



UNIVERSIDAD JORGE TADEO LOZANO
FACULTAD DE CIENCIAS NATURALES E INGENIERÍA
MAESTRÍA EN INGENIERÍA Y ANALÍTICA DE DATOS
MIAD

PREDICCIÓN PARA EL MERCADO DE
ACCIONES CON REDES NEURONALES
LSTM

Dennys Fabian Herrera Cofre
Maestría Ingeniería y Analítica de Datos

Universidad Jorge Tadeo Lozano
Facultad de Ciencias Naturales e Ingeniería
Bogotá, Colombia
22 de septiembre de 2020

UNIVERSIDAD JORGE TADEO LOZANO

FACULTAD DE CIENCIAS NATURALES E INGENIERÍA
MAESTRÍA EN INGENIERÍA Y ANALÍTICA DE DATOS
MIAD

PREDICCIÓN PARA EL MERCADO DE
ACCIONES CON REDES NEURONALES
LSTM

TRABAJO DE GRADO PRESENTADO COMO REQUISITO PARA OPTAR AL TÍTULO EN :
Magíster en Ingeniería y Analítica de Datos

Directores:

PhD. Olmer Garcia Bedoya ¹

PhD (c). Oscar Granados ²

Universidad Jorge Tadeo Lozano
Facultad de Ciencias Naturales e Ingeniería
Bogotá, Colombia 22 de septiembre de 2020

¹Docente de tiempo completo email olmer.garciab@utadeo.edu.co

²Docente de tiempo completo email oscarm.granadose@utadeo.edu.co

Dedicatoria...

Este trabajo está dedicado a mis padres, quienes siempre vieron a la educación como la mejor herencia que me pueden dejar, me enseñaron que el conocimiento es el arma mas poderosa para cambiar el mundo. También está dedicado a mi esposa, quien me enseñó que incluso la tarea más grande se puede lograr si se hace un paso a la vez.

Agradecimientos

Agradezco principalmente a Dios por darme la oportunidad, fuerza y capacidades para culminar esta hermosa etapa de mi vida, lo cual llena de alegría y orgullo a mi familia.

Como no agradecer la ayuda que muchas personas, colegas y compañeros me han prestado durante el proceso de desarrollo y redacción de este trabajo. Agradecimiento muy especial a mi esposa que me ha ayudado y apoyado en todo momento, a mis tutores por haberme orientado cuando necesité su ayuda.

Resumen

El objetivo de este trabajo es demostrar la capacidad de las redes neuronales para apoyar en la predicción del precio de las acciones. La capacidad predictiva estable y significativa de las redes neuronales utilizadas y su superioridad predictiva respecto por ejemplo a una regresión múltiple comparable, se puede ver que la memoria larga a corto plazo (LSTM) puede resolver varias tareas de series de tiempo que no pueden resolverse mediante métodos tradicionales, lo cual sugiere usar LSTM solo cuando los enfoques tradicionales más simples fallan. En este trabajo se pretende demostrar que el nivel de precios de las acciones es más que la suma de la información económica disponible en datos históricos, teniendo en cuenta que la mayoría de las decisiones se toman teniendo en cuenta investigaciones previas y las noticias del entorno que rodean a los inversionistas. En este sentido, la idea de este trabajo es contribuir mediante aprendizaje automático a la toma de estas decisiones para operar en bolsa, utilizando de una manera diferente la gran cantidad de información que se encuentra disponible en internet como lo son las noticias en un portal financiero e información histórica de precios.

Índice general

Agradecimientos	3
Resumen	4
1. Introducción	7
2. Preliminares	12
2.1. Series de Tiempo	12
2.2. Aprendizaje Automático	12
2.3. Aprendizaje automático y aprendizaje profundo	13
2.4. Redes Neuronales LSTM	14
3. Materiales y Métodos	16
3.1. Conjunto de datos	16
3.2. Análisis de precios históricos	17
3.3. Análisis de noticias económicas	18
3.4. Clasificación de noticias	19
3.5. Modelos Predictivos	20
3.5.1. Modelo 1	20
3.5.2. Modelo 2	21
3.5.3. Modelo 3	22
4. Resultados	24
5. Conclusiones	28
Bibliografía	29

Índice de figuras

2.1. Estructura de una Red LSTM	14
3.1. Curva Precios OHLC (Open, High, Low, Close) para APPLE	17
3.2. Curva Año 2020 para acción de Amazon	18
3.3. Modelo webscraper para Extraer Noticias	19
3.4. Arquitectura Librería Spacy	20
3.5. Modelo 1	21
3.6. Modelo 2	21
3.7. Modelo 3	22
4.1. Representación de valores obtenidos para entrenamiento	24
4.2. Curva de Predicción del Modelo 1	25
4.3. Curva de Predicción del Modelo 2	26
4.4. Curva de Predicción del Modelo 3	27

Capítulo 1

Introducción

El mercado de valores es conocido por su alta complejidad y los inversionistas siempre están buscando una forma efectiva de optimizar el intercambio de acciones. La predicción del mercado financiero ha sido tradicionalmente un propósito de la industria y la academia, pues puede ser ineficiente confiar únicamente en la experiencia personal, la intuición y el análisis técnico realizado por un inversionista para el análisis de una acción o un portafolio de estas. De esta forma, los inversionistas necesitan un método de investigación científico para dirigir el intercambio de acciones y la aplicación de la inteligencia artificial y el aprendizaje profundo se convierten en herramientas para lograr ese propósito.

Por lo general, los inversionistas tienen un mayor nivel de conocimiento de los mercados financieros frente a aquellos que no cuentan con el, lo que regularmente se hace es contratar los servicios de un asesor de cartera que gestione cada una de las inversiones a partir de los parámetros de riesgo y el propósito que tiene el capital de la inversión. Mientras a algunos les gusta reducir el riesgo, otros buscan una mayor rentabilidad y optan por instrumentos financiero que tienen mayor volatilidad. De esta forma, las características que definen a cada inversionista y el nivel de riesgo que se permita tolerar llevan a que tome una decisión final sobre la acción en la cual quiere invertir. Sin embargo, es importante decir que la mayoría de las decisiones no se toman sin ninguna información, siempre se tiene en cuenta las noticias del entorno que rodean a los inversionistas y los reportes de inversión de los diferentes agentes financieros.

