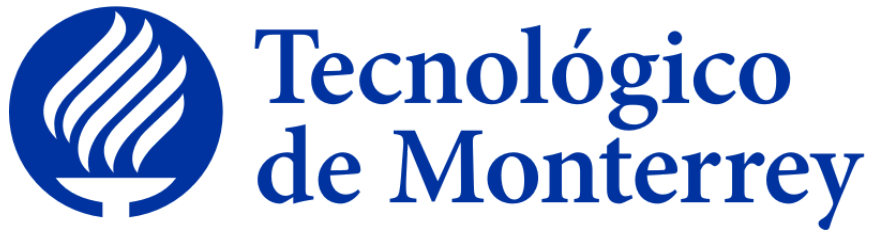


Instituto Tecnológico y de  
Estudios Superiores de Monterrey



**Análisis de Ventas de Don Colchón:**

**Modelos Predictivos con Aprendizaje Automático**

Escuela de Ingeniería y Ciencias

Análisis de Ciencia de Datos

TC 2004B.101

Equipo 2

Raúl Correa Ocañas	A01722401
Ricardo Kaleb Flores Alfonso	A01198716
Ricardo Marín Pérez	A01174384
Diego Armando Mijares Ledezma	A01722421
Daniel Ríos Zúñiga	A01174445

Docentes:

Prof. Rafael Martínez García Peña

Dr. Rasikh Tariq

16 de marzo de 2024

<b>1. Antecedentes.....</b>	<b>5</b>
1.1 Consumo Responsable y Producción.....	5
1.2 Panorama Global.....	5
1.3 Mal Manejo de Inventario.....	5
1.4 Relevancia de Analizar un Caso de Estudio.....	6
<b>2. Planificación.....</b>	<b>6</b>
2.1 Problema del Socio.....	6
2.2 Predicción de Ventas Erróneo.....	7
<b>3. Análisis de Requerimientos.....</b>	<b>7</b>
<b>4. Diseño.....</b>	<b>9</b>
4.1 Análisis Exploratorio Descriptivo.....	9
4.2 Declaración de Medidas de Éxito.....	11
4.2.1 Mean Absolute Percentage Error - MAPE.....	11
4.2.2 Mean Average Absolute Percentage Error - MAAPE.....	11
4.2.3 Mean Square Errors - MSE.....	12
4.2.4 Coeficiente de Determinación - R2.....	12
4.3 Marco Teórico.....	12
4.3.1 Exploración de la Literatura Científica.....	12
4.3.3 Teoría de Potenciación del Gradiente.....	14
4.3.4 Teoría de SARIMAX.....	17
<b>5. Desarrollo.....</b>	<b>19</b>
5.1 Librerías Principales.....	19
5.1.1 NumPy.....	19
5.1.2 Pandas.....	19
5.1.3 TQDM.....	19
5.1.4 Matplotlib.pyplot.....	20
5.1.5 Statsmodels (librería ACF).....	20
5.1.6 Sklearn preprocessing.....	20
5.1.7 Sklearn ensemble.....	21
5.1.8 Skforecast.....	21
5.1.9 Lightgbm.....	21
5.2 Validación Cruzada para Series de Tiempo.....	22
5.3 Implementación del Modelo 1: HistGradientBoost.....	23
5.4 Implementación del Modelo 2: SARIMAX.....	23
<b>6. Prueba.....</b>	<b>24</b>
6.1 Análisis de Resultados para Gradient Boosting.....	24
6.2 Análisis de Resultados para SARIMAX.....	25
6.3 Comparación de las Métricas.....	26
6.4 Indicadores Positivos.....	26

6.5 Indicadores Negativos.....	27
<b>7. Integración y Ejecución.....</b>	<b>28</b>
7.1 ¿Cómo son significativas las métricas?.....	28
7.2 Comparación con el Modelo del Socio Formador.....	28
7.3 Comparación del Desempeño de los Modelos Propuestos.....	28
7.4 ¿Cuándo usar cada modelo?.....	29
7.5 Posibles Mejoras Aplicables.....	29
<b>8. Conclusión.....</b>	<b>30</b>
8.1 Resumen.....	30
8.2 Hallazgos Principales.....	30
8.3 Implicación.....	31
8.4 Implementación.....	32
8.5 Ventaja competitiva.....	33
8.6 Implicación Global.....	33
<b>9. Apéndice.....</b>	<b>34</b>
<b>10. Referencias.....</b>	<b>34</b>

## **1. Antecedentes**

### **1.1 Consumo Responsable y Producción**

El progreso económico y social del último siglo ha provocado una degradación ambiental que pone en peligro a nuestro planeta, su desarrollo futuro, y la supervivencia del ser humano. Por ello, el Objetivo de Desarrollo Sustentable (ODS) 12, producción y consumo responsables, busca garantizar modalidades de consumo y producción sostenibles. El ODS 12 busca cambiar el modelo actual de producción y consumo, para lograr una gestión eficiente de recursos naturales, así eliminando grandes problemáticas provocadas por la producción y el superávit de materias primas. El uso ecológico de productos generados mundialmente de todo tipo, al igual que la disminución de los desechos que generan, lleva directamente al consumo responsable y producción que pretende taclear el ODS 12 (*Pacto Mundial*, 2023).

### **1.2 Panorama Global**

El desafío al que se encuentra el planeta es grave. Aunque uno pudiera intuitivamente vincular la sobreproducción, o el mal manejo de materiales e inventarios, a impactos negativos en el medio ambiente, es crucial resaltar que globalmente también están directamente relacionados con impactos negativos de magnitud considerable en la alimentación, sanidad, educación, paz y seguridad (*UNSD*, 2024).

### **1.3 Mal Manejo de Inventario**

El mal manejo del inventario de una empresa, es la ruta de todo problema que causa exceso o falta de producción. Desde tener que tirar alimentos de más, a gastar más dinero en espacios que contaminan con tal de guardar inventario, entre muchos otros ejemplos. Mientras que un buen manejo del inventario potencia el aprovechamiento de la materia primera y el flujo de las operaciones, al igual que reducción de gastos, y también auxilia en

cuidar al medio ambiente. Por otro lado, una mala administración provoca pérdidas de capital, e impacta directamente al planeta, generando materia prima de más, o en ocasión de generar menos, obligando a empresas a generar de manera ineficiente y polutante materias primas para complacer demandas inesperadas. El mal manejo de inventario frecuentemente se ve reflejado en el robo por empleados, ventas sin registrar, errores administrativos y falta de control, datos mínimos y no seguros, mala gestión de almacenamiento y un registro incorrecto de movimientos de la materia prima (Holguín, 2010).

#### 1.4 Relevancia de Analizar un Caso de Estudio

Es importante tomar un caso de estudio, para poder trabajar sobre un inventario que sufre de alguno de estos problemas, para poder analizar cómo se tacearía el problema de la ODS 12, en una escala realística y laboral. Por ello, se analizó un caso de estudio que se le presentó a los alumnos, con la finalidad de mejorar el manejo de inventario de un socio formador.

## **2. Planificación**

### 2.1 Problema del Socio

A través del Tecnológico de Monterrey, se presentó la oportunidad de colaborar con Don Colchón, una reconocida fabricante de colchones. El desafío del socio era aquello de un mal manejo de inventario, causando pérdidas económicas, y producción de cantidades erróneas de colchones. El socio comentó que era un problema grave, ya que en ocasiones no contaban con el espacio para guardar todo su inventario, y la empresa recorría a guardar inventario donde hubiera espacio, mientras que en otras ocasiones, producían de menos, provocando pérdida de posibles ganancias, y esfuerzos mayores a la hora de hacer ventas, por una mala anticipación de venta. Ambas instancias, la sobre preparación y la sub preparación

de inventario, no solo causan impacto negativo al socio, si no también contribuye a la separación del progreso que se ha hecho hacia el ODS 12 globalmente.

## 2.2 Predicción de Ventas Erróneo

Se analizó el problema del socio, quien introdujo el hecho de que la empresa contaba con un modelo de predicción de ventas semanal que presentaba un error del 60%. Esto evidentemente es lo que ocasiona un manejo ineficiente del inventario. Por ello, para solucionar el problema del socio formador y atender a la ODS 12, se propuso la siguiente pregunta. ‘¿Cómo se pueden predecir las ventas del socio?’, lo que se enlaza directamente con los contenidos de este reporte, donde se explora cómo se utilizaron varios modelos predictivos con técnicas de aprendizaje automático, para intentar reducir el error de predicción de ventas del socio.

