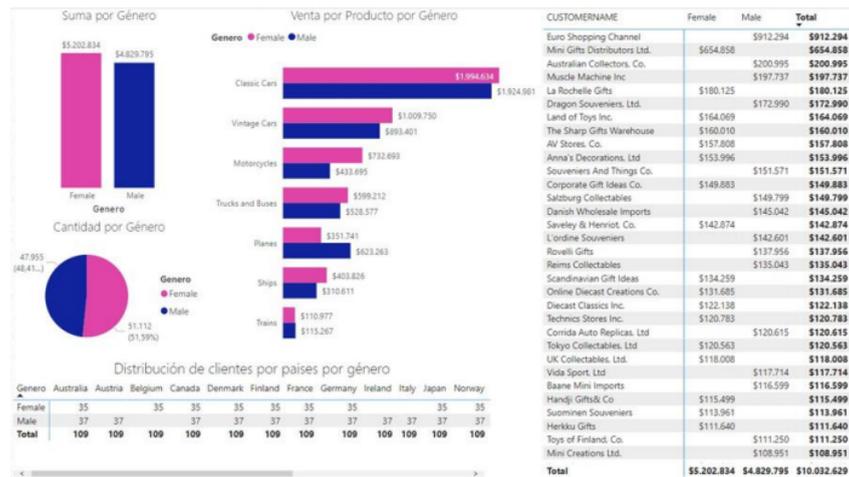


Figura 2.3: Resultado del análisis de datos [6]

## 2.2 Conclusión

Estos trabajos demuestran que los métodos de aprendizaje automático son una herramienta prometedora en una amplia gama de aplicaciones. A pesar de retos como la complejidad de la aplicación y la sensibilidad a los cambios en los datos [10], ofrecen oportunidades para conseguir mejores resultados en este campo.

# Capítulo 3

## Objetivo de la Investigación

El objetivo de esta investigación es reducir la brecha existente en la barrera de entrada para una planificación estratégica eficaz en las empresas. Para lograrlo, se ha establecido el siguiente objetivo: *Desarrollar un modelo de aprendizaje automático para realizar pronósticos en series de tiempo y determinar si es viable su uso en la toma de decisiones de pequeñas y medianas empresas.*

### 3.1 Justificación de la propuesta

La planificación estratégica es crucial para el éxito y la sostenibilidad de cualquier organización, ya sea una empresa con fines de lucro, una institución sin fines de lucro o una entidad gubernamental. Los pronósticos proporcionan información crucial para la planificación a largo plazo de una empresa. Permiten identificar tendencias y patrones en la demanda, lo que facilita la formulación de estrategias que se ajusten a las oportunidades y retos futuros del mercado [3].

Un pronóstico preciso de la demanda ayuda a evitar el exceso o la escasez de existencias. Esto permite a la empresa mantener un nivel de inventario óptimo que satisfaga las necesidades de los clientes y minimice los costes asociados al almacenamiento y la obsolescencia. La empresa puede planificar su producción y sus operaciones con mayor eficacia. Esto implica ajustar la capacidad de producción en función de los pronósticos de la demanda, optimizando los recursos. Al evitar el exceso de inventario, la empresa reduce los costes asociados al almacenamiento y la gestión. Además, una gestión más eficaz del inventario reduce la necesidad de reaccionar rápidamente a los cambios repentinos de la demanda, que pueden dar lugar a decisiones costosas y desorganizadas.

La proyección exacta de los ingresos y gastos futuros es crucial para que cualquier empresa logre la estabilidad financiera. Un pronóstico preciso permite establecer objetivos financieros realistas y garantiza la viabilidad financiera a largo plazo. Como sugieren Wilson & Koerber (1992), los métodos cuantitativos son útiles para obtener predicciones más precisas y ayudan en los procesos de toma de decisiones relativas a la empresa. Por lo tanto, es esencial poseer los conocimientos necesarios para realizar estos cálculos, ya sea manualmente o con programas informáticos.