El primer tropiezo que tiene alguien interesado en invertir en bolsa es el conocimiento, pues un mercado como el de renta variable tiene varias implicaciones y no es tan fácil arriesgar un capital sin conocer con precisión lo que sucede en el. En otras palabras, el universo de personas interesadas y sin conocimiento de como opera el mercado de valores es muy amplio. Por esta razón, el principal interrogante que tiene un inversionista es cómo saber si el activo que adquiere tiende a subir o a bajar, es por ello que obtener una predicción acertada para operar en la bolsa se ha vuelto un reto debido a que es muy difícil o casi imposible predecir con exactitud el precio de una acción, y si bien existen muchos robots hoy en día que ayudan a este tipo de tareas a los grandes inversionistas, no es muy común y sobre todo en América Latina que este tipo de algoritmos estén al alcance de un inversionista común que puede tener cierta cantidad de dinero para invertir pero que no tiene el conocimiento ni la ayuda necesaria. Es por ello que uno de los objetivos de esta investigación es ayudar al inversionista

a que pueda tomar una decisión a la hora de invertir en productos de renta variable que con ayuda de la tecnología logre tomar una decisión correcta.

En lugar de tener una persona, asesor o gestor que tome las decisiones sobre las inversiones o de ser uno mismo quien las gestione, se podría tener un sistema que a través de un algoritmo automatizado apoye a la toma de decisiones. Es la forma más sencilla de invertir, ya que gran parte de la decisión dependerá de sistemas libres de sentimientos. Obviamente siempre la decisión final la tomará el inversionista. Sin embargo, todo el que invierta a través de estas herramientas deberá tener una formación mínima sobre este tipo de inversiones y tener claro los riesgos relacionados.

Dado que la predicción del mercado de valores ha sido un tema que ha llamado la atención de investigadores a través del uso de diferentes metodologías, el desarrollo de métodos de aprendizaje profundo han ganado fuerza pero han surgido en gran parte por los avances de la economía conductual o del comportamiento [1]. Si bien su enfoque también es de aplicaciones de métodos de aprendizaje profundo, la predicción del mercado de valores solo podría ser un ejemplo de muchos problemas financieros. El enfoque surgió de la psicología y su relación con los mercados financieros, especialmente, la teoría de la retroalimentación, la cual establece que cuando los movimientos accionarios son exitosos para algunos inversionistas logran atraer la atención del público, promover el entusiasmo voz a voz y aumentar las expectativas de ganancia. Entonces, la economía y las finanzas conductuales se abren paso como una opción que se distancia de la lógica de los mercados eficientes [2–4].

En este contexto, Shiller hace un recuento histórico del mercado de valores y analiza los factores y mecanismos que han influido en la volatilidad de las acciones. La parte central del libro hace referencia a una frase usada por el presidente de la Reserva Federal en ese entonces Alan Greenspan, en donde uso el término de irracionalidad exuberante para describir el comportamiento de los inversionistas en los mercados de valores, lo cual generó un desplome de la mayoría de estos, pues las palabras fueron interpretadas como evidencia de la posible sobrevaloración de los mercados financieros [5].

Lamont y Thaler [6] describen la importancia de la ley del precio (el mismo activo no puede comercializarse a diferentes precios de manera simultánea) y el arbitraje (la práctica de tomar ventaja de una diferencia de precio entre dos o más mercados). En este sentido se destacan tres tipos de riesgo: primero, el riesgo del modelo malo o problema de la hipótesis conjunta, el cual consiste en comparar dos acciones por su precio, descuidando las diferencias en liquidez, riesgo o impuestos, diferencias que se reflejan en los precios existentes. Segundo, el riesgo fundamental, el cual es la exposición que tienen los inversionistas ante una situación que afecta parcial o totalmente a un mercado financiero. Estos riesgos son causados por situaciones súbitas que surgen de fenómenos naturales como terremotos, inundaciones, huracanes o fenómenos sociales como la inflación, el desempleo o una guerra. En tercer lugar, está el riesgo del ruido de mercado, el cual no es más que el movimiento que puede surgir en un activo o en un índice en un momento determinado, pero que no afecta a su tendencia real. Es decir, son los cambios súbitos del precio que afecta los costos de transacción y puede limitar el arbitraje [6]. El arbitraje se refiere a tomar ventaja de una diferencia de precio entre dos o más mercados. Existen ejemplos en el mercado de valores donde se puede ver este escenario, se han dado muchos casos en el que las empresas que cotizan en bolsa dividen su

patrimonio (escisión) para transferir cierto porcentaje de acciones y crear nuevas sociedades, dependiendo de la normativa de cada país. El punto está en como se fijan los precios para las acciones, por ejemplo, el aumento en el primer día en las acciones puede atraer la atención de los inversionistas con propósitos de arbitraje, ya que muchos buscan explotar la diferencia en la capitalización de mercado entre la nueva empresa y su empresa matriz. Algo similar sucede con la división de acciones (split) [7, 8].

Shiller [9] muestra que el nivel de volatilidad del mercado de valores no puede ser explicado con cualquier variante del modelo de mercados eficientes en el que los precios de las acciones están formados a partir de la evaluación del valor actual descontado los rendimientos futuros [10]. Se han hecho varios esfuerzos para capturar la volatilidad [11] de los precios de las acciones y para esto se argumenta que los inversionistas también tuvieron una gran cantidad de información sobre los cambios en los factores que influyen en las tasas de descuento. Después de todos los esfuerzos para defender la teoría de los mercados eficientes, todavía hay varias razones para pensar que, si bien los mercados no surgen como procesos ineficientes, si contienen un ruido que impactan en los movimientos del mercado en su conjunto [12]. El modelo de mercados eficientes aún requiere la vinculación de las fluctuaciones del mercado de acciones con fundamentos posteriores, por ejemplo como identificar los valores de las acciones que actualmente no se reflejan en los precios y de esta manera predecir rendimientos anormales como la probabilidad estimada de un aumento de las ganancias.

