

UNIVERSIDAD DE CONCEPCIÓN
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE ESTADÍSTICA



Tesis Para Optar al Título Profesional de Ingeniero Estadístico

**APLICACIÓN DE REDES NEURONALES RECURRENTE Y MODELOS
DE SERIES DE TIEMPO BAYESIANOS A LA PREDICCIÓN DE
RENTABILIDAD DE FONDOS DE PENSIONES**

Sergio Esteban Altamirano Pontigo
Concepción-Chile
2020

Profesor patrocinante: Sebastián Niklitschek Soto
Departamento de Estadística
Facultado de Ciencias Físicas y Matemáticas
Universidad de Concepción

UNIVERSIDAD DE CONCEPCIÓN
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE ESTADÍSTICA



Tesis Para Optar al Título Profesional de Ingeniero Estadístico



**APLICACIÓN DE REDES NEURONALES RECURRENTE Y MODELOS
DE SERIES DE TIEMPO BAYESIANOS A LA PREDICCIÓN DE
RENTABILIDAD DE FONDOS DE PENSIONES**

Profesor Patrocinante	: Sebastián Niklitschek Soto	Firma	
Profesor Colaborador	: Christopher Thraves Caro	Firma	
Profesor Consejero	: Guillermo Ferreira Cabezas	Firma	
	Nora Serdyukova	Firma	<i>Nora Serdyukova</i>
Nombre Memorante	: Sergio Altamirano Pontigo	Firma	



Dedicatoria

Dedicado a mis padres, por su sacrificio, apoyo y amor incondicional. A quienes debo principalmente mis logros. Gracias a ellos, soy la persona con valores en la que me convertí.



Agradecimientos

Agradezco a mi profesor guía, que, con su sabiduría, conocimiento y apoyo, motivaron a desarrollarme como persona y profesional. Agradezco a mis compañeros y amigos, que juntos pasamos una linda etapa en la universidad.

Agradezco a mi abuela, a mis tíos por permitirme llegar a esta familia y así poder estudiar acá en Concepción y por último, agradezco a mi pareja, Valentina, por su apoyo en todo momento, amor y respaldo, logrando así alcanzar mis objetivos.

Índice general

Índice de figuras	VII
Índice de tablas	IX
1 Introducción	1
2 Modelo de predicción Prophet	6
2.1 Modelos de tendencia	6
2.1.1 Crecimiento saturado no lineal	6
2.1.2 Tendencia lineal con puntos de cambio	7
2.2 Estacionalidad	8
2.3 Eventos	8
3 Redes neuronales recurrentes	10
3.1 Redes neuronales	10
3.2 Redes neuronales recurrentes	11
3.3 Long short-term memory	12
4 Estrategias de validación	19
4.1 Validación cruzada	19
4.2 Validación cruzada de avance	19
5 Proceso experimental	21
5.1 Conjuntos de datos	21
5.1.1 Valor Cuota	22
5.1.2 Rentabilidad real	22
5.2 Indicadores de evaluación experimental	22
5.3 Proceso de ajuste experimental	23
5.3.1 Prophet	23
5.3.2 Long short-term memory	36
5.4 Modelos Propuestos	51
6 Resultados experimentales	54
6.1 Valor cuota	54

6.2 Rentabilidad real	58
6.3 Discusión de resultados	62
7 Conclusión	64
Referencias	66



Índice de figuras

Figura 1.1	Rentabilidad real histórica de los fondos de pensiones	3
Figura 3.1	Esquema de una arquitectura sobre una red neuronal simple	11
Figura 3.2	Procesamiento secuencial de una red neuronal recurrente	12
Figura 3.3	Estructura de una red neuronal híbrida <i>Long short-term memory</i>	13
Figura 3.4	Unidad <i>Long short-term memory</i> con sus respectivos componentes	14
Figura 3.5	Modelo de red neuronal con Dropout	18
Figura 4.1	Validación de avance con k bloques	20
Figura 5.1	Estructura del proceso de optimización para los hiper-parámetros de Prophet	23
Figura 5.2	Valores RMSE y Sentido para distintos modelos de tendencia sobre el valor cuota	25
Figura 5.3	Valores RMSE y Sentido para distintas componentes estacionales sobre el valor cuota	26
Figura 5.4	RMSE y Sentido al incluir listado de feriados y eventos al modelo sobre el valor cuota	27
Figura 5.5	Valores RMSE y Sentido para distintas regresoras sobre el valor cuota	29
Figura 5.6	Valores RMSE y Sentido para distintos modelos de tendencia sobre la rentabilidad real	30
Figura 5.7	Valores RMSE y Sentido para distintas componentes estacionales sobre la rentabilidad real	32
Figura 5.8	Valores RMSE y Sentido al incluir fechas y eventos sobre la rentabilidad real	33
Figura 5.9	Valores RMSE y Sentido al incluir variables regresoras sobre la rentabilidad real	34
Figura 5.10	Valores RMSE y Sentido al modificar el porcentaje de la serie en la rentabilidad	35
Figura 5.11	RMSE y sentido de la tendencia al modificar el valor a priori τ	36
Figura 5.12	Promedio RMSE y Sentido de la tendencia al utilizar distintos números de días	37
Figura 5.13	Promedio RMSE y Sentido para distintos valores de neuronas e iteraciones sobre los valores cuota	39

Figura 5.14 Promedio RMSE y Sentido al utilizar distintos optimizadores sobre los valores cuota	40
Figura 5.15 Promedio RMSE y Sentido para distintas funciones de activación en la capa intermedia sobre los valores cuota	41
Figura 5.16 Promedio RMSE y Sentido para distintas tasas de <i>dropout</i> sobre los valores cuota	42
Figura 5.17 Promedio RMSE y Sentido al incluir cero o más capas densas sobre los valores cuota	43
Figura 5.18 Promedio RMSE y Sentido para distintas funciones de activación en la capa de combinación de salida sobre los valores cuota	44
Figura 5.19 Promedio RMSE y Sentido para distintos días como variable de entrada sobre la rentabilidad	45
Figura 5.20 Promedio RMSE y Sentido para distintos valores de neuronas e iteraciones sobre la rentabilidad	46
Figura 5.21 Promedio RMSE y Sentido al utilizar distintos optimizadores sobre la rentabilidad	47
Figura 5.22 Promedio RMSE y Sentido para distintas funciones de activación en la capa de combinación de salida sobre la rentabilidad	48
Figura 5.23 Promedio RMSE y Sentido para distintas tasas de <i>dropout</i> sobre la rentabilidad	49
Figura 5.24 Promedio RMSE y Sentido al incluir cero o más capas densas sobre la rentabilidad	50
Figura 5.25 Promedio RMSE y Sentido para distintas funciones de activación en la capa de combinación de salida sobre la rentabilidad	51
Figura 6.1 Ajuste de los modelos Prophet y <i>Long short-term memory</i> para los valores cuota en el fondo A	57
Figura 6.2 Ajuste de los modelos Prophet y <i>Long short-term memory</i> para los valores cuota en el fondo E	58
Figura 6.3 Ajuste de los modelos Prophet y <i>Long short-term memory</i> para la rentabilidad real fondo A	61
Figura 6.4 Ajuste de los modelos Prophet y <i>Long short-term memory</i> para la rentabilidad real fondo E	61

Índice de tablas

Tabla 1.1	Límites de inversión en instrumentos de renta variable	2
Tabla 1.2	Número de veces que los fondos fueron más rentable en el 2019	5
Tabla 5.1	Valores para el parámetro τ	36
Tabla 5.2	Número de días a probar como variable de entrada	37
Tabla 5.3	Cantidad de neuronas e iteraciones a probar en el modelo <i>Long short-term memory</i>	38
Tabla 5.4	Hiper-parámetros para los modelos <i>Long short-term memory</i> sobre el valor cuota y rentabilidad real	51
Tabla 5.5	Componentes de los modelos Prophet para los valores cuotas y rentabilidad real	52
Tabla 6.1	Resultados obtenidos con Prophet sobre los valores cuota utilizando la validación de avance	54
Tabla 6.2	Resultados obtenidos con <i>Long short-term memory</i> sobre los valores cuota utilizando la validación de avance	56
Tabla 6.3	Resultados obtenidos con Prophet sobre la rentabilidad real utilizando la validación de avance	59
Tabla 6.4	Resultados obtenidos con <i>Long short-term memory</i> sobre la rentabilidad real utilizando la validación de avance	60
Tabla 6.5	Promedio sentido de la tendencia para distintos fondos y técnicas estadísticas	62
Tabla 7.1	Rentabilidad real en porcentaje de los Fondos de Pensiones	64

Capítulo 1

Introducción

En una sociedad cada habitante posee el derecho a la Seguridad Social, donde el Estado es el encargado de brindar a su población prestaciones de salud, previsión social, de accidentes y enfermedades laborales, seguro de invalidez, desempleo, entre otros. Dentro de todos los beneficios que engloba este concepto, nos enfocaremos específicamente en uno, la previsión social. Esto considera principalmente las necesidades futuras, donde se desea brindar una protección a aquellas personas que están imposibilitadas de recibir un ingreso, contribuyendo a la cobertura de problemáticas sociales.

Chile, fue pionero en introducir medidas de provisiones sociales, el gobierno de Alessandri Palma fue el propulsor de estas "leyes sociales", donde se brindó seguros sociales que se separaban por segmentos dependiendo de la profesión que el individuo desempeñaba. Luego, en los años 60, se unificó el sistema de seguridad social sin privilegios y distinciones, creando un fondo único mantenido por los aportes de los trabajadores, los empleadores y el Estado. Sin embargo, con el transcurso de los años, un acelerado crecimiento demográfico y la variabilidad económica, ocasionó que el costo que invertía el Estado en el sistema llegaría a un tercio del gasto fiscal. Por lo que en 1980 se publica el Decreto Ley (DL) 3.500 y se dio inicio a una nueva etapa en la historia de la seguridad social.

En mayo de 1981, se da inicio a la Ley que establece un sistema de pensiones financiado mediante capitalización individual y el cual es administrado por el sector privado. Este sistema propone contribuciones basadas en planes individuales de ahorro obligatorio, en la que un individuo cotiza mensualmente el 10 % de su renta a su cuenta de ahorro obligatorio administrado por las Administradoras de Fondo de Pensiones (AFP). Es así como finalmente las pensiones de los individuos dependen de factores tales como el total del ahorro obligatorio que han realizado durante su vida laboral y el retorno de las inversiones que realizan las AFP con el ahorro obligatorio del individuo, las cuales en un principio prometieron una pensión igual al ingreso mensual que poseían las personas.

El 19 de abril de 2001, se aprobó la Reforma al Mercado de Capitales I. El sistema de capitalización individual se cambió al sistema de multifondos. Así, el 28 de febrero de 2002

fue introducido el Sistema de multifondos mediante la Ley 19.795 que modificó el DL 3.500 pero no fue hasta el 01 de agosto del mismo año, cuando se pone en marcha este sistema, que principalmente corresponde a la administración de 5 Fondos de Pensiones por parte de cada AFP. Estos Tipos de Fondos se diferencian a partir de la proporción de su portafolio invertido en títulos de renta variable, lo que significa diferentes niveles de riesgo y rentabilidad entre cada uno.

Tabla 1.1: Límites de inversión máximos y mínimos en instrumentos de renta variable

	Límites máximo permitido	Límite mínimo obligatorio
Fondo A - Más riesgoso	80 %	40 %
Fondo B - Riesgoso	60 %	25 %
Fondo C - Intermedio	40 %	15 %
Fondo D - Conservador	20 %	5 %
Fondo E - Más Conservador	5 %	0 %

Fuente: Superintendencia de Pensiones.

En el Tabla 1.1 se observan los límites permitidos de inversión de renta variable por tipo de fondo, donde el Fondo A, denominado el fondo más riesgoso, posee un límite máximo de inversión de un 80%. Mientras que el Fondo E, denominado el más conservador, solo posee un máximo de inversión permitida de 5% en renta variable.

La Asociación de AFP (2018) afirma que la rentabilidad de los fondos de pensiones debe evaluarse en el largo plazo. Si uno observa todo el rendimiento desde la creación de los multifondos hasta fines del año 2019, el Fondo A y el Fondo E ha rentado, 6,55% y 4,75%, respectivamente.

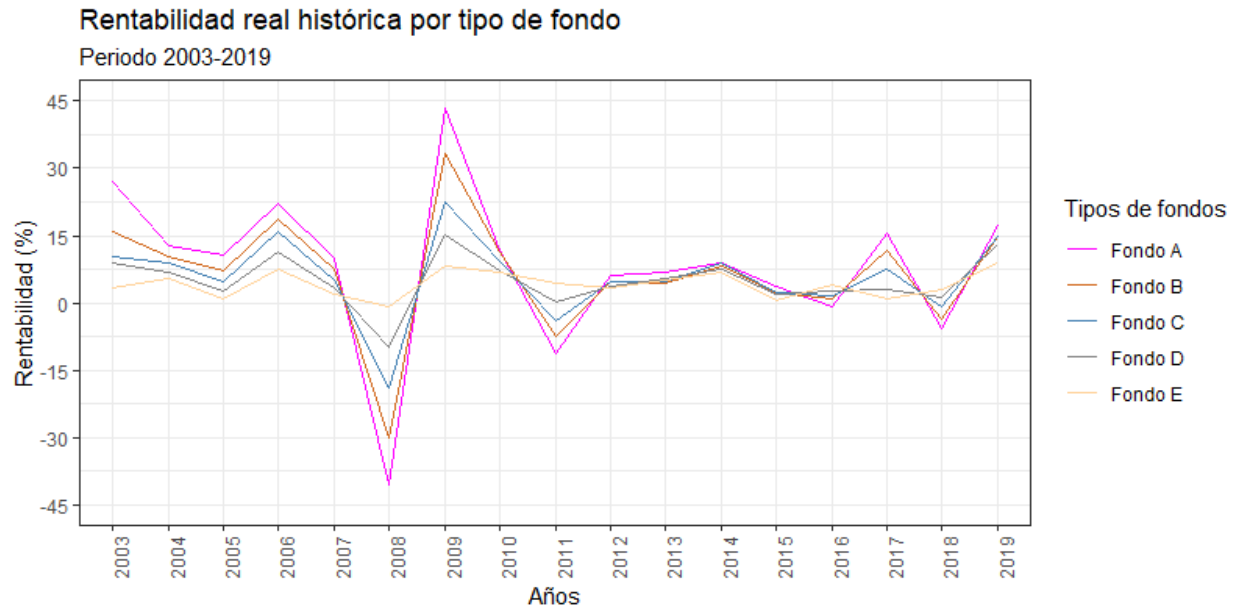


Figura 1.1: Serie de tiempo sobre la rentabilidad real de los cinco tipos de fondos desde el año 2003 hasta el 2019.

En la Figura 1.1 se presenta la rentabilidad real de los cinco tipos de fondos desde el año 2003 hasta el 2019. El año que presentó las rentabilidades más bajas fue el 2008, donde el Fondo A y el Fondo E tuvieron una rentabilidad negativa de 40,3 % y 0,9 %, respectivamente. Por otro lado, el año 2009 presentó uno de los resultados más sobresalientes de rentabilidad producto de la fase de recuperación de la crisis financiera internacional donde el Fondo A y el Fondo E tuvieron una rentabilidad de 43,5 % y 8,3 %, respectivamente.

Si nos centramos específicamente en los últimos años, destacados principalmente por el contexto de incertidumbre, generado por el conflicto comercial entre EE.UU. y China; la desaceleración económica local y externa; y la baja generalizada de las tasas de política monetaria, para el año 2019, los fondos de pensiones obtuvieron buenos retornos, gracias al desempeño de la inversión extranjera y de la renta fija local (Asociación de AFP, 2019). Pues en ese mismo año, a pesar del “estallido social” ocurrido en Chile, se obtuvo a nivel de sistema, el segundo mejor resultado después de 2009, llegando el Fondo A y el Fondo E, a rentar 17,26 % y 9,03 %, respectivamente.

El sistema de pensiones no ha estado exento de críticas, recientemente un estudio de la Fundación Sol describe que

La actual situación previsional es de extrema gravedad. Un país que pretende que el desarrollo humano y económico llegue a todos sus habitantes, no puede seguir sosteniendo un sistema de pensiones que obliga a sus adultos mayores a seguir trabajando y a endeudarse. (...) Es el sistema de pensiones el que debe adaptarse a la realidad de las personas que viven en un país y no al revés como ocurre actualmente en Chile. (Doniez and Kremerman, 2019, p.30)

A su vez, según datos de la Superintendencia de Pensiones, en el mes de diciembre del año 2019, el monto promedio recibido por los jubilados, fue de \$130.304 (4,56 UF). Datos preocupantes se dan a conocer con el reporte de febrero 2020 del Ministerio del Desarrollo Social, en él se expone que este monto está por debajo de la línea de la pobreza (\$169.883).

Teniendo en cuenta esto como punto de partida, las administradoras poseen distintas carteras de inversión en los fondos de pensión, que dependiendo de la situación, se generará mayor o menor rentabilidad. Es por esto que es de interés poder predecir el comportamiento de las rentabilidades de los distintos fondos de pensiones para así maximizar las ganancias generadas por las AFP y aumentar el monto de su cuenta individual, que a largo plazo mejoraría las pensiones.

Son muy pocas las herramientas disponibles que ayuden a la elección de un cambio de fondo, donde para su utilización, se necesitan conocimientos respecto a técnicas de mercado o modelos matemáticos. Existen solo los consejos que brinda tu AFP a la elección del fondo.

Actualmente en la literatura se ha intentado predecir la rentabilidad de los fondos de pensiones. Alarcón (2014) demuestra que las redes neuronales permiten predecir en forma estable el cambio de signo del valor cuota de los multifondos de la AFP CUPRUM, obteniendo un porcentaje de precisión del 64%. Por otro lado, Garrido (2014) y Ramos (2015) utilizan redes neuronales de base radial, pero con distintos algoritmos de aprendizaje, logrando una precisión de predicción del fondo A del 69,33% y 74,34%, respectivamente. Tras esto, son pocas las investigaciones que estudian la predicción de la rentabilidad de los fondos de pensiones, más aún usando redes neuronales recurrentes y modelos Bayesianos de series temporales.

Los cambios de fondo se demoran 4 días hábiles, por lo que, dependiendo del día, el trámite se realizará entre 6-7 días después. Al observar que fondo rentó más en el año 2019 (considerando que el cambio se ejecuta 7 días después), se tienen los siguientes resultados:

Tabla 1.2: Número de veces que los fondo fueron más rentable en el año 2019.

Tipos de fondos	Número de veces más rentable
Fondo A	111
Fondo B	6
Fondo C	13
Fondo D	33
Fondo E	93

Fuente: Elaboración propia.

Del total de 286 posibles cambios de fondos en el 2019, un 71,33 % corresponde al Fondo A o al Fondo E, mientras que el restante 28,67 % corresponde al Fondo B, C o D. Por lo que toma un protagonismo el Fondo A y el Fondo E, pues son los que abarcan las rentabilidades más altas un mayor número de veces en el año 2019, caracterizado por una disputa comercial entre EE. UU y China y por otro lado una crisis social vivido el 18 octubre en Chile.

Por lo que, en la siguiente investigación, se busca desarrollar un estudio comparativo sobre la aplicación de técnicas de predicción para series de tiempo a la rentabilidad de los fondos de pensiones. Específicamente, implementar y evaluar el rendimiento de las redes neuronales recurrentes y modelos bayesianos de series temporales sobre los fondos A y E con dos metodologías distintas.

Para poder llevar a cabo el estudio comparativo, se identificarán variables macroeconómicas que puedan influir en la rentabilidad de los fondos de pensiones, proponiendo índices bursátiles de los mercados financieros más representativos a nivel mundial. Por otro lado, se implementarán herramientas de web scraping para la extracción de información desde diferentes portales web con la finalidad de agilizar la extracción y actualización de nuestras bases de datos.

El trabajo se ordena de la siguiente forma: en el Capítulo 2 se presenta el modelo Bayesiano de series temporales y sus características esenciales, explicando cómo se estiman las componentes respectivas. En el Capítulo 3 se presentan el estado del arte de las redes neuronales y la evolución hasta llegar a las redes neuronales *Long short-term memory*, donde se explican la arquitectura específica a utilizar en este trabajo. En el Capítulo 4 se mencionan las técnicas de validación más utilizadas y la estrategia de validación a llevar a cabo. En el Capítulo 5 se explica el proceso experimental a llevar a cabo para ambas técnicas estadísticas, llegando a proponer los modelos óptimos de cada uno. En el Capítulo 6, se presenta un análisis de los resultados obtenidos y una discusión respecto a estos. Y, por último, en el Capítulo 7 se exponen las conclusiones obtenidas en esta investigación.

Capítulo 2

Modelo de predicción Prophet

Taylor and Letham (2018) proponen un modelo de pronóstico para series de tiempo capaz de representar las características más comunes de las series de tiempo comercial, es decir, que sean capaces de modelar la tendencia por partes, estacionalidad múltiple, efecto a la serie en fechas relevantes no periódicas y presencia de valores atípicos.

Es por esto, que se propone un modelo de serie temporal descompuesta (Harvey and Peters 1990) con tres componentes principales: tendencia, estacionalidad, y eventos. Estos combinados presentan la siguiente ecuación:

$$y(t) = g(t) + s(t) + h(t) + \epsilon_t, \quad (2.1)$$

donde $g(t)$ función de tendencia que modela los cambios no periódicos en el valor de la serie de tiempo, $s(t)$ representa cambios periódicos, $h(t)$ representa el efecto irregular producido por un evento durante uno o más días y ϵ_t representa cualquier cambio idiosincrásico que no sea explicado por el modelo.

2.1. Modelos de tendencia

Para modelar la tendencia se proponen dos modelos: modelo de crecimiento saturado y tendencia lineal a trozos.

2.1.1. Crecimiento saturado no lineal

Para representar el crecimiento, se usa una variación del modelo de crecimiento logístico, pues se considera que la capacidad de carga y la tasa de crecimiento no son constantes.

El modelo más básico del crecimiento logístico es:

$$g(t) = \frac{C}{1 + \exp(-k(t - m))}, \quad (2.2)$$

con C capacidad de carga, k tasa de crecimiento, y m parámetro de desplazamiento.