## 3. Análisis de Requerimientos

### 3.1 Dataset

El Socio Formador *Don Colchón* compartió el dataset con los registros de ventas en las distintas sucursales presentes en la república; las variables incluidas dentro del dataset son: “Índice”, “Fecha”, “Producto”, “Total \$”, “Cantidad”, “Total libre de impuestos” y “Cantidad ordenada”, siendo los “Total \$” y “Cantidad” repetidos para las 10 sucursales donde se pueden adquirir los productos (incluyendo el servicio online). Estas variables incluyen un índice único para identificar cada registro de compra, la fecha en la que se llevó a cabo la transacción, así como del producto del cual se realizó la compra y la cantidad, y el valor total del estado donde se compró (todas las demás columnas de estados se marcan con 0).

Al observar los datos, se pueden observar las ventas de productos distintivos, así como la sucursal donde se obtuvo la mayor cantidad de ventas. Después de este primer análisis, se observó que los productos con mayor demanda son el 273, 0, 1 y 5. Además, la ciudad de Monterrey representa la mayoría de las ventas realizadas. Estos resultados llevarán a que se realicen modelos enfocados principalmente en estos productos, ya que al obtener la mayor cantidad de ventas, será posible predecir con mejor precisión (de igual manera los modelos se realizarán para cada producto).

### 3.2 Preprocesamiento y limpia de datos

El dataset original no contaba con el formato adecuado para poder realizar modelos de aprendizaje automático, es por eso que se realizó un procesamiento de los datos para aplicar los modelos deseados. Primeramente, se descartó de la base de datos las columnas de “Total libre de impuestos” y “Cantidad ordenada”, puesto que para los modelos que se realizaron no fueron de importancia. Incluirlos hubiera generado ruido o afectado el rendimiento de los mismos. Una de las variables que se transformó fue la variable “Fecha”. Estaba en un formato en español, no compatible con la librería de pandas, por lo que se sustituyó con un formato en inglés. Finalmente, se crearon nuevas variables, sustituyendo las 20 columnas que especificaban la región, por una única, obteniendo un único data frame con las variables “Producto”, “Lugar”, “Precio unitario”, “Cantidad” y “Fecha”.

### 3.3 Limitaciones

Una de las principales limitaciones que existieron en la base de datos proporcionada por el Socio Formador, fueron la cantidad de datos que existían por producto. Siendo 273 productos, muy pocos contaban con ventas significativas, lo que producía que el modelado no pudiera ser considerado para una gran cantidad de productos. Una de las soluciones

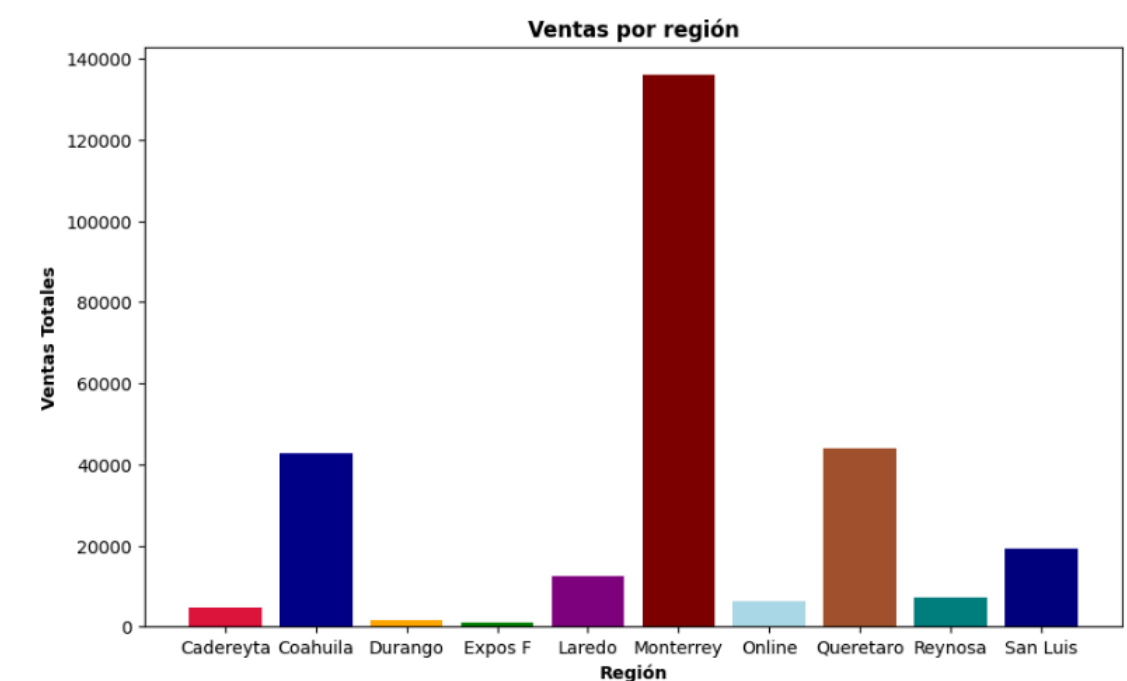
propuestas fue el remover del dataset los productos que tuvieran ventas mínimas, reduciendo considerablemente los datos para poder enfocarse en los productos de interés.

Después de haber realizado la limpieza de datos, se llevó a cabo el diseño de propuesta de implementación de modelos, y después dicha implementación.

## 4. Diseño

### 4.1 Análisis Exploratorio Descriptivo

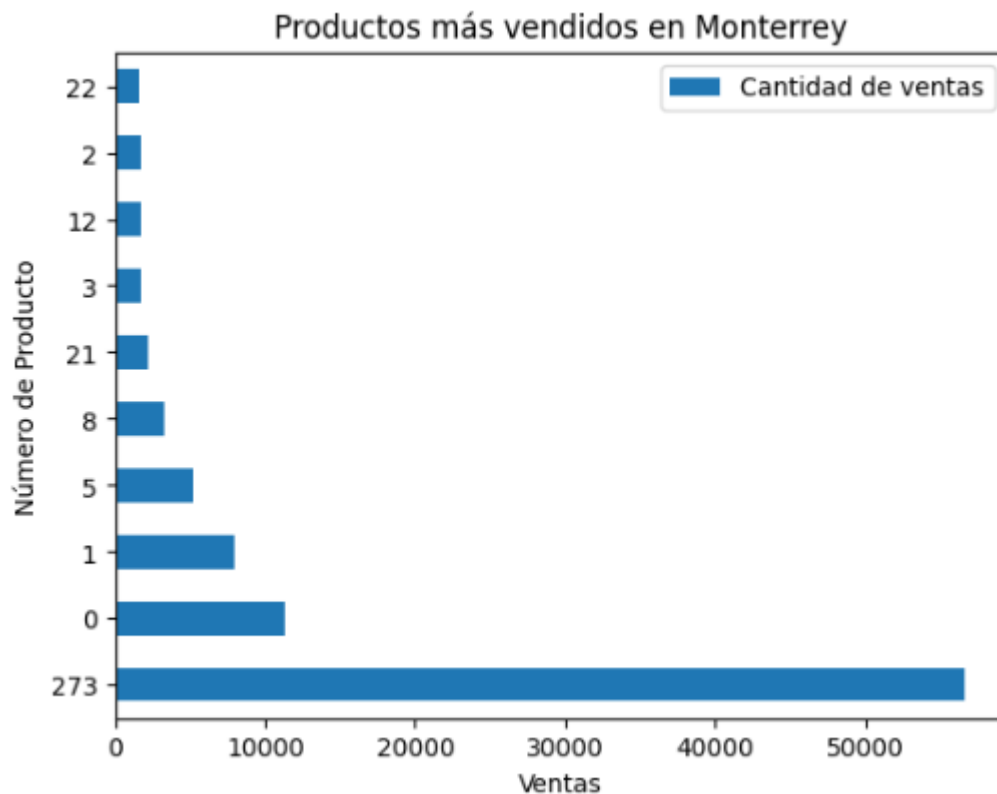
Para iniciar el análisis exploratorio se consideró fundamental conocer la región que más productos vende de cualquier tipo, ya que si una región vende pocos productos será más difícil generar predicciones acertadas acerca de las ventas futuras.