Por otro lado, la redes neuronales LSTM tiene una variedad de aplicaciones, que vienen desde la medicina con trabajos para la detección temprana del shock séptico [13] la diabetes [14] o la malaria [15], trabajos para pronosticar la carga de trabajo de los centros de datos en la nube [16], trabajos que permiten realizar análisis de sentimientos a partir de tomar las reseñas de películas para clasificarlas mediante una red LSTM [17] o características de personalidad en redes sociales [18]. Frente a los mercados financieros, Wanjawa & Muchemi [19] propone una red neuronal artificial utilizando un perceptrón multicapa de retroalimentación con retropropagación de error para predecir los precios de las acciones y sus resultados se ajustan a la posibilidad de predecir un mercado de valores típico. Zhang [20] desarrolla una red neuronal convolucional combinada (CNN) y una red neuronal recurrente (RNN) para proponer una nueva arquitectura conocida como la red neuronal de área profunda y amplia (DWNN). Los resultados de esta combinación muestran que este modelo puede reducir el error cuadrado medio previsto en un 30 % en comparación con el modelo RNN general.

Para este caso, se propuso un modelo híbrido de heterocedasticidad condicional autorregresiva generalizada (GARCH) combinado con redes LSTM para predecir las fluctuaciones del precio de las acciones, método que se ha usado para algunos otros activos financieros como índices de acciones [21, 22], futuros de bienes básicos como el petróleo [23], el oro [24] o el cobre [25]. CNN se utiliza para desarrollar una estrategia de selección de acciones cuantitativa que ayuden a determinar las tendencias de las acciones y luego predecir sus precios utilizando el modelo LSTM, lo cual permite definir un modelo de red neuronal híbrido para estrategias de sincronización cuantitativa que ayuden a aumentar las ganancias. Se agregó una función ponderada en el tiempo a una red neuronal LSTM y los resultados de algunos estudios han superado a los de otros modelos [26]. Jiang et al [27] utilizaron una red neuronal LSTM y RNN para construir modelos y descubrieron que las redes LSTM pueden aplicarse

mejor a la previsión de existencias. Jin et al [28] agregaron una tendencia al sentimiento del inversionista en el análisis del modelo e introdujo la descomposición modal empírica (EMD) combinada con LSTM para obtener pronósticos de precios. Finalmente, el trabajo de Qiu et al [29] es el más cercano a lo que se propone en esta investigación, debido a que además de utilizar redes neuronales LSTM se utiliza un concepto llamado mecanismos de atención, lo cual permite resaltar la información local importante al asignar la atención adecuada a información clave. Para el caso de predicción de precios en acciones la idea es centrar la atención y analizar los movimientos del mercado que se pueden dar por diferentes factores externos, entender esto e incluirlos como variables al algoritmo ayudará a tener predicciones más precisas.

El modelo LSTM basado en el mecanismo de atención es común en el reconocimiento de voz e imagen, pues tienen las características de selectividad, células de memoria, las redes neuronales LSTM son adecuadas para secuencias aleatorias no estacionarias, como las series temporales de precios de acciones [30, 31]. Aunque todavía existen algunas dificultades en las predicciones financieras que utilizan el aprendizaje profundo [32], lo que se busca es establecer un modelo de pronóstico confiable del mercado de renta variable. Se trata de un algoritmo para previsión de precios de acciones con una red LSTM basada en mecanismo de atención, el cual destaca la información local importante al asignar la atención adecuada a la información clave que fue obtenida a partir de los datos de las cotizaciones diarias en la plataforma de Yahoo Finance.

Entonces ¿cómo aportan estos modelos a lo que sucede en los mercados financieros de predicción de precios de apertura de acciones tecnológicas? Varias pueden ser las respuestas a esta pregunta, pero principalmente, porque para mitigar el impacto de la volatilidad se requieren herramientas que ayuden a identificar si la acción tiene una tendencia alcista o bajista a corto plazo. Luego está la información disponible en la red, si bien los algoritmos han evolucionado durante los últimos años, no podrían ser efectivos de no ser por la precisión en sus orígenes de datos (variables de entrada) y esto para el caso del mercado propuesto es información pública y disponible sobre el ciclo de vida de cada acción. Para esta investigación se escogieron diez acciones con alta bursatilidad y liquidez del sector tecnológico a nivel mundial. Pero, ¿está preparada la infraestructura tecnológica para este tipo de proyectos que requieren gran capacidad de cómputo? Para subsanar esto, la capacidad de la infraestructura on-premise y en la nube en los últimos años ha mejorado considerablemente, es así que existen procesadores creados específicamente para trabajar con redes neuronales, por lo que se puede encontrar fácilmente capacidad de cómputo con diferentes proveedores.

Las investigaciones mencionadas y específicamente las que tienen que ver con redes LSTM demuestran un alto grado de precisión en sus resultados. Así mismo, los estudios relacionados con economía del comportamiento ayudan a entender la parte sentimental y conductual que puede influir en los mercados bursátiles. Sin embargo, son investigaciones independientes, es por ello que lo que se propone es una predicción clásica de precios de acciones con datos históricos, pero con un componente adicional que son las noticias de un medio digital de tal manera que se pueda ver cómo influye una noticia en el precio de una acción. Así mismo, se tienen en cuenta indicadores técnicos como variables adicionales para identificar la dirección del mercado. El objetivo de este trabajo, entonces, es demostrar como los factores fuera del mercado de valores, como la tecnología y los sentimientos dan forma al comportamiento del

mercado. La idea es proponer modelos de aprendizaje automático que permitan predecir el precio de las acciones tecnológicas a nivel mundial con los componentes antes mencionados. Los modelos propuestos permiten comparar salidas de predicción teniendo en cuenta diferentes variables y capas de una red neuronal.

Para alcanzar este propósito, la tesis esta dividida de la siguiente forma. El capítulo 2 presenta algunos elementos teóricos necesario para desarrollar el modelo. El capítulo 3 presenta un análisis descriptivo y predictivo de los datos, así como los métodos usados. En el capítulo 4 se presentan algunos resultados y se termina con el capítulo 5 de conclusiones.

Capítulo 2

Preliminares

A continuación se presentan algunos conceptos que serán usados directa e indirectamente en el desarrollo de los siguientes capítulos.

2.1. Series de Tiempo

Las series de tiempo, son datos estadísticos que se recopilan, observan o registran en intervalos de tiempo regulares (diario, semanal, semestral, anual, entre otros) [33]. En finanzas, las series de tiempo son inherentemente ruidosas y no estacionarias, pues la característica no estacionaria implica que la distribución de series de tiempo financieras cambia con el tiempo. En el modelado de series de tiempo financieras, esto conducirá a cambios graduales en la dependencia entre las variables de entrada y salida [34]. Por lo tanto, el algoritmo de aprendizaje utilizado debe tener en cuenta esta característica, pues la información proporcionada por los puntos de datos recientes tiene más peso que la proporcionada por los puntos de datos distantes, ya que en las series de tiempo financieras no estacionarias los puntos de datos recientes podrían proporcionar información más importante que los puntos de datos distantes. Sin embargo, cuando se usan retrasos o se consideran horizontes de tiempo muy largos pueden presentarse pequeños errores debido a la naturaleza de los datos [35].