# Capítulo 4

## Solución

### 4.1 Tecnología

Esta solución fue desarrollada utilizando el lenguaje de programación Python por las siguientes razones:

- Es fácil de usar
- Extensas y poderosas librerías para el análisis de datos y visualización de información.

Por otro lado, se decidió utilizar XGBoost como algoritmo de aprendizaje. XGBoost es un algoritmo de aprendizaje automático que se utiliza para tareas de clasificación y regresión. Se basa en la construcción de árboles de decisión en secuencia, donde cada árbol se enfoca en corregir los errores del anterior. Las razones para usarlo son las siguientes:

- Eficiencia. Usa algoritmos de optimización para encontrar la configuración óptima de los parámetros del modelo. [\[13\]](#)
- Precisión. Uso de técnicas de regularización para evitar un sobreajuste a los datos de entrenamiento. [\[13\]](#)
- Versatilidad. Es capaz de adaptarse a cambios en los datos a medida que nuevos datos son recopilados. [\[13\]](#)

El método XGBoost tiene alta precisión y baja probabilidad de sobreajuste, además de que tiene mejor desempeño que otros métodos:

	RMSE	MAE	R <sup>2</sup>
XGBoost	17.298	11.774	0.9520
Random Forest (RF)	18.911	12.973	0.9426
Support Vector Machine for regression (SVMreg)	20.447	14.503	0.9329
Multiple Linear Regression (MLR)	21.854	16.489	0.9234
Decision Tree Regression (DTR)	25.921	15.917	0.8922

Figura 4.1: Comparación entre cinco modelos diferentes [13]

## 4.2 Conjunto de datos

El conjunto de datos es un factor importante para determinar el éxito de un método de aprendizaje automático. Ya que mientras más alta la calidad de los datos, más preciso y robusto puede ser el modelo [13]. Para esta investigación, se utilizó un conjunto de datos proporcionados por una empresa local de San Miguel de Allende, Gto. Los datos proporcionados son fechas por día a lo largo de diez años y la cantidad de ventas en cada día.

```

Fecha,Venta
01/01/2012,552
02/01/2012,483
03/01/2012,504
04/01/2012,584
05/01/2012,980
06/01/2012,312
07/01/2012,552
08/01/2012,490
09/01/2012,580
10/01/2012,518
11/01/2012,863

```

Figura 4.2: Conjunto de datos en formato csv

### 4.3 Desafíos

Inicialmente, el plan consistía en emplear modelos pre desarrollados a los que se pudiera acceder a través de una API. Sin embargo, estos modelos tenían un alto costo en relación con su precisión. A medida que se volvían más precisos, también se encarecían debido al mayor número de llamadas necesarias para gestionar los datos. En consecuencia, este enfoque resultó inviable.

### 4.4 Desarrollo

El modelo recibe el conjunto de datos a revisar de un archivo con formato csv. Podemos visualizar nuestros datos gracias a python.