En la gráfica se puede observar que la región que más productos vende, con una gran diferencia, es Monterrey. Por otro lado se observa que regiones como Durango, Expos F, Cadereyta y Online, venden muy poca cantidad comparadas con las otras regiones. Esto representa un problema, ya que estas regiones con pocos datos no son lo suficientemente significativas como para realizar modelos de predicción en estas regiones. Esto llega a



implicar dos opciones: Realizar las predicciones por producto a nivel nacional, o despreciar estas regiones mencionadas y aplicar los modelos únicamente a las otras regiones. Para elegir la mejor opción de entre estas dos, se analizará la cantidad de productos vendidos en Monterrey, que es la región que cuenta con la mayor cantidad de ventas.



Se graficaron únicamente los 10 productos más vendidos en Monterrey, y se puede notar que el producto 273 es el más vendido superando en gran medida al resto. Esta diferencia tan grande en la cantidad de ventas entre los productos resultará en modelos menos exactos para los productos con menos ventas. Por este motivo se considera que la mejor opción es realizar predicciones por producto a nivel nacional, ya que de esta manera, al juntar las ventas de todas las regiones, se tendrá la cantidad de datos necesarios para crear y entrenar modelos capaces de hacer los pronósticos adecuados de las ventas por producto.

## 4.2 Declaración de Medidas de Éxito

Para observar el desempeño de los modelos realizados, se utilizaron distintas métricas estadísticas, relacionadas con las que el modelo del Socio Formador trabaja, esto para observar con más detalle la diferencia entre los modelos realizados y el modelo ya existente, y de esta manera comprobar si los nuevos modelos predicen de mejor manera.

### 4.2.1 Mean Absolute Percentage Error - MAPE

MAPE fue la métrica de mayor importancia para el análisis, debido a que el Socio Formador utiliza esta métrica como su principal herramienta de análisis. Esta calcula el error porcentual promedio entre las predicciones y los valores reales.

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{Y_i - \hat{Y}_i}{Y_i} \right| \cdot 100$$

MAPE representa el porcentaje de error del valor real, lo que hace muy útil para comparar distintos modelos.

### 4.2.2 Mean Average Absolute Percentage Error - MAAPE

Similar a MAPE, esta métrica recopila los distintos valores obtenidos en MAPE para sacar un promedio de los mismos, óptimo para cuando se recopila varios productos y se necesita obtener un promedio de la eficiencia de los modelos. A diferencia de la formula de MAPE, se hace el cálculo del MAPE para  $k$  series de tiempo, y se prosigue a obtener el promedio de estas.

$$MAAPE = \frac{1}{k} \sum_{j=1}^k \frac{1}{n} \sum_{i=1}^n \left| \frac{Y_{ij} - \hat{Y}_{ij}}{Y_{ij}} \right| \cdot 100$$

#### 4.2.3 Mean Square Errors - MSE

Busca calcular el promedio de los cuadrados de las diferencias entre los valores predichos y los valores reales. Al cuantificar esta discrepancia también sirve para determinar qué modelo tiene un mejor ajuste a los valores reales. No obstante, cabe mencionar que tiene una alta sensibilidad a los valores atípicos.

$$MSE = \frac{1}{n} \sum_{i=1}^n (Y_i - \gamma_i)^2$$

#### 4.2.4 Coeficiente de Determinación - $R^2$

Esta métrica estadística describe el ajuste que tiene el modelo con respecto a los datos reales. Un valor cercano a 1 describe que el modelo explica correctamente gran parte de la variabilidad de la variable dependiente. Sin embargo,  $R^2$  no determina la calidad general de un modelo, sino que únicamente proporciona información sobre la proporción de variabilidad explicada.

$$R^2 = 1 - \frac{SSE}{SST}$$

### 4.3 Marco Teórico

#### 4.3.1 Exploración de la Literatura Científica

Al buscar casos similares como los que ha presentado el socio formador, se encontró la investigación dirigida a la comparación de métodos estadísticos y de aprendizaje automático para la predicción de demanda visto por SKU. En el documento, *Comparison of statistical and machine learning methods for daily SKU demand forecasting*, se propone una metodología en la que se estudia la clasificación de series de tiempo de demanda para proponer mejores modelos de predicción.

El equipo de investigadores de E. Spiliotis indica que existen métricas que ayudan a evaluar la naturaleza de una serie de tiempo. El coeficiente de variación cuadrado es una

medida de proporción de cuántas veces mayor es la desviación estándar a comparación de la media. Esta se utiliza como una medición de variabilidad de los datos. Por otra parte, el ADI (Average Inter-Demand Interval) mide el tiempo promedio entre ventas. Según como se clasifique cada serie de tiempo en estas dos métricas, se pueden categorizar en cuatro tipos de demanda: *Intermitente*, *Abultado*, *Liso* y *Errático*. Debido a que se tiene un bajo ADI y alto coeficiente de variación para series de tiempo erráticas, predecir demanda atípica es complicado. Por otra parte, productos como los *Intermitentes* y *Abultados* tienen un ADI alto, implicando que su naturaleza de predicción es complicada en cuanto a determinar cuándo sucederá la venta. Estos tipos de series de tiempo se alinean al caso presentado por el socio formador, y se espera que los datos correspondan a las categorías mencionadas (Spiliotis, E., 2022).

Tomando en cuenta esta clasificación, la investigación propone que para series de tiempo que se alinean a las categorías de *Intermitente* y *Abultado*, el mejor método en ambos resulta ser el modelo de bosques aleatorios, seguido del método de árboles potenciados por gradiente. Sin embargo, este segundo método es más efectivo para la totalidad de las series de tiempo, por lo que también puede ser considerado como una opción relevante.

Aunque la investigación propone que los métodos de aprendizaje automático suelen ser más eficaces contra modelos estadísticos, métodos variantes de ARIMA no fueron considerados. Tomando en cuenta esto, surgió el interés para comparar y evaluar cómo funcionan contra los previamente mencionados.

Adicionalmente, se consideraron modelos de aprendizaje profundo para evaluar su rendimiento, los cuales no fueron implementados debido a: Las conclusiones realizadas según la investigación, y la falta de tiempo para presentar una solución más amplia considerando el aprendizaje profundo como una opción.

### 4.3.2 Decisión de Modelos Predictivos

Después de tomar en cuenta la información necesaria para evaluar los modelos por probar, se decidió implementar SARIMAX y Gradient Boosting Regressor, un caso específico de Gradient Boosting Tree dirigido a la predicción de valores continuos. En el caso de Gradient Boosting Regressor, este funciona como un método base que se va a utilizar para comparar la alternativa, siendo SARIMAX. Una vez escogidos los modelos por evaluar, es crucial entender su funcionamiento para evaluar cómo son distintas las formas en la que se realizan las predicciones de cada modelo.