Dado las limitaciones de (2.2), se reemplaza la capacidad de carga C con una capacidad que dependa del tiempo $C(t)$, además de definir explícitamente los puntos de cambio donde se permita cambiar la tasa de crecimiento. Suponga que existe S puntos de cambio s_j , $j = 1, \dots, S$. Adicionalmente se define vector de ajuste $\delta \in \mathbb{R}^S$, donde δ_j es la tasa de cambio que ocurre en el momento s_j . La tasa en cualquier momento t es entonces la tasa en base k , más todos los ajustes hasta ese punto: $k + \sum_{j:t>s_j} \delta_j$. Tras esto, se define el vector $a(t) \in \{0, 1\}^S$ tal que

$$a(t) = \begin{cases} 1, & \text{si } t \geq s_j \\ 0, & \text{en otro caso.} \end{cases} \quad (2.3)$$

La tasa en el tiempo t es entonces $k + a(t)^T \delta$. Cuando se ajusta la tasa k , el parámetro de desplazamiento m también debe ajustarse para conectar los puntos finales de los segmentos. El ajuste correcto en el punto de cambio j se calcula fácilmente como

$$\gamma_j = \left(s_j - m - \sum_{l < j} \gamma_l \right) \left(1 - \frac{k + \sum_{l < j} \delta_l}{k + \sum_{l \leq j} \delta_l} \right). \quad (2.4)$$

Entonces el modelo crecimiento logístico por partes viene dado por

$$g(t) = \frac{C(t)}{1 + \exp(-(k + a(t)^T \delta)(t - (m + a(t)^T \gamma)))}. \quad (2.5)$$

2.1.2. Tendencia lineal con puntos de cambio

Para los problemas que no exhiben un crecimiento saturado, una tasa de crecimiento constante y por partes proporciona un modelo a menudo útil. Aquel modelo viene dado por

$$g(t) = (k + a(t)^T \delta)t + (m + a(t)^T \gamma), \quad (2.6)$$

donde k es la tasa de crecimiento, δ tiene los ajustes de la tasa, m parámetro de desplazamiento y γ_j se establece como $-s_j \delta_j$ para obtener una función continua.

Selección automática de puntos de cambio

Los puntos de cambio s_j pueden ser especificados por el analista utilizando fechas conocidas como lanzamientos de productos y otras alteraciones de crecimiento producto de algún evento. O bien, la selección automática de puntos de cambio puede formularse a través de (2.5) y (2.6) usando una priori pequeña para δ .

A menudo se especifica una gran cantidad de puntos de cambio, por lo que se utiliza $\delta_j \sim \text{Laplace}(0, \tau)$. El parámetro τ controla la flexibilidad del modelo.

2.2. Estacionalidad

Las series de tiempo a menudo tienen estacionalidad de periodos múltiples como resultado de los comportamientos humanos que representan. Por lo que se debe especificar modelos de estacionalidad que sean funciones periódicas de t .

Sea p el periodo regular que se espera de la serie de tiempo. Se puede aproximar los efectos de suavizado estacional con la estándar serie de Fourier.

$$s(t) = \sum_{n=1}^N (a_n \cos(\frac{2\pi nt}{p}) + b_n \sin(\frac{2\pi nt}{p})). \quad (2.7)$$

Ajustar la estacionalidad requiere $2N$ parámetros $\beta = [a_1, b_1, \dots, a_N, b_N]^T$. Por lo que se construye una matriz de vectores de estacionalidad ($X(t)$) para cada valor de t en nuestros datos históricos y futuros.

Entonces la componente estacional es

$$s(t) = X(t)\beta. \quad (2.8)$$

Para el modelo genérico, tomamos $\beta \sim N(0, \sigma_\beta^2)$ para imponer una priori de suavizado estacional.

2.3. Eventos

Los días festivos y los eventos proporcionan alta volatilidad y, a menudo no siguen un patrón periódico, por lo que su efecto no es bien modelado por un ciclo suave. Para cada evento i , sea D_i el conjunto de datos pasado y futuro para cada evento. Adicionalmente se incluye una función indicadora que representa si el tiempo t está durante el evento i .

$$Z(t) = [\mathbb{1}(t \in D_1), \dots, \mathbb{1}(t \in D_L)]. \quad (2.9)$$

Además se asigna a cada evento el parámetro κ , donde representa el cambio correspondiente al pronóstico. Entonces queda de la forma

$$h(t) = Z(t)\kappa. \quad (2.10)$$

Tal como con la componente estacional, se usa la priori $\kappa \sim N(0, \sigma_\kappa^2)$.



Capítulo 3

Redes neuronales recurrentes

3.1. Redes neuronales

Uno de los primeros trabajos que reporta una idea referente a las redes neuronales fue el de McCulloch and Pitts (1943) como un intento de modelar matemáticamente el comportamiento de una neurona biológica. De esto nacieron dos ramas de la investigación. Por un lado, se siguió profundizando sobre el funcionamiento de las neuronas biológicas, mientras que, por otro lado, nace un enfoque aplicado sobre modelos en el área de inteligencia artificial.

Hornik, Stinchcombe and White (1989) consideran a las redes neuronales como funciones aproximadoras universales, pudiendo encontrar y reproducir relaciones complicadas sin ingresar conocimiento directo entre variables de entrada y salida si se cuenta con suficientes datos de entrenamiento.

El modelo más común de redes neuronales consiste en tres capas. Una capa de entrada, una capa oculta y una capa de salida. En este tipo de redes la capa de entrada y de salida son independientes. Además, entre la capa de entrada y salida, una o más capas ocultas están conectadas por matrices de pesos, *bias* y varias funciones de activación.

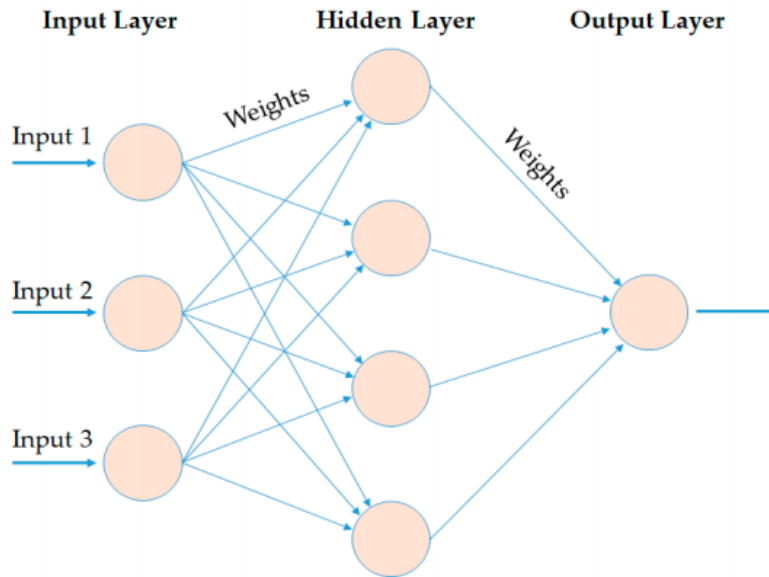


Figura 3.1: Esquema de una arquitectura sobre una red neuronal simple. Le, X., Ho, H., Lee, G., and Jung, S. (2019).

La gran diferencia entre los distintos tipos de redes neuronales involucra en la arquitectura de su red, relacionada a la cantidad de capas ocultas y la cantidad de neuronas en cada capa oculta.

Las redes neuronales presentan ventajas en las predicciones debido a su poder adaptativo no lineal. Pero al momento de hablar sobre series de tiempo, las redes neuronales tradicionales presentan falencias debido a que no son capaces de capturar la dependencia temporal que puedan tener las observaciones. Es por eso, que nacen las redes neuronales recurrentes (RNN), pues utilizan eventos previos para poder inferir eventos posteriores.

3.2. Redes neuronales recurrentes

La RNN se desarrolló por primera vez en la década de 1980 (Rumelhart, Hinton and Williams 1988). Su estructura consiste en una capa de entrada, una o más capas ocultas y una capa de salida. Las RNN tienen estructura con forma de cadena como módulos repetidos con la idea de usar estos módulos para almacenar información importante de los pasos de procesamiento anteriores. Gracias a esto las RNN han tenido éxito en el aprendizaje de secuencias.

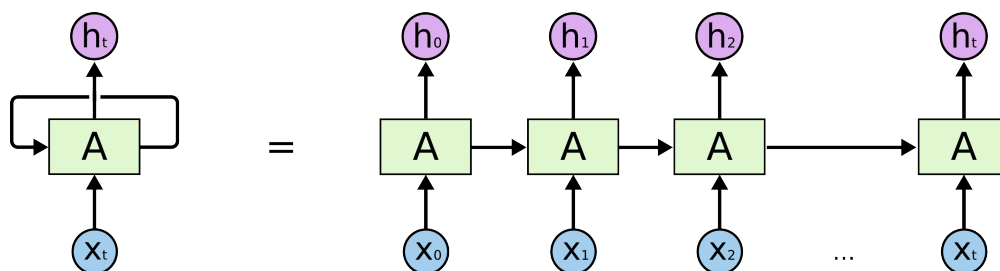


Figura 3.2: Procesamiento secuencial de una red neuronal recurrente. Olah, C. (2015)

En la Figura 3.2 se representa una simple RNN, con una unidad de entrada, una unidad de salida y una unidad oculta recurrente expandida en toda la red, donde X_t es la entrada en un tiempo t y h_t es la salida en un tiempo.

Durante el proceso de entrenamiento, las RNN utilizan un algoritmo de retro-propagación, frecuentemente aplicado en el cálculo del gradiente y el ajuste de las matrices que se ajustarán luego de modificar el proceso de retroalimentación. Es por esto, que este algoritmo también llamado comúnmente como retro-propagación a través del tiempo (BPTT). El BPTT utiliza un enfoque de trabajo hacia atrás, capa por capa, desde la salida hasta el comienzo de la red, ajustando los pesos de acuerdo con la porción calculada de las unidades de error. En consecuencia, si los datos de entrenamiento son largos, entonces los residuos de la red que necesita ser devuelto disminuirá exponencialmente, lo que afecta a una lenta actualización de los pesos de la red. Es decir, el RNN posee problema de gradiente, no es capaz de reflejar el efecto de la memoria al largo plazo en aplicaciones prácticas, por lo que se necesita de una componente adicional para almacenar esta memoria. Por lo tanto, se propuso una RNN especial, el modelo *Long short-term memory*, que presenta tanto memoria corta como a largo plazo.

3.3. Long short-term memory

Long short-term memory (LSTM) es una revolución de RNN, fue introducida por Hochreiter and Schmidhuber (1997) para abordar los inconvenientes antes mencionados del RNN al agregar interacciones adicionales por celda, siendo capaces de aprender dependencias a largo plazo y recordar información por periodos prolongados.

Según Olah (2015), el modelo LSTM está organizado en forma de estructura de cadena. Sin embargo, el módulo de repetición es diferente. A diferencia de la RNN estándar, la LSTM tiene cuatro capas interactivas con un método único de comunicación.

En particular, para este trabajo, la estructura de la red a utilizar motivado por Xiao and Yin (2019) es de la siguiente forma:

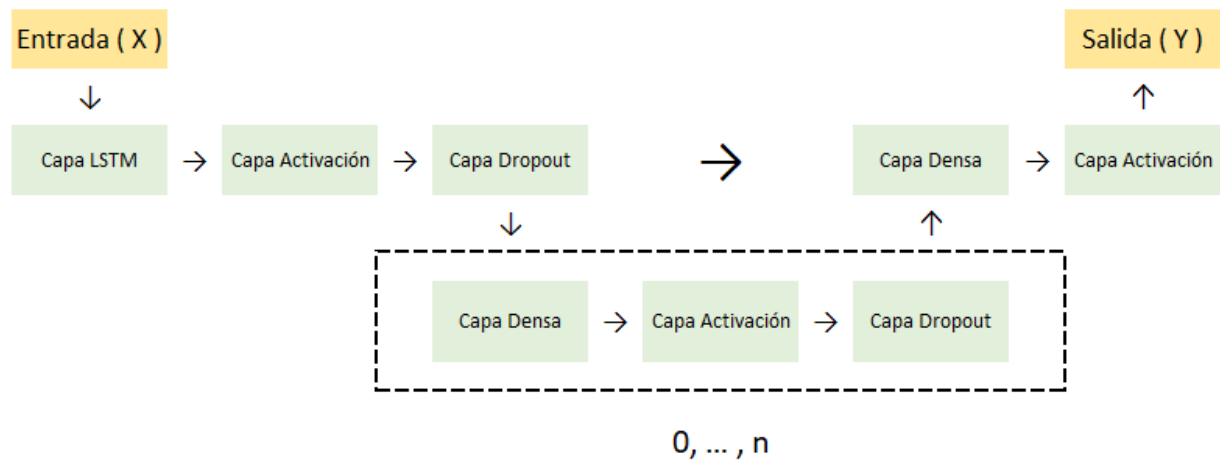


Figura 3.3: Estructura de una red neuronal híbrida LSTM.

En la Figura 3.3 se tiene la estructura de la red neuronal LSTM, que consta de una capa de combinación de entrada, cero o más capas de combinación intermedia y una capa de combinación de salida. La capa de entrada consta de una capa LSTM, capa de activación y una capa *Dropout*. Luego la capa de combinación intermedia consta de una capa densa, una capa de activación y una capa *Dropout*. Por último la capa de salida incluye una capa densa y una de activación.

Capa LSTM

La capa LSTM consta de una serie de unidades LSTM, llamadas modelo LSTM (Hochreiter and Schmidhuber 1997). Comparados con otras redes neuronales, el modelo LSTM tiene tres unidades multiplicativas, es decir, *Input gate*, *Output gate* y *Forget gate*. En resumen, *Input gate* es la encargada de memorizar la información del presente, *Output gate* es la encargada de abandonar información y *Forget gate* escoge que información olvidar del pasado.

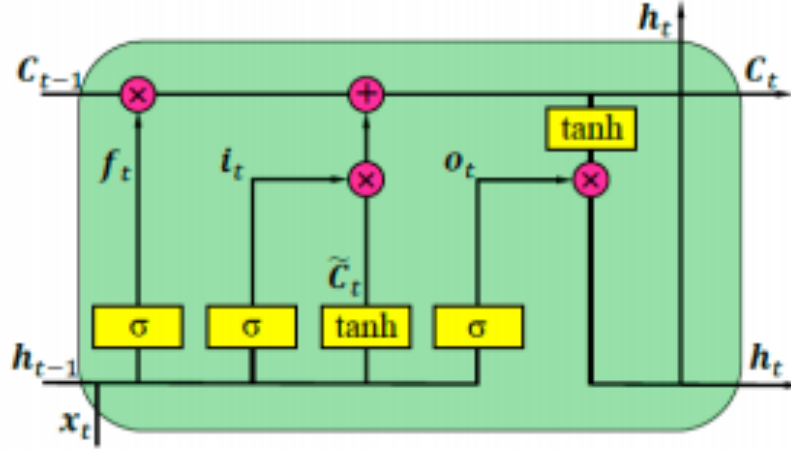


Figura 3.4: Unidad LSTM con sus respectivos componentes. Xiao and Ying (2019).

En la Figura 3.4, f_t es *forget gate*, i_t denota *input gate* y o_t es *output gate*, que juntos controlan el estado celular, C_t almacena información histórica y futura. x_t y h_t , respectivamente, representan la información de entrada y de salida de la unidad LSTM.

El primer paso de la celda LSTM es identificar la información que no se requiere para omitirla. Este proceso de identificación y exclusión es decidido por la función sigmoïdal, que toma la salida de la última unidad LSTM (h_{t-1}) en el tiempo $t - 1$ y la entrada actual (x_t) en el tiempo t . Además, la función sigmoïdal determina que parte de la salida anterior debe ser eliminada. Todo este proceso es llevado a cabo por *forget gate* (f_t), donde f_t es un vector de valores entre 0 y 1, correspondiente a cada número en el estado de celda, C_{t-1} .

$$f_t = \sigma(w_f [h_{t-1}, x_t] + b_f), \quad (3.1)$$

donde σ es la función sigmoïdal, w_f y b_f son las matrices de pesos y *bias*, respectivamente, del proceso *forget gate*.

El segundo paso es almacenar y decidir sobre la nueva entrada (x_t) entorno a la actualización del estado de la celda. Este proceso tiene dos partes, en primer lugar, se decide si la información debe actualizarse o ignorarse (0 o 1), encargado por la función sigmoïdal, y, en segundo lugar, dependiendo de los valores que pudieron continuar, la capa \tanh les otorga pesos a estos valores, según su nivel de importancia (-1 a 1). Los dos valores se multiplican para actualizar el estado de la celda. Esta nueva memoria se adiciona a la memoria anterior (C_{t-1}) dando como resultado C_t .

$$\begin{aligned} i_t &= \sigma(w_i [h_{t-1}, x_t] + b_i) \\ N_t &= \tanh(w_n [h_{t-1}, x_t] + b_n) \\ C_t &= C_{t-1}f_t + N_t i_t, \end{aligned} \quad (3.2)$$

donde C_{t-1} y C_t son las celdas de estado en el tiempo $t - 1$ y t , mientras que w y b son matrices de pesos y *bias*, respectivamente de las celdas de estados.

Por último, los valores de salida (h_t) son basados por el estado de la celda de salida (o_t), pero siendo una versión filtrada. Pues, una capa sigmoideal decide que parte de la celda de estado llega a la salida, luego *output gate* se multiplica por los nuevos valores creados por la capa tanh desde el estado de la celda (C_t), que varía entre -1 y 1.

$$\begin{aligned} o_t &= \sigma(w_o [h_{t-1}, x_t] + b_o) \\ h_t &= o_t \tanh(C_t), \end{aligned} \tag{3.3}$$

donde w_o y b_o corresponde a las matrices de pesos y *bias*, respectivamente, del *output gate*.

Capa de activación

La capa de activación es usada principalmente para mejorar la habilidad de ajuste del modelo. En la figura, cada capa combinada contiene una capa de activación.

Existen muchas funciones de activación que pueden ser usadas en la capa de activación de la entrada e intermedia, estas son la función tanh, función sigmoid, función Rectified linear Unit (ReLU), función SoftSign, función Softplus, función Leaky ReLU, función Parametric ReLU (PReLU), entre otros.

La función sigmoid, también conocida como la función logística, es una función no lineal, real y diferenciable limitada. Definida para valores de entrada real, con derivada positiva en todas partes con cierto grado de suavidad (Han and Moraga, 1995). La expresión matemática de la función sigmoid viene dada por:

$$f(x) = \frac{1}{1 + e^{-x}}. \tag{3.4}$$

La función tanh es completamente diferenciable. Esta función tiene imagen asimétrica y su centro es simétrico en el origen. La expresión matemática de la función tanh es la siguiente

$$f(x) = \frac{\sinh x}{\cosh x} = \frac{e^x - e^{-x}}{e^x + e^{-x}}. \tag{3.5}$$

Sin embargo, la función tanh al igual que funciones de activación tradicionales de redes neuronales, poseen el problema de gradiente de fuga. Esto es, al momento de que el valor de la función alcanza el área de saturación, el gradiente comienza a desaparecer, volviendo la velocidad de aprendizaje más lenta. (Karlik and Vehni, 2011)

Para aliviar este problema, investigadores propusieron una variante de la función tanh, la función softsign (Turian et al., 2009). Esta función es asimétrica, descentralizada, totalmente diferenciable y retorna valores entre -1 y 1. La gran diferencia es que la pendiente de la función softsign es menor que la función tanh en el área de activación, por lo que ingresa posterior al área de saturación. La expresión matemática es la siguiente:

$$f(x) = \frac{x}{1 + |x|}. \quad (3.6)$$

La función ReLU es una función de aprendizaje más rápida, ofreciendo mejor rendimiento y generalización en el aprendizaje profundo en comparación a la función sigmoid y tanh (Nair and Hinton, 2010). ReLU es una función casi lineal, logrando conservar las propiedades lineales de los modelos lineales, por lo que es fácil de optimizar con métodos del gradiente descendiente. La expresión matemática es la siguiente:

$$f(x) = \text{máx}(0, x) = \begin{cases} x_i, & \text{si } x_i \geq 0 \\ 0, & \text{en otro caso.} \end{cases} \quad (3.7)$$

La función PReLU es la función ReLU con parámetros. La expresión matemática es la siguiente

$$f(x) = \begin{cases} x_i, & \text{si } x_i \geq 0 \\ a_i x_i, & \text{en otro caso.} \end{cases} \quad (3.8)$$

donde, a_i es parámetro de control de la pendiente negativa, este parámetro se aprende durante el periodo de entrenamiento con retro-propagación(He et al., 2015). Cabe destacar que si en la ecuación (4.8) $a_i = 0$, se tiene la función ReLU.

La función LeakyReLU introduce una pequeña pendiente negativa al ReLU(Mass et al., 2013). El parámetro α se introdujo como una solución a los problemas de neuronas muertas de ReLU, de modo que los gradientes no serán ceros en ningún momento durante el entrenamiento. La expresión matemática es:

$$f(x) = \alpha x + x = \begin{cases} x_i, & \text{si } x_i \geq 0 \\ \alpha x_i, & \text{en otro caso.} \end{cases} \quad (3.9)$$

La función softplus es una forma suave de la función ReLU que tiene propiedades de suavizado y gradientes distintos de cero, mejorando así la estabilización y el rendimiento de la red neuronal (Dugas et al., 2001). La expresión matemática es

$$f(x) = \log(1 + e^x). \quad (3.10)$$

Capas Densa

La capa densa principalmente clasifica vectores de características y asigna las muestras del espacio de características a las etiquetas. Este proceso consiste en dos partes: parte lineal y parte no lineal.

La parte lineal principalmente realiza transformaciones lineales analizando los datos de entrada a través del cálculo de la suma ponderada lineal, donde la expresión matemática es de la siguiente forma

$$Z = WX + b, \quad (3.11)$$

donde Z es el vector de salida de la parte lineal expresada como $Z = [Z_1, \dots, Z_m]^T$. X representa el vector de entrada de la parte lineal, expresada como $X = [X_1, \dots, X_n]^T$. W matriz de pesos $m \times n$ y b vector de *bias*, expresada como $b = [b_1, \dots, b_m]^T$.

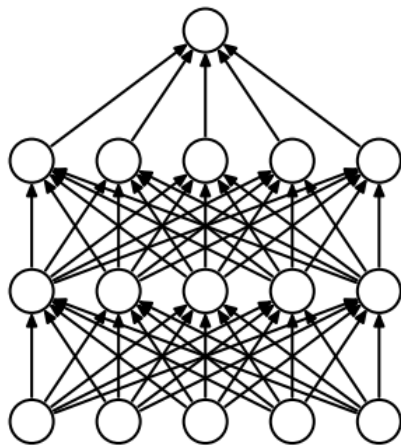
La parte no lineal realiza transformación no lineal. Esta operación consta de dos funciones:

- Normalización de los datos, es decir, no importa lo realizado en la parte lineal, todos los valores en la parte no lineal son limitados a un cierto rango de valores.
- Deshacer las relaciones lineales generadas anteriormente.

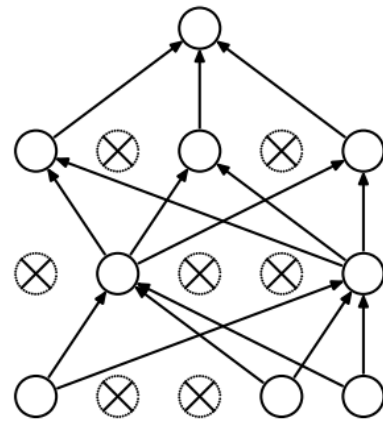
Capa Dropout

La capa *Dropout* principalmente evita el sobreajuste y proporciona una forma de combinar muchas arquitecturas de redes neuronales diferentes de manera eficiente. El término *Dropout* se refiere a eliminar temporalmente unidades (ocultas y visibles), junto con todas sus conexiones entrantes y salientes de la red.

La elección de que unidades eliminar es aleatoria. En el caso más simple, cada unidad se retiene con una probabilidad p independiente de otras unidades, donde p se puede elegir usando un conjunto de validación o simplemente establecer en 0.5.