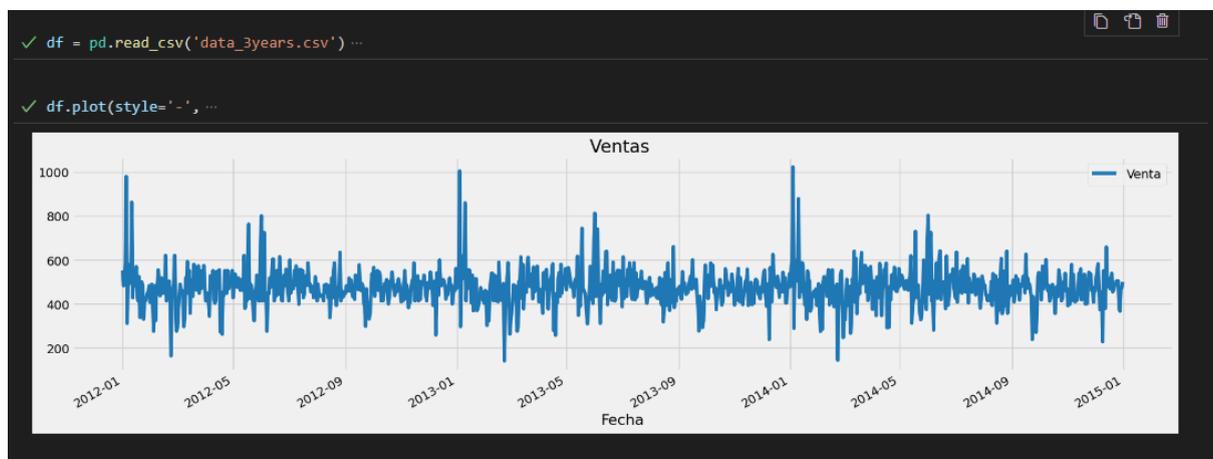
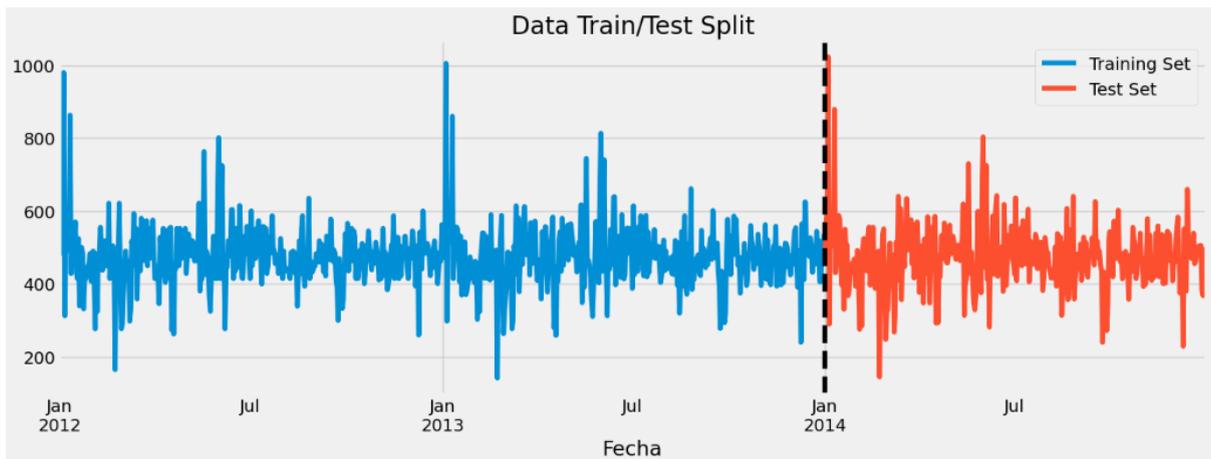


Figura 4.3: Datos históricos

Una vez que tenemos los datos, se define la forma en que vamos a dividirlos para entrenar el modelo y para probarlo.



*Figura 4.4: Datos de entrenamiento / Datos de prueba*

Para el reconocimiento de patrones, es necesario determinar características o 'features' a la información que el modelo va a recibir. En este caso, utilizamos las fechas como el índice, y de esas fechas obtendremos más información que será utilizada por el modelo.

```
df = df.copy()
df['dayofweek'] = df.index.dayofweek
df['quarter'] = df.index.quarter
df['month'] = df.index.month
df['year'] = df.index.year
df['dayofyear'] = df.index.dayofyear
df['dayofmonth'] = df.index.day
df['weekofyear'] = df.index.isocalendar().week
```

Figura 4.5: Features

Ya que se encuentran definidas estas características, comienza el aprendizaje del modelo, ajustando los parámetros necesarios y de acuerdo a nuestras necesidades.

```
reg = xgb.XGBRegressor(base_score=0.5, booster='gbtree',
                       n_estimators=500,
                       early_stopping_rounds=50,
                       objective='reg:linear',
                       max_depth=3,
                       learning_rate=0.01)
reg.fit(X_train, y_train,
        eval_set=[(X_train, y_train), (X_test, y_test)],
        verbose=100)
```

