



MEJORANDO LA GESTIÓN DE UNA PEQUEÑA EMPRESA UTILIZANDO TÉCNICAS DE ML: CASO DE UNA CAFETERÍA

POR: RAFAEL FARIÁS POBLETE

Proyecto presentado a la Facultad de Ingeniería de la Universidad del
Desarrollo para optar al grado de Magíster en Data Science

PROFESOR GUÍA:
DRA. LORETO BRAVO.

Octubre 2022
SANTIAGO

*Esta tesis está dedicada a mis padres, a mi esposa y a mi hija que viene en camino.
A mis padres por haber sacrificado su esfuerzo por mi felicidad, a mi esposa por
acompañarme con su gran amor y entendimiento y a mi hija por ya augurar
tiempos llenos de felicidad.*

*Primero agradezco a los más cercanos por permitirme no compartir gran parte de
mi tiempo libre con ellos y poder dedicarme a esto. Agradezco también la
oportunidad de seguir educándome y también agradezco a la UDD y a quienes me
ayudaron, por su comprensión en tiempos que fueron difíciles tanto personal como
profesionalmente.*

Resumen

Generar predicciones es algo de lo cual los negocios dependen en estos tiempos. El no poder pronosticar ya sea las venta diarias o mensuales o de la periodicidad que estratégicamente se determine, aumenta la probabilidad de fracaso o de menores rentabilidades. Algo que las grandes empresas ya han adoptado, sin embargo las pequeñas aún no.

Es por eso que este documento busca entender el sesgo en la aplicación de forecasting en tres cafeterías pequeñas y a través de los datos de sus ventas generadas el año 2022, buscar un modelo sencillo de ML para realizar pronósticos que permita obtener un MAPE¹ menor al 20 por ciento para que sea considerado como bueno y exacto. El objetivo general se enmarca en facilitar a personas sin conocimientos en la aplicación de forecasting utilizando Machine Learning. Dado que los tiempos son acotados dejamos fuera de alcance el analizar el verdadero impacto que tendrá este forecasting en las cafeterías estudiadas, sin embargo se se espera realizar en un trabajo futuro.

Los datos son obtenidos de tres cafetería y corresponden a las ventas diarias del año 2022.

Luego de finalizado el proyecto de puede concluir que las dificultades abordadas por las pequeñas empresas para la aplicación de ML en sus negocios son ciertas y que pueden en parte verse mejoradas debido a la existencia de nuevas herramientas que buscan facilitar su ejecución. Se pudo proponer la existencia de un modelo simple y capaz de pronosticar de manera buena y exacta los ingresos para los próximos 7 días en tres pequeñas cafeterías de la V región, obteniendo para la cafetería A un MAPE de 0.139, para la Cafetería B de 0.1806 y para la Cafetería C de 0.179.

¹Mean Abosolute Percentage Error

Es interesante mencionar que estas tres pequeñas empresas tienen diferencias en madurez en el mercado, ubicaciones, público objetivo, ingresos y ciclos semanales, por lo que se puede indicar que el modelo cumple en generalidad. Se logra mostrar además, la capacidad que tiene el modelo de mejorar a través de la optimización de hiperparámetros mejorando el resultado de la Cafetería B desde 0.186 a 0.179.

Contenido

1. Introducción	2
1.0.1. Encuesta Cámara de Comercio	4
2. Revisión Bibliográfica	5
3. Hipótesis y Objetivos	8
4. Descripción de los datos y Metodología	9
5. Resultados	21
6. Conclusiones y trabajo futuro	29
Anexo 1: Código Modelo sin Optimizar	32
Anexo 3: Búsqueda de hiperparámetros	34
Anexo 2: Código Modelo Optimizado	36

Índice de figuras

4.1. Representación gráfica de los Datasets y qué información será utilizada (Elaboración propia, 2020)	10
4.2. Distribución de los datos Cafetería A (Elaboración propia, 2022) . . .	11
4.3. Gráfico de BoxPlot Cafetería A (Elaboración propia, 2022)	12
4.4. Distribución de los datos Cafetería B (Elaboración propia, 2022) . . .	13
4.5. Gráfico de BoxPlot Cafetería B (Elaboración propia, 2022)	14
4.6. Distribución de los datos Cafetería C (Elaboración propia, 2022) . . .	15
4.7. Gráfico de BoxPlot Cafetería C (Elaboración propia, 2022)	16
4.8. Mean Absolute Percent Error	17
5.1. Pronóstico Cafetería A sin Optimización (Elaboración propia, 2022) .	22
5.2. Pronóstico Cafetería B sin Optimización (Elaboración propia, 2022) .	23
5.3. Pronóstico Cafetería C sin Optimización (Elaboración propia, 2022) .	24
5.4. Componentes Cafetería A (Elaboración propia, 2022)	25
5.5. Componentes Cafetería B(Elaboración propia, 2022)	26
5.6. Componentes Cafetería C(Elaboración propia, 2022)	27

Índice de tablas

5.1. Resultados de MAPE para las tres cafeterías.	21
---	----

Capítulo 1

Introducción

Las Pymes son un pilar fundamental del desarrollo económico sustentable, porque son generadoras de riqueza, además de ser entes dinámicos que identifican, explotan y desarrollan nuevas actividades productivas que no sólo benefician a los *shareholders*¹, sino que también a todos los *stakeholders*².

Lamentablemente son organizaciones que no se adaptan rápidamente a las nuevas tecnologías[3], a pesar de que su planeación y organización no requiere de mucha burocracia en comparación con las grandes empresas, lo que podría significar una ventaja, estas organizaciones tienen que perdurar en los mercados de alta competencia y para ello deben alcanzar un desarrollo empresarial que se los permita y para eso necesitan ir ligados a los cambios tecnológicos.

Una de las fuentes de este desarrollo empresarial y tal como lo indica Andrew Ng en su artículo *AI Transformation Playbook*[19], es la Inteligencia Artificial, la cual está destinada a transformar la gestión en toda la industria. Sin embargo Machine Learning no es realmente viable para pequeñas compañías en el actual ecosistema[9] debido principalmente a la falta de capacidades, adquisición de talentos y costos.

El aprendizaje automático puede mejorar la eficiencia de una empresa de muchas formas y en diferentes áreas de estas. Por ejemplo, la alta dirección puede obtener

¹Shareholders es aquella persona natural o jurídica que es propietaria de acciones de los distintos tipos de sociedades anónimas o comanditarias que pueden existir en el marco jurídico de cada país.

²Stakeholder es cualquier individuo u organización que, de alguna manera, es impactado por las acciones de determinada empresa.

mejor información y predicciones. Las estrategias de marketing se puede optimizar mediante la publicación de anuncios más personalizados para los clientes. El ML también puede ayudar con la participación del cliente mediante el uso de soluciones como chatbots [10]. Es así como varias tareas que antes se realizaban manualmente en una empresa se puede automatizar mediante el uso de ML, liberando así tiempo para que los empleados realicen otras tareas cruciales[7] para la empresa.

Los beneficios de incorporar ML ya han sido ampliamente probados por empresas tales como Google, Amazon, Microsoft entre otras grandes corporaciones, en donde han demostrado que su aplicación genera rentabilidad para sus negocios. Sin embargo, es preciso mencionar que ellos cuentan con los recursos necesarios para poder levantar proyectos en temas relacionados al Machine Learning o a la Inteligencia Artificial en general, en cambio, las pequeñas empresas, no cuentan con este nivel de recursos, así como tampoco con los conocimientos, por lo que la adopción se complejiza. Además, al no existir un mayor crecimiento en el uso y adopción de técnicas de Machine Learning en las pequeñas empresas; no se forman Data Scientist expertos en esta categoría de negocios, por lo que existen barreras no solo internas, sino que también el mercado arrastra que se forme esta brecha tecnológica.

En un estudio realizado en Estados Unidos y México sobre pequeñas empresas[10], se identificaron varios problemas comunes, entre ellos, desarrollo de estrategias financieras, comercializar productos, contratar empleados, mejorar el servicio al cliente, entre otros. Lo interesante de estos puntos es que todos pueden ser abordados desde una mirada del Machine Learning para mejorar la gestión aunque a sabiendas de que en una pequeña empresa no son muy aplicados en la actualidad.

Del párrafo anterior y sumado a lo explicado de la importancia de las PYMES en la economía de un país, se cree que el impacto que puede tener la aplicación de Machine Learning en una pequeña empresa no solo la beneficiaría a ella sino que además a la economía en su totalidad.

Es por todo lo explicado anteriormente que este documento busca dar a conocer la experiencia de una cafetería en la aplicación de Machine Learning, elaborando pronósticos que permitan alimentar la gestión y mejorar el resultado financiero, cuyo beneficio no sea sólo corto plazista sino que también apunte a aumentar el valor de una empresa en el largo plazo. Evitando grandes complejidades para que pueda ser replicable por otras pequeñas empresas. Esto se espera lograr abordando temas

como revisiones bibliográficas, descripción de los datos, metodologías utilizadas, resultados, conclusiones y limitaciones.

1.0.1. Encuesta Cámara de Comercio

Esta encuesta realizada en el marco de este escrito, tiene el objetivo de poder validar si efectivamente existe un desconocimiento de estas herramientas en las pequeñas empresas. La encuesta consideró tres preguntas siendo la primera de ellas ¿Qué sabes de Machine Learning? y se obtuvo un resultado bastante pobre en donde de un universo de 104 pequeñas empresas sólo el 9 por ciento ha escuchado de ML, sin embargo no la aplican en su negocio y el 91 por ciento restante jamás ha escuchado sobre esta tecnología.