(a) Red neuronal estándar.



(b) Después de aplicar Dropout.

Figura 3.5: Modelo de red neuronal con Dropout.

En la Figura 3.5 se observa un modelo de red neuronal con Dropout. En la Figura (a) se tiene un modelo de red neuronal con dos capas ocultas, mientras que en la Figura (b) se tiene un ejemplo de una red reducida, producto del *Dropout* aplicado a la red neuronal de la Figura (a).



Capítulo 4

Estrategias de validación

4.1. Validación cruzada

Al momento de validar modelos, existen diversas metodologías, como son la comparación de los parámetros obtenidos con los adquiridos mediante simulación, utilizar conjunto de datos adicional para comparar con los obtenidos o bien el uso de validación cruzada. (Kozak, 2003).

Las dos versiones de validación cruzada más utilizada son *hold-out* y *k-fold*. La primera de ellas separa el conjunto de datos en dos subconjuntos, uno utilizado para entrenar el modelo y generar datos de salida, luego compararlos con el segundo subconjunto para realizar la validación. Los estadísticos obtenidos representan la validez del modelo en términos del error. Una aplicación alternativa de *hold-out* es repetir el proceso un determinado número de veces, tomando un conjunto de datos de entrenamiento aleatorios de manera de calcular los estadísticos a partir de la media de cada una de las repeticiones.

La otra versión de validación cruzada, *k-fold*, divide el total de datos en k subconjuntos de manera de aplicar el método *hold-out* k veces, utilizando cada vez un subconjunto distinto para validar el modelo entrenado con el $k - 1$ subconjunto restante. (Jung and Hu, 2015). El error promedio obtenido de los k análisis nos proporciona el error cometido por el modelo, así también evaluar su validez.

4.2. Validación cruzada de avance

Aplicar las técnicas de validación cruzada mencionadas anteriormente a datos de series de tiempo origina dos problemas.

- Las series temporales están ordenadas por el tiempo, la aleatorización en el algoritmo estándar de validación cruzada no conserva el orden temporal.

- Las series temporales a menudo están fuertemente correlacionadas a lo largo del tiempo. La aleatorización afectará que, para cada muestra en el conjunto de validación, existen muestras fuertemente correlacionadas con el conjunto de entrenamiento.

Claramente, la fuente principal de estos problemas es la aleatorización. La validación cruzada de avance resuelve este problema al restringir el conjunto de muestras de manera diferente a cada división.

Bloques	Semana 1	Semana 2	Semana 3	Semana 4	Semana 5	Semana 6	Semana 7	Semana 8
1	Conjunto de entrenamiento			Validación				
2		Conjunto de entrenamiento			Validación			
3			Conjunto de entrenamiento			Validación		
⋮					⋮			
k					Conjunto de entrenamiento		Validación	

Figura 4.1: Validación de avance con k bloques. Elaboración propia.

En la Figura 4.1 se presenta un ejemplo de validación de avance. Primero se dividió el conjunto de datos en k bloques iguales de muestras contiguas y decidimos que el conjunto de entrenamiento constará de p bloques contiguos (1 semana). Por lo que las divisiones están dadas como:

- Bloque 1.**
 - Conjunto de entrenamiento: p a $p+2$
 - Conjunto de validación: $p+3$
- Bloque 2.**
 - Conjunto de entrenamiento: $p+1$ a $p+3$
 - Conjunto de validación: $p+4$
- ⋮
- Bloque k.**
 - Conjunto de entrenamiento: $p+4$ a $p+6$
 - Conjunto de validación: $p+7$

De esta manera, al “avanzar hacia adelante” en el conjunto de muestras completo, uno puede construir k divisiones. Este enfoque produce dos beneficios significativos. Primero, ofrece una visión realista de cómo se comportaría un modelo en particular con el tiempo y, en segundo lugar, brinda a la capacidad de aprovechar en mayor medida la disponibilidad de datos para validar modelos. De hecho, la metodología de validación no solo prueba un modelo en particular, sino que prueba todo el enfoque de modelado ya que se vuelven a parametrizar periódicamente a medida que entran nuevos datos. Así se puede obtener una idea más clara de cómo se mantendrá todo el enfoque del modelado a través de varios ciclos en el tiempo. (Stein, 2002).

Capítulo 5

Proceso experimental

Para la siguiente investigación, se proponen dos líneas metodológicas distintas, donde su principal diferencia es el dato a utilizar. Por una parte, se tiene los registros de valores cuotas asociados a cada fondo y, por otra parte, se tiene las rentabilidades reales respectiva de cada fondo.

Para el modelo Prophet y LSTM, el proceso de optimización se basará en el principio de maximización, aunque el proceso de entrenamiento será componente por componente y capa por capa, respectivamente.

5.1. Conjuntos de datos

Los conjuntos de datos relacionados al sistema de multifondos están disponible en el sitio web de la superintendencia de pensiones. Los registros que se proveen son los valores cuota y patrimonios diarios de cada AFP por tipo de fondo. Información disponible en <http://www.spensiones.cl>.

Se realizó una implementación propia de *web scraping* que consistía principalmente en dos partes. En primer lugar, descargar la información histórica por tipo de fondo, generar el valor cuota para cada fondo como promedio de los valores cuota de las AFP en ese instante de tiempo y guardar la tabla generada en un archivo Excel. En segundo lugar y en forma periódica, descargar los nuevos registros para cada fondo, generar el valor cuota y añadirlos al Excel generado en la primera parte.

Por otro lado, el conjunto de datos relacionado a la unidad de fomento (UF) fueron extraídos desde el sitio web del banco central. Información disponible en <https://si3.bcentral.cl/siete/secure/cuadros/home.aspx>.

Por último, los conjuntos de datos relacionados a la información histórica de las regresoras fueron extraídos desde Yahoo! Finance, un servicio que proporciona información financiera

en los mercados internacionales. Información disponible en <https://finance.yahoo.com/>.

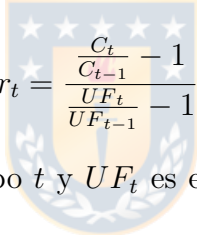
5.1.1. Valor Cuota

La cuota es una unidad de medida en la que se expresa los recursos acumulados en las cuentas de capitalización individual, o bien, el dinero de cada afiliado ha acumulado en su AFP corresponde a un número de cuotas. El valor cuota aumenta o disminuye diariamente dependiendo de la rentabilidad de las inversiones.

Para el caso del Fondo A, se cuenta con 6361 datos diarios, correspondiente al 8 agosto 2002 hasta 31 diciembre 2019. Por otro lado, para el Fondo E, se cuenta con 7245 datos diarios, correspondiente al 1 marzo 2000 hasta 31 diciembre 2019.

5.1.2. Rentabilidad real

El valor cuota en sí mismo no es relevante, sino su variación, pues representa la rentabilidad que están obteniendo sus recursos. La rentabilidad real se define como:


$$r_t = \frac{\frac{C_t}{C_{t-1}} - 1}{\frac{UF_t}{UF_{t-1}} - 1}, \quad (5.1)$$

donde C_t es el valor cuota en el tiempo t y UF_t es el valor de la unidad de fomento en un tiempo t .

Es importante destacar que la rentabilidad puede ser afectada por factores externos, como la inflación, impuestos, etc. Por lo que la rentabilidad real deflactada por la UF expresada en (5.1) permite observar el retorno obtenido solo por las ganancias y pérdidas productos de las inversiones realizadas en el fondo y no por factores externos.

Los datos del Fondo A y Fondo E para la rentabilidad real fueron calculados por (5.1). Para el Fondo A se cuentan con 6360 datos diarios, correspondiente al 9 agosto 2002 hasta 31 diciembre 2019. Por otro lado, para el Fondo E, se cuenta con 7244 datos diarios, correspondiente al 2 marzo 2000 hasta 31 diciembre 2019.

5.2. Indicadores de evaluación experimental

Como forma de comparar y evaluar el rendimiento de los pronosticadores del modelo, se implementaron casos de prueba, donde la bondad de ajuste de las predicciones se evaluaron por las siguientes métricas:

- Raíz del Error Cuadrático Medio (RMSE): Esta métrica mide la magnitud del error.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}, \quad (5.2)$$

donde y es el valor observado, \hat{y} es el valor estimado y n es el tamaño de la muestra.

- Error Absoluto Medio (MAE): Esta métrica mide la magnitud del error. La importancia de esta métrica es que no es tan sensible a los valores atípicos a diferencia de la métrica anterior.

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|, \quad (5.3)$$

donde y es el valor observado, \hat{y} es el valor estimado y n es el tamaño de la muestra.

- Error Medio Absoluto Porcentual (MAPE): Esta métrica mide el tamaño del error (absoluto) en terminos porcentuales.

$$MAPE = 100 \times \frac{\sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right|}{n}, \quad (5.4)$$

donde y es el valor observado, \hat{y} es el valor estimado y n es el tamaño de la muestra.

- Sentido de la tendencia (Sentido): Esta métrica determina el porcentaje de aciertos en el sentido de la tendencia, si el valor observado va al alza, la predicción debiese seguir ese comportamiento.

Es importante señalar que el criterio de selección de hiper-parámetros asociados a las técnicas estadísticas a utilizar en este trabajo será comparando la métrica del sentido de la tendencia, es decir, mientras mayor sea este valor, mejor captura el sentido de la curva.

5.3. Proceso de ajuste experimental

5.3.1. Prophet

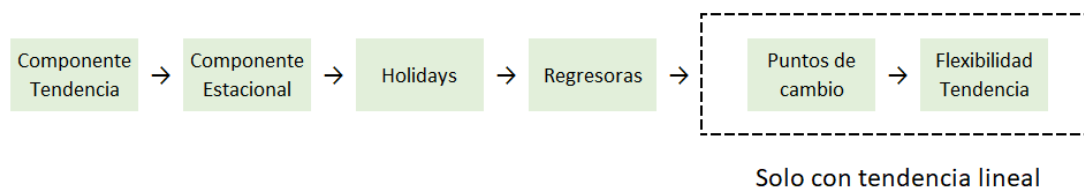


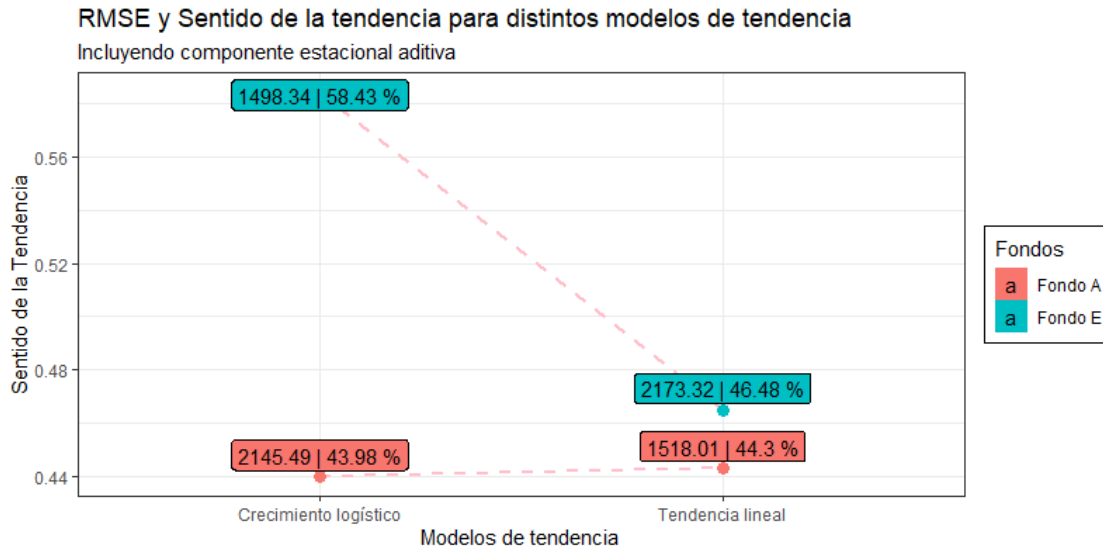
Figura 5.1: Estructura del proceso de optimización para los hiper-parámetros de Prophet.

El proceso de optimización de los hiper-parámetros de Prophet se llevará a cabo componentes por componente como se refleja en la Figura 6.1, donde se elegirán los que tengan un mayor valor en el sentido de la tendencia. Se utilizará el 80 % de los datos como entrenamiento y el restante 20 % como validación. Después de encontrar los hiper-parámetros óptimos se realizará la validación cruzada de avance.

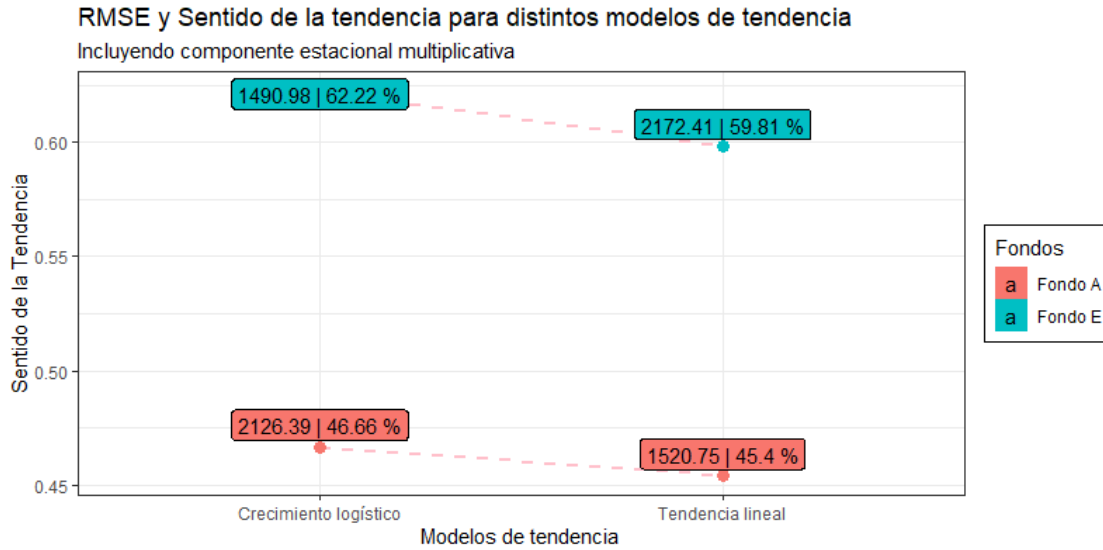
Valor Cuota

(1) Componente tendencia

Al momento de contemplar la tendencia, existen dos posibles modelos mencionados en Capítulo 3.1, por defecto, Prophet utiliza un modelo lineal para sus pronósticos, además de incluir componentes estacionales semanales y anuales, que pueden ser aditiva o multiplicativa a la tendencia. Las componentes adicionales se probarán en la componente próxima.



(a)



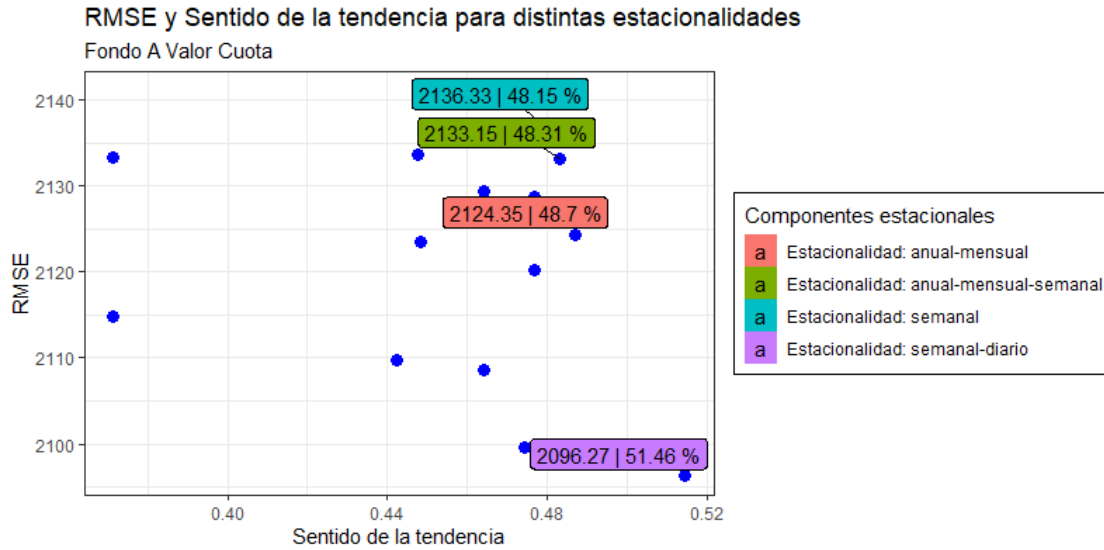
(b)

Figura 5.2: (a) Valores RMSE y Sentido para el modelo de tendencia lineal y crecimiento logístico con factor aditivo. (b) Valores RMSE y Sentido para el modelo de tendencia lineal y crecimiento logístico con factor multiplicativo.

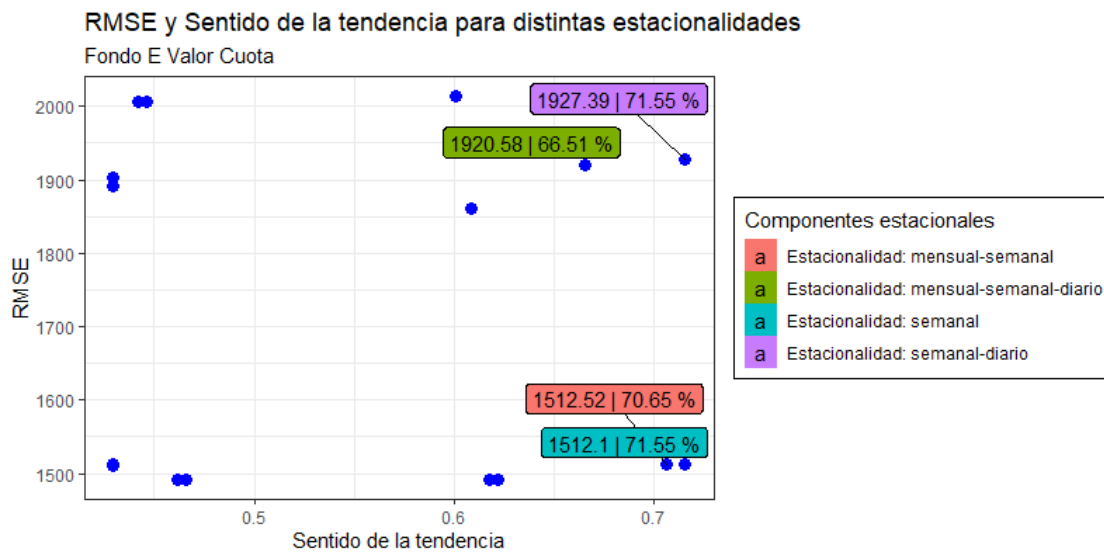
En la Figura 5.2 se tiene los valores RMSE y sentido de la tendencia para distintos modelos de tendencia, donde además se incluye componente estacional como factor aditivo o multiplicativo. Para el Fondo A y Fondo E, el mejor modelo respecto al sentido de la tendencia es el modelo de crecimiento logístico considerando la componente estacional multiplicativa respecto a la tendencia, pues alcanza un Sentido de 46,66 % y 62,22 %, respectivamente.

(2) Componentes estacionales

Como se mencionó anteriormente, Prophet ajusta por defecto estacionalidad semanal y anual, si la serie de tiempo tiene más de dos ciclos de duración, pero se pueden añadir otras estacionalidades como diaria, mensual, etc.



(a)



(b)

Figura 5.3: (a) Valores RMSE y Sentido para el Fondo A donde se incluyen distintas componentes estacionales. (b) Valores RMSE y Sentido para el Fondo E donde se incluyen distintas componentes estacionales.

En la Figura 5.3 se tiene los valores de RMSE y sentido de la tendencia para los Fondo A y Fondo E al incluir componentes estacionales. Se observa que mejoran los resultados entorno a la componente anterior, pues en el Fondo A, aumenta de 46,66 % a 51,46 % al incluir una

componente estacional diaria y semanal. Por otro lado, para el Fondo E, aumenta de 62,22 % a 71,55 % al incluir solo una componente estacional semanal.

(3) Fechas y eventos

Para el modelado de feriados y eventos especiales, se utilizó el paquete *holidays* de Python. Para esta investigación en particular, se utilizaron el listado de feriados y eventos de Chile, países Sudamericanos, EE.UU, países Europeos y países Asiáticos.

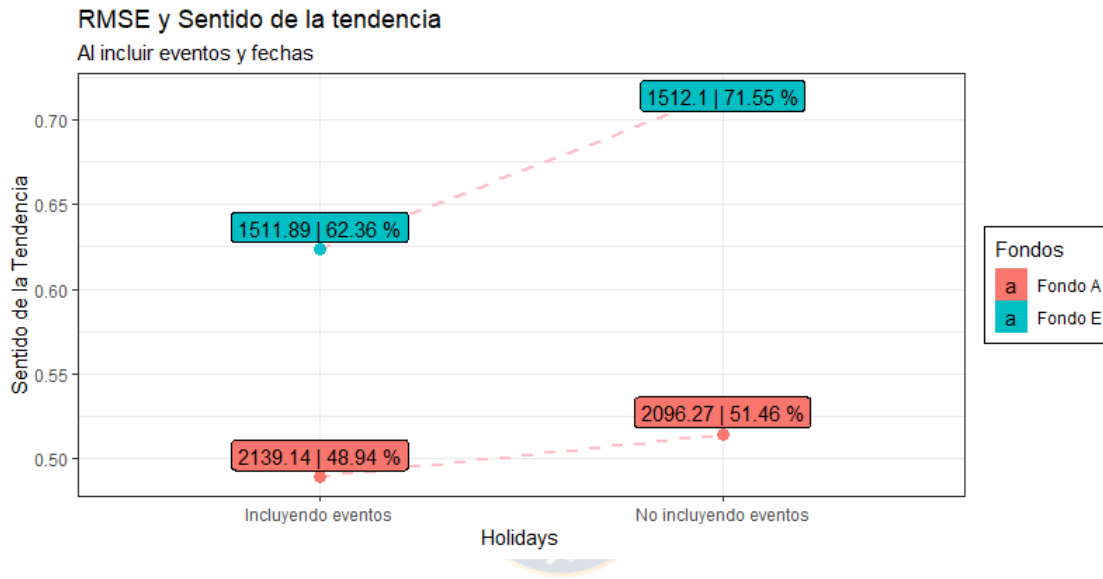


Figura 5.4: RMSE y Sentido de la tendencia al incluir listado de feriados y eventos al modelo.