Figura 4.6: Ajustes

```
[0] validation_0-rmse:476.73651 validation_1-rmse:476.62059
[100] validation_0-rmse:189.75604 validation_1-rmse:191.73967
[200] validation_0-rmse:98.52844 validation_1-rmse:105.88586
[300] validation_0-rmse:75.54377 validation_1-rmse:86.04262
```

Figura 4.7: Ejecución

Una vez terminada la ejecución, se hacen ajustes necesarios de acuerdo a las observaciones como el uso de las features que definimos. Podemos visualizarlo gráficamente gracias a python.

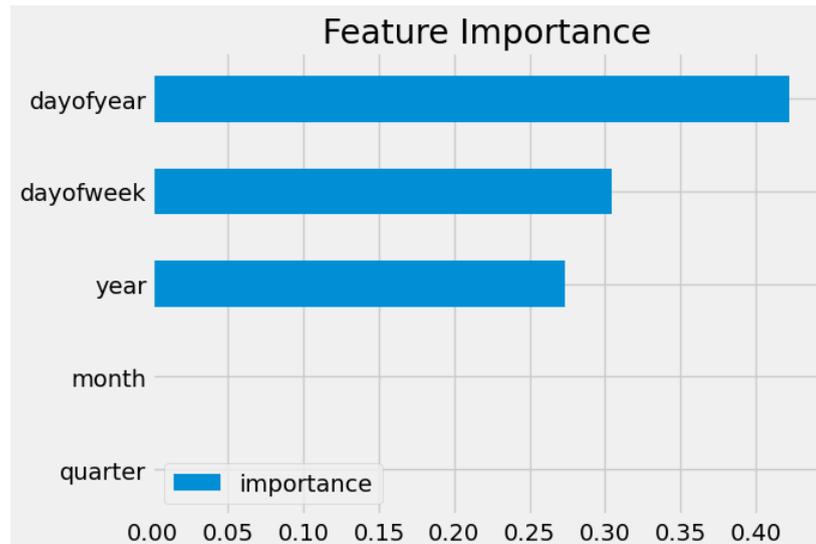


Figura 4.8: Características que más usó el modelo

Ahora sí, podemos observar el pronóstico realizado por el modelo, comparándolo con los datos históricos reales.

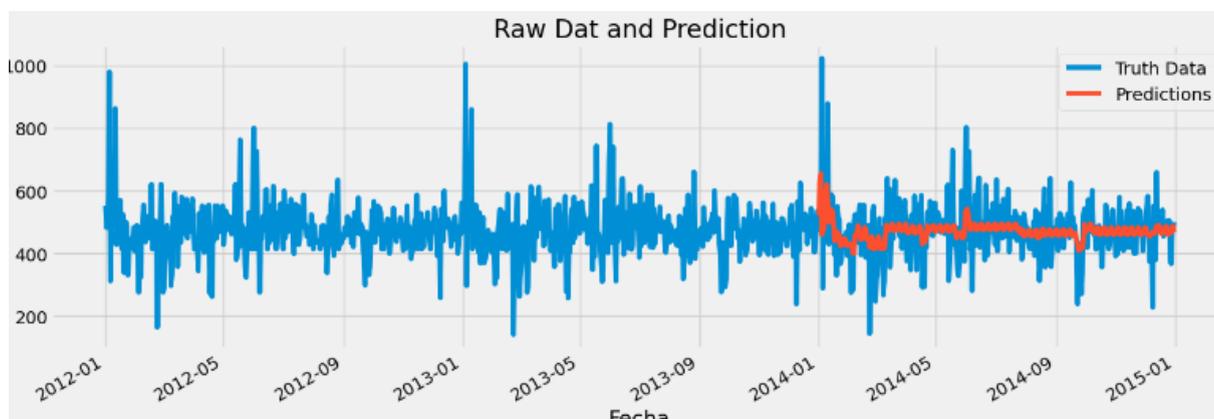


Figura 4.9: Comparación

## 4.5 Ajustes a los datos y al modelo

Debido a la naturaleza de los datos, a menudo existen valores atípicos, que son observaciones numéricamente distantes del resto de los datos. Por lo tanto, para obtener un mejor resultado del