La segunda pregunta tuvo relación con ¿Aplicas pronósticos en tu empresa? en donde el 74 por ciento ha escuchado pero no aplica, el 23 por ciento no ha escuchado nunca y sólo un 3 por ciento lo aplica. Por último se les pidió que se auto evaluaran con una escala de notas de 1 a 7 respecto de la utilización de tecnologías en sus negocios en donde se obtuvo una sorprendente mala calificación promedio de 3.89.

A través de esta encuesta se pudo demostrar que efectivamente las pequeñas empresas carecen de acceso y conocimientos sobre Machine Learning, no realizan técnicas de gestión necesarias para proyectar sus ventas y además tienen una muy baja adopción de las tecnologías en general.

Capítulo 2

Revisión Bibliográfica

Como se indica en el artículo de Daniel Fagella [9], Machine Learning no es realmente viable para pequeños negocios en el actual ecosistema. La principal razón que se indica en este artículo es la falta de recursos tales como cantidad representativa de datos, experiencia en temas de Machine Learning, conocimientos profundos del negocio y dinero para hacer inversiones en este tipo de herramientas, dificultan que una pequeña empresa pueda escalar en este tipo de soluciones. Si bien todos los recursos nombrados pueden ser escasos en una empresa pequeña, según lo indicado, la experiencia en temas de Machine Learning puede ser la más difícil de obtener, algo que también comparte el estudio realizado por IBM [8].

La experiencia requerida para un proyecto de ML incluye gran parte de las herramientas de la ciencia de datos [15]. Sin embargo para las pequeñas empresas, obtener experiencia en este tipo de herramientas puede resultar bastante difícil, sobre todo en Chile; debido principalmente a que por un lado existe una brecha digital importante, y por otro a que los actuales Data Scientist existentes son captados por las grandes empresas, haciendo que tanto por escasez como por valor, sea prácticamente imposible para una pequeña empresa llegar a ellos [9].

Quizás una de las maneras a través de la cual una pequeña empresa puede acceder a desarrollar soluciones de ML, es contratar un proveedor como Google¹, Amazon² o Microsoft³, sin embargo incluso las actuales tarifas pueden ser inalcanzables para

¹<https://cloud.google.com/>

²<https://aws.amazon.com/es/>

³<https://azure.microsoft.com/es-es/>

una pequeña empresa. A pesar de esta gran limitante, parecen ser otros los mayores problemas de adquirir este tipo de soluciones [9], por un lado es necesario contar con data organizada, así como también también aprender a integrar el sistema propio con estos sistemas en la nube; por lo tanto se llega nuevamente a la necesidad de contar con un data scientist para poder realizar estas operaciones.

Por ejemplo, en un estudio realizado por Amershi et al. [1] se busca describir el proceso que realiza un equipo de Microsoft para levantar proyectos de ML. En este estudio los participantes fueron encuestados sobre cómo son resueltos los problemas relacionados al ML en su equipo. De aquí se pudo presentar un modelo con nueve estaciones en donde se trabaja con los datos además de las modelaciones respectivas. También fue presentando un nuevo indicador que permite medir que tan bien realizan el trabajo. Otros estudios similares también proveen una guía sobre como realizar un proceso de ML, entre ellos el de Ishikama et al. [13], Nascimento [18], Hill et al. [11] y Saleema Amershi et al. [1]. Estos estudios son importantes para este documento ya que en él se indican la importancia de tener una metodología de desarrollo por etapas para llevar a cabo un proyecto de ML. Sin embargo estos estudios y como la gran mayoría, están apuntados a una empresa de gran tamaño y con grandes recursos, que son capaces de articular múltiples áreas y etapas para un proyecto, contrario a lo que se busca con este escrito que es simplificar el proceso para que cualquier pequeña empresa lo pueda implementar.

Otro artículo interesante es el propuesto por Alfredo de Massis et al. [4] en donde propone ciertas metodologías de investigación en las pequeñas empresas para encontrar insights que permitan mejorar las decisiones, sin embargo en este artículo no se hace mención a la aplicación de técnicas de Machine Learning, por lo que se cree que tiene un sesgo tecnológico importante.

Otro trabajo es el realizado por Nascimento et al. [18]. En éste se discute sobre el desafío de la aplicación de ML en tres pequeñas empresas para analizar como el ML es usado, también se indican los problemas a los que se ven enfrentados en la realización de estas tareas. Entre los principales problemas se destaca el manejo de los datos y la estructura con que debe contar la base de datos. Sin embargo a pesar de ser un estudio interesante, éste es realizado en empresas que ya cuentan con experiencia en temas relacionados al ML. Esto difiere de este trabajo ya que está enfocado en una empresa sin experiencia en ML y espera ser replicado a otras pequeñas empresas carentes de conocimiento en el área.

Los siguientes estudios [5, 16, 17, 20] que se refieren a que los problemas financieros son de los temas que más enfrenta una pequeña empresa y que impiden su rápido crecimiento, no consideran herramientas de Machine Learning para mejorar estos indicadores.

Por ultimo existe una basto número de estudios para hacer forecast en pequeñas empresas como por ejemplo [22, 2, 14], sin embargo carecen de la simpleza que se busca para que el modelo sea replicable por otras pequeñas empresas.

Capítulo 3

Hipótesis y Objetivos

Sin duda alguna el poder plasmar el rumbo de este proyecto es de suma importancia y para esto se describe una hipótesis y un objetivo general.

La principal hipótesis es: *¿Existe un modelo sencillo de ML para realizar Forecasting de ventas para los próximos siete días y que permita obtener un MAPE menor al 20 por ciento?*

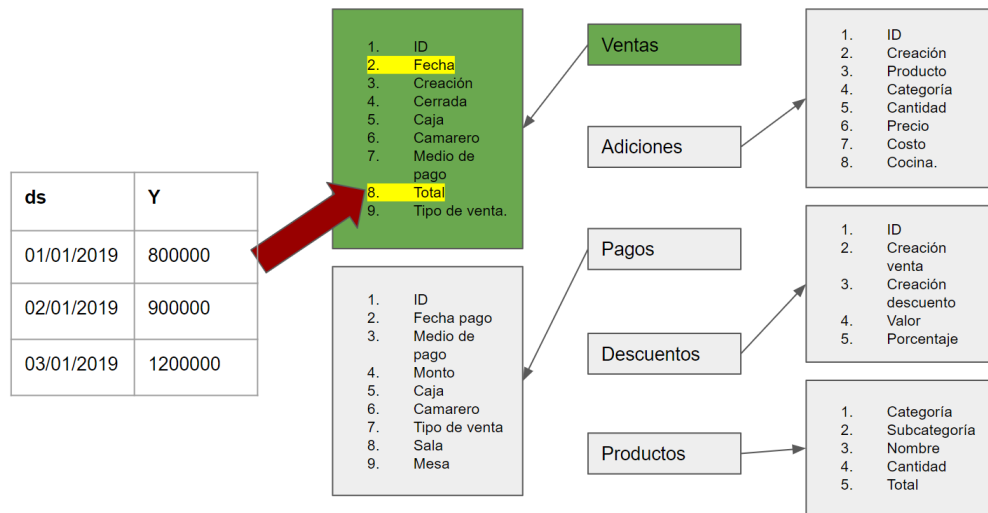
Como objetivo general, se pretende demostrar la existencia de una herramienta de ML sencilla que pueda ser utilizada por personas sin conocimientos, al mismo tiempo que por aquellos que si los tienen, para poder pronosticar datos de ventas diarias en pequeñas cafeterías.

Capítulo 4

Descripción de los datos y Metodología

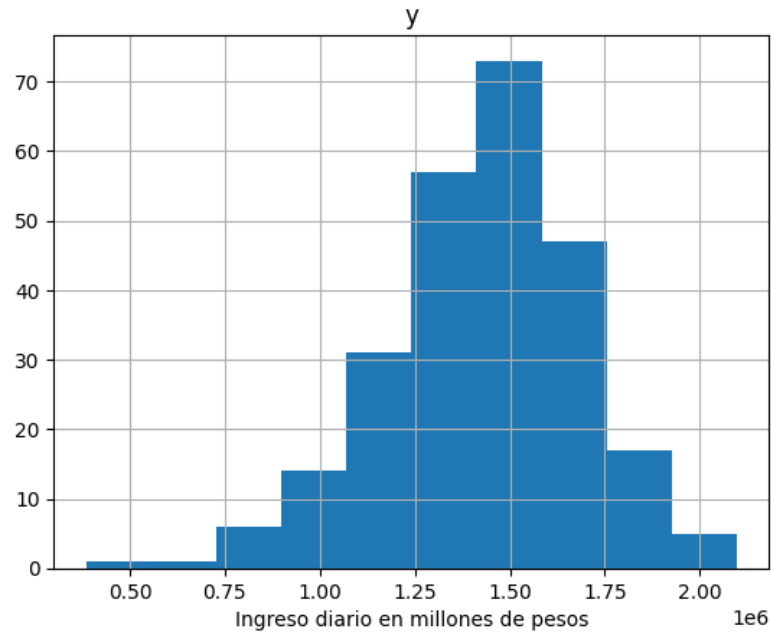
Los datos a utilizar en este proyecto corresponden a la información obtenida de tres pequeñas cafeterías, denominadas Cafetería A, B y C respectivamente, en donde como se aprecia en la Figura 3.1, del dataset ventas se utilizan los ingresos diarios totales desde el 16 de Febrero de 2022 al 27 de Octubre de 2022 con 252 registros.