2.2. Aprendizaje Automático

El aprendizaje automático es un tipo de inteligencia artificial (AI) que proporciona a las computadoras la capacidad de aprender sin ser programadas explícitamente. El aprendizaje automático se centra en el desarrollo de programas informáticos que pueden cambiar cuando se exponen a nuevos datos [36]. Es decir, el aprendizaje automático como área de las ciencias computacionales aplicadas desarrolla algoritmos capaces de tomar datos numéricos y alfanuméricos almacenados en un computador [37]. Por lo tanto, para poder predecir el comportamiento de un portafolio de inversión asociado a una cartera eficiente, es necesario utilizar modelos de regresión que permitan mostrar el mejor camino de inversión. Los modelos de regresión son modelos matemáticos que buscan determinar la relación entre una variable dependiente (Y) con respecto a otras variables llamadas explicativas o independientes (X). “Las técnicas de aprendizaje automático tienen como objetivo aprender y reconocer

automáticamente patrones en grandes cantidades de datos” [36]. A continuación, se describen algunas de estas técnicas.

- Support Vector Machine. Es un algoritmo de aprendizaje automático supervisado que se puede utilizar para problemas de clasificación o regresión. Utiliza una técnica llamada truco del núcleo para transformar sus datos y luego, basándose en estas transformaciones, encuentra un límite óptimo entre las posibles salidas [33].
- Support Vector Regression. Es una técnica que realiza una regresión lineal en el espacio de características de alta dimensión utilizando una pérdida insensible y al mismo tiempo, intenta reducir la complejidad del modelo minimizando [38].
- Redes Neuronales. Es un algoritmo basado en una red de alimentación de múltiples capas entrenadas inspirado en las neuronas del cerebro humano, estos sistemas aprenden y se forman a sí mismos ya que sus neuronas artificiales están conectadas, en lugar de ser programados de forma explícita [39].
- Modelos No Lineales. Es un método a partir del cual se puede encontrar un modelo lineal. Se utilizan a menudo para una variedad de propósitos, uno de sus principales usos es para pronóstico y simulación [40].
- Modelos de Dependencia Dinámica. Son extensiones de modelos dinámicos multirregresivos Los MDM incorporan la estructura del modelo gráfico dirigido en una serie de tiempo multivariante, lo que permite que los valores de algunas series univariadas aparezcan como predictores de otras series [41].

2.3. Aprendizaje automático y aprendizaje profundo

El propósito del aprendizaje automático es poder extraer unas características o parámetros de interés de los datos analizados para generar unos datos de salida que permitan interpretar la información que se está ingresando en un sistema y el aprendizaje profundo se desprende de esta misma estructura. Para mostrar la diferencia entre los dos se puede hacer a través del siguiente ejemplo. Un sistema de asistencia (ej.: SIRI, ALEXA, Google Home) que entiende la voz humana, convierte esas grabaciones en sus datos de entrada. Un sistema de aprendizaje automático lo que hace es, primero, segmentar esas señales de audio, por ejemplo, el inicio y final de cada una de las palabras y después se incluye un proceso de extracción de características, entonces la idea es que de cada una de esas palabras detectadas dentro de esas señales de audio logren extraer unos parámetros que representen de la mejor forma posible esa señal y permitan diferenciar una palabra de otra. Posteriormente ese sistema lleva a cabo un proceso de clasificación, el cual determina a que corresponde cada una de las palabras pronunciadas por la persona y luego como esas palabras están relacionadas entre si para conformar una frase relacionada. Finalmente, el sistema genera una acción, por ejemplo “Búsqueme el restaurante cercano”.

En un sistema de aprendizaje profundo no se tiene un bloque de segmentación o extracción de características. Es decir, estos sistemas están en capacidad de que sin necesidad de

que un ser humano le especifique durante la programación del algoritmo como segmentar el audio o como determinar cuáles son las características más relevantes. En este caso, se deben generar unos datos de entrada, definir como están interconectados los elementos internos dentro de este sistema de aprendizaje profundo y a través de un proceso que se conoce como entrenamiento y aprendizaje genera la segmentación y los datos de salida.

¿Por qué profundo? La base del aprendizaje profundo es una neurona artificial, la cual contiene coeficientes y realiza operaciones, mientras en el dato resultante realiza algún tipo de comparación. Para ejemplos más complejos ya no es suficiente con una sola neurona, por lo cual nace el concepto de red neuronal. En sistemas complejos las redes neuronales son complejas dependiendo del escenario, para ello existe el concepto de redes neuronales profundas. ¿Por qué aprendizaje? Para sistemas complejos no se tiene el coeficiente, por lo general el coeficiente se calcula de forma autónoma sin necesidad de que el humano ayude y en redes neuronales profundas pueden ser millones de coeficientes. Entonces el hecho de calcular estos coeficientes se convierte en un proceso de aprendizaje.

Para que una red profunda no tenga ninguna influencia humana en su proceso de aprendizaje, se deben mostrar miles de ejemplos para que poco a poco la red a través de ese proceso de entrenamiento identifique los patrones correspondientes y aprenda también a interconectar los datos. En este sentido el crecimiento acelerado en la cantidad de datos disponibles permite el fortalecimiento de esta técnica de inteligencia artificial. La idea es que al final del proceso de aprendizaje se tengan los coeficientes para la representación y asociado a estos conceptos de representación y aprendizaje, está el entrenamiento, que surge de las iteraciones que progresivamente permiten encontrar la representación más adecuada de los datos, es decir, aprender paulatinamente los valores de los coeficientes.