pronóstico, puede ser necesario eliminar estos valores que sólo añaden ruido al analizar los datos, ya que no se producen con la frecuencia suficiente como para formar parte de un patrón.

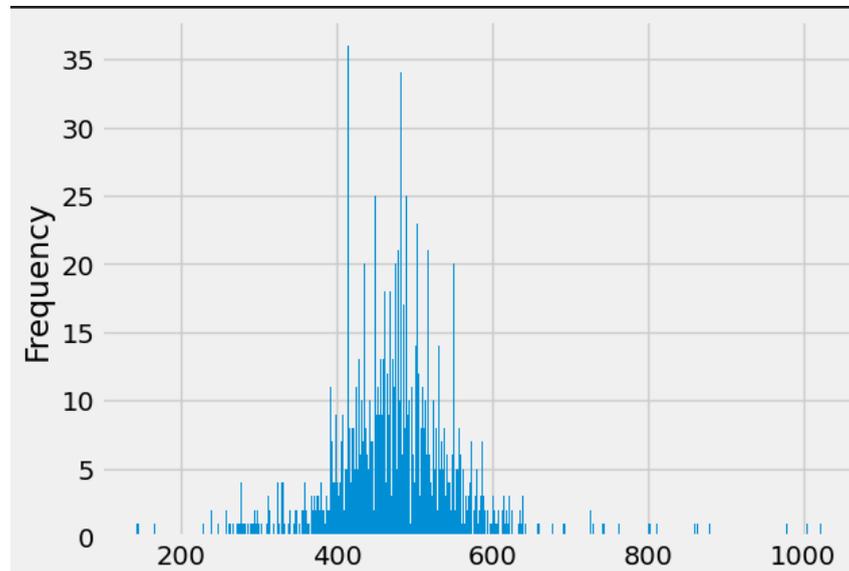


Figura 4.10: Frecuencia de los datos

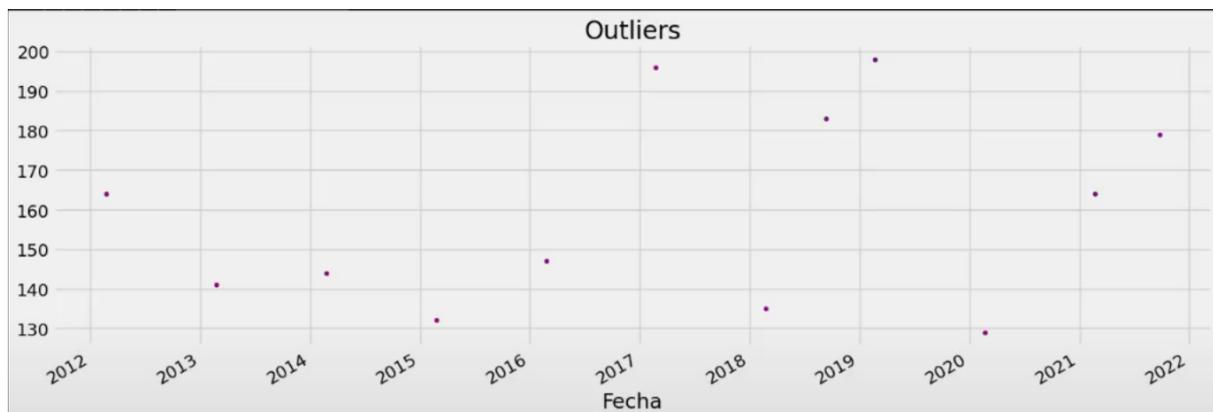


Figura 4.10: Datos atípicos

# Capítulo 5

## Evaluación

### 5.1 Diseño de la evaluación

Para validar los resultados, se llevó a cabo un proceso de comparación entre los resultados obtenidos por el modelo entrenado y los datos reales. Para ello se utilizaron dos métricas principales.