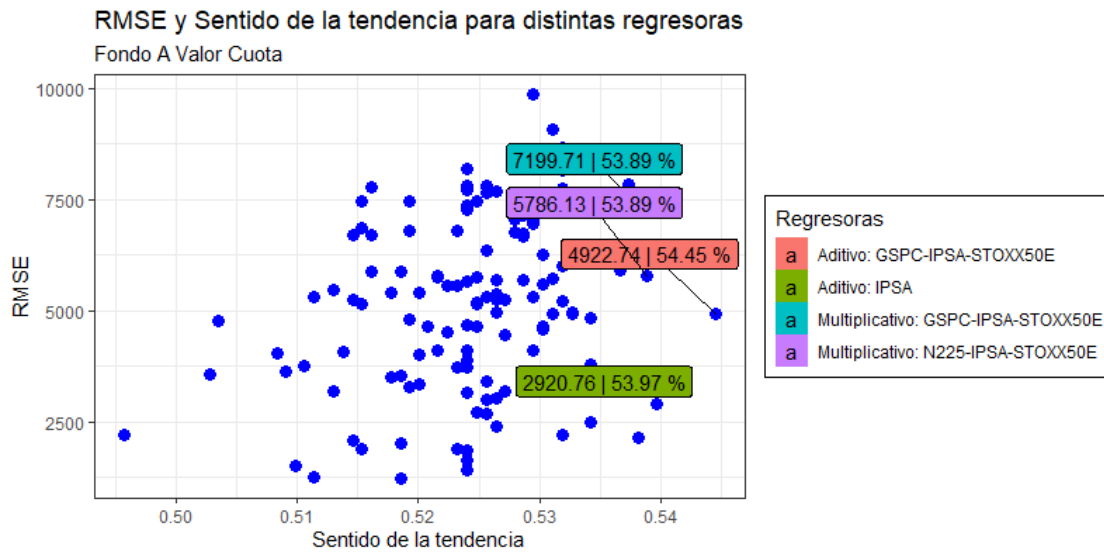
En la Figura 5.4 se observan los valores RMSE y Sentido de la tendencia de los Fondo A y Fondo E al incluir fechas y eventos. En ambos casos se destaca una disminución del Sentido de la tendencia, por lo que no se incluirá este listado de fechas.

(4) Regresoras

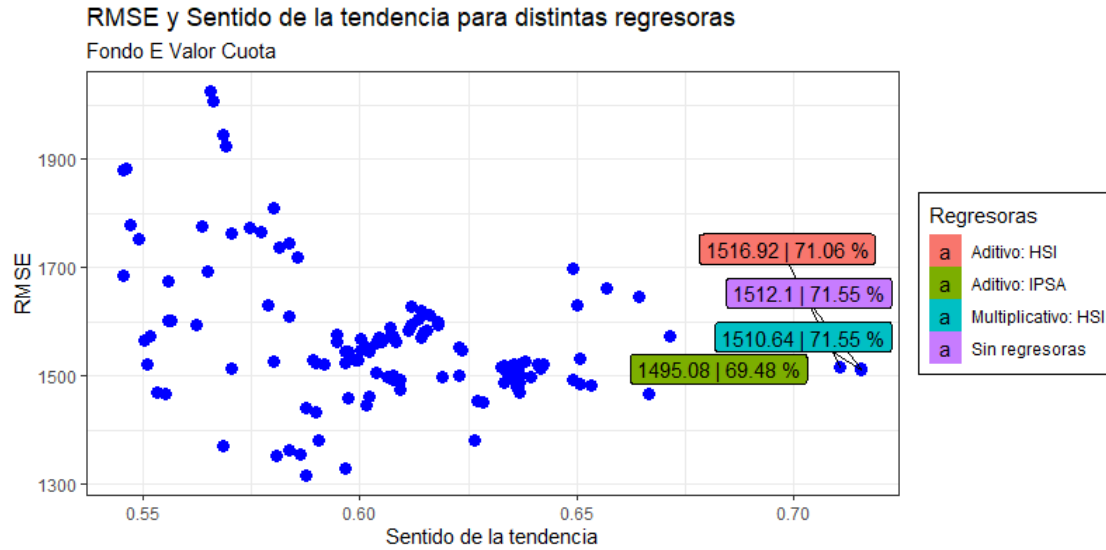
Adicionalmente, se pueden agregar regresoras adicionales en la parte lineal del modelo. Las regresoras a utilizar son índices bursátiles que posiblemente podrían influir en el comportamiento de la serie. Cabe destacar que Prophet no admite valores perdidos en las regresoras, por lo que se imputaron los datos con método lineal cuando fuese necesario. Por lo cual, las regresoras son:

- **HSI:** Denominado Hong Seng Index. Es el principal índice bursátil chino de Hong Kong.
- **GSPC:** También conocido como S&P500. Es uno de los índices bursátiles más importantes de EE.UU., pues se basa en la capitalización bursátil de 500 grandes empresas.

- **N225**: Comúnmente denominado índice Nikkei. Es el índice bursátil más popular del mercado Japonés, lo componen las 225 empresas que cotizan en la bolsa de Tokio.
- **KS11**: Comúnmente denominado Kospi. Es el índice bursátil de Corea del Sur compuesto por todas las compañías negociadas en el mercado de valores de Corea.
- **IPSA**: Denominado índice de precios selectivo de acciones. Es el principal índice bursátil de Chile, elaborado por la bolsa de comercio de Santiago.
- **STOXX50E**: Denominado Dow Jones Euro Stoxx50. Es un índice bursátil que representa el rendimiento de las 50 empresas más grandes en términos de capitalización de mercado en 11 países de la Eurozona.



(a)



(b)

Figura 5.5: (a) Valores RMSE y Sentido para distintos modelos que incluyen las regresoras mencionadas en el Fondo A. (b) Valores RMSE y Sentido para distintos modelos que incluyen las regresoras mencionadas en el Fondo E.

En la Figura 5.5, se observan el RMSE y Sentido de la tendencia para distintos modelos, donde se prueban todas las combinaciones posibles de las regresoras, incluyendo si se incorporan aditiva o multiplicativa. A modo de comprensión, solo se destaca los modelos que obtuvieron mejores resultados.

Para el caso del Fondo A, mejoraron los resultados respecto a los ítems anteriores, pues de un 51,46 % aumenta a 54,45 % incluyendo los índices GSPC, IPSA y STOXX50E como factores aditivos. Por otro lado, para el Fondo E, no se obtuvo mejores resultados en el Sentido de la tendencia, pero se redujo el RMSE al incluir el índice HSI como factor multiplicativo, bajo de 1512,1 a 1510,64.

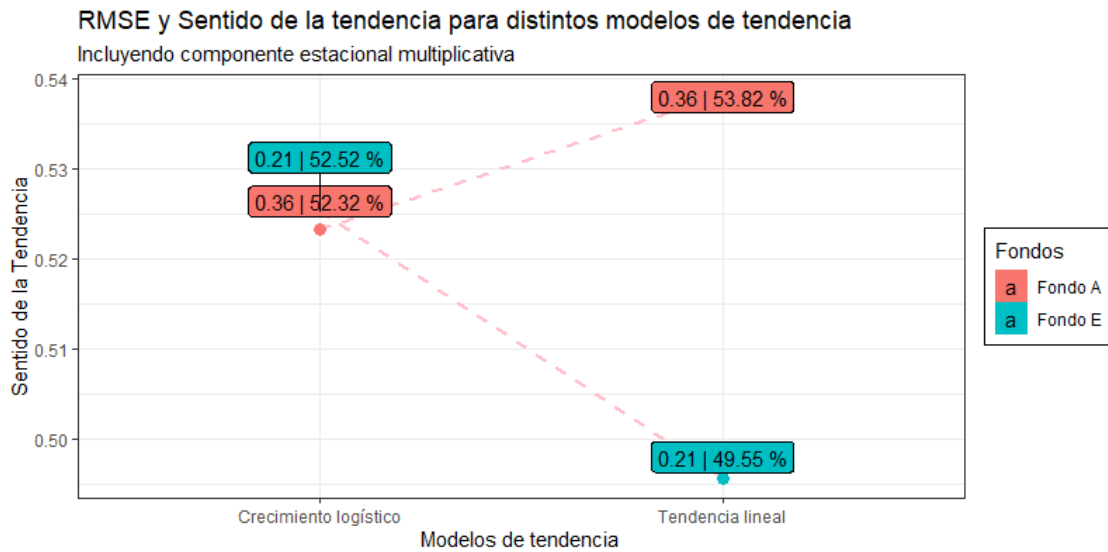
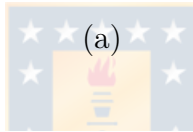
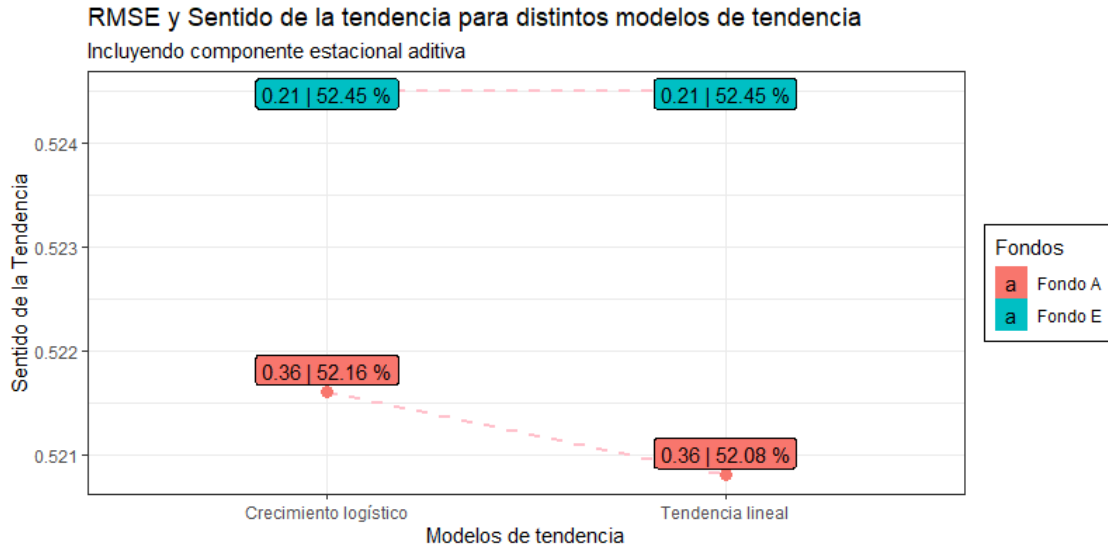
Rentabilidad real

Los pasos a seguir que se utilizaron con los valores cuota del Fondo A y Fondo E se replicaran con la rentabilidad real asociada a ambos fondos.

(1) Componente tendencia

Para la tendencia, tenemos dos modelos en específico, la tendencia lineal con puntos de cambio y el crecimiento logístico. Recordar que por defecto, Prophet incluye componente estacional semanal y anual, donde se pueden incluir aditiva o multiplicativa respecto a la

tendencia.



(b)

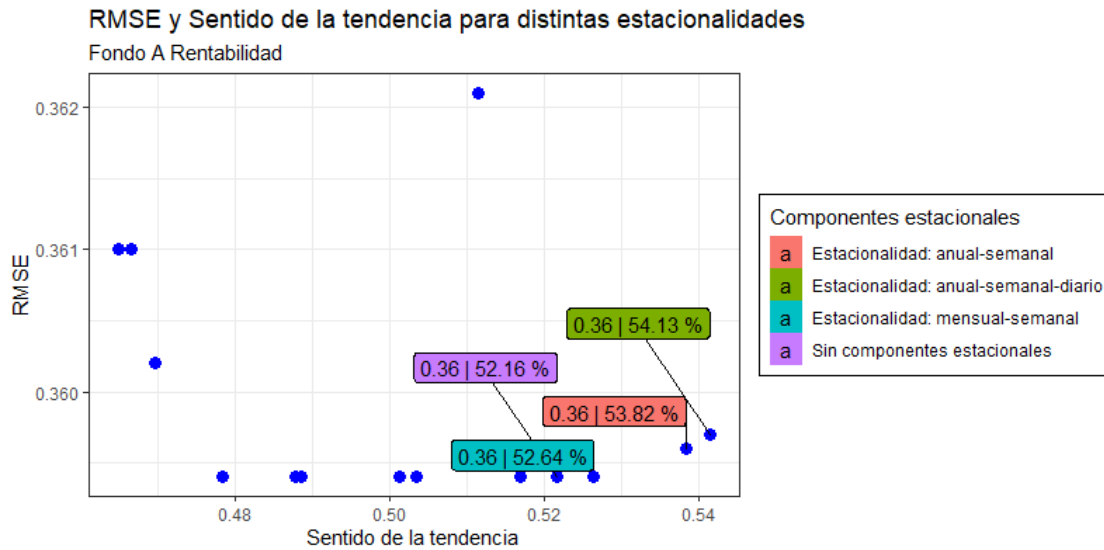
Figura 5.6: (a) Valores RMSE y Sentido para el modelo de tendencia lineal y crecimiento logístico con factor aditivo. (b) Valores RMSE y Sentido para el modelo de tendencia lineal y crecimiento logístico con factor multiplicativo.

En la Figura 5.6 se tiene los valores RMSE y sentido de la tendencia para distintos mo-

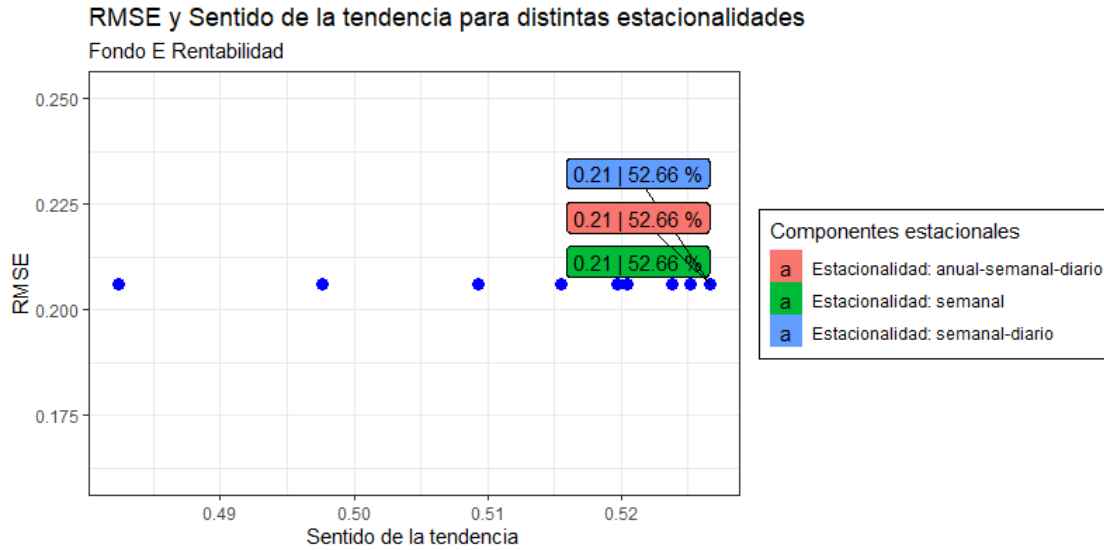
delos de tendencia, donde además se incluye componente estacional como factor aditivo o multiplicativo. Al comparar los resultados de modelado de la tendencia respecto al valor cuota, se presenta una diferencia, pues solo el Fondo E le corresponde el modelo de crecimiento logístico con componente estacional multiplicativo, donde llega un valor de sentido de la tendencia a 52,52%. Por otro lado, para el Fondo A, el mejor modelo es usando la tendencia lineal con componente estacional multiplicativo, llegando alcanzar un 53,82% en el sentido de la tendencia.

(2) Componente estacionales

Además de contemplar la estacionalidad anual y semanal, se probarán otras estacionalidades, como puede ser mensual y diaria.



(a)



(b)

Figura 5.7: (a) Valores RMSE y Sentido para el Fondo A donde se incluyen distintas componentes estacionales. (b) Valores RMSE y Sentido para el Fondo E donde se incluyen distintas componentes estacionales.

En la Figura 5.7 se tiene los valores de RMSE y sentido de la tendencia para los Fondo A y Fondo E al incluir componentes estacionales. Al comparar las componentes estacionales respecto al valor cuota, para el Fondo A, se le añade una componente adicional, de sólo considerar estacionalidad semanal y diaria, ahora se le incluye la estacionalidad anual. Por otro lado, para el Fondo E se tienen las mismas componentes.

Se observa que hubo una mejoría mínima en ambos fondos respecto al resultado de la componente anterior. La mejora es menor a un 1% alcanzando en el sentido de la tendencia para el Fondo A 54,13%, mientras que para el Fondo E se tiene un 52,66%.

(3) Fechas y eventos

Una de las fortalezas de Prophet es la inclusión manual de fechas y eventos que podrían afectar en la estabilidad de la serie de tiempo. Específicamente, los países a considerar en el listado de fechas fueron mencionados anteriormente.

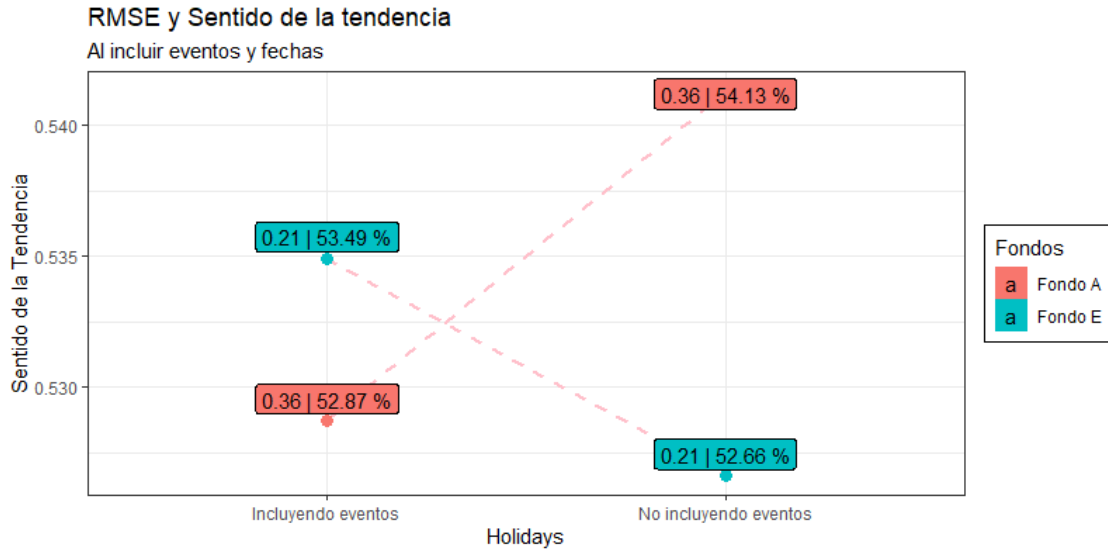


Figura 5.8: RMSE y Sentido de la tendencia al incluir listado de feriados y eventos al modelo.

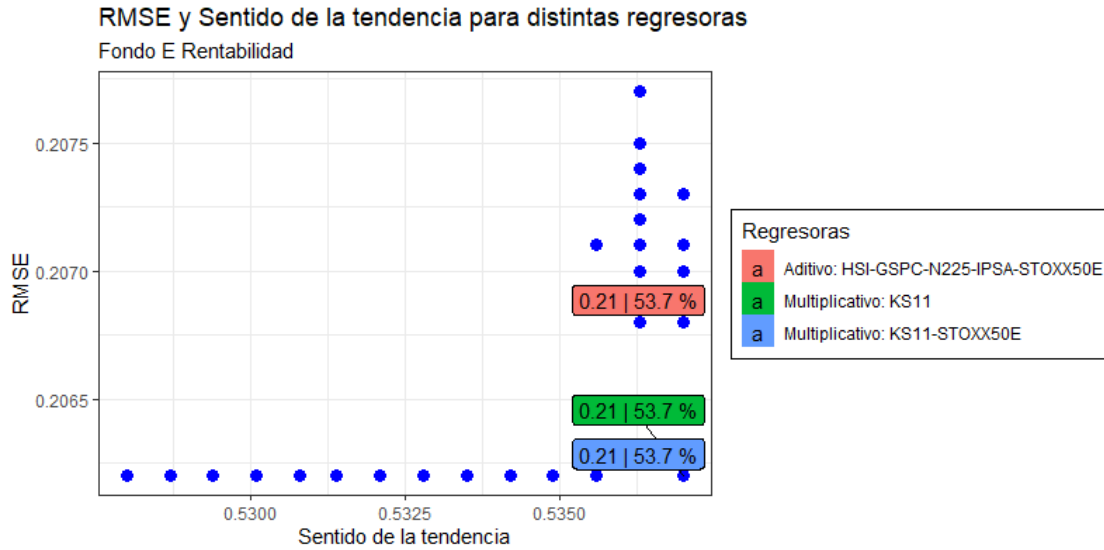
En la Figura 5.8 se observan los valores RMSE y Sentido de la tendencia de los Fondo A y Fondo E al incluir fechas y eventos. Solo para el Fondo E hubo una mejoría respecto al sentido de la tendencia, pasó de un 52,66 % a 53,49 %.

(4) Regresoras

La descripción de las regresoras a incorporar en la parte lineal de modelo fue mencionada anteriormente.



(a)



(b)

Figura 5.9: (a) Valores RMSE y Sentido para distintos modelos que incluyen las regresoras mencionadas en el Fondo A. (b) Valores RMSE y Sentido para distintos modelos que incluyen las regresoras mencionadas en el Fondo E.

En la Figura 5.9, se observan el RMSE y Sentido de la tendencia para distintos modelos, donde se prueban todas las combinaciones posibles de las regresoras, incluyendo si se incorporan aditiva o multiplicativa. A modo de comprensión, solo se destaca los modelos que obtuvieron mejores resultados.

Al comparar los resultados con los valores cuota, existen diferencias entre las regresoras a utilizar respecto al mejor modelo. En primer lugar, para el Fondo A en el valor cuota, el mejor modelo tenía tres regresoras aditivas, estas son el GSPC, IPSA y STOXX50E, por otro lado, para la rentabilidad real, las regresoras son factores multiplicativos y considera solo el índice HSI e IPSA. En segundo lugar, para el Fondo E, al observar la rentabilidad, se tiene tres modelos con los mismos resultados, pero se considerará el que tenga menor número de parámetros, por lo que solo queda el índice KS11 como factor multiplicativo, distinto al valor cuota, que considero como mejor modelo el índice HSI como factor multiplicativo.

Las mejoras respecto a las componentes anteriores son mínimas, siendo mejorías menor al 1% alcanzando para el Fondo A un 54,76% en el sentido de la tendencia y para el Fondo E un 53,70%.

(5) Puntos de cambio en la tendencia

En este caso y en la siguiente componente, solo se llevará a cabo con el Fondo A, pues para la tendencia se usará el modelo de tendencia lineal, explicado en (2.1.2).

Prophet detecta los puntos de cambio al especificar, primero una gran cantidad de potenciales puntos de cambio donde la tendencia puede cambiar, luego usará la menor cantidad posible de ellos. Por defecto, especifica 25 potenciales puntos de cambio que se colocan de manera uniforme en el 80 % de la serie con la finalidad de aprender de estos y proyectar la tendencia hacia adelante y evitar fluctuaciones de sobreajuste al final de la serie. En esta componente se probaron desde el 80 % hasta el 99 % de las series de tiempo.