2.4. Redes Neuronales LSTM

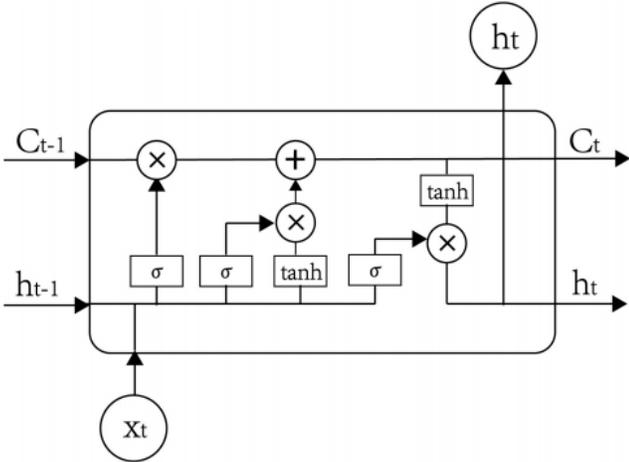


Figura 2.1: Estructura de una Red LSTM [29]

Las redes LSTM fueron desarrolladas a partir de redes neuronales recurrentes (RNN) y adicional a su diversidad de aplicaciones, evitan los problemas de dependencia a largo plazo debido a su estructura de unidad de almacenamiento única y ayudan a predecir series de tiempo financieras. La red neuronal en el aprendizaje profundo se ha convertido en un predictor popular debido a su capacidad de aproximación no lineal y a su autoaprendizaje adaptativo [42].

Comparado con una celda de red recurrente básica, la celda LSTM (Ver Figura 2.1) tiene una entrada (C_{t-1}) y una salida (C_t) adicionales, este elemento se conoce como celda de estado la cual es la clave del funcionamiento de las redes LSTM, es como una banda transportadora a la que se pueden añadir o de donde se pueden remover datos que se desean sacar de la memoria de la red. Para añadir datos a esta memoria usamos tres compuertas:

- **Forget Gate.** Permite eliminar elementos de la memoria
- **Update Gate.** Permite añadir nuevos elementos a la memoria
- **Output Gate.** Permite crear el estado oculto actualizado

Estas compuertas funcionan como válvulas: totalmente abiertas permiten el paso de información y totalmente cerradas lo bloquean por completo. Cada una de estas compuertas esta conformada por tres elementos: una red neuronal, una función sigmoïdal y un elemento multiplicador.

La variable (X_t) corresponde a la información de entrada, la variable (h_{t-1}) corresponde a la salida anterior y finalmente la variable (h_t) corresponde a la nueva salida.

Capítulo 3

Materiales y Métodos

En esta investigación se tomó como punto de partida la información histórica de precios de un grupo de acciones del sector tecnológico y se complementó con algunos indicadores técnicos de esta información además de noticias económicas de mercado de valores. A continuación, se presentan algunos análisis de los datos para posteriormente desarrollar los modelos de predicción.

3.1. Conjunto de datos

Para esta investigación se analizan las acciones ordinarias de diez empresas de tecnología, siendo algunas las más importantes a nivel mundial. Luego de un breve estudio, se tomaron en cuenta algunas características que las hacen interesantes para los inversionistas, así como por el volumen de negociación y su liquidez. Las acciones seleccionadas son:

- **Amazon (AMZN)**. En los últimos años se ha convertido en una de las empresas con mayor capitalización bursátil por su expansión en diferentes negocios más allá del comercio electrónico. Su volumen de negociación promedio diario está en 4.858.353 acciones.
- **Apple (APPL)**. La empresa de Cupertino, California goza de una alta capacidad de generación de caja. Su volumen de negociación promedio diario está en 25.949.554 acciones.
- **Alphabet (GOOG)**. Es la empresa matriz de Google y su volumen de negociación promedio diario está en 1.821.271 acciones.
- **Microsoft (MSFT)**. Conocida históricamente como una empresa pionera de tecnología y con un alto número de innovaciones, tiene un volumen de negociación promedio diario de 35.811.771 acciones.
- **Twitter (TWTR)**. Red social que ha incrementado sus interacciones en los últimos años. Su volumen de negociación promedio diario está en 19.844.987 acciones.
- **Facebook (FB)**. Es una empresa clave en los mercados de comunicación con marcas como WhatsApp e Instagram. Su volumen de negociación promedio diario está en 25.949.554 acciones.

- **Synopsys (SNPS)**. Empresa dedicada al desarrollo de software especializado para el diseño de circuitos integrados complejos y de tecnologías de inteligencia artificial. Su volumen promedio diario está en 861.667 acciones.
- **Tesla (TSLA)**. Es la compañía líder en fabricación de autos eléctricos. Su volumen de negociación promedio diario está en 76.797.309 acciones.
- **PayPal (PYPL)**. Pese a existir muchas formas de pago en línea esta empresa sigue siendo líder en su mercado de pagos en línea. Su volumen promedio diario está en 7.635.779 acciones.
- **Nvidia (NVDA)**. Empresa dedicada al procesamiento gráfico para videojuegos y diseño. Su volumen de negociación promedio diario está en 1.821.271 acciones.

3.2. Análisis de precios históricos

El conjunto de datos principal corresponde a los precios que han tenido el grupo de acciones desde el año 2000. Para ello se analizan las características que se deben aplicar como por ejemplo la frecuencia mínima (información intradía) la cual se puede encontrar en la página de Yahoo Finance. Obtener información diaria es común y se puede encontrar en diferentes fuentes de internet de manera libre. Sin embargo, para encontrar la información intradía en diferentes temporalidades, es decir la información de los movimientos de precios de acciones en un día, se usaron diferentes orígenes de datos probando con diferentes librerías en Python para obtener esa frecuencia temporal de los datos. Además, se pudo obtener información de los dividendos, lo cual es una buena entrada para el algoritmo considerando que hay un pago de dividendos cada cierto tiempo para algunas acciones y eso puede tener incidencia en el precio de las mismas. Para la información intradía de cada 5 minutos o cada hora sólo se obtuvieron 60 y 720 días atrás respectivamente.