Figura 4.1: Representación gráfica de los Datasets y qué información será utilizada (Elaboración propia, 2020)



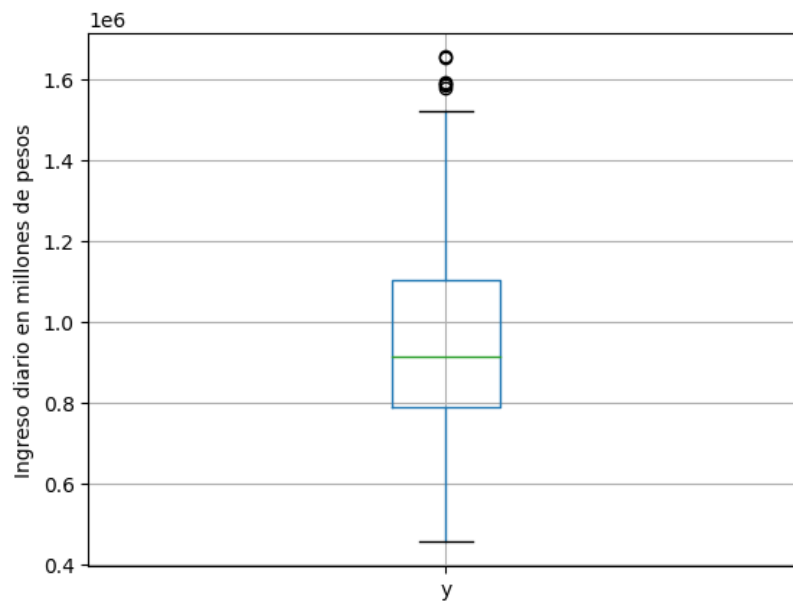
En la información concerniente a datos de ventas, se cuenta con una variable ordinal que es la fecha (ds) y una variables continua que es el monto total de ingreso diario (y).

Figura 4.2: Distribución de los datos Cafatería A (Elaboración propia, 2022)



Los datos tienen una distribución de característica principalmente log-normal, tal cual se aprecia en la Figura 3.2 en donde el eje horizontal indica el monto de ingreso en millones de pesos. En la Figura 3.3 se aprecia la existencia de valores alejados del promedio, esto debido principalmente a disminución no programada en el horario de atención, así como también aquellos valores cercanos al límite superior, se debe principalmente a días de alta afluencia de público por ser considerados días especiales.

Figura 4.3: Gráfico de BoxPlot Cafetería A (Elaboración propia, 2022)



Para la cafetería B también se tiene, según lo muestra la Figura 3.4 en donde el eje horizontal indica el monto de ingreso en millones de pesos, una distribución característica log-normal, al igual que valores alejados del promedio según lo muestra la Figura 3.5. Las causas son similares a la cafetería A, es decir, bajas en horarios de atención, así como alzas de ingresos en días especiales.

Figura 4.4: Distribución de los datos Cafatería B (Elaboración propia, 2022)

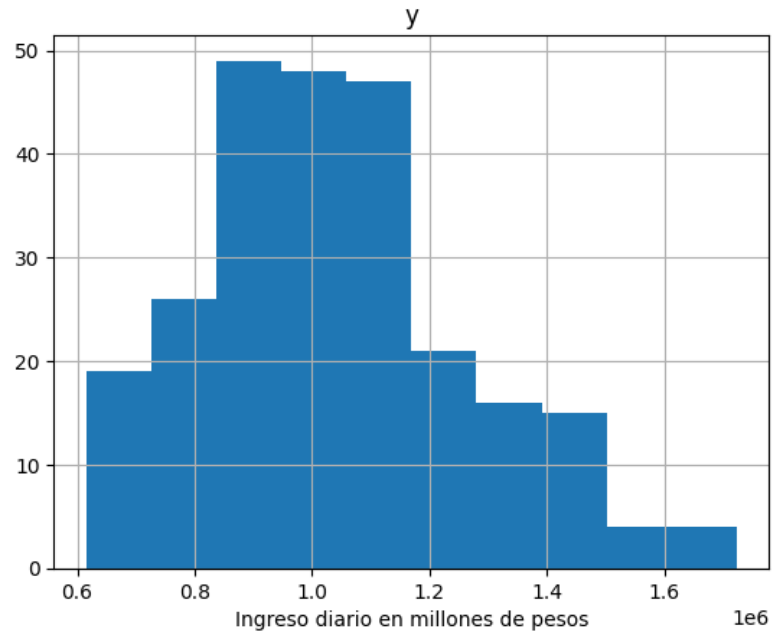
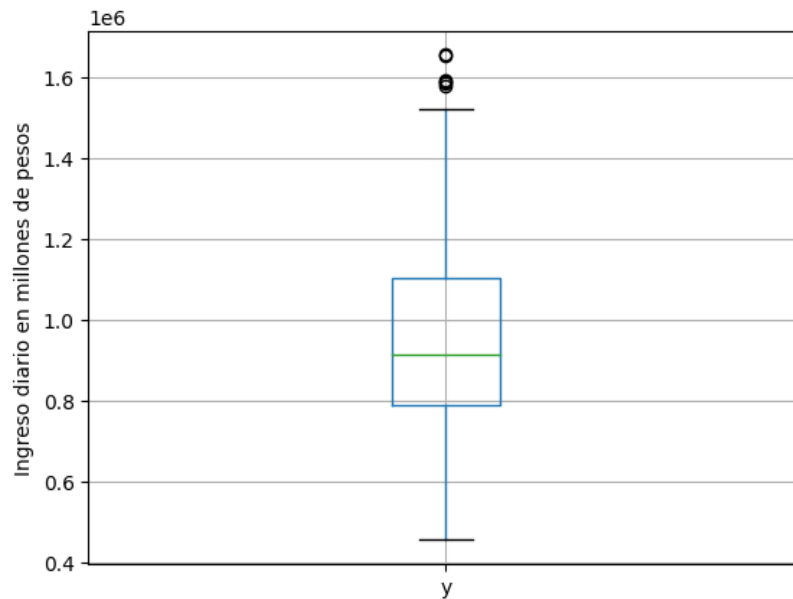


Figura 4.5: Gráfico de BoxPlot Cafetería B (Elaboración propia, 2022)



La Cafetería C también muestra, según la Figura 3.6 en donde el eje horizontal indica el monto de ingreso en millones de pesos, una distribución con característica log-normal, al igual que valores alejados del promedio según lo muestra la Figura 3.7. Las causas principales son que al ser una Cafetería recién inaugurada, sus valores tienden a tener una mayor fluctuación, siendo al inicio más altos que al final del período estudiado.

Figura 4.6: Distribución de los datos Cafatería C (Elaboración propia, 2022)

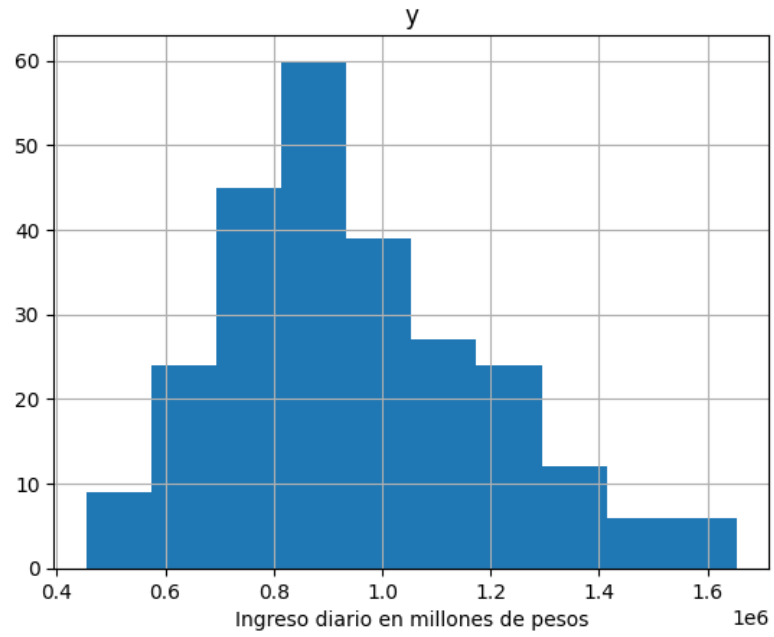
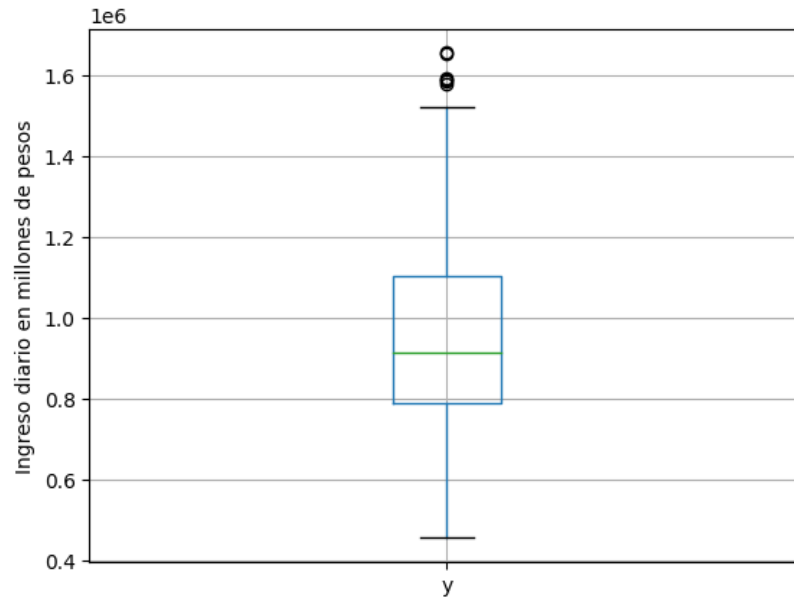


Figura 4.7: Gráfico de BoxPlot Cafetería C (Elaboración propia, 2022)



Para describir la metodología es preciso mencionar que los datos pertenecen a tres cafeterías de la quinta región de Chile. Luego de esto se realiza la exploración de los datos para determinar si hay valores nulos, su distribución o si bien existe una presencia importante de outliers.