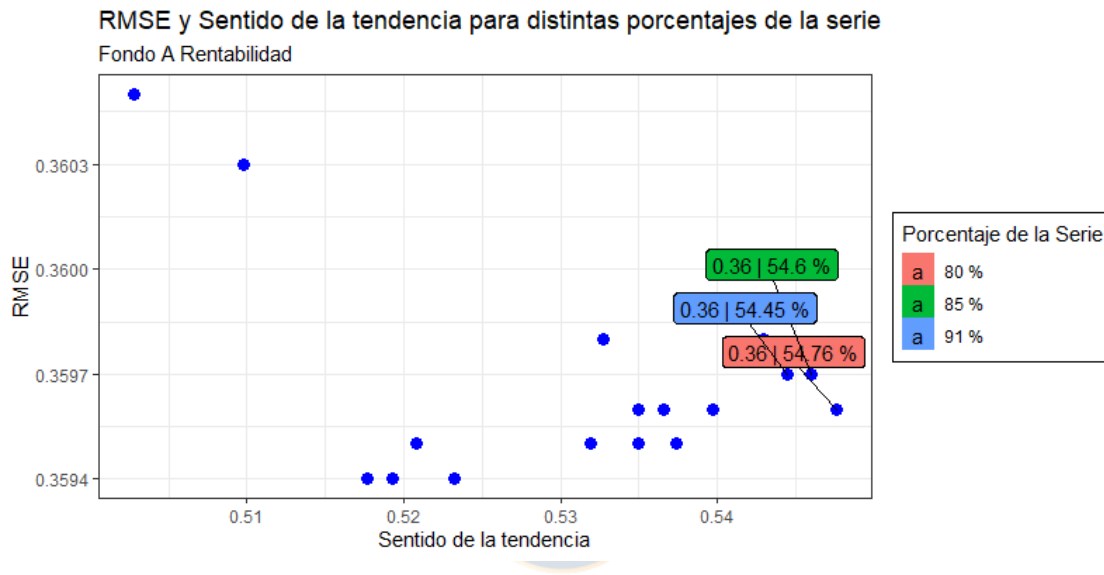


Figura 5.10: RMSE y Sentido de la tendencia al modificar el porcentaje de la serie en la rentabilidad del Fondo A.

En la Figura 5.10 se tiene el RMSE y Sentido de la tendencia para todos los modelos posibles tomando distintos porcentajes de la serie para especificar los potenciales puntos de cambio. No existe mejoría respecto al ítem anterior, usando el valor por defecto alcanza el máximo sentido de la tendencia.

(6) Flexibilidad en la tendencia

Los analistas pueden especificar los puntos de cambio. Como se mencionó en el Capítulo 2.1.2, δ_j es la tasa de cambio en que ocurre en un momento s_j , donde $\delta_j \sim Laplace(0, \tau)$. El parámetro τ controla la flexibilidad del modelo.

Por defecto, τ es 0,05. Al disminuir este valor, se hará que la tendencia sea menos flexible, en caso contrario, al aumentar este valor, la tendencia será más flexible. Los valores a utilizar son:

Tabla 5.1: Valores para el parámetro τ

Parámetro	Valores
τ	0.001, 0.01, 0.02, 0.03, 0.04, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1

Fuente: Elaboración propia.

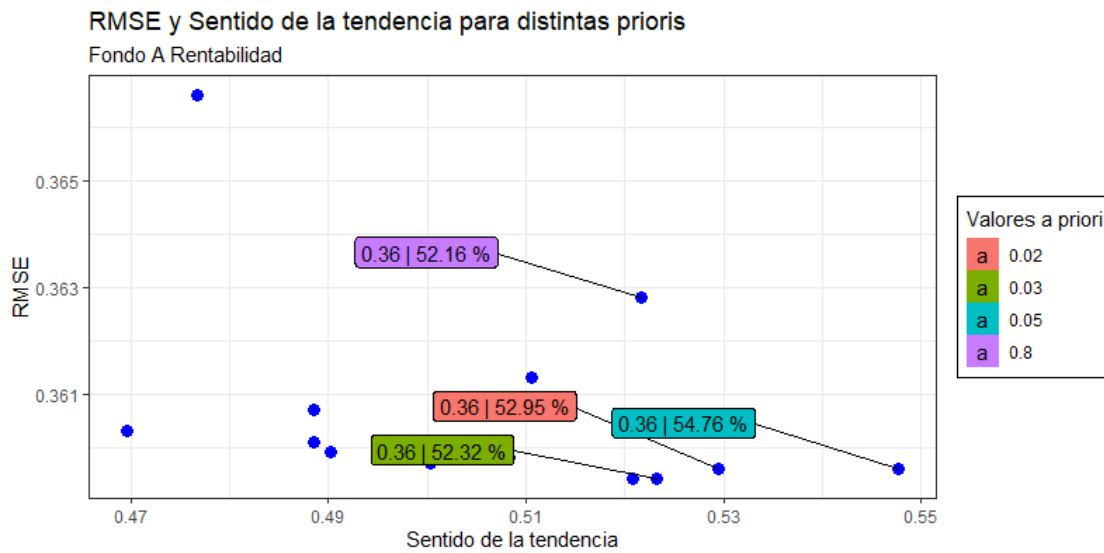


Figura 5.11: RMSE y sentido de la tendencia al modificar el valor a priori τ .

En la Figura 5.11 se tiene el RMSE y sentido de la tendencia para distintos valores de τ , esto implica volver a la tendencia más o menos flexible. Al igual que en la componente anterior, no existe mejoría de resultados a probar distintos prioris, utilizando el valor por defecto se alcanza el Sentido máximo, esto es 54,76 %.

5.3.2. Long short-term memory

El proceso de optimización de los hiper-parámetros de la red LSTM se llevará a cabo capa por capa como se refleja en la Figura 4.3, donde se elegirán los que tengan un mayor valor en el sentido de la tendencia. Se aplicará un pre-procesamiento de los datos y, a diferencia de Prophet, se aplicará la validación cruzada de avance directamente, donde se usará el 80 % de los datos como entrenamiento y la validación se realizará con la semana siguiente. Específicamente, el modelo será entrenado para generar predicción de 7 días. Es importante señalar, que cada componente tuvo 10 repeticiones y los resultados expuestos es el promedio de ellos. Sólo para el caso de las neuronas e iteraciones, hubo 3 repeticiones.

Valor Cuota

(1) Pre-procesamiento de los datos

Antes de optimizar los hiper-parámetros de la red, es necesario un pre-procesamiento de los datos. Como no existen valores negativos en las cuotas, se normalizaron los valores entre 0 y 1, manteniendo las relaciones existentes entre los patrones de la entrada original. Además, la serie temporal se transformará en un problema de aprendizaje supervisado. Esto es, organizar los datos en patrones de entrada y salida, donde la observación en el tiempo anterior se utiliza para pronosticar la observación en el tiempo actual.

Entonces hay que definir el número de días deseados, para poder convertir el vector de datos a una versión desplazada, con el objetivo de obtener una matriz de regresión de la variable explicativa. Los días a probar son:

Tabla 5.2: Cantidad de días a probar como variable de entrada.

Parámetro	Número de días
Timestep	3, 7, 10, 14, 17, 21, 28

Fuente: Elaboración propia.

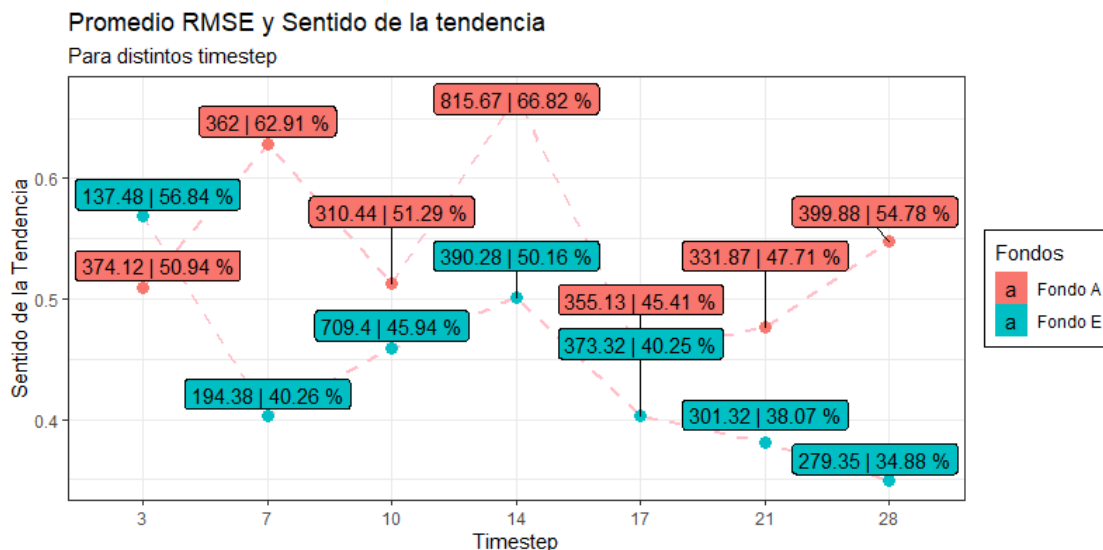


Figura 5.12: Promedio RMSE y Sentido de la tendencia al utilizar distintos números de días.

En la Figura 5.12 se tiene los resultados promedio de RMSE y sentido de la tendencia al utilizar distintos números de días como variable de entrada. Es importante señalar que estos resultados se obtuvieron utilizando un modelo simple LSTM, con 150 neuronas y 25

iteraciones.

Para ambos fondos se tienen resultados muy variables. En primer lugar, para el Fondo A se tiene un menor Sentido al utilizar 17 días como variable de entrada, alcanzado un 45,41 %, pero si se utiliza 14 días, existe una mejoría de más de 20 unidades porcentuales en el Sentido, específicamente, un 66,82%. Una de las razones de este resultado debe ser producto de una componente estacional semanal. En segundo lugar, para el Fondo E, sus valores están específicamente entre un 34 % y 57 %, pero alcanza el sentido de la tendencia más alto utilizando solo 3 días como entrada.

(2) Parámetros básicos de la red

Los parámetros básicos a optimizar son el tamaño del lote (es decir, *batch size*), número de iteraciones (*epoch*) y número de neuronas. El tamaño del lote define el número de muestras que se propagan a través de la red. En esta investigación, se establecerá con valor 1, lo que hace que el modelo tenga un aprendizaje en línea. (Xiao and Yin, 2019).

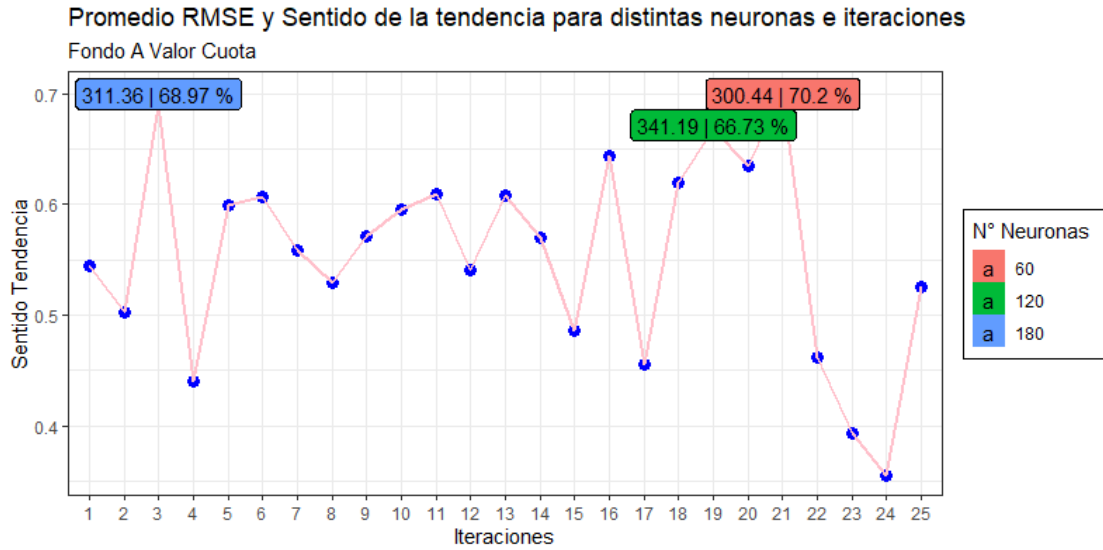
Para el caso de las neuronas e iteraciones, se probaron los siguientes valores:

Tabla 5.3: Cantidad de neuronas e iteraciones a probar

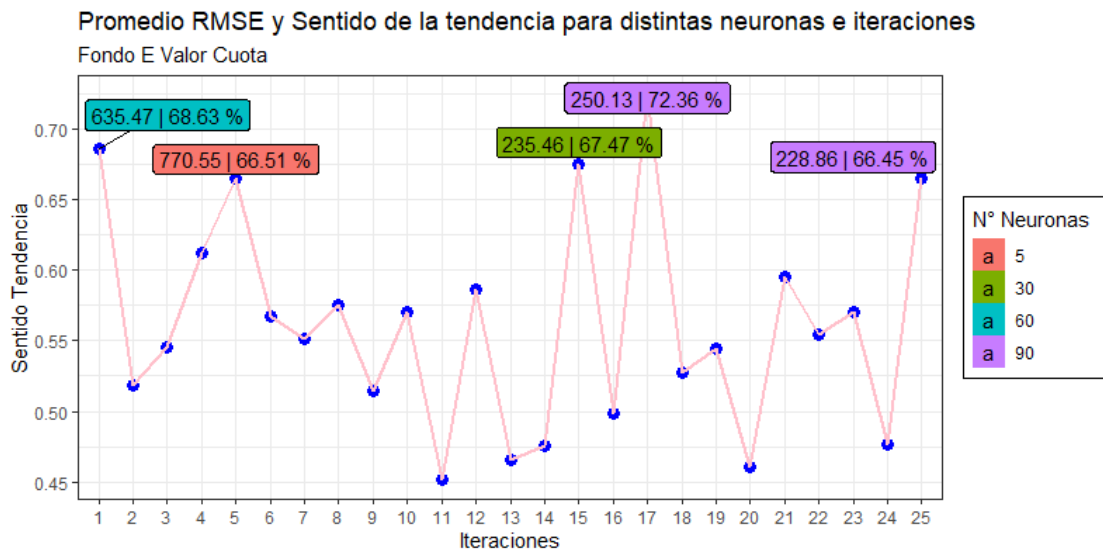
Parámetro	Valores
Neuronas	5, 10, 15, 30, 45, 60, 90, 120, 150, 180
Iteraciones	1, 2, 3, ..., 25

Fuente: Elaboración propia.

Se probaron todas las combinaciones posibles entre iteraciones y neuronas con 3 repeticiones. Los resultados del experimento son:



(a)



(b)

Figura 5.13: (a) Promedio RMSE y Sentido para distintos valores de neuronas e iteraciones para el Fondo A. (b) Promedio RMSE y Sentido para distintos valores de neuronas e iteraciones para el Fondo E.

En la Figura 5.13 se tienen los promedios de RMSE y Sentido de las 3 repeticiones con todas las combinaciones posibles de neuronas e iteraciones, donde se destacan solo los mejores resultados. Para el caso del Fondo A, la cantidad de neuronas e iteraciones que se obtiene un mayor sentido de la tendencia es con 90 neuronas en 17 iteraciones alcanzando un 70,2%,

aumentando cerca de un 4 % respecto del ítem anterior.

Por otro lado, para el Fondo E, usando 60 neuronas en 21 iteraciones alcanza un 72,36 %, que, comparado con el ítem anterior, aumenta cerca de un 16 % en el sentido de la tendencia.

(3) Optimizador

El problema de optimización es encontrar los parámetros para una función objetivo determinada, esto es que minimice el error cuadrático medio (MSE). Los optimizadores más comunes son el descenso del gradiente estocástico (SGD), Adam, Adamax, Nadam, Adadel-ta, Adagrad, RMSprop, etc. (Keras Documentation, 2020).

Para probar los distintos optimizadores, se utilizaron los mejores hiper-parámetros encontrados hasta el momento, el resto se deja por defecto.

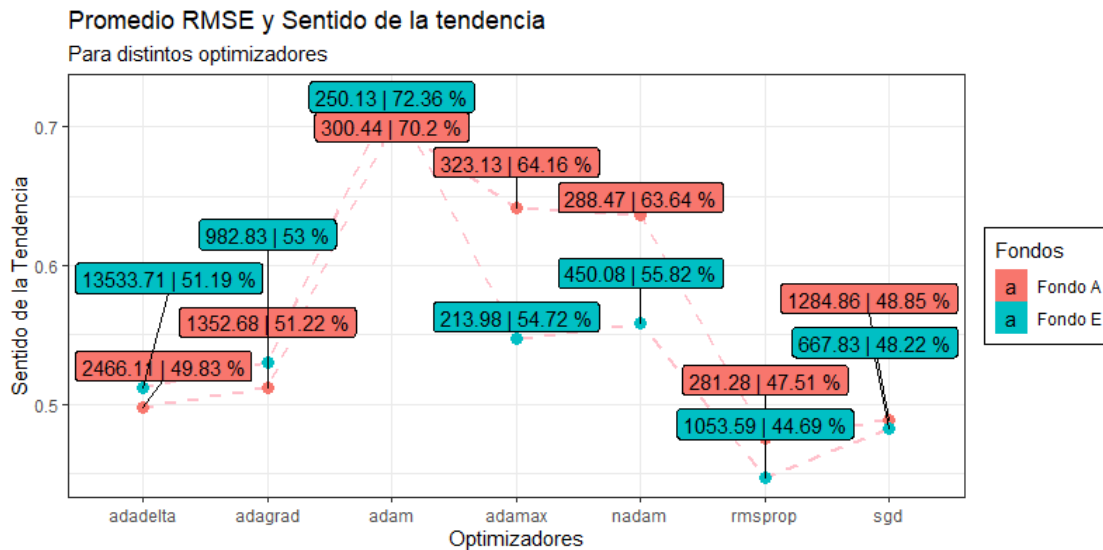


Figura 5.14: Promedio RMSE y Sentido de la tendencia al utilizar distintos optimizadores en el modelo.

En la Figura 5.14 se tiene el promedio de RMSE y Sentido utilizando distintos optimizadores para el Fondo A y el Fondo E. Por defecto, se utiliza el optimizador adam, por lo que no existe mejoría respecto al ítem anterior.

(4) Función activación capa intermedia

Las funciones de activación en la capa intermedia corresponden a las funciones de activación de la capa de activación, que conecta la capa de combinación de entrada como la capa de combinación intermedia, representada en la Figura 3.3.

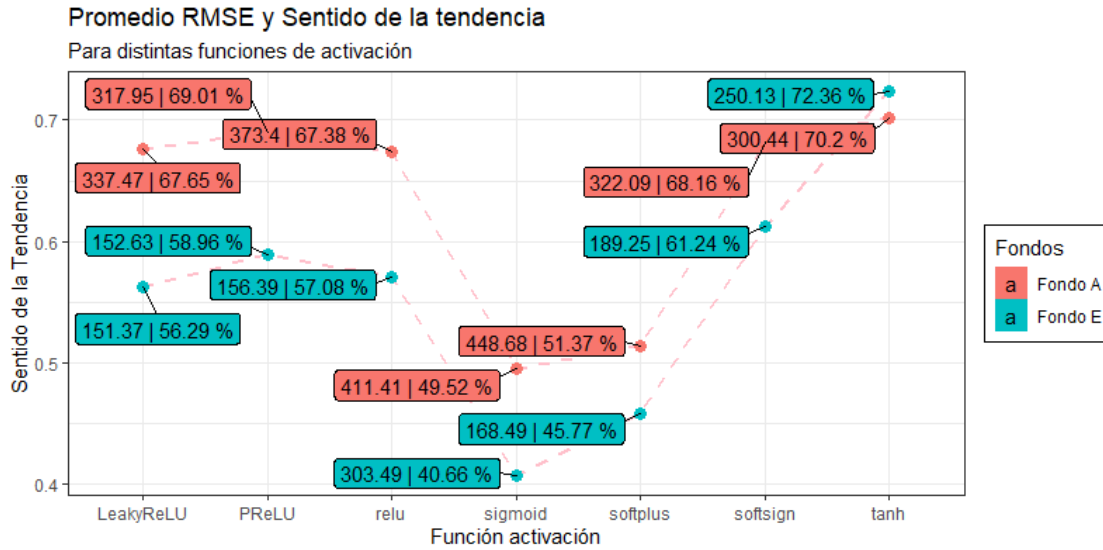


Figura 5.15: Promedio RMSE y Sentido de la tendencia para distintas funciones de activación en la capa intermedia.

En la Figura 5.15 se tiene el promedio de RMSE y Sentido para distintas funciones de activación en la capa intermedia para el Fondo A y Fondo E. Por defecto, la LSTM utiliza la función de activación tanh, siendo está la que se obtiene mejor Sentido, al igual que con los optimizadores, con un 70,2% y 72,36%, respectivamente.

(5) Dropout

Para la estructura de la red, la selección de *dropout* abarca en dos capas, en la capa que conecta la capa de entrada con la combinación de capas intermedia y las capas intermedias, sí y sólo sí, hay más de una capa densa. Para el *dropout* de las capas intermedias se fijará en 0,5 (Srivastava et al., 2014)

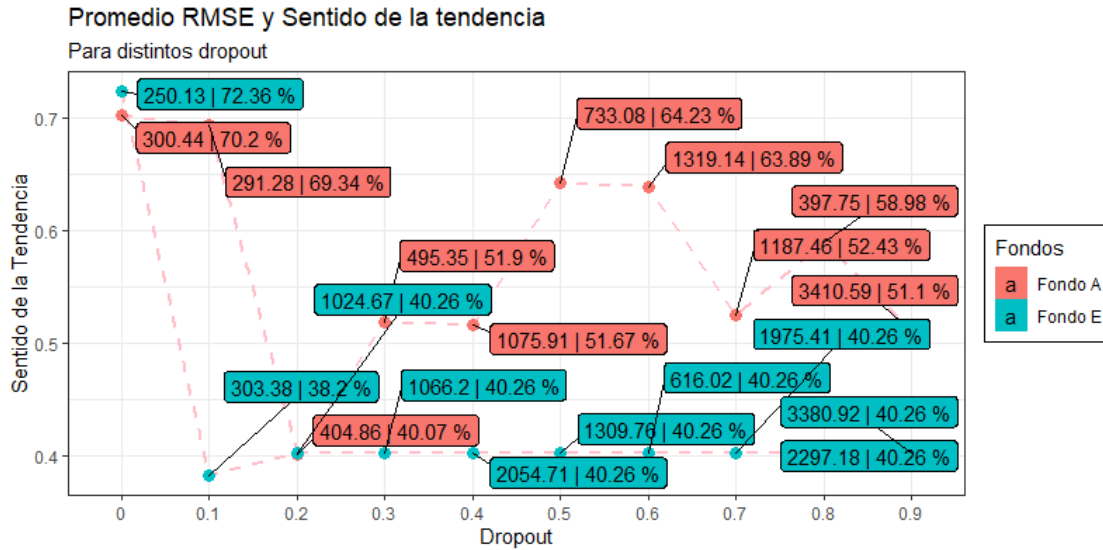


Figura 5.16: Promedio RMSE y Sentido de la tendencia para distintas tasas de *dropout*.