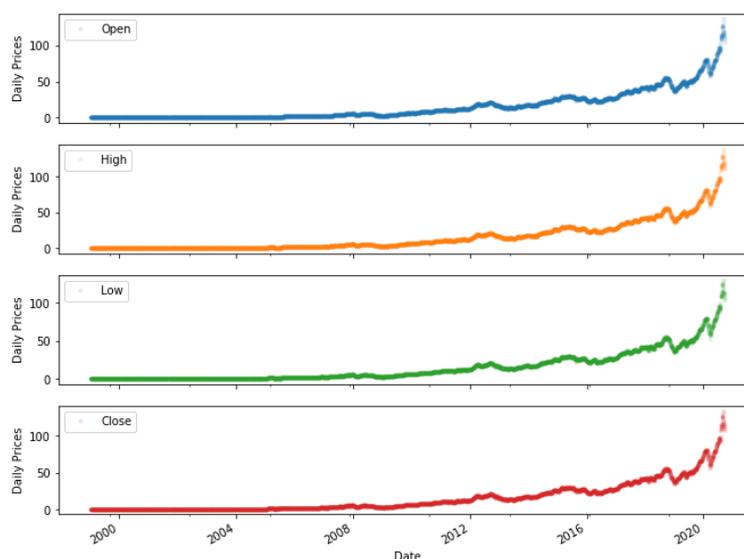


Figura 3.1: Curva Precios OHLC (Open, High, Low, Close) para APPLE (Elaboración Propia)

Para el caso de Apple (APPL) así como de las demás acciones, los datos parecen estar correlacionados en el tiempo, cada uno con cinco valores diarios que corresponden a la apertura, el máximo, el mínimo y el cierre de manera uniforme en cada jornada. Esta es la forma habitual en la que se usan los datos por parte de los inversionistas para entender visualmente la información y el comportamiento de un activo financiero durante una jornada, creando una visualización del comportamiento en un periodo específico [43]. Sin embargo, hay ruido en todas las mediciones y es difícil ver alguna relación que sea diferente. Para ello es importante ver por separado cada una de los precios diarios (Figura 3.1). Finalmente, el conjunto de datos analizado también nos arroja información de dividendos, es decir, las fechas en que estos se pagaron. Este dato ayuda a identificar cual fue el comportamiento de las acciones durante las fechas de pago.

3.3. Análisis de noticias económicas

Si se analiza la variación de precios para la acción de Amazon (AMZN) durante el año 2020, se puede ver una caída durante finales de febrero y principios del mes de marzo, esto se dio justamente por la pandemia que iniciaba en esas fechas. Esto es un ejemplo de muchos posibles de cómo influyen las noticias económicas y, en general, las diferentes noticias en los mercados financieros, conociéndose como análisis fundamental de los mercados

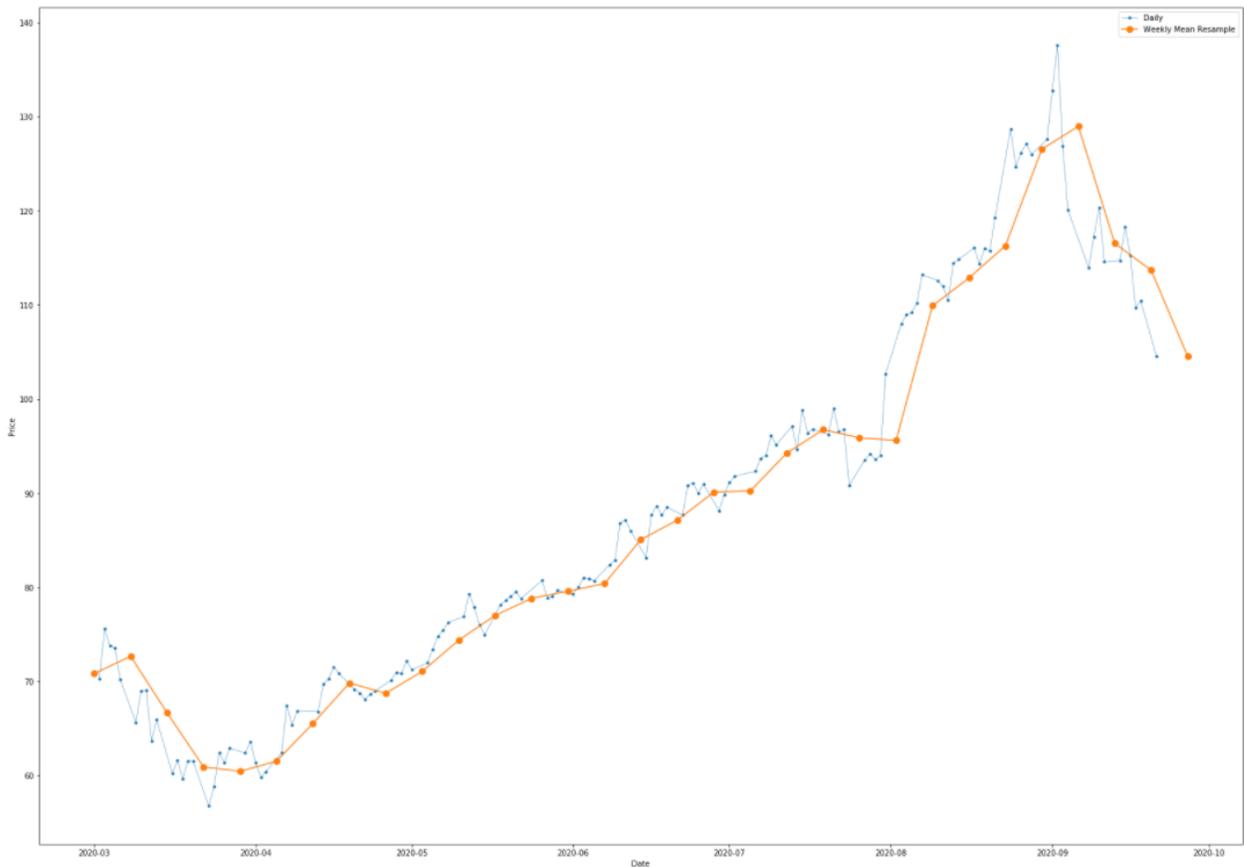


Figura 3.2: Curva Año 2020 para acción de Amazon
(Elaboración Propia)

Para el caso de las noticias se usó una página con noticias económicas y financieras de acceso abierto (investing.com) y se aplicó web scraping para obtener el texto. Una vez almacenado se usó procesamiento de lenguaje natural para el análisis de sentimientos e identificar si una noticia es buena o mala. Las noticias se obtienen en línea con webscraper del sitio <https://www.investing.com/news/stock-market-news>.

3.4. Clasificación de noticias

A través de los algoritmos de búsqueda de web scraping se puede rastrear centenares de webs para extraer sólo aquella información que se necesita. Para este proyecto se utilizó webscraper io una extensión de Google Chrome que hace más fácil la extracción de datos desde la fuente permitiendo navegar por todos los niveles y categorías.