### 4.3.3 Teoría de Potenciación del Gradiente

El algoritmo comienza identificando una métrica como función de pérdida. El modelo tradicional utiliza la métrica de MSE, por lo que la función es equivalente a la media del cuadrado de las diferencias.

$$L(y_i, \gamma) = (y_i - \gamma)^2 \quad (1.1)$$

Como predicción inicial, se encuentra el valor de gamma que minimice la función.

$$F_0(x) = \arg \min_{\gamma} \sum_{i=1}^n L(y_i, \gamma) \quad (1.2)$$

Para encontrar el valor mínimo de  $L(y_i - \gamma)$ , se busca su punto crítico.

$$\frac{\partial}{\partial \gamma} \sum_{i=1}^n L(y_i, \gamma) = \frac{\partial}{\partial \gamma} \sum_{i=1}^n (y_i - \gamma)^2 \quad (1.3)$$

$$\frac{\partial}{\partial \gamma} \sum_{i=1}^n L(y_i, \gamma) = -2 \sum_{i=1}^n (y_i - \gamma) \quad (1.4)$$

$$\frac{\partial}{\partial \gamma} \sum_{i=1}^n L(y_i, \gamma) = 2n\gamma - 2 \sum_{i=1}^n y_i \quad (1.5)$$

$$0 = 2n\gamma - 2 \sum_{i=1}^n y_i \implies n\gamma = \sum_{i=1}^n y_i \quad (1.6)$$

$$\gamma = \frac{1}{n} \sum_{i=1}^n y_i = \bar{y} \quad (1.7)$$

$$\therefore F_0(x) = \bar{y} \quad (1.8)$$

De esto se infiere que la primera aproximación es equivalente a la media de la variable predecida. Existen modelos en el que este primer paso se inicializa como un valor aleatorio, el cual después es optimizado iterativamente. Se denomina a  $m \in M$ ,  $M$  siendo el conjunto de árboles generados y  $m$  el índice correspondiente a cada árbol. Para cada árbol  $m$ , se calcula el residuo entre la función predictora y el valor real. Si  $r_{im}$  corresponde al residuo de la predicción  $i$  del árbol  $m$ , entonces el residuo es:

$$r_{im} = \frac{\partial L(y_i, F_{m-1}(x))}{\partial F_{m-1}(x)} = \frac{\partial (y_i - F_{m-1}(x))^2}{\partial F_{m-1}(x)} = -2(y_i - F_{m-1}(x)) \quad (1.9)$$

$$r_{im} := y_i - F_{m-1}(x) \quad (1.10)$$

La constante multiplicando la diferencia entre el valor real y la función predictora puede ser descartada.

Se entrena un árbol de decisión  $m$  tomando una matriz  $x$  de características para ajustarse al gradiente de residuos. Se denota  $j \in J_m$ ,  $j$  siendo el índice de cada hojas de predicción generada por el árbol, y  $J_m$  el conjunto de todas las hojas de predicción. A continuación, se continúa con averiguar un valor de gamma para cada hoja que minimice la función de pérdida aplicada al valor real y la función predictora.

$$\gamma_{jm} = \arg \min_{\gamma} \sum_{x_i \in R_{jm}}^n L(y_i, F_{m-1}(x) + \gamma) \quad (1.11)$$

$$\frac{\partial}{\partial \gamma} \sum_{x_i \in R_{jm}}^n (y_i - F_{m-1}(x) - \gamma)^2 = -2 \sum_{x_i \in R_{jm}}^n (y_i - F_{m-1}(x) - \gamma) \quad (1.12)$$

$$\frac{\partial}{\partial \gamma} \sum_{x_i \in R_{jm}}^n (y_i - F_{m-1}(x) - \gamma)^2 = 2n\gamma - 2 \sum_{x_i \in R_{jm}}^n (y_i - F_{m-1}(x)) \quad (1.13)$$

$$\gamma = \frac{1}{n} \sum_{x_i \in R_{jm}}^n (y_i - F_{m-1}(x)) \quad (1.14)$$

$$\therefore \gamma_{jm} = \frac{1}{n} \sum_{x_i \in R_{jm}}^n r_{im} \quad (1.15)$$

De esta forma, se ha encontrado una gamma para cada división hecha en el nuevo árbol, por lo que la función predictora  $F_m$  tiene un valor de gamma para diferentes segmentos de  $x$ . Se concluye que la función predictora tendrá la forma:

$$F_m(x) = F_{m-1}(x) + v \sum_{j=1}^{J_m} \gamma_{jm} \quad (1.16)$$

Donde la suma de cada gamma corresponde a la suma de las divisiones de cada nodo, y la constante  $v$  es la tasa de aprendizaje. Con esto se tiene la primera iteración de potenciación de gradiente, la cual continúa hasta iterar sobre todos los árboles (*Towards Data Science*, 2022).

En el caso específico del modelo, se hace una optimización para identificar los mejores puntos de decisión. Esta consiste en tomar cada variable y tratarla como un histograma, lo cual disminuye los valores de decisión por revisar.

#### 4.3.4 Teoría de SARIMAX

SARIMAX, como es descrito anteriormente, consiste en un modelo autorregresivo que toma la media móvil, variables exógenas y periodicidad como predictores. Para plantear la construcción matemática del modelo, se comienza describiendo la ecuación de un modelo autorregresivo.

$$y_t = \beta + \varepsilon_t + \sum_{i=1}^p \theta_i y_{t-i} \quad (2.1)$$

Se toman los últimos  $p$  registros previos al tiempo de predicción de la serie de tiempo  $y$  como predictores para  $y_t$ . Cada uno de estos registros se asocia a un peso obtenido al tratar la predicción como una minimización de cuadrados (OLS). Esta regresión tiene una constante asociada,  $\beta$  conocida como el sesgo del modelo. Finalmente, se cuenta con un ruido el cual no puede ser explicado con la regresión.

$$y_t = \beta + \varepsilon_t + \sum_{i=1}^q \phi_i \varepsilon_{t-i} \quad (2.2)$$

Después se considera un modelo de media móvil, el cual en vez de hacer una regresión basada en los  $p$  valores pasados, se toman los errores del valor con la media en los últimos  $q$  registros como predictores. Con este modelo, el ruido entre predicción y realidad es tomado como un factor el cual aporta a la predicción de  $y_t$ .

$$y_t := y_t - y_{t-1} \quad (2.3)$$

Adicionalmente, se busca que los supuestos estadísticos en series de tiempo sean cumplidos para respaldar adecuadamente las predicciones hechas. Uno de estos supuestos es la estacionariedad, la cual asume que la tendencia de la serie de tiempo es constante. Para poder trabajar con una combinación del modelo autoregresivo y el modelo de media móvil, es necesario aplicar un componente de diferenciación. Este se repite sobre la serie original hasta



que se cumpla el supuesto, por lo que se denomina al número de veces necesario para cumplir esto como  $d$ .

$$y_t = \beta + \varepsilon_t + \sum_{i=1}^P \Theta_i y_{t-is} \quad (2.4)$$

$$y_t = \beta + \varepsilon_t + \sum_{i=1}^Q \Phi_i \varepsilon_{t-is} \quad (2.5)$$

Al tomar en cuenta la posibilidad de periodicidad en los datos, se toman las ecuaciones 2.1 y 2.2 y se ajustan a tomar los últimos  $s$  ciclos.

$$y_t = \beta + \varepsilon_t + X_t \Omega \quad (2.6)$$

Por último, se considera agregar una regresión con variables exógenas a la serie de tiempo. Esta se representa como un vector  $X$  con sus pesos asociados en el vector transpuesto omega. Los pesos de esta regresión, así como todas las utilizadas para este modelo, se encuentran como soluciones de problemas OLS.

Al implementar cada una de estas regresiones para predecir  $y_t$ , se tiene un modelo autorregresivo con media móvil, variables exógenas, periodicidad y supuesto de estacionariedad verificado. Sumando las ecuaciones 2.1, 2.2, 2.4, 2.5 y 2.6, se obtiene la ecuación de regresión:

$$y_t = \beta + \varepsilon_t + \sum_{i=1}^p \theta_i y_{t-1} + \sum_{i=1}^P \theta_i y_{t-1} + \sum_{i=1}^q \phi_i \varepsilon_{t-1} + \sum_{i=1}^Q \phi_i \varepsilon_{t-1} + X_t \Omega \quad (2.6)$$

## 5. Desarrollo

### 5.1 Librerías Principales

#### 5.1.1 NumPy

NumPy es una biblioteca de Python que proporciona un objeto de matriz multidimensional de un alto rendimiento, al igual que herramientas para trabajar con dichas matrices. Es fundamental para la computación científica. Más allá de sus usos obvios, también se utiliza como contenedor multidimensional eficiente de datos genéricos (*GeeksforGeeks*, 2023). En los modelos de predicción de ventas, se utiliza para realizar una gran variedad de operaciones matemáticas, para calcular predicciones y el error de las mismas (Li, 2021).

#### 5.1.2 Pandas

Pandas es una biblioteca de Python que proporciona estructura de datos y funciones que permiten manipular y analizar datos de manera eficiente. También es una herramienta fundamental, en esta ocasión para ciencias de datos y en general proyectos que manejen datos estructurados en Python (*GeeksforGeeks*, 2023). Se construyó sobre la biblioteca NumPy, por ello muchas veces son dependientes una de la otra (McIntire et al., 2023). En los modelos de predicción de ventas, se utiliza para limpiar, transformar y analizar los datos (Bisman Preet Singh, 2020).