La primera métrica utilizada fue el Error Cuadrático Medio (MSE), que es una medida de la dispersión del error en el pronóstico. Sin embargo, esta medida maximiza el error elevando al cuadrado, castigando los periodos en los que la diferencia era mayor en comparación con otros. Por tanto, esta medida es más útil para periodos con desviaciones menores.

$$MSE = \frac{\sum \text{Error de pronóstico}^2}{n}$$

*Figura 5.1: Error cuadrático medio*

La segunda métrica utilizada es el Error Medio Porcentual Absoluto (MAPE). Esta medida proporciona la desviación en términos porcentuales, en lugar de en unidades como la anterior. Es la media del error absoluto, o lo que es lo mismo, la diferencia entre la demanda real y el pronóstico.

$$MAPE = \frac{\sum_{i=1}^n 100 |Real_i - Pronóstico_i|}{n \cdot Real_i}$$

*Figura 5.2: Error porcentual medio absoluto*

Con estos datos, es posible determinar la precisión del modelo desarrollado.

# Capítulo 6

## Resultados y análisis

Para fines de esta investigación, se realizaron pruebas de entrenamiento con diferentes periodos del conjunto de datos, para determinar cual nos provee un mejor resultado.

### 6.1 Resultados

Conjunto de datos	Resultados	Observaciones
10 años	MSE: 91.6896 MAPE: 15%	Train: 9 años Test: 1 año
10 años	MSE: 83.9364 MAPE: 15%	Train: 9 años Test: 1 año Quitando valores atípicos
10 años	MSE: 90.88 MAPE: 16%	Train: 7 años Test: 3 años
10 años	MSE: 86.48 MAPE: 15%	Train: 4 años Test: 6 años
3 años	MSE: 71.62 MAPE: 12%	Train: 2 años Test: 1 año
10 años	MSE: 67.54 MAPE: 11%	Train: 9 años Test: 1 año Quitando valores atípicos Ajustando modelo

## 6.2 Análisis

Al concluir las pruebas, se observa que el uso de un gran conjunto de datos no es necesariamente mejor si no se conocen las técnicas adecuadas para entrenar el modelo. Los modelos que se entrenaron utilizando conjuntos de datos de diez años son significativamente menos precisos que el que utiliza sólo tres años. Al analizar los datos individuales del modelo entrenado con diez años (Figuras 6.1 y 6.2), se observa que los días en los que fue más preciso son menos consistentes.

date	
2014-12-17	0.627533
2014-07-25	0.643402
2014-07-27	0.851013
2014-12-23	0.884399
2014-10-14	1.357605
2014-11-27	1.671936
2014-11-30	1.862061
2014-04-24	1.869293
2014-07-20	1.939423
2014-12-22	2.345795

Figura 6.1: 10 años - Días más precisos

date	
2014-01-04	562.875916
2014-01-10	306.141418
2014-06-01	275.473572
2014-02-22	275.472534
2014-05-18	248.064484
2014-12-09	241.125183
2014-06-04	219.435669
2014-06-07	204.726898
2014-09-23	189.972931
2014-01-05	183.017303

Figura 6.2: 10 años - Días menos precisos

date	
2021-11-04	0.094513
2020-06-05	0.190979
2020-10-26	0.612091
2021-01-08	0.773529
2020-09-17	0.795715
2021-08-30	0.815674
2021-09-13	0.818878
2021-06-23	0.863770
2021-09-27	1.242554
2021-02-07	1.302460