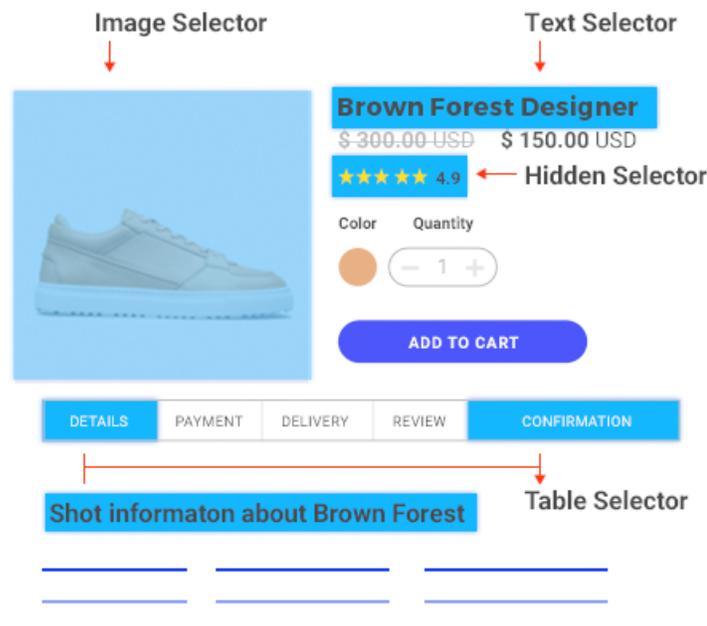


Figura 3.3: Modelo webscraper para Extraer Noticias

El problema de clasificación de noticias puede ser un modelo de clasificación binaria, teniendo en cuenta esto, la solución propuesta permite detectar si las noticias son positivas o negativas al determinar si la información obtenida de internet es correcta sobre la base de cuantificar el contenido de una noticia en el título. Esta sección describe la estructura del modelo y el método de procesamiento de datos utilizado para el experimento. El modelo propuesto entrena el conjunto de datos para identificar si la noticia es positiva o negativa. En este caso se utilizó el clasificador de texto de spaCy de Python que utiliza una red neuronal convolucional para su modelo.

Esta librería permite agregar modelos de categorización de texto a spaCy. El modelo admite la clasificación con múltiples etiquetas que no se excluyen mutuamente, por lo que

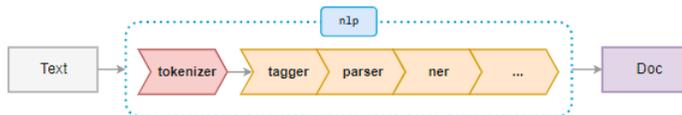


Figura 3.4: Arquitectura Librería Spacy [44]

se pueden aplicar varias a la vez. La precisión del rendimiento del modelo se hace utilizando varias métricas de evaluación disponibles, lo cual depende del conjunto de entrenamiento asignado. Para verificar la precisión de la predicción si el artículo de noticias se clasifica como positivo o negativo, se utiliza la precisión como métrica de evaluación lo cual se puede ver como un porcentaje.

3.5. Modelos Predictivos

A partir de los análisis de datos, se seleccionan y aplican las técnicas de modelado pertinentes al problema y se calibran sus parámetros a valores óptimos. En este caso se obtienen los datos de entrada que serían la información histórica de precios y a través de un proceso de entrenamiento y aprendizaje se genera la segmentación y datos de salida. Para lograr esto, se plantearon tres modelos que se describen a continuación.

3.5.1. Modelo 1

Se toman bloques de 60 datos consecutivos y el siguiente dato, es decir el número 61, será usado como salida. Antes de crear el modelo se ajusta el tamaño del conjunto de entrenamiento para indicar que cada ejemplo de entrada a la red será un vector de 60×1 . En el caso de la salida simplemente se especifica que su tamaño será igual a 1 correspondiente a la predicción.

El modelo que se propone es una red con 50 neuronas, así mismo para el entrenamiento se usa el método RMSPROP”, el cual funciona de manera similar al método de gradiente descendiente. La función loss (error cuadrático medio) se usó con lotes de 32 ejemplos y un total de 20 iteraciones (epoch). Para que la red LSTM pueda ser entrenada es necesario garantizar que los valores de la acción se encuentran en un rango definido. Para ello, se utiliza la librería Scikit Learn para normalizar estos valores en el rango de 0 a 1 y una vez se haya entrenado la red, se hará la transformación inversa sobre los datos mencionados anteriormente para obtener valores en escala real.

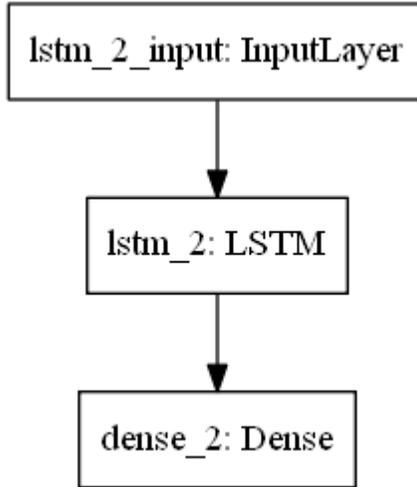


Figura 3.5: Modelo 1

3.5.2. Modelo 2

En este modelo se propone una RNN que tiene una capa densa como capa de salida con una neurona. Este modelo requiere una secuencia de características (sequence length) y se usan 50 o 100 días consecutivos para generar un valor único que indica el precio del siguiente día. El número de capas de la RNN es 3, la tasa de abandono dropout después de cada capa RNN es 0.4 y el número de neuronas es de 256. Además se usó el método *adam* y la función de loss (huber loss) en el que se usaron lotes de 64 ejemplos y un total de 400 iteraciones.

Layer (type)	Output Shape	Param #
lstm_54 (LSTM)	(None, None, 256)	365568
dropout_54 (Dropout)	(None, None, 256)	0
lstm_55 (LSTM)	(None, None, 256)	525312
dropout_55 (Dropout)	(None, None, 256)	0
lstm_56 (LSTM)	(None, 256)	525312
dropout_56 (Dropout)	(None, 256)	0
dense_18 (Dense)	(None, 1)	257

Figura 3.6: Modelo 2

3.5.3. Modelo 3

Para este modelo se definen dos capas de entrada. La capa de la derecha tiene forma (history points, 5), ya que cada punto de datos de entrada es una matriz. El modelo tiene 50 celdas LSTM en la primera capa, una capa de caída para evitar el sobreajuste y luego algunas capas densas para unir todos los datos de LSTM.