#### 5.1.3 TQDM

TQDM es una biblioteca de Python que ofrece herramientas muy útiles en cuanto a la visualización de progreso para ciclos, bucles o procesos largos en general, mediante una barra de progreso y estimaciones de tiempo durante la ejecución de estos procesos. Una barra de

progreso es un elemento de control gráfico utilizado para visualizar el progreso de una operación computacional extensa, como una descarga, transferencia de archivo o instalación (*GeeksforGeeks*, 2020). En cuestión de predicción de ventas es útil para ver el progreso del entrenamiento y la predicción de los modelos, ya que son los procesos que más tiempo y recursos computacionales emplean.

#### 5.1.4 Matplotlib.pyplot

Matplotlib.pyplot es una interfaz basada en estados para Matplotlib. Proporciona una manera implícita de trazar gráficos. Así mismo, abre figuras en pantalla y administra la interfaz gráfica de las figuras generadas por el usuario (*Matplotlib*, 2023). En la predicción de ventas, es útil para visualizar los datos y resultados del modelo.

#### 5.1.5 Statsmodels (librería ACF)

Este módulo de Python proporciona clases y funciones para la estimación de muchos modelos estadísticos diferentes, con el fin de realizar pruebas estadísticas y exploración de datos estadísticos (*Statsmodels*, 2023). Para el análisis de series temporales, existe la función de autocorrelación (ACF). Esta mide cómo los valores de una serie temporal se correlacionan con sus propios valores pasados (*Medium*, 2020). Para la predicción de ventas, permite identificar patrones y tendencias en los datos de ventas con función al tiempo (*Analytics Vidhya*, 2020).

#### 5.1.6 Sklearn preprocessing

La librería Sklearn.preprocessing de Python proporciona varias funciones de utilidad comunes y funciones para transformar los datos brutos, en formatos que sean más adecuados para los modelos de aprendizaje automático. En el ámbito de predicción de ventas, esta

librería es útil debido a las funciones que incorpora para transformar los datos a una distribución normal (*Scikit-learn*, 2023).

#### 5.1.7 Sklearn ensemble

Sklearn ensemble es una biblioteca de Python que combina las predicciones de varios estimadores base construidos con un algoritmo de aprendizaje dado, con el fin de mejorar la generalización y resistencia sobre un solo estimador (*Scikit-learn*, 2023). Para la predicción de ventas, permite la implementación de Histogram Gradient Boosting Regressor (HGB), que es más veloz que Gradient Boosting Regressor (*Scikit-learn*, 2023).

#### 5.1.8 Skforecast

Por otro lado, Skforecast también es una biblioteca de Python. facilita el uso de los regresores de scikit-learn, como predictores de un solo paso y de varios. En Skforecast, se puede implementar Seasonal Auto-Regressive Integrated Moving Average with eXogenous factors (SARIMAX). Este, como ya se cubrió, es un modelo diseñado para capturar y predecir patrones subyacentes, tendencias y estacionalidad en datos de series temporales (*GeeksforGeeks*, 2023). También, cabe mencionar que Skforecast proporciona herramientas para la búsqueda de hiper parámetros, facilitando la optimización de modelos de predicción de ventas (*Medium*, 2023).

#### 5.1.9 Lightgbm

Lightgbm es un módulo de Python que implementa algoritmos basados en árboles de decisión para aumentar la eficiencia de modelos de aprendizaje automático y reducir el uso de memoria (*GeeksforGeeks*, 2023). Para problemas de predicción de ventas esta librería es de utilidad para optimizar los procesos de entrenamiento de modelos y búsqueda de hiper

parámetros, entre otros, lo que reduce los recursos computacionales necesarios para aplicar estos modelos.

## 5.2 Validación Cruzada para Series de Tiempo

Teniendo la base de datos limpia y lista para comenzar con la implementación de modelos de aprendizaje automático, el siguiente paso consiste en separar los datos de tal forma en la que se pudieran entrenar, validar y probar los modelos. Teniendo aproximadamente dos años de datos, se decidió utilizar los primeros 22 meses como conjunto de entrenamiento, dejando un mes de validación de datos y otro mes de prueba.

Conjunto	Fecha Inicio	Fecha Fin	# Datos
Entrenamiento	2022-01-02	2023-10-30	667
Validación	2023-10-30	2023-11-30	32
Prueba	2023-11-30	2023-12-31	32

Tabla de fechas delimitadoras entre conjuntos.

La función de un conjunto de entrenamiento consiste en entrenar un modelo con algunos parámetros predeterminados. Utilizando el conjunto de validación, se hace una búsqueda de hiper parámetros que maximicen su rendimiento en una métrica. Finalmente, una vez que se tenga seguro los hiper parámetros seleccionados, se toman los datos de entrenamiento y validación como uno solo para evaluar el rendimiento del modelo en datos nunca vistos (conjunto de prueba). Este conjunto de datos también es utilizado una vez entrenado todos los modelos, de tal forma que se pueda utilizar como una muestra aleatoria en la que los resultados obtenidos sean comparables.

### 5.3 Implementación del Modelo 1: HistGradientBoost

#### 5.3.1 Búsqueda de Hiper Parámetros

Para el modelo de HistGradientBoost (HGB) se realizó un Grid Search como método de búsqueda de los mejores hiper parámetros. Este método se basa en la creación, entrenamiento y evaluación del modelo, con todas las combinaciones posibles de hiper parámetros de un espacio predefinido de estos. Para este modelo, los hiper parámetros que se deseaban encontrar eran ‘max iter’, ‘max depth’ y ‘learning rate’, por lo que se le asignaron 3 valores a cada uno de estos y se aplicó el Grid Search. Una vez entrenados y evaluados todos los modelos, el método despliega los valores de los hiper parámetros del modelo que tuvo un mejor desempeño en el conjunto de validación.

#### 5.3.2 Evaluación en Conjunto de Prueba

Una vez que se hayan encontrado los mejores hiper parámetros, se prosigue a crear una lista en la que se guardarán las configuraciones hechas al ajustar las series de tiempo de cada producto en el modelo. Aunque el rendimiento del modelo será principalmente evaluado semanalmente, se hacen predicciones para dentro de una semana y para dentro de un mes. Tomando las métricas que se definieron como las principales para evaluar el rendimiento del modelo, se decide calcularlas ordenadas según las ventas correspondientes a cada producto.

### 5.4 Implementación del Modelo 2: SARIMAX

#### 5.4.1 Búsqueda de Hiper Parámetros

Para encontrar los mejores hiper parámetros se utilizó el algoritmo de AutoSARIMAX, que a su vez emplea el algoritmo AutoARIMA para seleccionar automáticamente el modelo ARIMA base que mejor se ajusta a la serie temporal sin tener en cuenta la estacionariedad ni las variables exógenas. AutoARIMA utiliza técnicas como la

búsqueda en cuadrícula (Grid Search) para encontrar el modelo ARIMA óptimo para el conjunto de datos proporcionado.

#### 5.4.2 Evaluación en Conjunto de Prueba

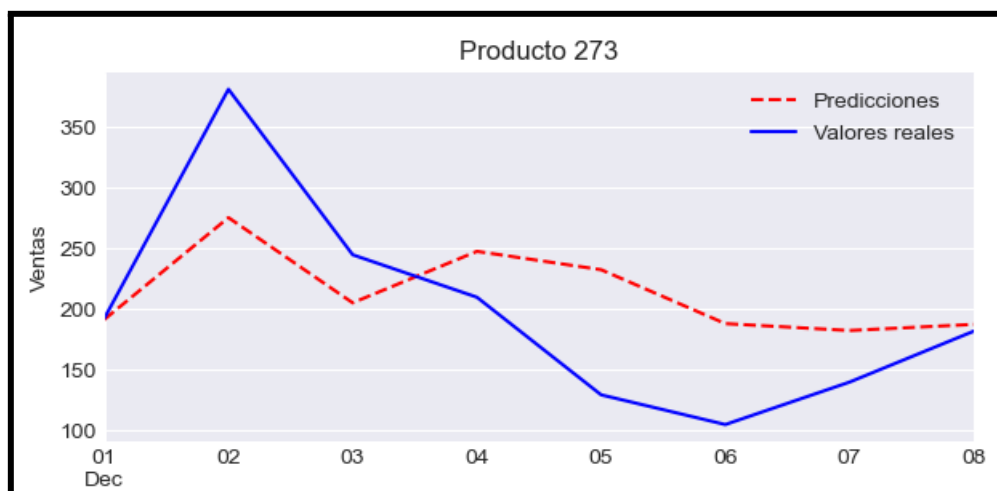
Nuevamente, se utiliza el mismo código utilizado para evaluar el conjunto de prueba con Histogram Gradient Boost para evaluar el modelo SARIMAX. Este paso es esencial para garantizar que las condiciones para comparar ambos modelos sean lo más similares posible.