Luego de revisada la literatura en lo que respecta a modelos de forecasting, se opta por la utilización de Prophet¹[21] debido principalmente a su simpleza ya que indican que es aplicable por no expertos y además posee una excelente documentación y que por lo tanto permitiría una mayor replicabilidad en pequeñas empresas. Sumado a lo anterior y según sus autores, es un modelo robusto frente a la presencia de outliers, lo que facilita aún más su aplicación.

Una vez escogido el modelo se opta por hacer 4 experimentos. Tres de ellos prueban el modelo sin la optimización de hiperparámetros (un modelo para cada cafetería), además se realiza una prueba relacionada a la optimización de los hiperparámetros de la Cafetería B con el fin de descubrir si el resultado es mejor o no. En todos los casos lo que se busca es un MAPE, cuya formulación se explica en la Figura 3.8, me-

¹<https://facebook.github.io/prophet/>

nor a 20 por ciento, ya que según Clark es considerado como un forecasting bueno y exacto[12]. Para la obtención del MAPE, se utiliza un método de cross validation de k-fold ², una herramienta que posee el modelo, en donde se setean tres parámetros: initial, que es la cantidad de días de entrenamiento, horizon que indica los días que tendrá cada fold y period modifica que tan a menudo comienza un nuevo fold.

Figura 4.8: Mean Absolute Percent Error

$$\frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \check{y}_i}{y_i} \right|$$

Para el caso de los modelos sin optimización de hiperparámetros se siguen las siguientes instrucciones:

1. Se importa la data con el formato requerido una columna en datetime que tenga la fecha y que se titule ds y otra columna que contenga el monto de ingreso diario y que se denomine y.
2. Se inicia el modelo Prophet de Facebook que utiliza una metodología de descomposición aditiva no lineal para modelar series de tiempo. Matemáticamente es lo siguiente:

$$y(t) = g(t) + s(t) + h(t) + e(t)$$

Donde:

$y(t)$ es el valor observado de la serie de tiempo en el tiempo t . $g(t)$ es la tendencia de la serie de tiempo en el tiempo t . $s(t)$ es la componente estacional de la serie de tiempo en el tiempo t . $h(t)$ es la componente de días festivos de la serie de tiempo en el tiempo t . $e(t)$ es el término de error o residuo de la serie de tiempo en el tiempo t .

La tendencia $g(t)$ se modela utilizando un modelo de regresión no lineal, que se puede expresar de la siguiente manera:

$$g(t) = m(t) + k(t)x$$

²https://scikit-learn.org/stable/modules/cross_validation.html

Donde:

$m(t)$ es la tendencia subyacente de la serie de tiempo en el tiempo t . $k(t)$ es la tasa de crecimiento de la tendencia en el tiempo t . x es una variable binaria que toma el valor 0 o 1 dependiendo si el tiempo t está antes o después de un punto de inflexión.

La componente estacional $s(t)$ se modela utilizando una función de Fourier, que se puede expresar de la siguiente manera:

$$s(t) = \text{sum}(a(n)\cos(2\pi nt/P) + b(n)\sin(2\pi nt/P))$$

Donde:

$a(n)$ y $b(n)$ son los coeficientes de Fourier que definen la forma de la estacionalidad en la serie de tiempo. P es el período de la estacionalidad.

La componente de días festivos $h(t)$ se modela como una variable binaria que toma el valor 1 en los días festivos y 0 en los demás días.

La estimación de los parámetros y coeficientes de Fourier se realiza utilizando un algoritmo de optimización bayesiana que maximiza la verosimilitud de los datos de entrenamiento.

En resumen, el modelo es la descomposición aditiva no lineal de la serie de tiempo utilizando una combinación de modelos de regresión no lineal y funciones de Fourier, que se ajustan a los datos utilizando un algoritmo de optimización bayesiana para producir un pronóstico lo más preciso posible de los ingresos diarios en el futuro.

3. Se realiza un fit [6] al modelo, es decir, se crea el objeto modelo y luego se llama el método fit el cual se utiliza para ajustar el modelo a los datos de la serie de tiempo, utilizando un algoritmo de optimización bayesiana para encontrar los valores óptimos de los hiperparámetros del modelo. Una vez que el modelo ha sido ajustado, se pueden generar pronósticos futuros y evaluar la precisión del modelo utilizando diversas métricas de evaluación.
4. Se crean las nuevas fechas que contendrán el forecast, es decir, se crea un nuevo objeto denominado future, el cual contiene las fechas a pronosticar.

5. Se crea un nuevo objeto denominado `forecast`, al cual es la predicción del objeto `future` creado anteriormente, a través del método `predict` de Prophet el que devuelve un `DataFrame` que contiene los valores pronosticados para las fechas futuras, junto con los intervalos de confianza. El usuario puede extraer los valores de interés, como los valores medios pronosticados, los límites inferiores y superiores de los intervalos de confianza, y otros estadísticos de interés para su análisis.

En resumen, la función `predict` de Prophet utiliza el modelo ajustado para generar pronósticos futuros para una serie de tiempo dada. Los pronósticos se generan utilizando una combinación de la tendencia, la estacionalidad y los efectos de días festivos del modelo ajustado, y se proporcionan intervalos de confianza para cuantificar la incertidumbre en los pronósticos.

Para el caso del modelo con optimización de hiperparámetros se utiliza la guía entregada por Prophet ³ en el apartado `hyperparameter tuning`, en donde se optimizan los siguientes parámetros:

1. `Changepoint prior scale`: Probablemente es el parámetro que más impacto tiene. Determina la flexibilidad de la tendencia, en otras palabras, indica como esa tendencia afecta los puntos de cambio de la serie. Sus valores pueden ir entre 0.001 y 0.5.
2. `Seasonality prior scale`: Este parámetro controla la flexibilidad de la estacionalidad. De manera similar, un valor grande permite que la estacionalidad se ajuste a grandes fluctuaciones, un valor pequeño reduce la magnitud de la estacionalidad. El valor predeterminado es 10, que básicamente no aplica ninguna regularización. Un rango razonable para ajustarlo es desde 0.01 hasta 10.

En palabras propias, simples y resumidas de lo descrito en el escrito de Taylor SJ [21], la matemática detrás del hiperparámetro `seasonality prior scale` en Prophet es la siguiente:

En el modelo Prophet, la estacionalidad de la serie de tiempo se modela utilizando una suma de funciones sinusoidales de diferentes periodos. Cada función sinusoidal se ajusta a los datos utilizando una amplitud y una fase específicas.

El hiperparámetro `seasonality prior scale` controla la prioridad asignada a la

³<https://facebook.github.io/prophet/docs/diagnostics.html>

suavidad de las amplitudes de las funciones sinusoidales. Un valor más alto de seasonality prior scale permite que las amplitudes varíen más suavemente, lo que hace que el modelo sea más flexible. Por otro lado, un valor más bajo de seasonality prior scale penaliza las variaciones en las amplitudes, lo que hace que el modelo sea menos flexible.

En otras palabras el hiperparámetro seasonality prior scale se utiliza para ajustar las amplitudes de las funciones sinusoidales en la serie de tiempo. Específicamente, el hiperparámetro seasonality prior scale se utiliza como una escala para la prioridad asignada a la suavidad de las amplitudes de las funciones sinusoidales.

La prioridad para las amplitudes se modela utilizando una distribución Laplaciana⁴, que penaliza las variaciones abruptas en las amplitudes. La escala de la distribución Laplaciana se establece como el inverso del hiperparámetro seasonality prior scale. Un valor más alto de seasonality prior scale corresponde a una escala más baja de la distribución Laplaciana, lo que significa que se permiten variaciones más abruptas en las amplitudes. Por otro lado, un valor más bajo de seasonality prior scale corresponde a una escala más alta de la distribución Laplaciana, lo que penaliza las variaciones en las amplitudes y hace que el modelo sea menos flexible.

3. Holidays prior scale: Este parámetro controla la flexibilidad para hacerle fit a los días marcados como festivos o importantes. También tiene valores que van desde el 0.01 hasta el 10.
4. Change point range: Esta es la proporción de la data histórica en donde la tendencia tiene permitido cambiar. Este valor predeterminado es 0.8, 80 por ciento del historial, lo que significa que el modelo no se ajustará a ningún cambio de tendencia en el último 20 por ciento de la serie de tiempo. Los valores pueden ir desde 0.8 hasta 0.95.

Además de lo indicado anteriormente, todos los modelos son mejorados con la inclusión de un nuevo regresor, los días festivos de Chile, a través del método add country holidays, el cual crea nuevas columnas con todos los festivos y rellena el valor según sea la variación respecto de la media.

⁴La distribución Laplaciana es una distribución de probabilidad que se utiliza para modelar la incertidumbre en los datos y para ajustar modelos que penalizan las variaciones abruptas en los parámetros del modelo.