Figura 6.3: 3 años - Días más precisos

date	
2020-11-01	446.832581
2021-11-01	441.717804
2021-05-06	273.356110
2021-02-06	258.165924
2020-05-18	243.454559
2021-06-01	240.268524
2020-08-26	239.095093
2020-05-31	237.787872
2020-06-01	237.512573
2021-06-15	236.669312

Figura 6.4: 3 años - Días menos precisos

Sin embargo, el mejor resultado obtenido en las pruebas fue utilizando un conjunto de datos de diez años al que se ajustaron tanto los datos como las métricas de entrenamiento, lo que dio lugar a un porcentaje de error del 11%. La diferencia entre estos dos resultados se debe al ruido de los datos generado a lo largo de los años, así como al uso del error cuadrático medio como métrica, ya que

tiende a penalizar los casos en los que el error fue mayor. Por tanto, utilizar un conjunto de datos de diez o más años nos da un mejor resultado siempre que se hagan los ajustes necesarios.

Asimismo, el porcentaje de error obtenido por el conjunto de datos de tres años también es viable, ya que requiere menos ajustes. Utilizando el programa "Minitab", un programa para realizar funciones estadísticas avanzadas, se hizo un pronóstico utilizando el conjunto de datos de tres años y obtuve los siguientes resultados:

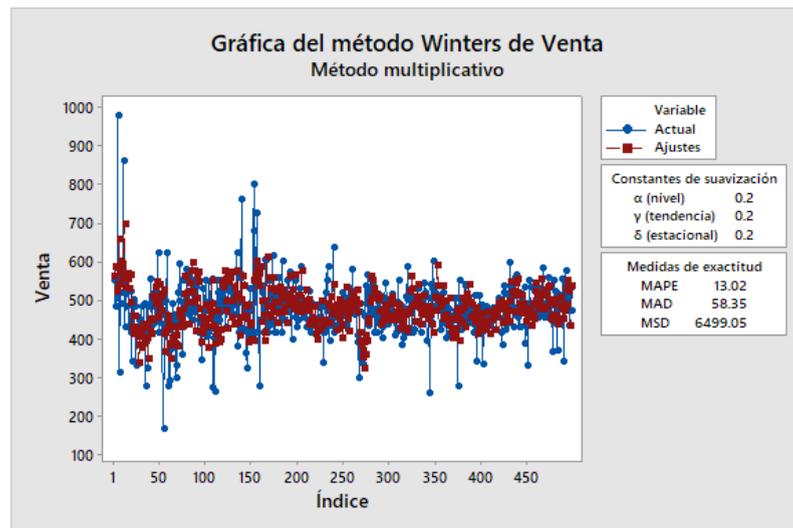


Figura 6.5: Pronóstico en Minitab

Como puede verse, el error porcentual medio absoluto (MAPE) es comparable a los resultados del modelo entrenado.

# Capítulo 7

## Conclusión y trabajo futuro

La investigación realizada junto con la solución desarrollada, afirman que los métodos de aprendizaje automático son precisos cuando se utilizan correctamente. Hay que tener en cuenta distintas variables para garantizar su eficacia en la amplia gama de aplicaciones posibles.

Dicho esto, también podemos afirmar que se ha alcanzado el objetivo de esta investigación. Se ha podido determinar el uso de métodos de aprendizaje automático para la toma de decisiones en pequeñas y medianas empresas en función de sus necesidades y tipo de negocio. Antes de utilizarlos, las empresas deben definir si el porcentaje de error obtenido es algo que puede absorberse sin que represente pérdidas significativas o un cambio radical en el funcionamiento.

### 7.2 Trabajo futuro

A pesar de los resultados obtenidos, esta investigación tiene varias áreas de oportunidad. Es posible continuar este trabajo con las siguientes mejoras del modelo:

1. Selección de características para considerar los días festivos o cualquier acontecimiento que pueda influir en la serie.
2. Experimentar con otros algoritmos, como ARIMA o SARIMA, en función del análisis de la naturaleza de los datos.
3. Utilizar diferentes conjuntos de datos con el mismo modelo. En esta investigación, se utilizó la misma base de datos durante todas las pruebas. Es posible obtener resultados más detallados si se utilizan conjuntos con comportamientos diferentes.
4. Por último, lo que considero más importante, utilizar los resultados del modelo para las operaciones cotidianas de una empresa, y evaluar así los impactos que puede tener en la toma de decisiones.