Al igual que en HGB, se evalúan las métricas de MAPE, MAAPE, MSE,  $R^2$ , entre otras. Una vez obtenidos los resultados para ambos modelos, es posible realizar comparaciones y redactar conclusiones de la efectividad de los modelos para predecir los datos.

### 6. Prueba

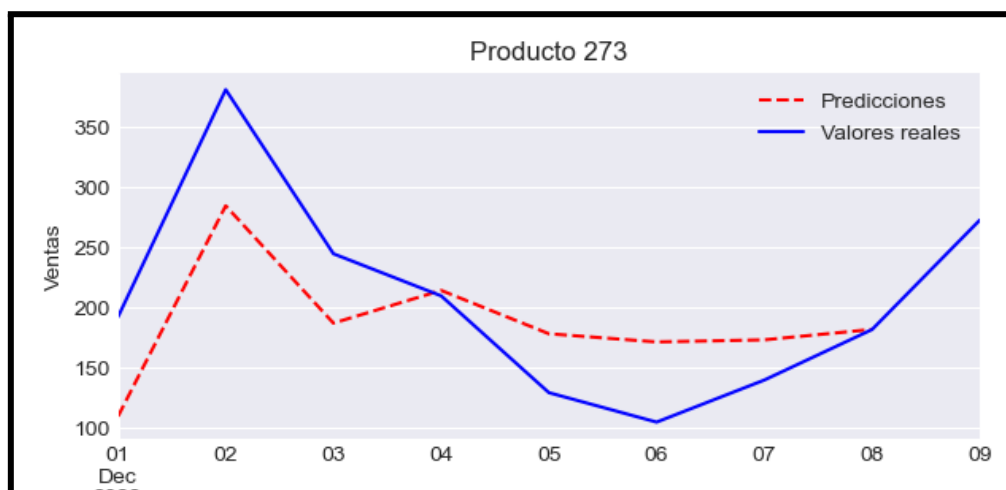
Para analizar gráficamente el desempeño de ambos modelos se tomó la gráfica de predicción contra el valor real del producto 273, ya que este fue el producto con más ventas en 2023. A su vez se seleccionó la predicción semanal sobre la mensual, ya que esta puede ser más útil para el socio formador.

#### 6.1 Análisis de Resultados para Gradient Boosting



En la gráfica anterior se observa de rojo la línea correspondiente a las predicciones realizadas por el modelo, y de color azul las ventas reales en esa semana. Se puede notar que el modelo identificó correctamente el aumento de ventas en el segundo día y el decrecimiento para el tercer día. Sin embargo, tuvo ciertas complicaciones para las predicciones de los días siguientes, ya que el modelo predijo ventas un poco por encima de las reales. Por último, en el día ocho, las líneas de la predicción y la venta real se unen, por lo que se puede decir que el modelo de regresión es bueno, aunque podría mejorar en ciertos aspectos.

## 6.2 Análisis de Resultados para SARIMAX



Nuevamente se analiza la gráfica de las predicciones contra el valor real. Es posible notar que este modelo también fue capaz de predecir la subida de ventas para el segundo día, y se acercó más al valor real, por lo que se nota más clara la tendencia en forma de pico para ese día. De igual manera tuvo una predicción muy precisa para el día 4, y a pesar de que también predijo más ventas para los siguientes días, su pronóstico tuvo un porcentaje de error más bajo para este producto en específico.



### 6.3 Comparación de las Métricas

Para hacer una correcta comparación de los resultados de ambos modelos, se considerarán las métricas de error descritas anteriormente (MAAPE, MSE y Coeficiente de Determinación) para los 15, 10 y 5 productos más vendidos, usando las predicciones mensuales y semanales hechas por cada modelo.

Tabla de métricas de los 15 productos más vendidos:

<i>Modelo</i>	<i>Tiempo de Predicción</i>	<i>MAAPE</i>	<i>R<sup>2</sup></i>	<i>MSE</i>
SARIMAX	Mensual	30.65%	-0.67	1321.99
SARIMAX	Semanal	50.24%	-0.7	464.19
HGB	Mensual	20.42%	-0.09	1118.93
HGB	Semanal	31.81%	-0.2	502.17

Tabla de las métricas de los 10 productos más vendidos:

<i>Modelo</i>	<i>Tiempo de Predicción</i>	<i>MAAPE</i>	<i>R<sup>2</sup></i>	<i>MSE</i>
SARIMAX	Mensual	18.46%	-0.18	1966.13
SARIMAX	Semanal	22.03%	-0.12	680.13
HGB	Mensual	9.09%	-0.02	1668.18
HGB	Semanal	14.61%	-0.04	743.23

Tabla de las métricas de los 5 productos más vendidos:

<i>Modelo</i>	<i>Tiempo de Predicción</i>	<i>MAAPE</i>	<i>R<sup>2</sup></i>	<i>MSE</i>
SARIMAX	Mensual	9.71%	-0.14	3842.39
SARIMAX	Semanal	8.19%	0.04	1303.11
HGB	Mensual	4.33%	0	72.07
HGB	Semanal	6.26%	0.06	852.89

### 6.4 Indicadores Positivos

Se puede notar que el indicador MAAPE, que es el que el socio formador Don Colchón utiliza para evaluar sus propios modelos de predicción, tuvo valores realmente bajos, indicando poco porcentaje de error. Los porcentajes de error más bajos se presentan para los 5 productos más vendidos donde el error más grande fue de 9.71%. Mientras más productos se

toman, es notable que el porcentaje de error también aumenta. Para los 10 productos más vendidos alcanza un valor de 22.03% para SARIMAX y 14.61% para Histogram Gradient Boosting, lo cuál también es un porcentaje bajo.

Para los 15 modelos más vendidos los porcentajes de error aumentan considerablemente, sin embargo, esto se debe a que a pesar de ser los 15 productos más vendidos, algunos de estos cuentan con muy pocas ventas, por lo que resulta realmente difícil realizar una predicción exacta de las ventas futuras.

## 6.5 Indicadores Negativos

En las tablas se puede observar que las predicciones de los 15 y 10 productos más vendidos cuentan con coeficientes de determinación ( $R^2$ ) negativos, lo que es algo inusual en los problemas de regresión. Esta métrica indica la variabilidad de los datos que es explicada por el modelo, y tiene un valor máximo de 1, sin embargo un coeficiente de determinación negativo no necesariamente indica que el modelo es malo. Si bien esto sí puede estar relacionado con un modelo ineficiente, también podría explicarse por una gran cantidad de datos atípicos y una alta variabilidad en los datos, que puede observarse en la base de datos proporcionada por el socio formador. Muchos de los productos tenían muchos registros sin ventas y algunos pocos con ventas muy altas, lo que ocasiona esta alta variabilidad en los datos, que a la vez genera los valores negativos del coeficiente de determinación. Es por esto, que para evaluar los modelos, también es necesario considerar las otras métricas.

Por otro lado, el valor del MSE puede parecer realmente alto para algunas predicciones, pero esto es debido a que algunos productos cuentan con ventas muy altas, por lo que a pesar de que el modelo falle por un porcentaje bajo, el valor de la diferencia es considerable alto. Por último, también es importante recordar que esta métrica eleva los

valores al cuadrado, por lo que se castigan mucho los errores. Por esta razón puede ser más significativo tomar en cuenta el MAAPE, ya que se trata de un error porcentual.