En la Figura 5.16 se tiene el promedio de RMSE y Sentido para distintas tasas de *dropout* en el Fondo A y Fondo E, correspondiente a la conexión de capas de entrada con la capa intermedia. Para ambos casos, el mejor resultado, tomando en consideración el sentido de la tendencia, es no aplicar *dropout*, pues repite el resultado del ítem anterior para ambos fondos.

(6) Capas densas

En la Figura 3.3 tenemos la arquitectura de la red LSTM a utilizar, donde se tiene una capa de combinación de entrada, cero o más capas de combinación intermedia y una capa de combinación de salida, donde los parámetros de la capa de combinación de entrada son los que se han descubierto hasta el momento. Para el caso de las capas densas, se utilizará la función de activación *tanh*, número de neuronas igual a la capa de entrada, *dropout* de 0.5. Por lo que los resultados son los siguientes:

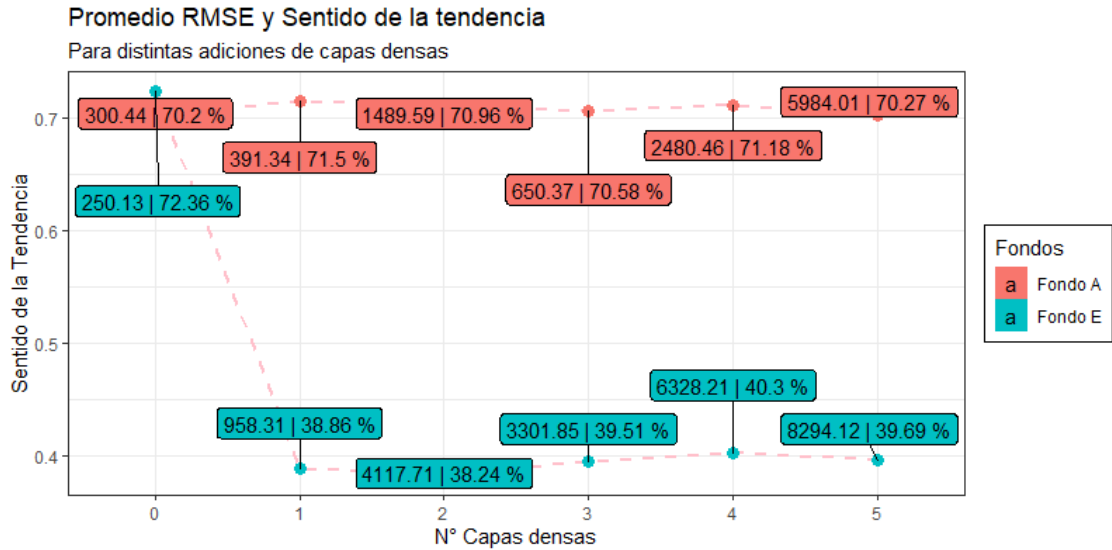


Figura 5.17: Promedio RMSE y Sentido de la tendencia al incluir cero o más capas densas al modelo.

En la Figura 5.17 se presenta el promedio de RMSE y Sentido al incluir cero o más capas densas al modelo del Fondo A y Fondo E. En primer lugar, para el Fondo A existe una mejora al incluir una capa densa adicional, logrando aumentar 1.3 unidad porcentual en el sentido de la tendencia, alcanzando un 71,5%. En segundo lugar, para el Fondo E no existe mejora al incluir capas densas, pues los resultados bajan drásticamente, teniendo una pérdida de 35-40 unidades porcentuales.

(7) Función de activación capa de salida

La función de activación de salida se refiere a la capa de función de activación conectada con la capa de combinación de salida. Los resultados son los siguientes:

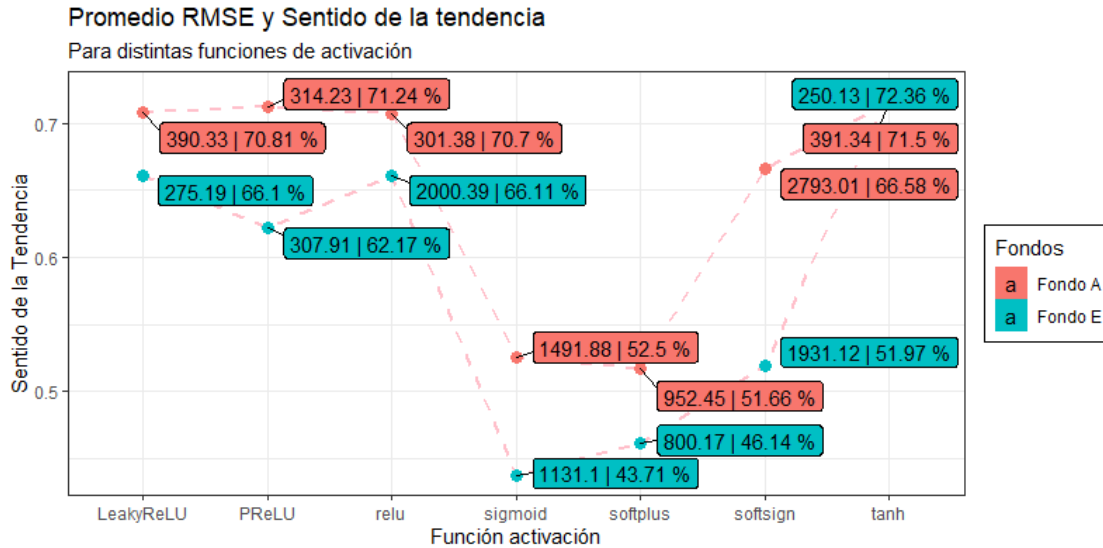


Figura 5.18: Promedio RMSE y Sentido de la tendencia para distintas funciones de activación en la capa de combinación de salida.

En la Figura 5.18 se presentan los promedios de RMSE y Sentido para distintas funciones de activación relacionada con la capa de combinación de salida para el Fondo A y el Fondo E. En ambos casos, no existe mejora respecto al ítem anterior, pues la función de activación por defecto, tanh, es la que obtuvo los mejores resultados, al igual que en la función de activación capa intermedia.

Rentabilidad real

(1) Pre-procesamiento de los datos

Antes de optimizar los hiper-parámetros de la red, es necesario un pre-procesamiento de los datos. Para la rentabilidad, hay valores tanto negativos como positivo, por lo que se normalizarán los valores entre -1 y 1.

Además, la serie temporal se transformará en un problema de aprendizaje supervisado. Al igual que el valor cuota, se probarán las mismas cantidades de días.

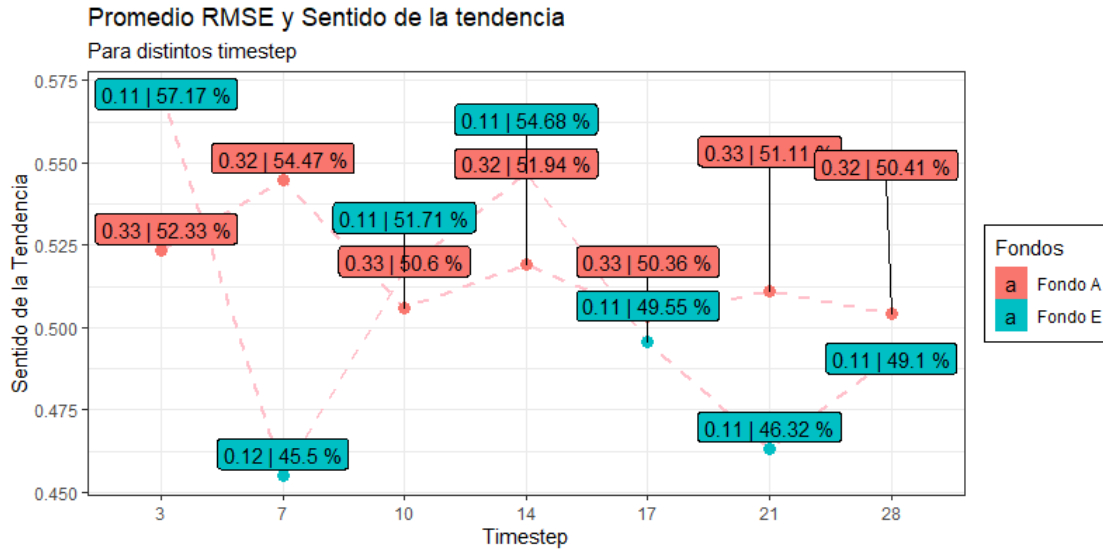


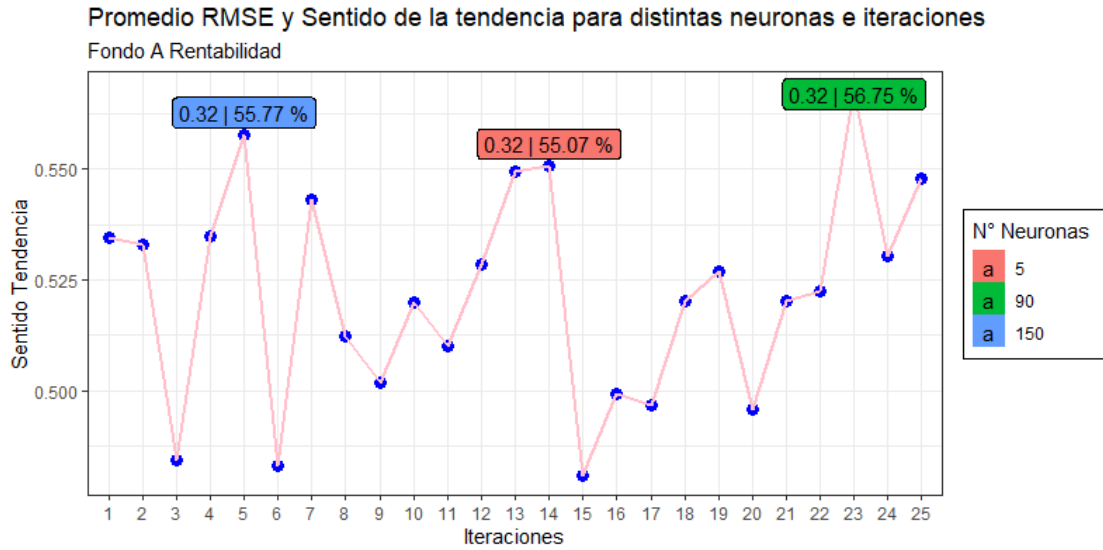
Figura 5.19: Promedio RMSE y Sentido de la tendencia distintos días como variable de entrada.

En la Figura 5.19 se tiene los resultados promedio de RMSE y sentido de la tendencia al utilizar distintos números de días como variable de entrada. Es importante señalar que estos resultados se obtuvieron utilizando un modelo simple LSTM, con 150 neuronas y 25 iteraciones.

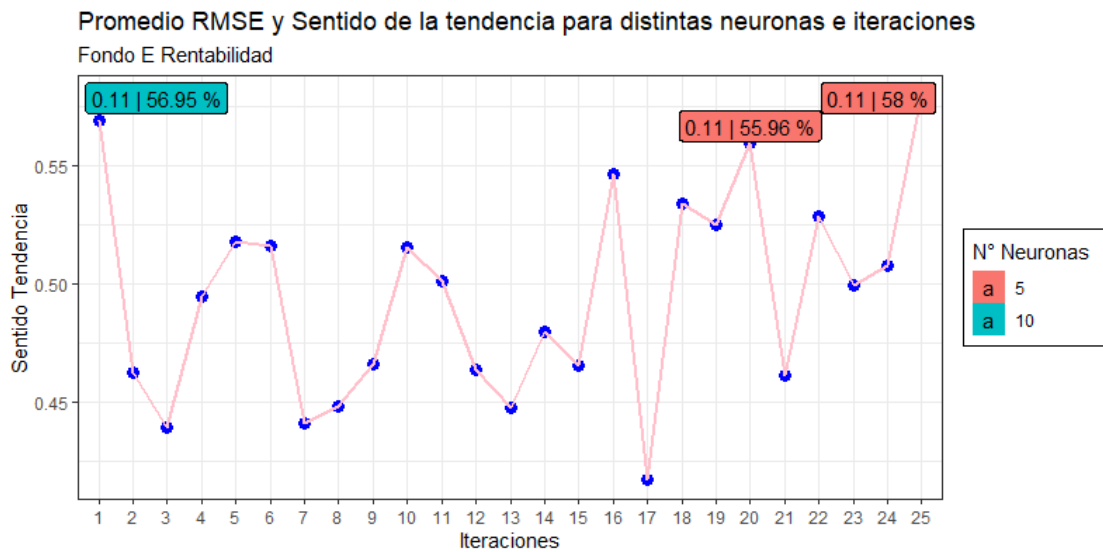
Se presentan diferencias respecto al valor cuota en el Fondo A, pues la cantidad óptima de días como variable entrada pasa a ser de 14 a 7 días, alcanzando un 54,47% en el sentido de la tendencia. Por otro lado, para el Fondo E se conserva el número de días comparado con el valor cuota, alcanzando un 57,17% en el Sentido.

(2) Parámetros básicos de la red

Los parámetros básicos a optimizar son el tamaño del lote (es decir, *batch size*), número de iteraciones (*epoch*) y número de neuronas. Al igual que con el valor cuota, se define el tamaño del lote a 1, mientras que para el número de neuronas e iteraciones se probarán todas las combinaciones posibles, generando los siguientes resultados:



(a)



(b)

Figura 5.20: (a) Promedio RMSE y Sentido para distintos valores de neuronas e iteraciones para el Fondo A. (b) Promedio RMSE y Sentido para distintos valores de neuronas e iteraciones para el Fondo E.

En la Figura 5.20 se tienen los promedios de RMSE y Sentido de las 3 repeticiones con todas las combinaciones posibles de neuronas e iteraciones, donde se destacan solo los mejores resultados. Para el caso del Fondo A, la cantidad de neuronas e iteraciones con las que se obtiene un mayor sentido de la tendencia es con 90 neuronas en 23 iteraciones alcanzando

un 56,75 %, aumentando cerca de un 3 % respecto del ítem anterior. Por otro lado, para el Fondo E, usando 5 neuronas en 25 iteraciones alcanza un 58 %, que, comparado con el ítem anterior, aumenta menos de un 1 %, respecto al sentido de la tendencia.

(3) Optimizador

Se probarán los mismos optimizadores usados en el valor cuota, además de utilizar los hiper-parámetros seleccionados hasta el momento.

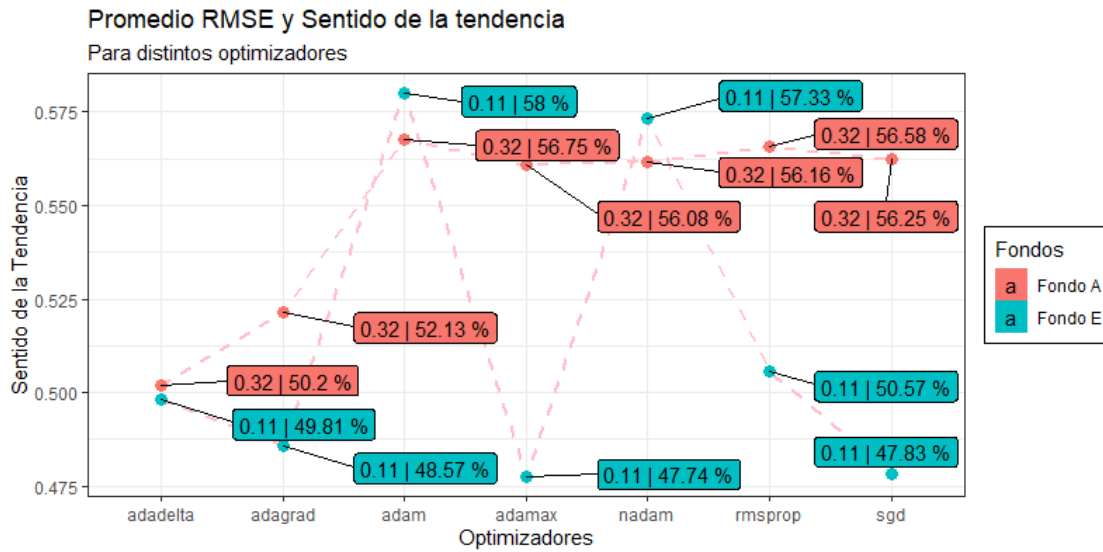


Figura 5.21: Promedio RMSE y Sentido de la tendencia al utilizar distintos optimizadores en el modelo.

En la Figura 5.21 se tiene el promedio de RMSE y Sentido utilizando distintos optimizadores para el Fondo A y el Fondo E. Por defecto, se utiliza el optimizador adam, por lo que no existe mejoría respecto al ítem anterior.

(4) Función activación capa intermedia

Las funciones de activación en la capa intermedia corresponden a las funciones de activación de la capa de activación, que conecta la capa de combinación de entrada como la capa de combinación intermedia, representada en la Figura 4.3

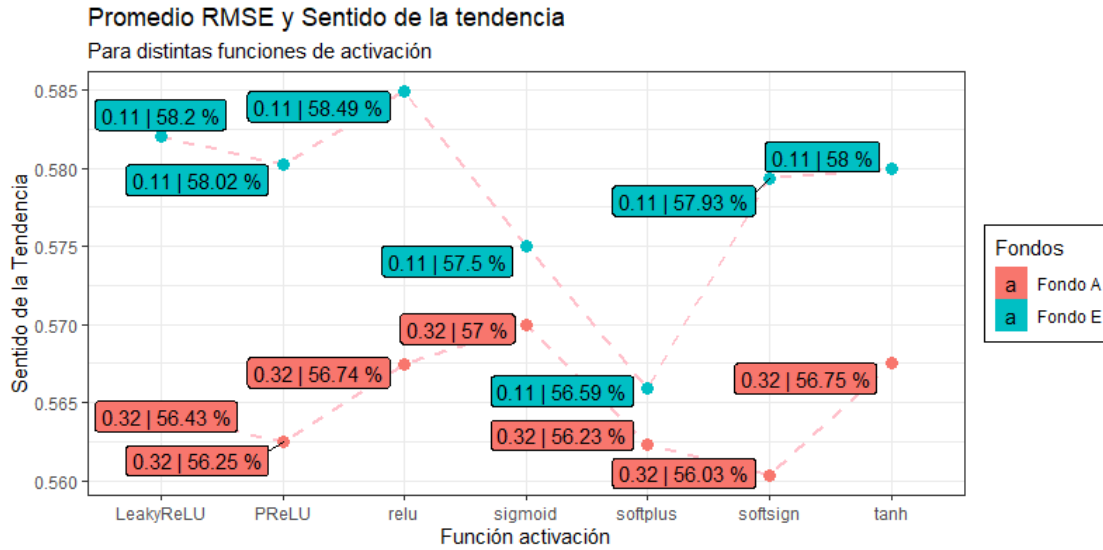


Figura 5.22: Promedio RMSE y Sentido de la tendencia para distintas funciones de activación en la capa de combinación de salida.

En la Figura 5.22 se tiene el promedio de RMSE y Sentido para distintas funciones de activación en la capa intermedia para el Fondo A y Fondo E. Existe un aumento del Sentido respecto al ítem anterior. Pues para el Fondo A, se aumenta un 0,25 unidad porcentual respecto al Sentido al utilizar la función *sigmoid*, alcanzando un 57%. Por otro lado, para el Fondo E, existe un aumento de 0,5 unidad porcentual al utilizar la función de activación 58,49%.

(5) Dropout

Para la estructura de la red, la selección de *dropout* abarca en dos capas, en la capa que conecta la capa de entrada con la combinación de capas intermedia y las capas intermedias, si y solo si, hay más de una capa densa.

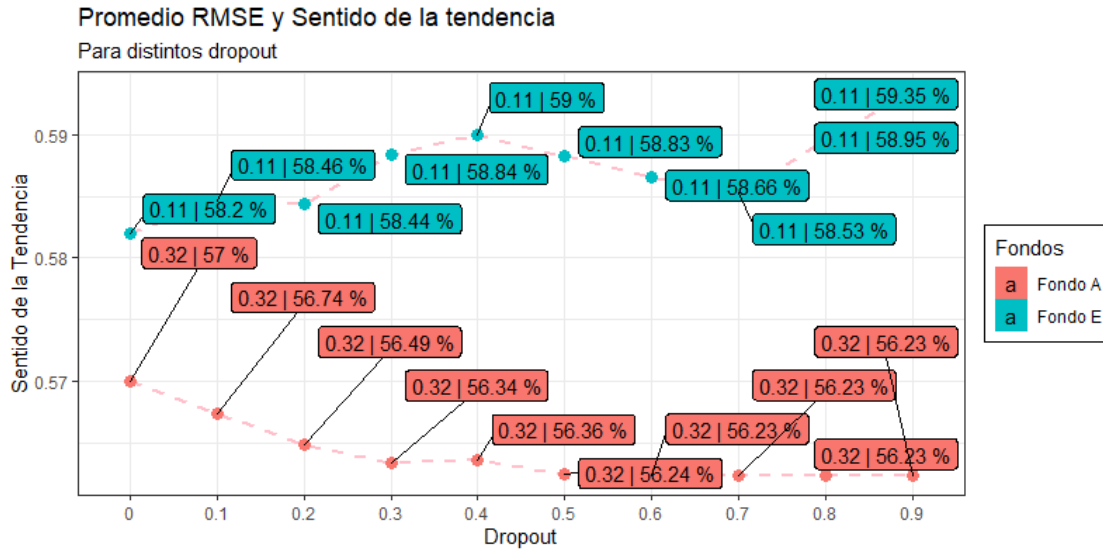


Figura 5.23: Promedio RMSE y Sentido de la tendencia para distintas tasas de *dropout*.

En la Figura 5.23 se tiene el promedio de RMSE y Sentido para distintas tasas de *dropout* en el Fondo A y Fondo E, correspondiente a la conexión de capas de entrada con la capa intermedia. Para el caso del Fondo A, no existe mejoría respecto al ítem anterior, pues al no aplicar *dropout*, se obtiene el mejor resultado. Por otro lado, para el Fondo E, existe un aumento en el sentido de la tendencia, cercana al 1%, pues al tener una tasa de *dropout* de 0,9, se alcanza un 59,35.