Capítulo 5

Resultados

Los resultados indicados en el Cuadro 5.1, corresponden a los valores obtenidos de MAPE en las tres cafeterías. Se puede notar que en todas se obtiene un MAPE bajo el 20 por ciento lo que permite demostrar que el modelo es bueno y exacto. Esto se debe a que Prophet se basa principalmente en las estacionalidades de los datos y en el caso de las cafeterías existen estacionalidades diarias, semanales, mensuales y anuales. Además de lo anterior el modelo con optimización de hiperparámetros obtuvo un mejor resultado en el caso de la Cafetería B, lo que también demuestra el poder de flexibilización del modelo para adaptarse a operaciones más complejas en beneficio de obtener mejores resultados.

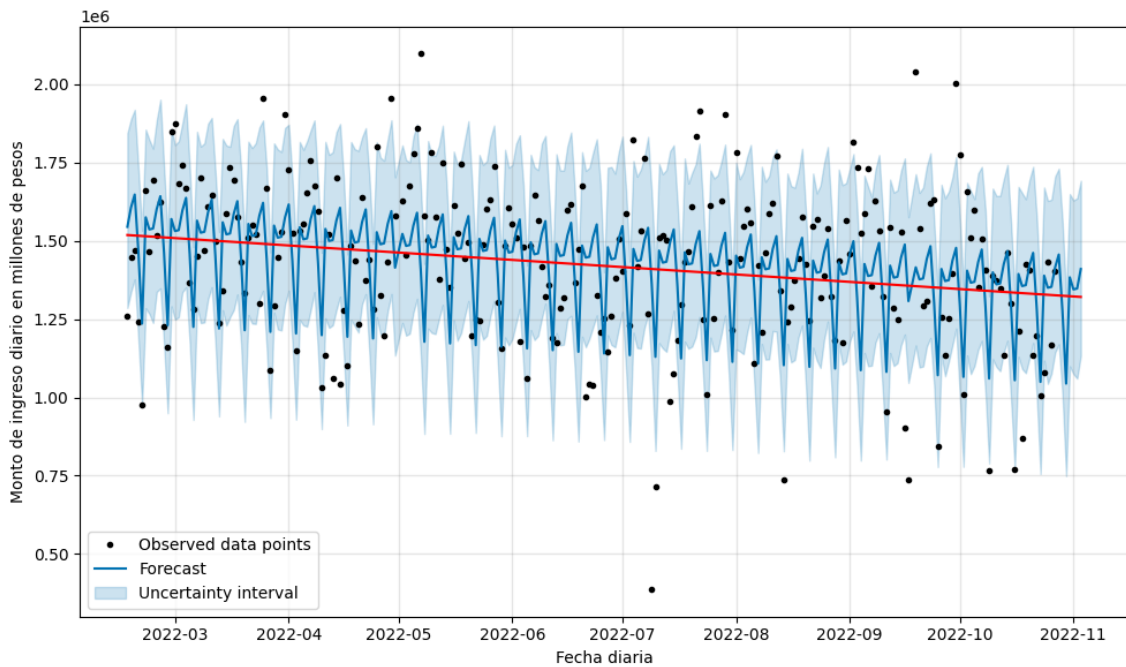
Cafetería A	Cafetería B	Cafetería B Optimizado	Cafetería C
0.139	0.186	0.179	0.187

Cuadro 5.1: Resultados de MAPE para las tres cafeterías.

Las siguientes Figuras 5.1, 5.2 y 5.3 muestran gráficamente los pronósticos para las cafeterías A, B y C del modelo sin optimización. La línea roja constante destaca la tendencia, las líneas punteadas de color rojo indican la presencia de changepoints, la línea azul muestra el pronóstico y los puntos negros son los datos históricos de ventas. En el modelo no optimizado se consideran los hiperparámetros por defecto, es decir, un changepoint prior scale de 0,05 lo que indica que el modelo variará en un 0.5 por ciento la tendencia en el punto de cambio (líneas punteadas rojas); un changepoint range de 0.8, es decir, que el modelo no calzará ningún cambio en la tendencia en el último 20 por ciento de la serie de tiempo; un seasonality prior scale de 10 indica que el modelo carece de regularización, indicando que no hay ningún

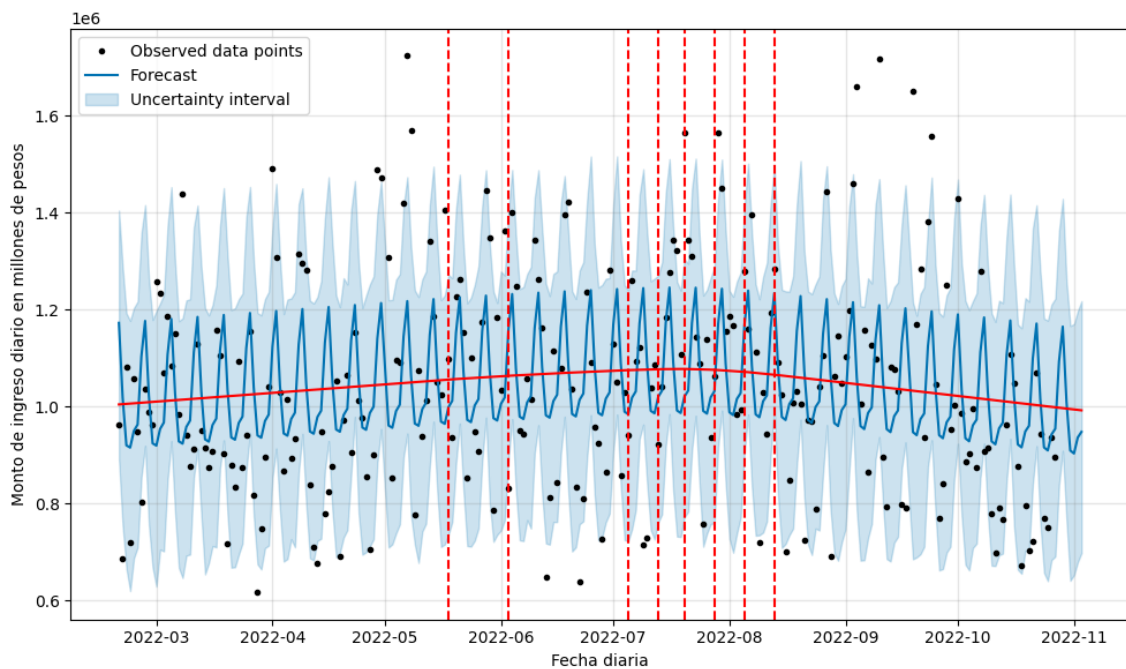
control sobre la flexibilidad en la estacionalidad y un holidays prior scale de 10 que al igual que el parámetro anterior indica que no tiene ningún tipo de regularización.

Figura 5.1: Pronóstico Cafetería A sin Optimización (Elaboración propia, 2022)



Del gráfico de la Figura 5.1 se desprende principalmente una tendencia a la baja, aunque constante y sin mayores cambios bruscos en esta. Esto se debe a que la Cafetería A es la de mayor trayectoria de las tres y que por lo tanto tiene menores fluctuaciones.

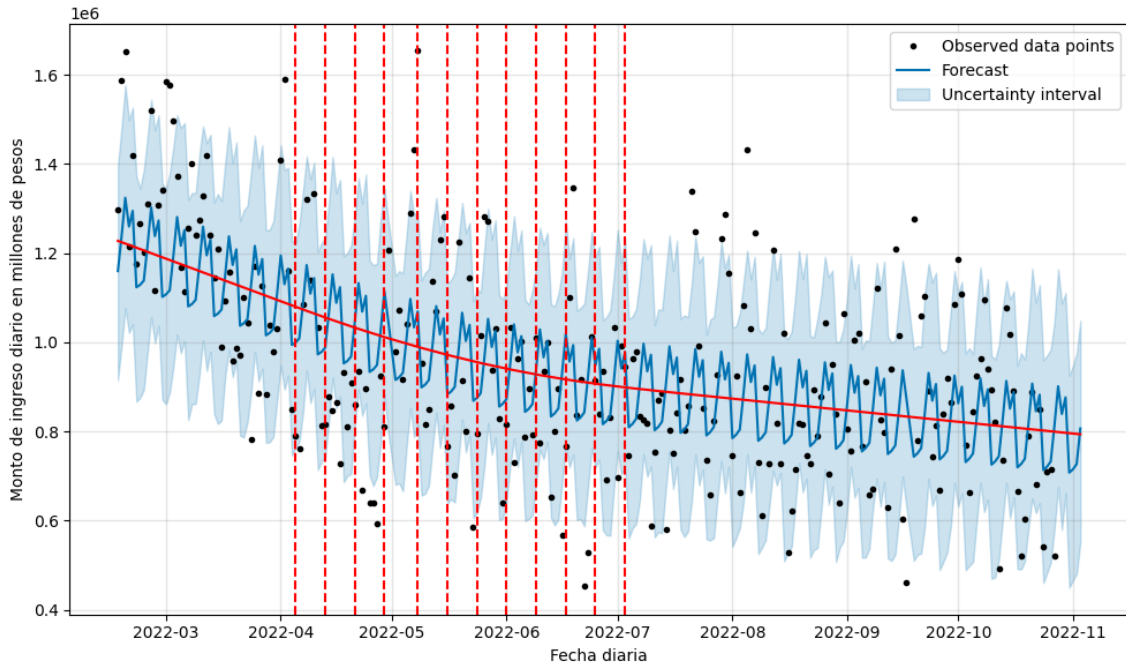
Figura 5.2: Pronóstico Cafetería B sin Optimización (Elaboración propia, 2022)



El gráfico de la Figura 5.2 muestra también una tendencia a la baja, con dos cambios notorios, uno producido en Mayo, que se explica principalmente por el cierre de una cafetería del sector y otro en Agosto en línea a la falta de liquidez, falta de ahorro y mayor endeudamiento en los hogares chilenos, tal como lo muestra el último Informe de Política Monetaria del Banco Central de Chile.¹

¹<https://www.bcentral.cl/web/banco-central/areas/politica-monetaria/informe-de-politica-monetaria>

Figura 5.3: Pronóstico Cafetería C sin Optimización (Elaboración propia, 2022)



El gráfico de la Figura 5.3 es el que tiene mayor changepoints debido principalmente al ajuste desde ser una cafetería recién abierta a ya tener 8 meses en el mercado, es decir, mayores ventas al inicio para luego ir tendiendo a períodos de más bajos ingresos.