# Bibliografía

- [1] Morandín-Ahuerma, F. (2022). What is artificial intelligence? *International Journal of Research Publication and Reviews*, 03(12), 1947-1951. <https://doi.org/10.55248/gengpi.2022.31261>
- [2] ]Torres, C. Á. F., Pérez, V. H. C., & Saucedo, C. I. T. (2022). Análisis comparativo de modelos tradicionales y modernos para pronóstico de la demanda: enfoques y características. *RIDE revista iberoamericana para la investigación y el desarrollo educativo*, 12(24). <https://doi.org/10.23913/ride.v12i24.1203>
- [3] Leandro, A. A. (2017). Modelos de planeación estratégica en las empresas familiares (Models of Strategic Planning in the family business). *TEC empresarial*, 11(1), 23. <https://doi.org/10.18845/te.v11i1.3093>
- [4] Wilson, J., Koerber, D. (1992). Combining Subjective and Objective Forecasts Improves Results. *Journal of Business Forecasting*, 11(3), 12-16.
- [5] Paredes, A. P., de Los Ángeles, J. A. C., Villalobos, A. M. G., & Fonseca, V. J. (2018). Importancia de los pronósticos en la toma de decisiones en las MIPYMES. *Revista geon*, 5(1), 97-114. <https://doi.org/10.22579/23463910.17>
- [6] Vera, J. S. R., & Rosero, A. L. B. (2023). Análisis de datos profundo mediante herramienta de inteligencia artificial para la generación de un dashboard gerencial. *Revista Científica Sapientia Technological*, 4(1). <https://doi.org/10.58515/010rspt>
- [7] Baqae, D., & Farhi, E. (2020). Supply and demand in disaggregated Keynesian economies with an application to the COVID-19 crisis. *Social Science Research Network*. <https://doi.org/10.2139/ssrn.3595275>
- [8] Majama, N. S., & Magang, T. I. ' (2017). Strategic Planning in Small and Medium Enterprises (SMEs): A case study of Botswana SMES. *Journal of management and strategy*, 8(1), 74. <https://doi.org/10.5430/jms.v8n1p74>
- [9] Shaveta. (2023). A review on machine learning. *International Journal of Science and Research Archive*, 9(1), 281-285. <https://doi.org/10.30574/ijrsra.2023.9.1.0410>

- [10] Gilliland, M. (2020). The value added by machine learning approaches in forecasting. *International Journal of Forecasting*, 36(1), 161-166. <https://doi.org/10.1016/j.ijforecast.2019.04.016>
- [11] Pavlyshenko, B. M. (2019). Machine-Learning Models for Sales time Series Forecasting. *Data*, 4(1), 15. <https://doi.org/10.3390/data4010015>
- [12] Ptotic, M. P., Stojanović, M. B., & Popović, P. (2022). A review of Machine Learning Methods for Long-Term Time Series Prediction. *2022 57th International Scientific Conference on Information, Communication and Energy Systems and Technologies (ICEST)*. <https://doi.org/10.1109/icest55168.2022.9828618>
- [13] Pan, B. (2018). Application of XGBOOST algorithm in hourly PM2.5 concentration prediction. *IOP conference series*, 113, 012127. <https://doi.org/10.1088/1755-1315/113/1/012127>
- [14] Das, P. (2019). Time series forecasting. En *Springer eBooks* (pp. 439-453). [https://doi.org/10.1007/978-981-32-9019-8\\_14](https://doi.org/10.1007/978-981-32-9019-8_14)

# Anexos

## Repositorio

<https://github.com/EmBP19/Tesina>