## **7. Integración y Ejecución**

### **7.1 ¿Cómo son significativas las métricas?**

Las métricas usadas para medir el desempeño de los modelos son indispensables para tener una mejor idea de qué tan confiables son estos modelos. Ayudan a determinar la precisión de las predicciones realizadas. Esto es de suma importancia, ya que la planeación y gestión de inventario de la empresa se basa en estas predicciones, por lo que tener un dato erróneo o poco confiable podría resultar en una pérdida importante de dinero. Por esto es necesario conocer que mide cada métrica y cómo interpretarla de manera correcta.

### **7.2 Comparación con el Modelo del Socio Formador**

Semanalmente, para los productos más vendidos, el modelo SARIMAX tuvo un porcentaje de error del 22.03%, y el Histogram Gradient Boosting tuvo un 14.61% de error, que al compararse con el 60% que presentaron las predicciones de Don Colchón en los últimos meses, se puede notar una gran mejoría en los resultados.

### **7.3 Comparación del Desempeño de los Modelos Propuestos**

Tras analizar las métricas de error de ambos modelos tanto para las predicciones semanales como las predicciones mensuales, fue evidente que el modelo de Histogram Gradient Boosting (HGB) presentó porcentajes de error más bajos. No obstante, si se analiza cada gráfica individualmente de los valores reales y valores de predicción, es posible concluir que el modelo SARIMAX se adapta mejor a las tendencias de venta, prediciendo de una

manera más acertada los picos y las caídas en las ventas. Por este motivo, cada modelo podría ser mejor que el otro dependiendo de la problemática específica que se quiera resolver.

#### 7.4 ¿Cuándo usar cada modelo?

Como se mencionó anteriormente, el modelo SARIMAX se adapta mejor a la tendencia de las ventas, por lo que si se buscara predecir las ventas en un día en específico, o la tendencia dentro de cada semana, el mejor modelo sería SARIMAX. Por otro lado, si se quiere predecir las ventas totales de una semana o de un mes, el modelo que se recomienda usar es HGB, ya que este modelo presenta menor error en periodos de tiempo más extensos.

#### 7.5 Posibles Mejoras Aplicables

Existen varias áreas de oportunidad en el proyecto que podrían hacer que el análisis de los datos y la predicción de los datos futuros fuera más rápida y precisa. Por un lado, la forma de almacenar los datos podría ser más eficiente, si en lugar de registrar la ventas totales por producto en cada región en un día determinado, se hiciera un registro de cada venta individual, indicando únicamente: el producto vendido, la sucursal donde se realizó la venta, la fecha y el precio unitario. Esto podría agilizar de gran manera la limpieza de datos, y este formato permite filtrar los datos por región, fecha o producto, de manera que se puedan agrupar en base a cierta característica y se logren hacer predicciones precisas.

De igual manera, para reducir el porcentaje de error en el pronóstico, es indispensable entrenar los modelos con una mayor cantidad de datos. El registro de dos años de ventas sí permite generar modelos de predicción, sin embargo, estos modelos tienen pocas referencias de eventos pasados en la misma fecha para basarse en al momento de hacer la regresión. Por lo que aumentar la cantidad de datos, incluyendo más años atrás de ventas, permitiría

adaptarse de una mejor manera a la tendencia por época del año, y consecuentemente, obtener mejores resultados.

En cuanto a la implementación del modelo, una mejora posible es identificar los valores predecidos que se encuentran por debajo del valor mínimo de ventas en un día (3.5 para los datos codificados, 1 para los datos reales) y convertir esos valores en el entero más cercano, que en la mayoría de los casos es cero. Esto es debido a que no tiene sentido una predicción por debajo del valor mínimo, y este cambio podría disminuir significativamente el porcentaje de error obtenido.

## **8. Conclusión**

### **8.1 Resumen**

Todo lo hecho se puede resumir simplificando el proceso a lo que sería un *pitch*, como el que se le hizo al socio formador Don Colchón. Primero se estudió el problema a taclear, en esta ocasión, el MAPE de 60% del socio formador, que impide un buen manejo de inventario y provoca pérdida de dinero. Después, se planificaron los modelos que se utilizarían, posterior a mucha investigación. Se escogió SARIMAX y HGB (inclusive después de probar otros). Después se estableció la métrica que se utilizaría para medir el error de nuestros modelos, que en este caso fue MAAPE, para semejar la métrica del socio y poder comparar. Después se llevaron a cabo los modelos SARIMAX y HGB. Una vez con los resultados de los dos modelos, se compararon entre sí y con el modelo del socio formador. Finalmente, se llevó a cabo un análisis para determinar áreas de oportunidad.

### **8.2 Hallazgos Principales**

Para el modelo SARIMAX, al predecir los 5 productos con más ventas en la base de datos del socio formador, se obtuvo un error MAAPE en predicciones de ventas de tan solo

8.19% semanalmente, y 9.71% mensualmente. Para los 10 productos más vendidos, se obtuvo un error de tan solo 22.03% semanalmente, y 18.46% mensualmente.

Para el modelo HGB, al predecir los 5 productos con más ventas en la base de datos del socio formador, se obtuvo un error MAAPE en predicciones de ventas de tan solo 6.26% semanalmente, y 4.33% mensualmente. Para los 10 productos más vendidos, se obtuvo un error de tan solo 14.61% semanalmente, y 9.01% mensualmente.

Se logra observar como el modelo que funciona mejor es el HGB, ya que tiene menor error. No obstante, es interesante ver como para SARIMAX, en los top 5 su rendimiento fue mejor en predicción semanal pero en el top 10 fue mensual. Así mismo, SARIMAX predice mejor los picos. Por otro lado, para HGB, para ambas evaluaciones de predicciones, top 5 y top 10, su rendimiento fue mejor en las predicciones mensuales. No obstante, también cabe mencionar, que en todos estos escenarios y en más no considerados dentro de hallazgos principales, se obtuvo mucho mejor rendimiento en la predicción de ventas que el error de 60% con el que contaba el modelo de predicción del socio Don Colchón.

### 8.3 Implicación

A partir de los resultados obtenidos, se concluye que la elección del modelo de predicción es crucial para minimizar el error en las predicciones de ventas. El modelo HGB demostró ser más eficaz, pero cabe resaltar que el rendimiento de SARIMAX también fue muy bueno, lo que subraya la importancia de considerar múltiples modelos y situaciones diferentes al realizar predicciones de ventas. Así mismo, aunque el socio maneje sus predicciones de manera semanal, se puede notar el valor en considerar otros rangos de tiempo, ya que las predicciones mensuales rindieron mejores resultados.

Adicionalmente, estos hallazgos demuestran que es posible mejorar significativamente la precisión de predicciones en comparación a modelos anteriores, por lo

que es importante actualizar con cierta frecuencia los modelos que utiliza una empresa. Esto, con el fin de que el impacto positivo de predicción de ventas siga siendo significativo para la empresa, su planificación y estrategia de negocio, a largo plazo.

#### 8.4 Implementación

Todos estos hallazgos, métodos y resultados son aplicables a lo largo de múltiples, si no todos los sectores. La predicción de ventas para el manejo de inventario se puede implementar en el comercio minorista, la manufactura, la agricultura y los servicios de alimentos. Así como se analizó cómo en un comercio minorista es posible optimizar el inventario, a través del caso de estudio de Don Colchón, es de gran valor también observar la aplicación de predicción de ventas en otros sectores.

En la agricultura, esta metodología puede apoyar a los agricultores a planear sus cosechas y manejar sus suministros de manera más eficiente. Con el objetivo de disminuir el desperdicio de recursos, buscar generar producto más fresco para la etapa en la que se demanda, y no producir de más, se taclea un problema real y con un impacto significativo en la alimentación. En caso de producir producto de más, no solo es desastroso para el agricultor, sino para toda la economía agrícola, llevando a descuentos y eliminación de las cosechas antes de que mermen, perdiendo gran inversión pasada de tiempo, esfuerzo y dinero.