(6) Capas densas

En la Figura 3.3 tenemos la arquitectura de la red LSTM a utilizar, donde se tiene una capa de combinación de entrada, cero o más capas de combinación intermedia y una capa de combinación de salida, donde los parámetros de la capa de combinación de entrada son los que se han descubierto hasta el momento. Para el caso de las capas densas, se utilizará la función de activación tanh, número de neuronas igual a la capa de entrada, *dropout* de 0,5. Por lo que los resultados son los siguientes:

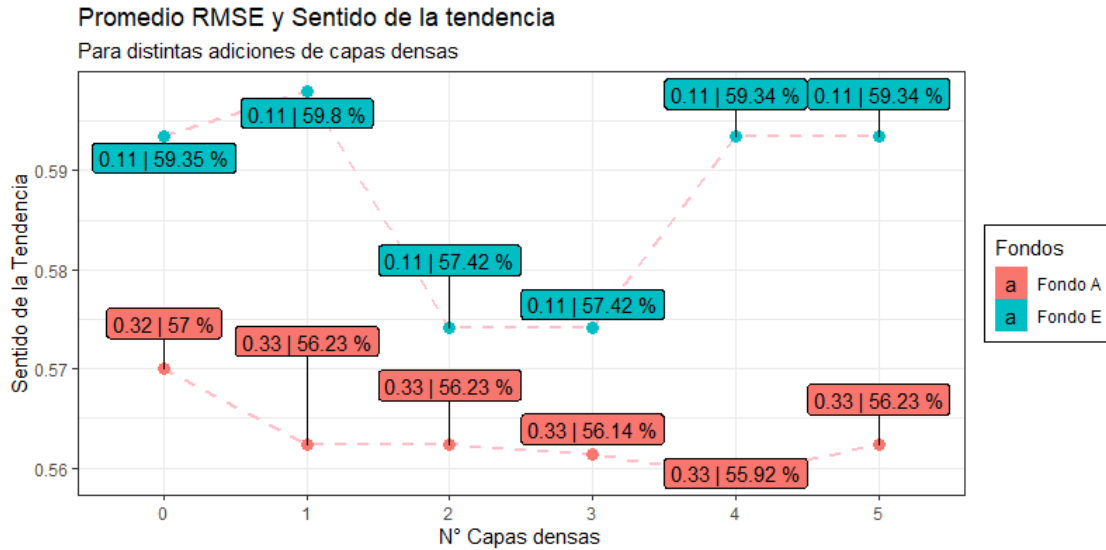


Figura 5.24: Promedio RMSE y Sentido de la tendencia al incluir cero o más capas densas al modelo.

En la Figura 5.24 se presenta el promedio de RMSE y Sentido al incluir cero o más capas densas al modelo del Fondo A y Fondo E. En primer lugar, para el Fondo A no existe una mejora al incluir capas densas adicionales. En segundo lugar, para el Fondo E existe una mejora al incluir una capa densa adicional, aumentando cerca de 0,5 unidad porcentual, logrando así un 59,8 % en el sentido de la tendencia.

(7) Función de activación capa de salida

La función de activación de salida se refiere a la capa de función de activación conectada con la capa de combinación de salida. Los resultados son los siguientes:

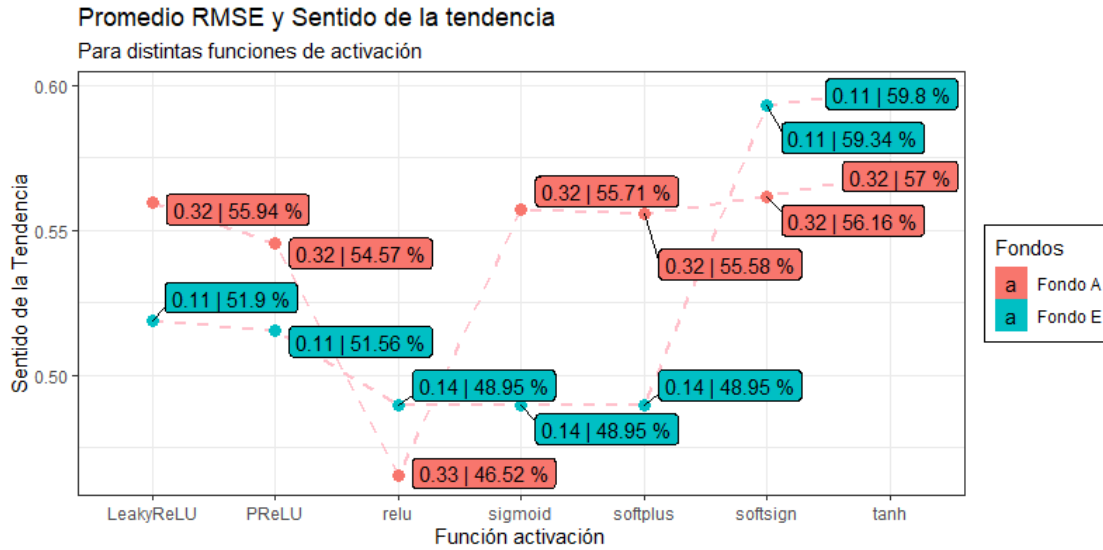


Figura 5.25: Promedio RMSE y Sentido de la tendencia para distintas funciones de activación en la capa de combinación de salida.

En la Figura 5.25 se presenta los promedios de RMSE y Sentido para distintas funciones de activación relacionada con la capa de combinación de salida para el Fondo A y el Fondo E. En ambos casos, no existe mejora respecto al ítem anterior, pues la función de activación por defecto, tanh, es la que obtuvo los mejores resultados.

5.4. Modelos Propuestos

Después del proceso experimental, se proponen modelos para los valores cuotas y rentabilidad real, correspondiente a los mejores resultados en el sentido de la tendencia. Para el modelo LSTM se tienen los siguientes hiper-parámetros.

Tabla 5.4: Hiper-parámetros para los modelos LSTM de valor cuota y rentabilidad real.

Componentes	Modelo LSTM			
	Valores Cuota		Rentabilidad Real	
	Fondo A	Fondo E	Fondo A	Fondo E
Timestep	14	3	7	3
Cantidad neuronas	60	90	90	5
Iteraciones	21	17	23	25
Optimizador	Adam	Adam	Adam	Adam
Función activación capa intermedia	Tanh	Tanh	ReLU	Sigmoid
Dropout	0	0	0	0.9
Número de capas densas	1	0	1	0
Función de activación capa de salida	Tanh	Tanh	Tanh	Tanh

Para los valores cuota, no existe mucha diferencia respecto a los hiper-parámetros de los modelos, conserva el optimizador, ambas funciones de activación y el *dropout*. Por otro lado, existe diferencia en el timestep, pues para el Fondo A el valor óptimo es utilizar 14 días como variables de entrada y para el Fondo E, solo necesita 3 días. Además, en el Fondo A se necesitan 60 neuronas con 21 iteraciones y una capa densa, para así alcanzar un 71,5% en el Sentido. Mientras que el Fondo E, se necesitan 90 neuronas con 17 iteraciones, logrando alcanzar un 72,36% en el Sentido.

Por otro lado, para la rentabilidad real se mantiene el optimizador, el número de capas densas y la función de activación en la capa de salida. Al contemplar el timestep, para el Fondo A se reduce el número de días, pasando de 14 días a 7, mientras que se mantiene en el Fondo E. Por último, para el Fondo A, se necesita 90 neuronas con 23 iteraciones y en la capa intermedia usando la función de activación *ReLU*, para así obtener un 57% en el sentido de la tendencia. En cambio, para el Fondo E, se necesita 5 neuronas en 25 iteraciones, una tasa de *dropout* de 0,9 y una función de activación de la capa intermedia *sigmoid* para alcanzar un 59,8% en el sentido de la tendencia.

Por otro lado, para los modelos Prophet, se tiene las siguientes componentes del modelo:

Tabla 5.5: Componentes de los modelos Prophet para los valores cuotas y rentabilidad real.

Componentes	Modelos Prophet			
	Valores Cuota		Rentabilidad Real	
	Fondo A	Fondo E	Fondo A	Fondo E
Componente tendencia	Crecimiento logístico	Crecimiento logístico	Tendencia lineal	Crecimiento logístico
Componentes estacionales	Multiplicativo: - Semanal - Diario	Multiplicativo: - Semanal	Multiplicativo: - Anual - Semanal - Diario	Multiplicativo: - Semanal
Holidays	No incluir fechas y eventos	No incluir fechas y eventos	No incluir fechas y eventos	Incluir fechas y eventos
Regresoras	Aditivas: - GSPC - IPSA - STOXX50E	Multiplicativa: - HSI	Multiplicativas: - HSI - IPSA	Multiplicativa: - KS11

No se hará mención a los puntos de cambio y los valores a priori para la tendencia, pues solo un modelo (Fondo A rentabilidad) posee componente de tendencia lineal y los valores por defecto, fueron los más óptimos.

Para los valores cuota, no existe mucha diferencia respecto a las componentes de los modelos, usan el crecimiento logístico para la componente tendencia, ambas consideran la componente estacional semanal como factor multiplicativo, aunque además para el Fondo A se le adiciona una componente diaria. Para ambos Fondos no se considera incluir fechas y eventos especiales y la gran diferencia son en las regresoras a incluir. Por una parte, para el Fondo A, se le agrega como variables aditivas el índice GSPC, IPSA Y STOXX50E, alcanzando un 54,45 % en el sentido de la tendencia, mientras que para el Fondo E solo se agrega como variable multiplicativa el índice HSI, alcanzando un 71,55 % en el Sentido.

Por otro lado, para la rentabilidad real, se mantiene diferencias entre los fondos, pues para la tendencia, el Fondo A propone una tendencia lineal, mientras que para el Fondo E propone crecimiento logístico. En las componentes estacionales, son similares a los valores cuotas, para el Fondo A se le adiciona una componente anual como factor multiplicativo y al Fondo E se mantiene la componente semanal. Además, solo al Fondo E se tiene una mejoría al incluir fechas y eventos especiales. Por último, para el Fondo A, se le agrega como variables regresoras multiplicativas el índice HSI e IPSA, alcanzando un 54,76 % en el sentido de la tendencia, mientras que para el Fondo E se le adiciona como regresoras solo el índice KS11 como factor multiplicativo, alcanzando un 53,7 % en el Sentido.

Capítulo 6

Resultados experimentales

6.1. Valor cuota

En el ítem anterior se llegó a la propuesta de modelos relacionados al Fondo A y Fondo E, donde para el modelo Prophet se buscaron las componentes que maximizarán el sentido de la tendencia. Ahora, usando esos modelos se ratificarán con la validación cruzada de avance utilizando siempre el 80% de los datos como entrenamiento, logrando un total de 181 predicciones de una semana para el Fondo A y 207 para el Fondo E. Los resultados del proceso de validación son:

Tabla 6.1: Resultados obtenidos con Prophet sobre los valores cuota utilizando la validación de avance.

Prophet	Valor cuota				Rentabilidad		
	RMSE	MAE	MAPE	Sentido	RMSE	MAE	Sentido
Fondo A							
Máximo RMSE	5533.555	5493.600	10.537	0.333	1.316	0.683	0.833
Mínimo RMSE	35.534	31.193	0.070	0.667	0.074	0.059	0.667
Promedio	1389.384	1375.965	3.216	0.714	0.322	0.228	0.596
Desv. estandar	1038.654	1043.151	2.273	0.167	0.161	0.108	0.197
Fondo E							
Máximo RMSE	5155.449	5154.563	10.081	0.5	2.279	1.528	0.667
Mínimo RMSE	14.721	11.646	0.081	0.667	0.016	0.011	1
Promedio	845.914	842.861	1.820	0.738	0.108	0.076	0.680
Desv. estandar	1166.519	1166.805	2.318	0.184	0.177	0.120	0.171

Fuente: Elaboración propia.

En la Tabla 6.1 están los resultados obtenidos por la validación cruzada de avance para el modelo Prophet. Por otro lado, se registran los resultados al estudiar las variaciones de las predicciones del valor cuota (Rentabilidad) para ambos fondos.

Para el caso del valor cuota Fondo A, del total de 181 predicciones realizadas, la predicción con el RMSE más alto es de 5533,6, y con el RMSE más bajo es de solo un 35,5, teniendo así un promedio de 1389,4 y una desviación estándar de 1038,7. Por otro lado, si se observa el MAPE, existe una diferencia de 10,467 entre el valor máximo y el mínimo, teniendo un promedio de 3,216 y una desviación estándar de 2,273. Es importante recordar que, en el proceso de optimización del modelo, en su proceso de validación alcanzó un 54,45% en el sentido de la tendencia, pero al aplicar la validación cruzada de avance, el promedio de las 181 predicciones en el Sentido aumentó a 71,4% teniendo una desviación estándar de 16,7%.

Ahora, si las predicciones obtenidas se transforman a rentabilidad, el promedio del sentido de la tendencia disminuye cercano al 12%, alcanzando solo un 59,6%.

Por otro lado, se tienen los resultados relacionados al Fondo E. Dentro de todas las predicciones realizadas en la validación cruzada de avance, el máximo RMSE obtenido es de 5155,5 y un mínimo de solo 14,7, alcanzando un promedio de 845,9 y una desviación estándar de 1166,5. Además, se obtiene en promedio un MAPE menor al Fondo A, pues es solo de un 1,8 con una desviación de 2,3. Al observar el sentido de la tendencia, en el proceso experimental, el Fondo E obtuvo un 71,55%. Pero al aplicar la validación de avance, en promedio se obtuvo un Sentido de 73,8%, aumentando un 2,3 unidades porcentuales con una desviación estándar de 18,4%.

Si las predicciones obtenidas se transforman a rentabilidad, a diferencia del Fondo A que disminuyó cerca de un 12%, el Fondo E solo disminuye cerca de un 5%, pues obtiene un promedio de Sentido de un 68%.

Para el caso del modelo LSTM, el proceso experimental ya llevaba a cabo la validación cruzada de avance, a diferencia de Prophet. Por lo que, los resultados promedios son los mismos. En resumen, se tiene los siguientes resultados:

Tabla 6.2: Resultados obtenidos con LSTM sobre los valores cuota utilizando la validación de avance.

LSTM	Valor cuota				Rentabilidad		
	RMSE	MAE	MAPE	Sentido	RMSE	MAE	Sentido
Fondo A							
Máximo RMSE	1453.286	1279.250	2.441	0.333	1.376	0.831	0.500
Mínimo RMSE	65.035	53.141	0.129	0.833	0.106	0.090	0.833
Promedio	391.335	363.372	0.858	0.715	0.475	0.319	0.638
Desv. estandar	211.070	204.114	0.470	0.165	0.233	0.136	0.200
Fondo E							
Máximo RMSE	1956.600	1608.492	3.495	0.667	2.683	2.087	0.667
Mínimo RMSE	39.414	28.324	3.495	0.500	0.049	0.048	0.833
Promedio	214.623	204.188	0.456	0.7351	0.172	0.139	0.655
Desv. estandar	212.069	198.008	0.406	0.187	0.214	0.162	0.185

Fuente: Elaboración propia.

Para el caso del Fondo A, del total de 181 predicciones realizadas, la predicción con el RMSE más alto es de 1453,3, y con el RMSE más bajo es de solo un 65, teniendo así un promedio de 391,3 y una desviación estándar de 211,1. Por otro lado, si se observa el MAPE, existe una diferencia de 2,3 entre el valor máximo y el mínimo, teniendo un promedio de 0,9 y una desviación estándar de 0,5. Al contemplar el sentido de la tendencia, en promedio, dentro de todas las predicciones, se obtuvo un 71,5 %, con una desviación estándar de 16,5 %.

Ahora, si las predicciones obtenidas se transforman a rentabilidad, el promedio del sentido de la tendencia disminuye cercano al 7,7 %, alcanzando 63,8 %, con una desviación estándar de 20 %.

Por otro lado, considerando al Fondo E, el máximo RMSE obtenido es de 1956,6 y un mínimo de solo 39,4, alcanzando un promedio de 214,6 y una desviación estándar de 212,1. Además, se obtiene en promedio un MAPE menor al Fondo A, pues es solo de un 0,5 con una desviación de 0,5. Al observar el sentido de la tendencia, en promedio se obtuvo un Sentido de 73,5 %, 2 unidades porcentuales más que el Fondo A, con una desviación estándar de 18,7 %.

Si las predicciones obtenidas se transforman a rentabilidad, se tiene una disminución cercana al Fondo A, pues disminuye un 8 %, alcanzo un Sentido de 65,5 % con una desviación estándar de 18,5 %.

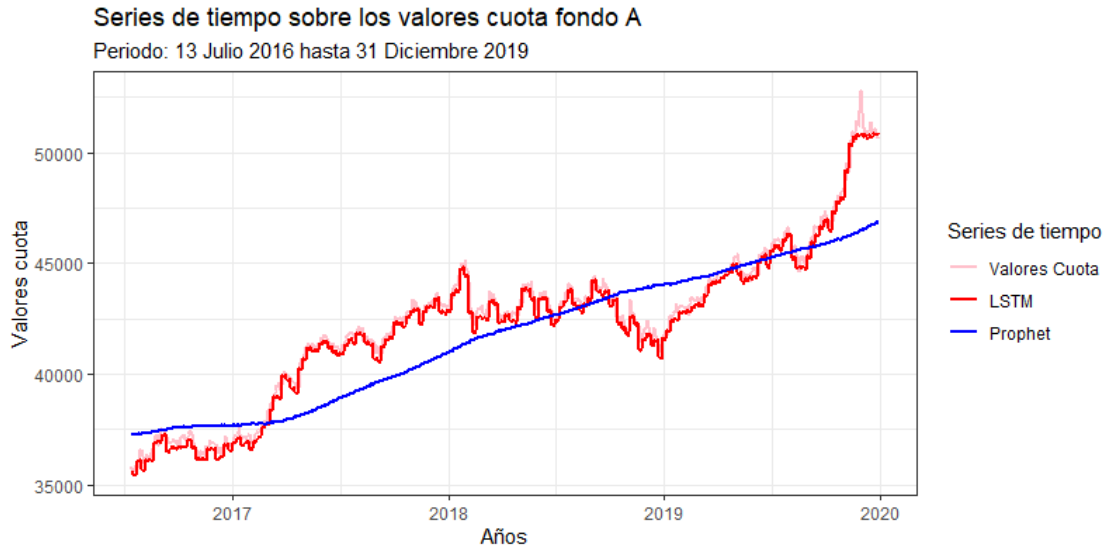


Figura 6.1: Ajuste de los modelos Prophet y LSTM para los valores cuota en el fondo A.

En la Figura 6.1 se presenta los ajustes de los distintos modelos al valor cuota del fondo A en el periodo de validación, para realizar el gráfico, se unieron todas las predicciones de 7 días en la validación cruzada de avance, así poder determinar el horizonte total del periodo de validación. Los valores cuota del fondo A presentan oscilaciones durante el periodo de validación, pero destaca una tendencia al alza, marcada principalmente en el periodo del año 2019. Por su parte, el modelo LSTM captura de buena manera la serie del valor cuota, por lo que se obtiene un promedio de RMSE bajo, siendo de solo 391,3. En cambio, Prophet captura solo la tendencia con unas pequeñas oscilaciones, no representa de buena manera el comportamiento de la tendencia, pues en promedio, su RMSE es de 1389,4.

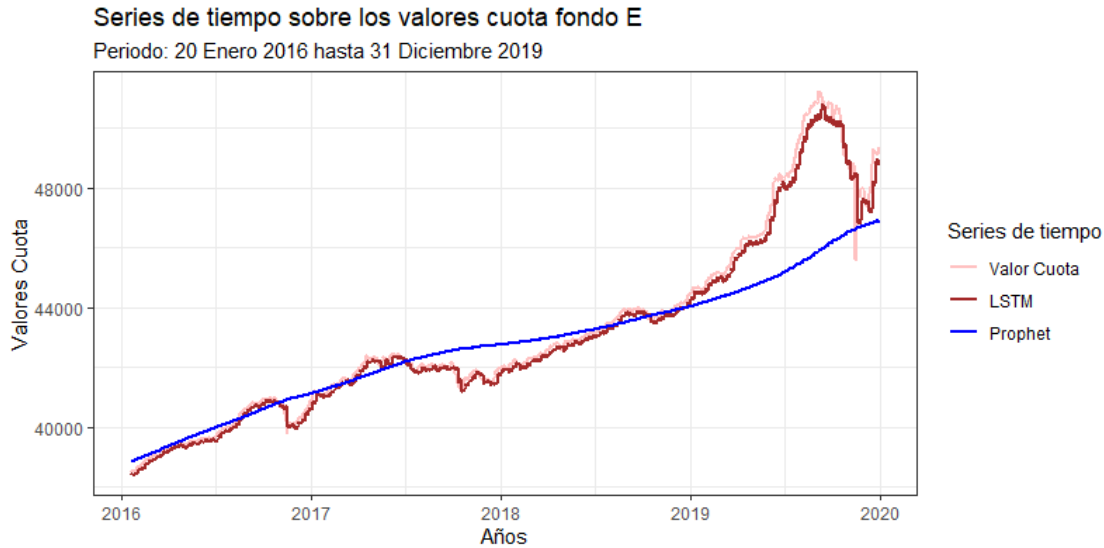


Figura 6.2: Ajuste de los modelos Prophet y LSTM para los valores cuota en el fondo E.

En la Figura 6.2 se presenta los ajustes de los distintos modelos al valor cuota del fondo E en el periodo de validación, al igual que en el fondo A, se unieron las predicciones de 7 días del proceso de validación. Los valores cuota del fondo E presentan una tendencia al alza con unos pequeños declives. El año más característico es el 2019, donde más de la mitad del año posee un alza considerable hasta el “estallido social”, donde se refleja una caída cercana a 2000 valores cuota, que luego de unas semanas vuelve al alza. El modelo LSTM captura de buena manera la serie, siguiendo la tendencia casi a la perfección, pronosticando esta alza y caída producto del año 2019. No así Prophet, que tiene un comportamiento de recta creciente con pequeñas fluctuaciones.

6.2. Rentabilidad real

Durante la investigación, se expresó que se aplicarían ambas técnicas estadísticas a dos tipos de datos distintos, al valor cuota y la rentabilidad real. En esta sección se presentarán los resultados asociados a la rentabilidad real, una obtenida por la variación de los valores cuota, pero deflactada por la UF, y la otra por el aprendizaje directamente desde la rentabilidad real.

Los resultados para Prophet son los siguientes:

Tabla 6.3: Resultados obtenidos con Prophet sobre la rentabilidad real utilizando la validación de avance.

Prophet	Rentabilidad real (usando Vc)			Rentabilidad real		
	RMSE	MAE	Sentido	RMSE	MAE	Sentido
Fondo A						
Máximo RMSE	1.316	0.683	0.833	1.321	0.730	0.500
Mínimo RMSE	0.074	0.059	0.667	0.049	0.041	0.5
Promedio	0.322	0.228	0.546	0.322	0.232	0.537
Desv. estandar	0.161	0.108	0.202	0.158	0.109	0.215
Fondo E						
Máximo RMSE	2.279	1.527	0.500	1.679	1.051	0.667
Mínimo RMSE	0.016	0.011	1	0.017	0.013	0.667
Promedio	0.108	0.076	0.620	0.110	0.077	0.547
Desv. estandar	0.177	0.120	0.192	0.172	0.100	0.211

Fuente: Elaboración propia.

En el proceso experimental del Fondo A, para la rentabilidad real se obtuvo un sentido de la tendencia de 54,76 %, pero al validar el modelo de validación cruzada de avance, en promedio, el Sentido se reduce cercano a la unidad porcentual, alcanzando un 53,7 %, con una desviación estándar de 21,5 %. Pero, deflactando las variaciones de rentabilidad obtenida por los valores cuota, en promedio se obtiene un 54,6 %, con una desviación estándar de 20,2 %.

Por otro lado, para el Fondo E, en el proceso experimental se obtuvo un 53,7 % en el sentido de la tendencia, pero al validar por validación cruzada de avance, en promedio, el Sentido es de 54,7 %, aumentando una unidad porcentual respecto al proceso experimental, además de obtener una desviación estándar de 21,1 %. Ahora, si se consideran los resultados del valor cuota, el promedio de la rentabilidad real aumenta al 62 %.