Los gráficos anteriores muestran los pronósticos para las tres cafeterías, sin embargo, es preciso mencionar que para que el modelo sea capaz de crear este pronóstico, necesita de la adición de tres componentes: la tendencia, que es una línea que busca ser trazada lo más central posible respecto de las variaciones, los holidays o días festivos, es decir que tanto afectan positiva o negativamente el promedio de ingresos en esos días y el ciclo semanal el cual indica las diferencias promedios en los ingresos entre los distintos días de la semana.

Las siguientes Figuras 5.4, 5.5 y 5.6 muestran el comportamiento de estos componentes para cada una de las cafeterías.

Figura 5.4: Componentes Cafetería A (Elaboración propia, 2022)

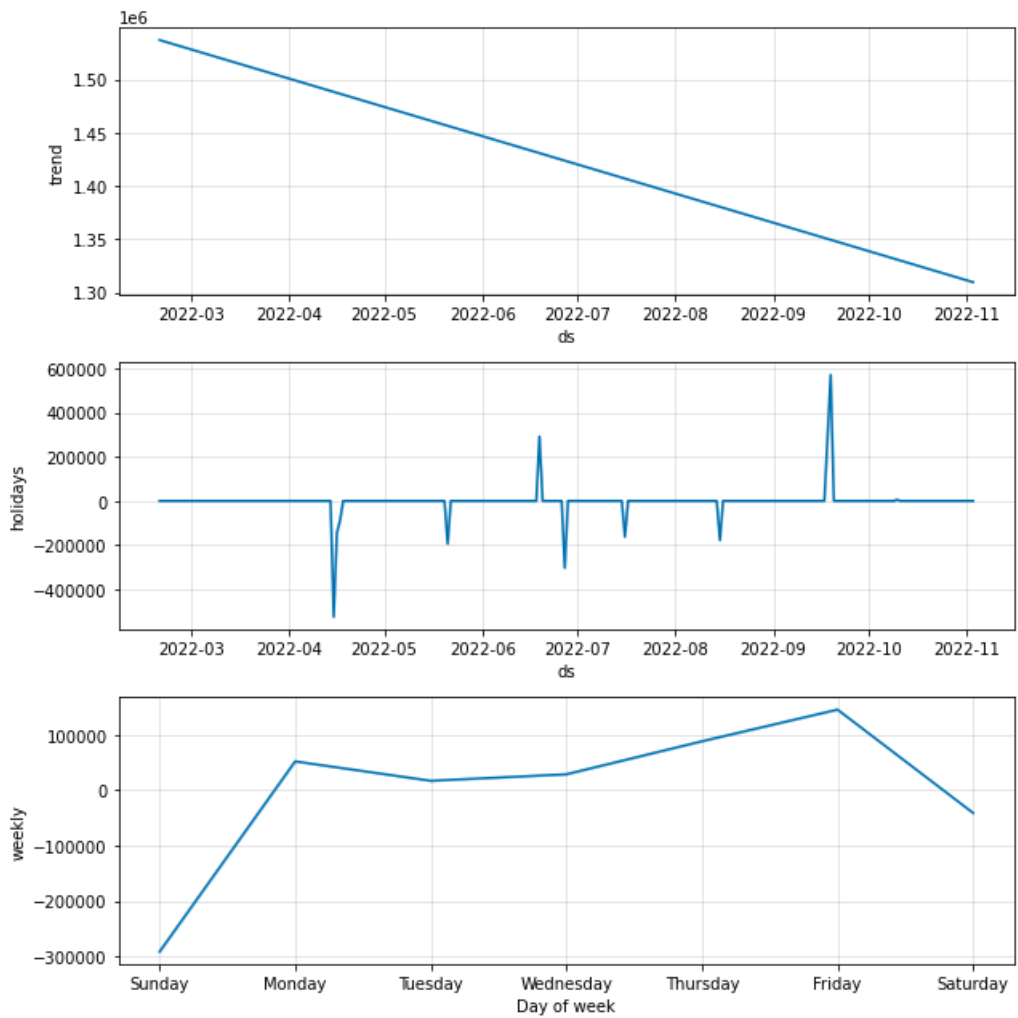


Figura 5.5: Componentes Cafetería B(Elaboración propia, 2022)

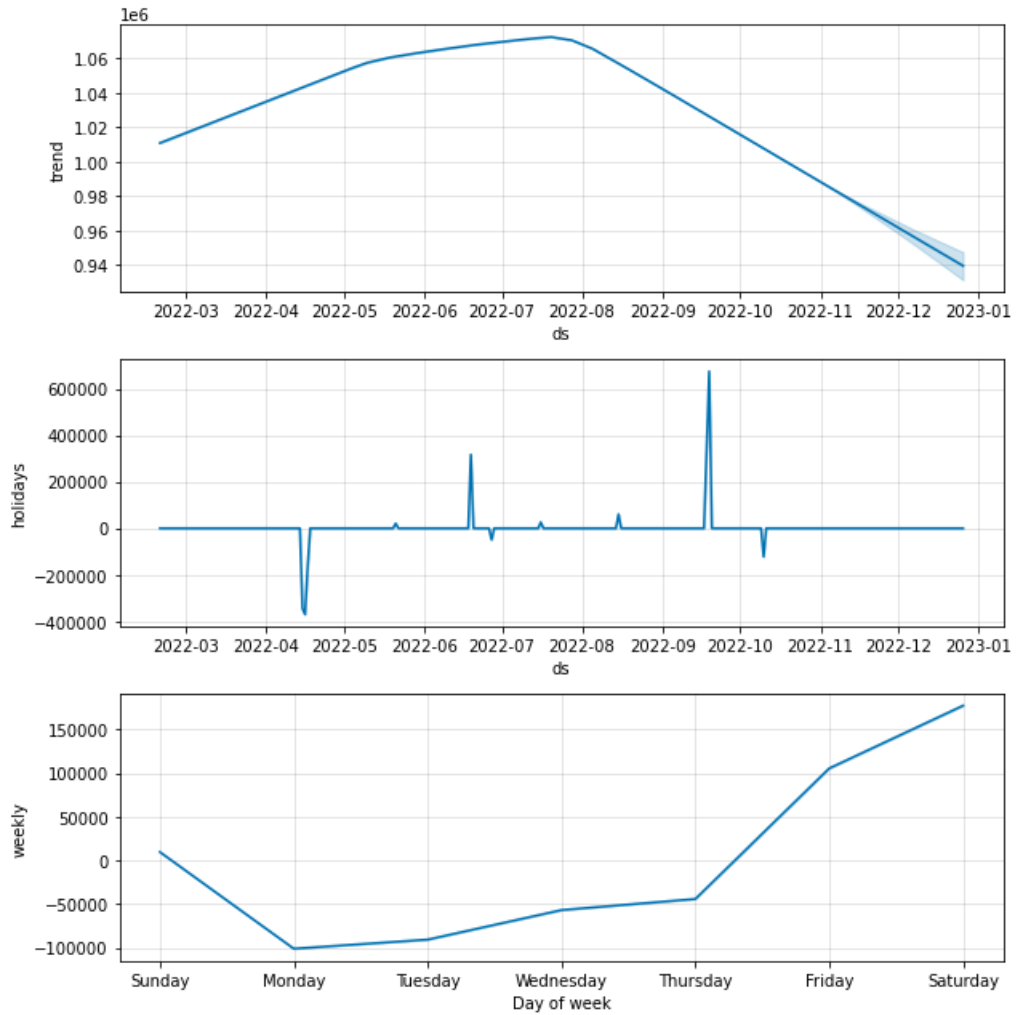
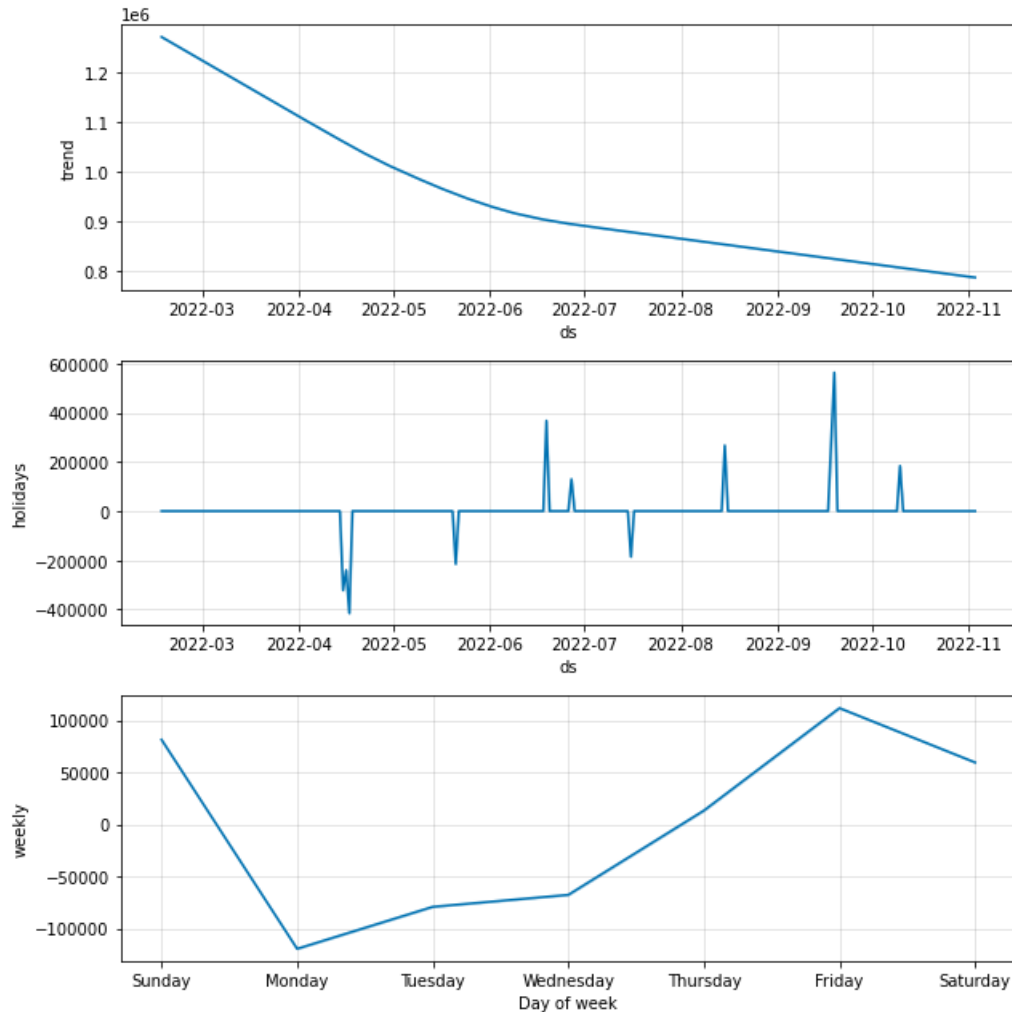