Así mismo, las técnicas de predicción no se limitan a ventas e inventario, y son aplicables en otros sectores. En la energía, permiten predecir la demanda de electricidad, en la salud se puede predecir la extensión de una enfermedad, en el clima se predice la temperatura y lluvia con semanas de anticipación, y en el sector financiero se utiliza para predecir tendencias en el mercado, entre otras cosas. Los modelos de predicción son herramientas muy

útiles, con alta versatilidad, y logran tener un gran impacto, no importa en qué industria se implementen.

### 8.5 Ventaja competitiva

El conocimiento adquirido a través de la implementación y evaluación de modelos de predicción es una herramienta poderosa en el espacio comercial. Al optimizar la precisión de predicciones relacionadas al negocio, se pueden anticipar tendencias, comportamientos y resultados futuros, permitiendo a las empresas tomar decisiones informadas y siempre estar un paso adelante de la competencia. El buscar predecir un entorno, impulsa a cualquier negocio hacia la innovación, optimización, y mejora en la toma de decisiones.

### 8.6 Implicación Global

Los hallazgos presentados tienen una implicación directa con el ODS 12 de producción y consumo responsable. Al optimizar el inventario, las empresas mejoran su rentabilidad y también reducen su impacto ambiental, contribuyendo a una economía más sostenible. Adicionalmente, a través del rol que puede jugar la predicción con modelos en general en otros sectores, se puede ver como tiene implicaciones potentes globales. Los modelos de predicción pueden llegar a apoyar la solución de todos los ODS, no solo el 12.

En conclusión, la ciencia de datos tiene el potencial de abordar una variedad de problemas globales, desde manejo de inventario hasta problemas ambientales. Por ello, se le extiende una invitación formal a los lectores para que continúen aprendiendo y aplicando modelos de predicción en sus respectivos campos. Se debe tener en cuenta siempre la implicación de la ciencia de datos en problemas mundiales, debido a que cuenta con todas las herramientas necesarias para apoyar a solucionar cualquier problemática, de una u otra forma.



La ciencia de datos no solo es una herramienta para el crecimiento empresarial y la ganancia de capital, sino también un medio para lograr un mundo más sostenible, equitativo y feliz.

## 9. Apéndice

Para más información, libretas trabajadas, variantes trabajadas, entre otros archivos relevantes al desarrollo de esta propuesta, se puede acceder al repositorio del proyecto aquí:

<https://github.com/Racoo203/colchoncitos>.

## 10. Referencias

Analytics Vidhya. (2020). Predicting Sales: Time Series Analysis & Forecasting with Python.

<https://medium.com/analytics-vidhya/predicting-sales-time-series-analysis-forecasting-with-python-b81d3e8ff03f>

Bisman Preet Singh. (2020). Prediciendo Ventas: Análisis de Series Temporales y Pronóstico con Python.

<https://medium.com/analytics-vidhya/predicting-sales-time-series-analysis-forecasting-with-python-b81d3e8ff03f>

Bind. (2020, 8 de octubre). Técnicas y métodos para el control de inventarios.

<https://bind.com.mx/blog/control-de-inventarios/tecnicas-y-metodos-para-el-control-de-inventarios>

Contreras, A., Atziry, C., Martínez, J. & Sánchez, D. (2016) Análisis de series de tiempo en el pronóstico de la demanda de almacenamiento de productos perecederos.

<https://doi.org/10.1016/j.estger.2016.11.002>

Friedman, J. H. (2001). Greedy Function Approximation: A Gradient Boosting Machine.

<https://jerryfriedman.su.domains/ftp/trebst.pdf>

GeeksforGeeks. (2023). Complete Guide To SARIMAX in Python.

<https://www.geeksforgeeks.org/complete-guide-to-sarimax-in-python/>

GeeksforGeeks. (2023). Introducción a Pandas.

<https://www.geeksforgeeks.org/introduction-to-pandas-in-python/>

GeeksforGeeks. (2023). LightGBM (Light Gradient Boosting Machine).

<https://www.geeksforgeeks.org/lightgbm-light-gradient-boosting-machine/>

GeeksforGeeks. (2023). Python Numpy. <https://www.geeksforgeeks.org/python-numpy/>

GeeksforGeeks. (2020). Python - Create progress bar using tqdm module.

<https://www.geeksforgeeks.org/python-create-progress-bar-using-tqdm-module/>

Gigola, C. (2017). Notas del curso de pronósticos e inventarios. Instituto Tecnológico Autónomo de México.

[https://gente.itam.mx/gigola/Curso\\_pron\\_inv/Notas%204\\_2017.pdf](https://gente.itam.mx/gigola/Curso_pron_inv/Notas%204_2017.pdf)

Hernandez, S. (2008). APLICACIONES DE LAS SERIES DE TIEMPO EN MODELOS DE TRÁFICO PARA UNA RED DE DATOS.

<https://www.redalyc.org/articulo.oa?id=84903806>

Holguín, C. J. V., & del Valle, U. (2010). Fundamentos de control y gestión de inventarios.

<https://doi.org/10.25100/PEU.48>

IBM. (2021). ARIMA. <https://www.ibm.com/docs/es/spss-modeler/saas?topic=series-arima>

Li, D., Lin, K., Li, X., Liao, J., Du, R., Chen, D., & Madden, A. D. (2022). Improved sales time series predictions using deep neural networks with spatiotemporal dynamic pattern acquisition mechanism. Information Processing And Management, 59(4), 102987. <https://doi.org/10.1016/j.ipm.2022.102987>

Li, K. (2021). Building Sales Prediction Models.

<https://kenpyfin.com/blog/2021/02/03/building-sales-prediction-models/>

Masui, T. (2023). All You Need to Know about Gradient Boosting Algorithm – Part 1.

Regression Algorithm explained with an example, math, and code.

<https://towardsdatascience.com/all-you-need-to-know-about-gradient-boosting-algorithm-part-1-regression-2520a34a502>

McIntire, G. (2023). Tutorial de Python Pandas: Una introducción completa para

principiantes. <https://www.learndatasci.com/tutorials/python-pandas-tutorial-complete-introduction-for-beginners/>

Pacto Mundial. (2023). ODS 12 Producción y consumo responsables.

<https://www.pactomundial.org/ods/12-produccion-y-consumo-responsables/>

Phosgene89. (2020). SARIMA. GitHub. <https://phosgene89.github.io/sarima.html>

PyPI. (2023). Skforecast. <https://pypi.org/project/skforecast/>

Matplotlib. (2023). matplotlib.pyplot — Matplotlib 3.8.3 documentation.

[https://matplotlib.org/stable/api/ pyplot\\_summary.html](https://matplotlib.org/stable/api/ pyplot_summary.html)

Medium. (2020). ACF (autocorrelation function) — simple explanation with Python

example.

<https://medium.com/@krzysztofdrzelczuk/acf-autocorrelation-function-simple-explanation-with-python-example-492484c32711>

SafetyCulture. (2023, 9 de febrero). Control de inventarios: qué es, tipos, ejemplos y más.

<https://safetyculture.com/es/temas/manejo-de-inventario/control-de-inventarios/>

Scikit-learn. (2023). Ensembles: Gradient boosting, random forests, bagging, voting,

stacking. <https://scikit-learn.org/stable/modules/ensemble.html>

Scikit-learn. (2023). HistGradientBoostingRegressor.

<https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.HistGradientBoostingRegressor.html>

Scikit-learn. (2023). 6.3. Preprocessing data — scikit-learn 1.4.1 documentation.

<https://scikit-learn.org/stable/modules/preprocessing.html>

Spiliotis, E., Spyros Makridakis, Artemios-Anargyros Semenoglou, & Vassilios

Assimakopoulos. (2020). Comparison of statistical and machine learning methods for daily SKU demand forecasting. *Operational Research*, 22(3), 3037–3061.

<https://doi.org/10.1007/s12351-020-00605-2>

Statsmodels. (2023). statsmodels 0.14.1. <https://www.statsmodels.org/stable/index.html>

Telesto. (2023, 13 de febrero). Pronóstico de demanda de inventario.

<https://www.telesto.app/blog/2023/02/13/pronostico-de-demanda-de-inventario/>

UNSD. (2024). Informe sobre los progresos en el cumplimiento de los ODS.

<https://www.un.org/sustainabledevelopment/es/progress-report/>