Por último, se tiene los resultados del modelo LSTM asociado a la rentabilidad real y la variación deflactada por la UF de las predicciones del valor cuota. Los resultados son:

Tabla 6.4: Resultados obtenidos con LSTM sobre la rentabilidad real utilizando la validación de avance.

LSTM	Rentabilidad real (usando Vc)			Rentabilidad real		
	RMSE	MAE	Sentido	RMSE	MAE	Sentido
Fondo A						
Máximo RMSE	1.376	0.831	0.500	1.297	0.729	0.667
Mínimo RMSE	0.106	0.089	0.833	0.052	0.044	0.500
Promedio	0.475	0.319	0.590	0.322	0.233	0.571
Desv. estandar	0.233	0.137	0.213	0.155	0.107	0.189
Fondo E						
Máximo RMSE	2.682	2.086	0.667	2.286	1.537	0.667
Mínimo RMSE	0.049	0.0480	0.667	0.017	0.028	1
Promedio	0.172	0.139	0.592	0.109	0.079	0.598
Desv. estandar	0.214	0.162	0.205	0.175	0.119	0.194

Fuente: Elaboración propia.

Para el caso del Fondo A, del total de 181 predicciones realizadas, la predicción con el RMSE más alto es de 1,30, y con el RMSE más bajo es de solo un 0,05, teniendo así un promedio de 0,32 y una desviación estándar de 0,16. Al contemplar el sentido de la tendencia, en promedio, dentro de todas las predicciones, se obtuvo un 57,1 %, con una desviación estándar de 18,9 %.

Ahora, si las predicciones obtenidas del valor cuota se transforman a rentabilidad real, el promedio del sentido de la tendencia aumenta cercano al 2 %, alcanzando 59 %, con una desviación estándar de 21,3 %.

Por otro lado, considerando al Fondo E, el máximo RMSE obtenido es de 2,29 y un mínimo de solo 0,02, alcanzando un promedio de 0,11 y una desviación estándar de 0,18. Al observar el sentido de la tendencia, en promedio se obtuvo un Sentido de 59,8 %, 2,7 unidades porcentuales más que el Fondo A, con una desviación estándar de 19,4 %.

Ahora, si las predicciones obtenidas del valor cuota se transforman a rentabilidad real, el promedio del sentido de la tendencia disminuye cercano al 0,6 %, alcanzando 59,2 %, con una desviación estándar de 20,5 %.

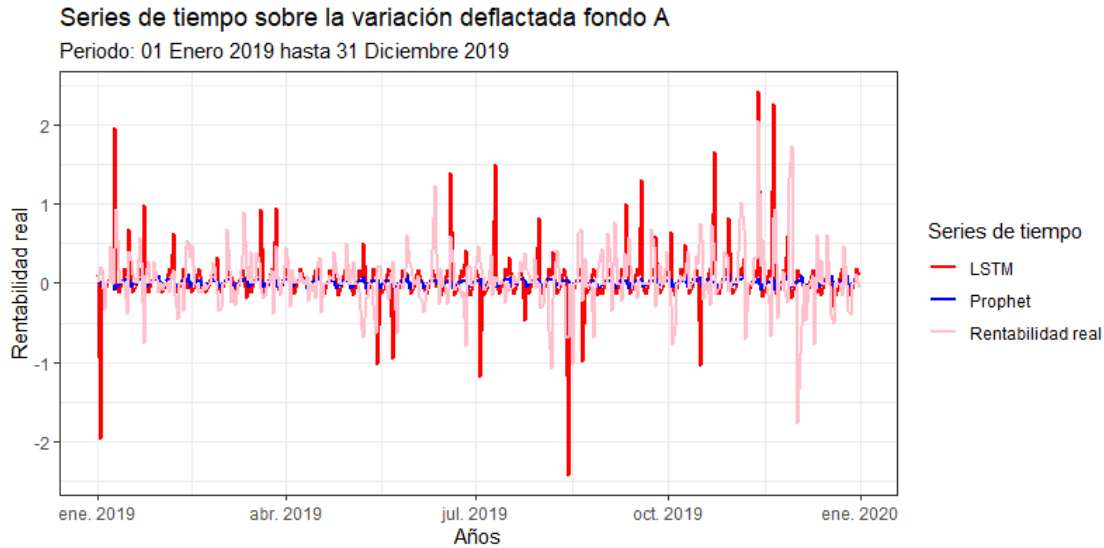


Figura 6.3: Ajuste de los modelos Prophet y LSTM para la rentabilidad real fondo A.

En la Figura 6.3 se presenta el ajuste de los modelos y la rentabilidad real del fondo A, durante el año 2019. El comportamiento de Prophet es regular, oscila en torno al cero sin mayores riesgos. En cambio, el modelo LSTM presenta un comportamiento distinto, tiene una postura más arriesgada donde en la mayoría de los casos sigue las oscilaciones de la serie original, sobre todo en el periodo septiembre-diciembre.

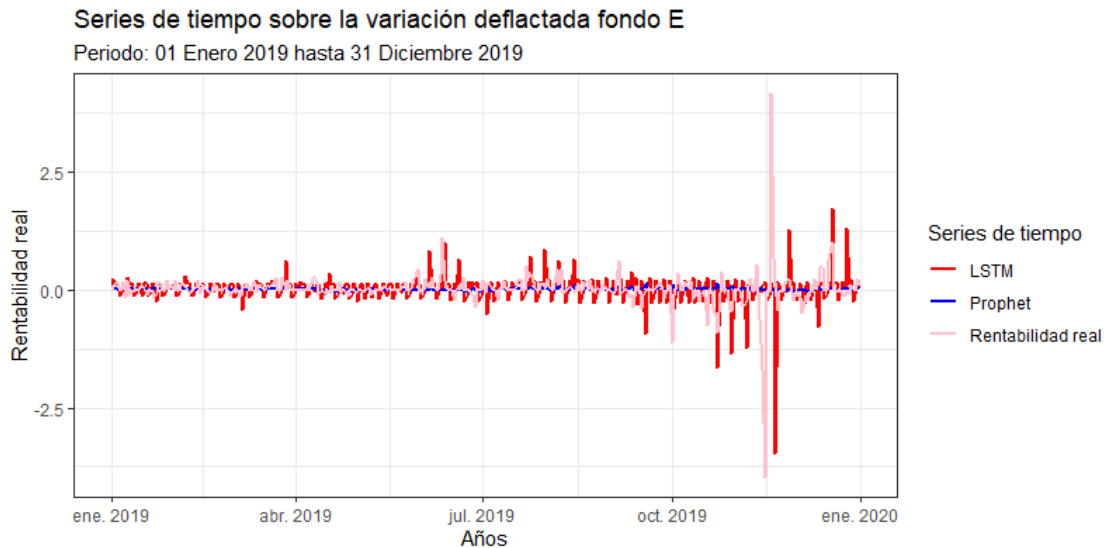


Figura 6.4: Ajuste de los modelos Prophet y LSTM para la rentabilidad real fondo E.

En la Figura 6.4 se presenta el ajuste de los modelos y la rentabilidad real del fondo E, durante el año 2019. El comportamiento de la rentabilidad real es bastante regular, se

mantiene cercano al cero durante la mayoría del periodo. Ya alcanzando octubre cambia el rumbo de la serie, presentando una inestabilidad mayormente a la pérdida producto de la crisis social. Al igual que en el fondo A, Prophet mantiene su comportamiento regular, oscilando cercano a cero durante todo el periodo. No así el modelo LSTM, que pronostica la pérdida ocurrida en octubre, pero sobrestima el alza producida semanas después.

6.3. Discusión de resultados

En base a los resultados obtenidos, se pueden percibir diferencias y similitudes respecto al tipo de fondo, dato a utilizar y técnica estadística a aplicar.

Si uno considera sólo dato, indiferente a la técnica a utilizar o el respectivo fondo, en todos los casos se obtuvieron mejores o similares resultados en el aprendizaje sobre el valor cuota y aplicar la transformación de la rentabilidad real a diferencia que los modelos se ajusten directamente de esta rentabilidad. El caso más significativo fue utilizando Prophet en el Fondo E, donde hubo una diferencia cercana del 8% en el Sentido.

Existe un detalle importante al ocupar el valor cuota, pues para calcular la rentabilidad real del fondo, es necesario deflactar por la UF. Entonces si se observa la rentabilidad obtenida por la variación del valor cuota y la rentabilidad deflactada, en todos los casos se redujo, en promedio el sentido de la tendencia, entre un 5 – 7%. Por lo que la variación de la UF afecta considerablemente los resultados.

Ahora, si se compara solo entre fondos, en promedio, el Fondo E siempre obtuvo mejor predicción en el sentido de la tendencia, esto debido principalmente a las características del fondo, pues, desde su creación, la finalidad del Fondo E es proteger a los afiliados pensionados y los próximos a pensionarse, pues al poseer principalmente inversión de renta fija, tiene menos fluctuaciones que el fondo A.

Por último, se tienen los promedios del sentido de la tendencia producto de todas las predicciones obtenidas en la validación cruzada de avance.

Tabla 6.5: Promedio sentido de la tendencia para distintos fondos y técnicas estadísticas.

Tipo de Fondo	Modelo Prophet		Modelo LSTM	
	Variación Vc deflactada	Rentabilidad real	Variación Vc deflactada	Rentabilidad real
Fondo A	54.6	53.7	59.0	57.1
Fondo E	62.0	54.7	59.2	59.8

Si se observa los resultados expresados en la Tabla 6.5, en la gran mayoría de los casos, el modelo LSTM obtuvo mejores resultados en la precisión del sentido de la tendencia, aunque

la mayor precisión la obtuvo el modelo Prophet con el Fondo E utilizando las variaciones del valor cuota deflactada, alcanzando un 62 % de precisión en el Sentido. Por lo que definir qué modelo obtiene mejores resultados, principalmente se debe a otras métricas, como pueden ser el RMSE, MAPE, tiempo de ejecución, dificultad para llevar a cabo la técnica, etc.

Ahora, si uno compara los resultados obtenidos con los de la literatura se tiene lo siguiente:

Alarcón (2014) demostró que la red neuronal Ward permite predecir de forma estable el cambio de signo del valor cuota. Pero la metodología implementada es distinta a la realizada en este trabajo, por lo que no se puede comparar.

Garrido (2014) obtuvo para el pronóstico de 5 días del valor cuota del fondo A de la AFP CUPRUM en promedio, un 69,33 % en el sentido de la tendencia y un MAPE de 0,5312 %. Garrido (2014) obtiene un menor MAPE que Prophet y LSTM, 3,216 % y 0,858 %, respectivamente. Pero en el caso del sentido de la tendencia, se obtienen, en este trabajo, mejores resultados, 71,4 % y 71,5 %, respectivamente.

Ramos (2015) obtuvo para el pronóstico de 7 días del fondo A de la AFP CUPRUM, en primer lugar, para el valor cuota, en promedio un 63,20 % para el Sentido y un MAPE de 3,009 %. En segundo lugar, para la rentabilidad, en promedio se tiene un 74,34 % para el Sentido. Si consideramos el valor cuota, Prophet posee una MAPE similar a Ramos (2015), pero un Sentido mayor, presentando una mejora cercana a 7,2 unidades porcentuales, por otro lado, LSTM supera en ambas métricas, pues posee 2,2 unidades porcentuales menos en el MAPE y 7,3 % más en el Sentido. Ahora si consideramos la rentabilidad, el sentido de la tendencia de Ramos (2015) es superior a Prophet y LSTM, siendo superados en más de 10 unidades porcentuales.

Capítulo 7

Conclusión

Todo el tema asociado a los fondos de pensiones no debería tomarse a la ligera, pues hay que considerar que nuestras pensiones son productos, en primer lugar, por el capital reunido durante todo nuestro periodo laboral y en segundo lugar, por las rentabilidades de las inversiones generadas por las AFP. Esto es crucial, ya que, dentro del total acumulado en nuestra cuenta individual, solo el 27 % corresponde a lo acumulado laboralmente, mientras que el restante 73 %, es debido a las administradoras de los fondos de pensiones.

Entonces uno podría pensar que las AFP de alguna manera realizan bien sus administraciones, pues más de la mitad del capital acumulado es producto de ellos. Generalmente es así, pero a lo largo de la historia ocurren eventos a nivel mundial, que afectan rotundamente a las inversiones y todos los fondos generan retornos negativos, como lo puede ser la crisis Subprime de 2008 o la actual crisis producto de la expansión del COVID-19.

Según la Superintendencia de pensiones, debido a la expansión del COVID-19 y la emergencia sanitaria a nivel global, han producido una caída drástica de los mercados mundiales, produciendo perdidas no solo de instrumento de renta fija, sino también de renta variable. La rentabilidad del primer trimestre del año 2020 es:

Tabla 7.1: Rentabilidad real en porcentaje de los Fondos de Pensiones.

Fondos de Pensiones	Marzo 2020	Acumulado Ene2020-Mar2020
Fondo A	-16.49	-17.02
Fondo B	-13.53	-14.22
Fondo C	-10.20	-11.64
Fondo D	-7.22	-9.64
Fondo E	-2.54	-4.81

Fuente: Superintendencia de pensiones.

Como se mencionó en la introducción, el año 2019 se caracteriza porque a nivel de sis-

tema, fue el segundo año en obtener los mejores resultados, después del año 2019. Donde el Fondo A y el Fondo E, rentaron 17,26 % y 9,03 %, respectivamente. Por otro lado, en la Tabla 8.1 se expone en porcentaje la rentabilidad de los 5 fondos, durante marzo del año 2020 y el primer trimestre del mismo año. Si uno considera los Fondos más riesgosos (A y B), marzo destacó por las caídas de un 16,49 % y 13,53 %, respectivamente. Con esto, entonces el primer trimestre registro una pérdida de 17,02 % y 14,22. Esta crisis en particular, también ha afectado a los fondos que poseen la menor renta variable, como son los Fondos D y E. Mientras el fondo D perdió 7,22 % en el mes y 9,64 % en el primer trimestre, el fondo E retrocedió 2,54 % en el mes y 4,81 % en este primer trimestre.

Si uno observa las grandes ganancias obtenidas en el año 2019 y las pérdidas obtenidas en el 2020, más aún, los fondos que se denominan como seguro también están generando rentabilidades negativas, demuestra que el tema de los fondos de pensiones es un tema no menor, pues está en juego tu futura jubilación. Existen ventajas respecto del sistema de pensiones, pues uno puede escoger en que fondo situarse, con excepción de restricción de edad a los fondos más riesgosos. Por lo que escoger en que fondo estar en un respectivo tiempo, puede generar bastante ventajas, pues cuando los fondos vallan al alza, situarse en los fondos más riesgosos para generar la mayor rentabilidad posible, o bien, cuando los fondos vallan a la pérdida, poder refugiarse en los fondos más conservadores, así obtener la menor pérdida posible.

Este tipo de consejos respecto a que fondo situarse, los realizan algunas AFP y otros servicios pagados, pero sería interesante poder complementar esta información con otras herramientas. El tema de predicción asociado a los fondos de pensiones es un tema relativamente nuevo, a pesar de los años que tiene el sistema, existen pocas investigaciones al respecto. Es por esto el interés de probar nuevas herramientas respecto a los fondos de pensiones, como son Prophet, un modelo de serie de tiempo bayesiano, caracterizado por una manipulación sencilla, poder incluir regresoras, fechas relevantes y un tiempo de ejecución bajo. O bien la red LSTM, una adaptación de las redes recurrentes, donde su fuerte son las series de tiempo por su capacidad adaptativa.

En esta investigación se probaron ambas herramientas al fondo A y fondo E, usando dos tipos de datos distintos. En primer lugar, los registros del valor cuota, donde las respectivas predicciones de los modelos se transformaron a rentabilidad real, comparándose con las rentabilidades originales, o bien, que los modelos se ajusten directamente de la rentabilidad real.

Si solo se considera el tipo de dato, utilizar el valor cuota se obtienen resultados iguales o mejores que utilizar la rentabilidad real directamente, donde el caso más significativo es el Fondo E con Prophet, donde hubo una diferencia de un 8 % en el sentido de la tendencia.

Además, al ocupar el valor cuota como el dato a utilizar, es necesario deflactar por la UF las variaciones de las predicciones, para obtener la rentabilidad real respectiva. En todos los casos, al aplicar esta transformación, la precisión en el sentido de la tendencia, disminuye

entre un 5 y 7 unidades porcentuales, por lo que la variación de la UF afecta considerablemente a los resultados.

Si se considera solo el fondo a estudiar, en promedio, el Fondo E siempre obtuvo mejor predicción en el sentido de la tendencia, esto debido a la naturaleza del mismo, al ser un fondo con mínima renta variable, esta menos expuesto a los vaivenes del mercado bursátil.

Por último, si se compara entre herramientas, el modelo LSTM en la gran mayoría de los casos obtuvo mejores resultados que Prophet en el sentido de la tendencia, pero, hubo un caso único y particular, donde Prophet obtuvo mejor resultado, específicamente en el Fondo E, donde alcanzó un 62 % en el Sentido, siendo el único caso que superó los 60 %. Entonces definir que herramienta obtiene los mejores resultados, debe definirse por otras métricas.

Si se tuviera que elegir un modelo, específicamente para la predicción del sentido de la tendencia, la elección sería Prophet, debido a su fácil manipulación, el extra de incluir fechas y eventos de manera manual y la inclusión de variables regresoras lo convierte en una herramienta a considerar. Además, que no requiere tanto nivel computacional, ya sea al momento de programar como especificaciones técnicas del ambiente de trabajo.

Como trabajo futuro, se propone dos líneas de trabajo. En primer lugar, perfeccionar lo realizado, como por ejemplo utilizar otras estrategias de pronósticos como de validación, o bien, proponer otras variables que podrían afectar a la serie, buscando que mejoren los resultados.

En segundo lugar, no interesarse en el sentido de la tendencia, sino en la rentabilidad misma, poder adicionar un clasificador que manifieste a que fondo cambiarse, no considerando solo el fondo A y fondo E, sino los 5 fondos. Tomando en consideración los costos adicionales producto de la comisión cobrada por las AFP y de la cantidad de cambios de fondo permitidos.

Finalmente, se espera que esta investigación se convierta en una base para futuros trabajos, pues existen pocos trabajos relacionados al tema y se exponen resultados con metodología y herramientas distintas. A pesar de que los resultados obtenidos no ayudan a la elección del fondo, ya que se necesitan pronósticos más precisos y lo ideal de un horizonte mayor, puede ser utilizada como un complemento a otros modelos de pronósticos, con el fin de ayudar a la obtención de una mayor pensión en tu jubilación.

Bibliografía

- [1] ALARCÓN JAQUE, M. (2014). *Modelos de Redes Neuronales aplicado a la predicción del signo de los fondos de Afp Cuprum*. Universidad de Chile. Santiago
- [2] ASOCIACIÓN DE AFP. (2018). *Boletín multifondos*. N°55.
- [3] DONIEZ, V. AND KREMERMAN, M. (2019). *Análisis Crítico de la Propuesta de Pensiones del Gobierno de Sebastián Piñera*. Estudios de la Fundación SOL.
- [4] ELMAN, J.L. (1990). *Finding structure in time*. Cogn. Sci., 14, 179–211.
- [5] GARRIDO VARGAS, F. (2014). *Redes Neuronales Artificiales de base radial con algoritmo de aprendizaje PSO para el pronóstico del valor cuota del Fondo A de Afp Cuprum*. Pontificia Universidad Católica de Valparaíso. Chile.
- [6] HAN, J. AND MORAGA C. (1995). *The influence of the sigmoid function parameters on the speed of backpropagation learning*. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, 1995, pp. 195–201.
- [7] HE, K., ZHANG X., REN, S. AND SUN J. (2015). *Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification*. arXiv, 2015.
- [8] HOCHREITER, S.; SCHMIDHUBER, J. (1997). *Long short-term memory*. Neural Comput. 9, 1735–1780.
- [9] HORNIK, K., STINCHCOMBE, M., WHITE, H. (1989). *Multilayer feedforward networks are universal approximators*. Neural networks, 2(5), 359-366.
- [10] KARLIK, B. AND VEHBI, A. (2011). *Performance Analysis of Various Activation Functions in Generalized MLP Architectures of Neural Networks*. International Journal of Artificial Intelligence and Expert Systems (IJAE), vol. 1, no. 4, pp. 111–122, 2011.
- [11] KERAS DOCUMENTATION. Disponible en: <https://keras.io/> (Acceso el 12 Enero 2020).
- [12] LE, X. H., HO, H. V., LEE, G., JUNG, S. (2019). *Application of long short-term memory (LSTM) neural network for flood forecasting*. Water, 11(7), 1387.
- [13] MAAS, A., HANNUN, A. AND NG, A. (2013). *Rectifier Nonlinearities Improve Neural Network Acoustic Models*. in International Conference on Machine Learning (icml), 2013.

- [14] MINISTERIO DE DESARROLLO SOCIAL. (2020). *Valor de la Canasta Básica de Alimentos y Líneas de Pobreza. Informe Mensual-Febrero 2020.*
- [15] McCULLOCH, W. S. AND W. H. PITTS (1943). *A Logical Calculus of the Ideas Immanent in Nervous Activity*. Bulletin of Mathematical Biophysics 7, 115–133. Reprinted in McCulloch 1964, pp. 16–39.
- [16] NAIR V. AND HINTON G. E. (2010). *Rectified linear units improve restricted boltzmann machines*. Haifa, 2010, pp. 807–814.
- [17] OLAH, C. *Understanding LSTM Networks*. Disponible en: <http://colah.github.io/posts/2015-08-Understanding-LSTMs/> (Acceso el 15 Marzo 2020).
- [18] RAMOS PERONI, J. (2015). *Redes Neuronales Artificiales con algoritmo de aprendizaje cuckoo search para el pronóstico del valor cuota del Fondo A de Afp Cuprum*. Pontificia Universidad Católica de Valparaíso. Chile.
- [19] RUMELHART, D.E.; HINTON, G.E.; WILLIAMS, R.J. (1986). *Learning representations by back-propagating errors.* Nature, 323, 533–536.
- [20] STEIN, R. M. (2002). *Benchmarking default prediction models: Pitfalls and remedies in model validation*. Moody's KMV, New York, 20305.
- [21] SUPERINTENDENCIA DE PENSIONES. (2010). *El sistema chileno de pensiones*. Santiago de Chile.
- [22] SUPERINTENDENCIA DE PENSIONES. (2019). *Inversión de los fondos de pensiones*. Reporte Trimestral.
- [23] SUPERINTENDENCIA DE PENSIONES. (2019). *Nuevas personas pensionadas y pensiones pagadas en el mes por tramo de años cotizados*. Reporte Trimestral.
- [24] TAYLOR, S.J. AND LETHAM, B. (2018). *Forecasting at Scale*. The American Statistician, 72(1):37-45.
- [25] TURIAN, J., BERGSTRÄ J. AND BENGIO Y. (2009). *Quadratic features and deep architectures for chunking*. in Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics, vol. Companion Volume:, 2009, pp. 245–248.
- [26] XIAO, Y., YIN, Y. (2019). *Hybrid LSTM neural network for short-term traffic flow prediction*. Information, 10(3), 105.