Figura 5.6: Componentes Cafetería C(Elaboración propia, 2022)



En las imágenes anteriores se pueden observar principalmente los componentes de tendencia subyacente de la serie de tiempo, que es la dirección general del cambio en los ingresos diarios ignorando las fluctuaciones y variaciones aleatorias de corto plazo. Se puede ver si la tendencia es lineal, no lineal, creciente, decreciente, estable o cambiante a lo largo del tiempo. Para este caso en todas se ve una tendencia lineal decreciente a lo largo del tiempo. Por otro lado tenemos la estacionalidad de la serie de tiempo, que es la variación periódica de los ingresos diarios que se repite regularmente. Se puede ver si la estacionalidad es mensual, semanal, diaria o en otros

periodos, y si es positiva o negativa. Para estos casos la estacionalidad es únicamente semanal, es decir indica el comportamiento diario de ventas durante la semana. Es interesante notar la similitud entre la cafetería B y C, aunque con pequeños cambios como el fin de semana. Por último se tienen los holidays que muestra los efectos de los días festivos en la serie de tiempo. Se pueden identificar los días festivos más importantes y ver cómo afectan a los ingresos diarios. En resumen, el gráfico de componentes proporciona información valiosa sobre la tendencia, la estacionalidad y los días festivos y la contribución de cada componente a los ingresos diarios, lo que puede ser útil para la toma de decisiones y la mejora de los pronósticos.

Capítulo 6

Conclusiones y trabajo futuro

Luego de finalizado el proyecto se puede concluir que las dificultades abordadas por las pequeñas empresas para la aplicación de ML en sus negocios son ciertas y que pueden en parte verse mejoradas debido a la existencia de nuevas herramientas que buscan facilitar su ejecución, como es el caso del modelo Prophet.

Se pudo descubrir que el modelo en su simpleza, es capaz de pronosticar de manera buena y exacta los ingresos para los próximos 7 días en tres pequeñas cafeterías de la V región, obteniendo para la cafetería A un MAPE de 0.139, para la Cafetería B de 0.186 y para la cafetería C de 0.179. Es interesante mencionar que estas tres pequeñas empresas tienen diferencias en madurez en el mercado, ubicaciones, público objetivo, ingresos y ciclos semanales, por lo que se puede indicar que el modelo cumple en generalidad.

Se logra mostrar además, la capacidad que tiene el modelo de mejorar a través de la optimización de hiperparámetros mejorando el resultado de la Cafetería B desde 0.186 a 1.179.

Por el lado de la planificación futura se pretende lograr lo siguiente:

1. Añadir nuevos regresores al modelo con el fin de seguir optimizándolo.
2. Crear una herramienta que permita cargar una base de datos con dos columnas con el formato que sea y que el código sea capaz de transformarlo al formato necesario para iniciar la modelación.
3. Medir el impacto real que tiene un forecasting en una pequeña empresa.

Bibliografía

- [1] Saleema Amershi et al. “Software Engineering for Machine Learning: A Case Study”. En: *2019 IEEE/ACM 41st International Conference on Software Engineering: Software Engineering in Practice (ICSE-SEIP)*. 2019, págs. 291-300. DOI: 10.1109/ICSE-SEIP.2019.00042.
- [2] Clemens; Bauer Markus; van Dinther y Daniel Kiefer. “Machine Learning in SME: An Empirical Study on Enablers and Success Factors”. En: *AMCIS 2020 Proceedings* 3.1 (2020), págs. 15-29. DOI: <https://doi.org/10.1016/j.jfbs.2014.01.007>.
- [3] Enrique José De la Hoz Domínguez, Tomás José Fontalvo Herrera y Adel Alfonso Mendoza Mendoza. “Aprendizaje automático y PYMES: Oportunidades para el mejoramiento del proceso de toma de decisiones”. En: *Investigación e Innovación en Ingenierías* 8.1 (ene. de 2020), págs. 21-36. DOI: 10.17081/invinno.8.1.3506. URL: <https://revistas.unisimon.edu.co/index.php/innovacioning/article/view/3506>.
- [4] Alfredo De Massis y Josip Kotlar. “The case study method in family business research: Guidelines for qualitative scholarship”. En: *Journal of Family Business Strategy* 5.1 (2014), págs. 15-29. DOI: <https://doi.org/10.1016/j.jfbs.2014.01.007>.
- [5] Marc Deloof y Maurizio La Rocca. “Local financial development and the trade credit policy of Italian SMEs”. En: *Small business economics* 44.4 (2015), págs. 905-924. ISSN: 0921-898X.
- [6] Amazon Web Services Expert Panel. *Amazon Machine Learning: Developer Guide*. 2021. URL: <https://docs.aws.amazon.com/machine-learning/latest/dg/model-fit-underfitting-vs-overfitting.html>.

- [7] Forbes Technology Council Expert Panel. *15 Smart Ways Machine Learning Helps Businesses And Entrepreneurs*. 2019. URL: <https://www.forbes.com/sites/forbestechcouncil/2019/09/11/15-smart-ways-machine-learning-helps-businesses-and-entrepreneurs/?sh=45a3934c7f67>.
- [8] IBM Expert Panel. *IBM, Global AI Adoption Index, 2021*. 2021. URL: <https://newsroom.ibm.com/IBMs-Global-AI-Adoption-Index-2021>.
- [9] Daniel Faggella. *How Machine Learning will become accessible to small companies*. 2019. URL: <https://emerj.com/ai-future-outlook/how-machine-learning-will-become-accessible-to-small-businesses/>.
- [10] John P. Hayes, Sudhir K. Chawla y Yunus Kathawala. “A comparative study of problems encountered in the development of small businesses in the U.S. and Mexico”. En: *The Journal of Developing Areas* 49.3 (2015), págs. 395-406. ISSN: 0022037X. URL: <http://www.jstor.org/stable/24737328>.
- [11] Charles Hill et al. “Trials and tribulations of developers of intelligent systems: A field study”. En: *2016 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC)*. 2016, págs. 162-170. DOI: 10.1109/VLHCC.2016.7739680.
- [12] Clark Hu, Ming Chen y Shiang-Lih Chen McCain. “Forecasting in Short-Term Planning and Management for a Casino Buffet Restaurant”. En: *Journal of Travel & Tourism Marketing* 16.2-3 (2004), págs. 79-98. DOI: 10.1300/J073v16n02_07. eprint: https://doi.org/10.1300/J073v16n02_07. URL: https://doi.org/10.1300/J073v16n02_07.
- [13] Fuyuki Ishikawa y Nobukazu Yoshioka. “How Do Engineers Perceive Difficulties in Engineering of Machine-Learning Systems? - Questionnaire Survey”. En: *2019 IEEE/ACM Joint 7th International Workshop on Conducting Empirical Studies in Industry (CESI) and 6th International Workshop on Software Engineering Research and Industrial Practice (SERIP)*. 2019, págs. 2-9. DOI: 10.1109/CESSER-IP.2019.00009.
- [14] Sacide Kalaycı, Mustafa Kamasak y Seçil Arslan. “Credit risk analysis using machine learning algorithms”. En: *2018 26th Signal Processing and Communications Applications Conference (SIU)*. 2018, págs. 1-4. DOI: 10.1109/SIU.2018.8404353.
- [15] Miryung Kim et al. “Data Scientists in Software Teams: State of the Art and Challenges”. En: *IEEE Transactions on Software Engineering* 44.11 (2018), págs. 1024-1038. DOI: 10.1109/TSE.2017.2754374.

- [16] Neil Lee, Hiba Sameen y Marc Cowling. “Access to finance for innovative SMEs since the financial crisis”. En: *Research Policy* 44.2 (2015), págs. 370-380. ISSN: 0048-7333. DOI: <https://doi.org/10.1016/j.respol.2014.09.008>. URL: <https://www.sciencedirect.com/science/article/pii/S0048733314001644>.
- [17] Farah Margaretha y Nina Supartika. “Factors Affecting Profitability of Small Medium Enterprises (SMEs) Firm Listed in Indonesia Stock Exchange”. En: 4.2 (2016), págs. 132-137. DOI: 10.7763/JOEBM.2016.V4.379.
- [18] Elizamary de Souza Nascimento et al. “Understanding Development Process of Machine Learning Systems: Challenges and Solutions”. En: *2019 ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM)*. 2019, págs. 1-6. DOI: 10.1109/ESEM.2019.8870157.
- [19] Andrew Ng. *AI Transformation Playbook How to lead your company into the AI era*. 2018. URL: <https://landing.ai/resources/ai-transformation-playbook/>.
- [20] Peter Quartey et al. “Financing the growth of SMEs in Africa: What are the constraints to SME financing within ECOWAS?” En: *Review of Development Finance* 7.1 (2017). Financial Sector Development in Africa: An Overview, págs. 18-28. ISSN: 1879-9337. DOI: <https://doi.org/10.1016/j.rdf.2017.03.001>. URL: <https://www.sciencedirect.com/science/article/pii/S1879933717300362>.
- [21] Sean J. Taylor y Benjamin Letham. “Forecasting at Scale”. En: *The American Statistician* 72.1 (2018), págs. 37-45. DOI: 10.1080/00031305.2017.1380080. URL: <https://doi.org/10.1080/00031305.2017.1380080>.
- [22] You Zhu et al. “Predicting China’s SME Credit Risk in Supply Chain Finance Based on Machine Learning Methods”. En: *Entropy* 18.5 (2016). ISSN: 1099-4300. DOI: 10.3390/e18050195. URL: <https://www.mdpi.com/1099-4300/18/5/195>.

Código Modelo sin Optimizar

```
1 # Importamos las librerías necesarias.
2
3 import pandas as pd
4 from prophet import Prophet
5 from prophet.plot import add_changepoints_to_plot
6 import matplotlib.pyplot as plt
7 from prophet.diagnostics import performance_metrics,
8   cross_validation
9 from prophet.plot import plot_cross_validation_metric
10 import random
11
12 random.seed(1234)
13
14 # Se abre el archivo
15 file_path = 'file.csv'
16 df = pd.read_csv(file_path)
17
18 # Creacion del df .
19 def crear_df(df):
20     df.rename(columns={"Date": "ds", "Credit": "y"}, inplace=True)
21     df['ds'] = pd.to_datetime(df['ds'], format='%d/%m/%Y')
22     df = df.sort_values(by=['ds'], ascending=True)
23     df.reset_index(drop=True, inplace=True)
24     df.drop_duplicates(['ds'], keep='last', inplace=True)
25     return df
26
27
28 df = crear_df(df)
29
30
31 # Modelo Prophet.
32 m2 = Prophet()
```

```

33
34 # Se agregan los festivos de Chile.
35 m2.add_country_holidays(country_name='Chile')
36 m2.fit(df)
37
38 # Se indica el periodo a pronosticar.
39 future2 = m2.make_future_dataframe(periods=7)
40 future2.tail()
41
42 # Prediccion
43 forecast2 = m2.predict(future2)
44 resumen_forecast = forecast2[['ds', 'yhat', 'yhat_lower',
45                               'yhat_upper']].
46     tail(14)
47 forecast2[['yhat_upper']].sum()
48
49 # Se grafican los componentes del forecast (trend, weekly, yearly)
50 fig6 = m2.plot_components(forecast2)
51 # plt.title('Componentes del forecast')
52 plt.show()
53
54 # Se grafica el pronostico.
55 fig7 = m2.plot(forecast2)
56 a2 = add_changepoints_to_plot(fig7.gca(), m2, forecast2)
57 # plt.title('Forecasting')
58 plt.show()
59
60 # Crossvalidation para encontrar el mejor MAPE
61 df_cv2 = cross_validation(m2, initial='150 days', horizon='7 days',
62                           period='1 days', parallel='processes')
63
64 df_p2 = performance_metrics(df_cv2, rolling_window=1)
65
66 fig8 = plot_cross_validation_metric(df_cv2, metric='mape')
67 plt.title('MAPE')
68 plt.show()
69
70 print('MAPE: %.3f' % df_p2['mape'])

```

Listing 6.1: Código en Python Modelo sin optimización.

Búsqueda de hiperparámetros

```
1 import pandas as pd
2 from prophet import Prophet
3 from prophet.diagnostics import performance_metrics,
   cross_validation
4 import itertools
5 import numpy as np
6 import random
7
8 random.seed(1234)
9
10 # IO
11
12
13 file_path = 'file.csv'
14 df = pd.read_csv(file_path)
15
16
17 # Creacion del df
18 def crear_df(df):
19     df.rename(columns={"Date": "ds", "Credit": "y"}, inplace=True)
20     df['ds'] = pd.to_datetime(df['ds'], format='%d/%m/%Y')
21     df = df.sort_values(by=['ds'], ascending=True)
22     df.reset_index(drop=True, inplace=True)
23     df.drop_duplicates(['ds'], keep='last', inplace=True)
24     return df
25
26
27 df = crear_df(df)
28
29 param_grid = {
30     'changepoint_prior_scale': [0.01, 0.1, 0.4],
31
32     'changepoint_range': [0.8, 0.90, 0.95],
```

```

33         'seasonality_prior_scale': [0.1, 0.5, 1, 5, 10],
34
35         'holidays_prior_scale': [0.1, 0.5, 1, 5, 10]
36     }
37 }
38
39 # Generando toda la combinacion de parametros.
40 all_params = [dict(zip(param_grid.keys(), v))
41               for v in itertools.product(*param_grid.values())]
42 mapes = [] # Guarda los MAPES para cada parametro probado.
43
44 # Usamos crossvalidation para probar estos parametros.
45 for params in all_params:
46     m = Prophet(**params)
47     m.add_country_holidays(country_name='Chile')
48     m.fit(df)
49     df_cv = cross_validation(m, initial='150 days', horizon='7 days'
50                             ,
51                             period='1 days', parallel='processes')
52     df_p = performance_metrics(df_cv, rolling_window=1)
53     mapes.append(df_p['mape'].values[0])
54
55 # Encontrando el mejor parametro.
56 tuning_results = pd.DataFrame(all_params)
57 tuning_results['mape'] = mapes
58
59 # Se imprime el mejor parametro.
60 best_params = all_params[np.argmin(mapes)]
61 print(best_params)

```

Listing 6.2: Código en Python Búsqueda de Hiperparámetros.

Código Modelo Optimizado

```
1 # Importamos las librerías necesarias.
2
3 import pandas as pd
4 from prophet import Prophet
5 from prophet.plot import add_changepoints_to_plot
6 import matplotlib.pyplot as plt
7 from prophet.diagnostics import performance_metrics,
8     cross_validation
9 from prophet.plot import plot_cross_validation_metric
10 import random
11
12 random.seed(1234)
13
14 # Se abre el archivo
15 file_path = 'file.csv'
16 df = pd.read_csv(file_path)
17
18 # Creacion del df .
19 def crear_df(df):
20     df.rename(columns={"Date": "ds", "Credit": "y"}, inplace=True)
21     df['ds'] = pd.to_datetime(df['ds'], format='%d/%m/%Y')
22     df = df.sort_values(by=['ds'], ascending=True)
23     df.reset_index(drop=True, inplace=True)
24     df.drop_duplicates(['ds'], keep='last', inplace=True)
25     return df
26
27
28 df = crear_df(df)
29
30
31 # Modelo Prophet Optimizado.
32 m2 = Prophet(changepoint_prior_scale=0.1, changepoint_range=0.8,
```

```

33         seasonality_prior_scale=1, holidays_prior_scale=1)
34
35 # Se agregan los festivos de Chile.
36 m2.add_country_holidays(country_name='Chile')
37 m2.fit(df)
38
39 # Se indica el periodo a pronosticar.
40 future2 = m2.make_future_dataframe(periods=7)
41 future2.tail()
42
43 # Prediccion
44 forecast2 = m2.predict(future2)
45 resumen_forecast = forecast2[['ds', 'yhat', 'yhat_lower',
46                               'yhat_upper']].
47     tail(14)
48 forecast2[['yhat_upper']].sum()
49
50 # Se grafican los componentes del forecast (trend, weekly, yearly)
51 fig6 = m2.plot_components(forecast2)
52 # plt.title('Componentes del forecast')
53 plt.show()
54
55 # Se grafica el pronostico.
56 fig7 = m2.plot(forecast2)
57 a2 = add_changepoints_to_plot(fig7.gca(), m2, forecast2)
58 # plt.title('Forecasting')
59 plt.show()
60
61 # Crossvalidation para encontrar el mejor MAPE
62 df_cv2 = cross_validation(m2, initial='150 days', horizon='7 days',
63                           period='1 days', parallel='processes')
64
65 df_p2 = performance_metrics(df_cv2, rolling_window=1)
66
67 fig8 = plot_cross_validation_metric(df_cv2, metric='mape')
68 plt.title('MAPE')
69 plt.show()
70
71 print('MAPE: %.3f' % df_p2['mape'])

```

Listing 6.3: Código en Python Modelo con optimización.