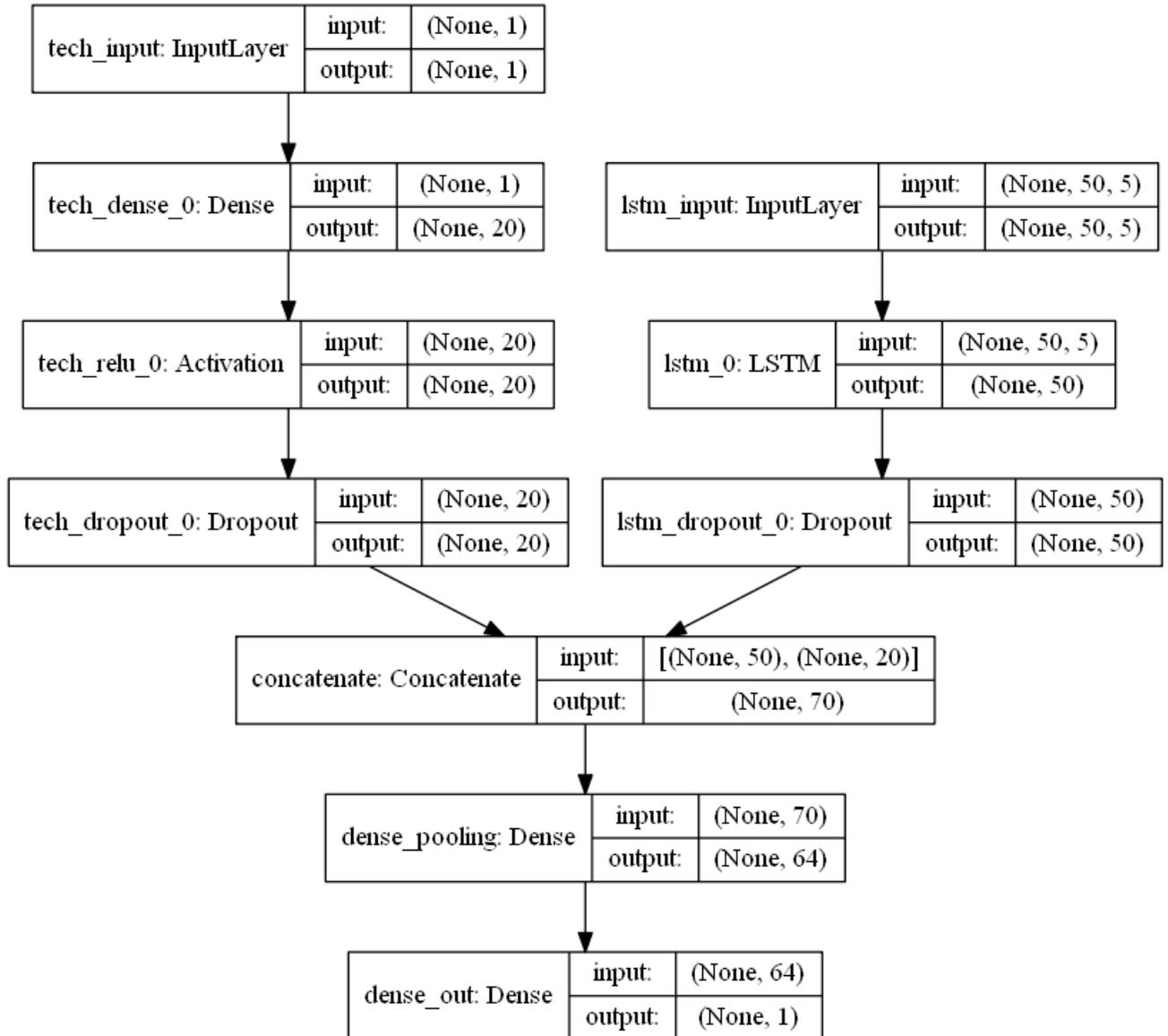


Figura 3.7: Modelo 3

Una métrica utilizada por los analistas del mercado de valores es el análisis técnico, el cual surge de diferentes operaciones matemáticas y estadísticas realizadas en el historial del precio de las acciones y con la visualización de los resultados se identifica la posible dirección y oscilación que puede tener el activo financiero [45]. Al modelo se pueden incorporar algunos de estos indicadores técnicos a través de una rama de entrada secundaria. Es así como se

usó el promedio móvil simple (SMA) como una entrada adicional en la red. Para calcular este promedio, se toma el promedio del precio de cierre de la acción durante los últimos n periodos de tiempo. Esto funciona muy bien, ya que se está tratando con bloques de pasos de tiempo fijos del historial de precios (Figura 3.7).

Posteriormente, se recorre cada bloque de datos de 50 precios y se calcula la media de la tercera columna, el precio de cierre, y se agrega ese valor a una lista de indicadores técnicos. Seguidamente, la lista pasa por las mismas transformaciones que el resto de los datos y se escala para ajustarse a los valores de 0 a 1. Luego se cambia la declaración de retorno para devolver los indicadores técnicos. Dado que el SMA no es un dato de serie temporal, no se debe pasar a través de la red LSTM. En su lugar, lo que se hace es mezclarlo antes de que se haga la predicción final, es decir, se ingresa en la penúltima capa densa de 64 nodos. Finalmente, el modelo propuesto es un modelo con dos entradas, una capa de concatenación y una salida. Finalmente, se intentó incluir el conjunto de datos de noticias positivas y negativas previamente procesado en este modelo, sin embargo la curva de predicción no fue óptima por lo que se decidió finalmente no incluirla debido a que la predicción con este modelo es bastante aceptable y no es necesario otra variable. Además hace falta un conjunto de datos de información histórica de noticias que se pueda integrar durante el periodo de análisis.

Capítulo 4

Resultados

Es importante comparar el modelo obtenido con los objetivos. En este sentido se hicieron algunas ejecuciones calibrando parámetros para lograr una mayor eficiencia. La salida de predicción que se muestra en la Figura 4.1 es resultado de la calibración de la red neuronal para mejorar la confiabilidad del algoritmo. El objetivo es predecir los valores máximos de una acción usando el comportamiento en años anteriores, además de disminuir la correlación incluyendo las noticias positivas y negativas en un día para validar cual es la eficiencia del algoritmo.

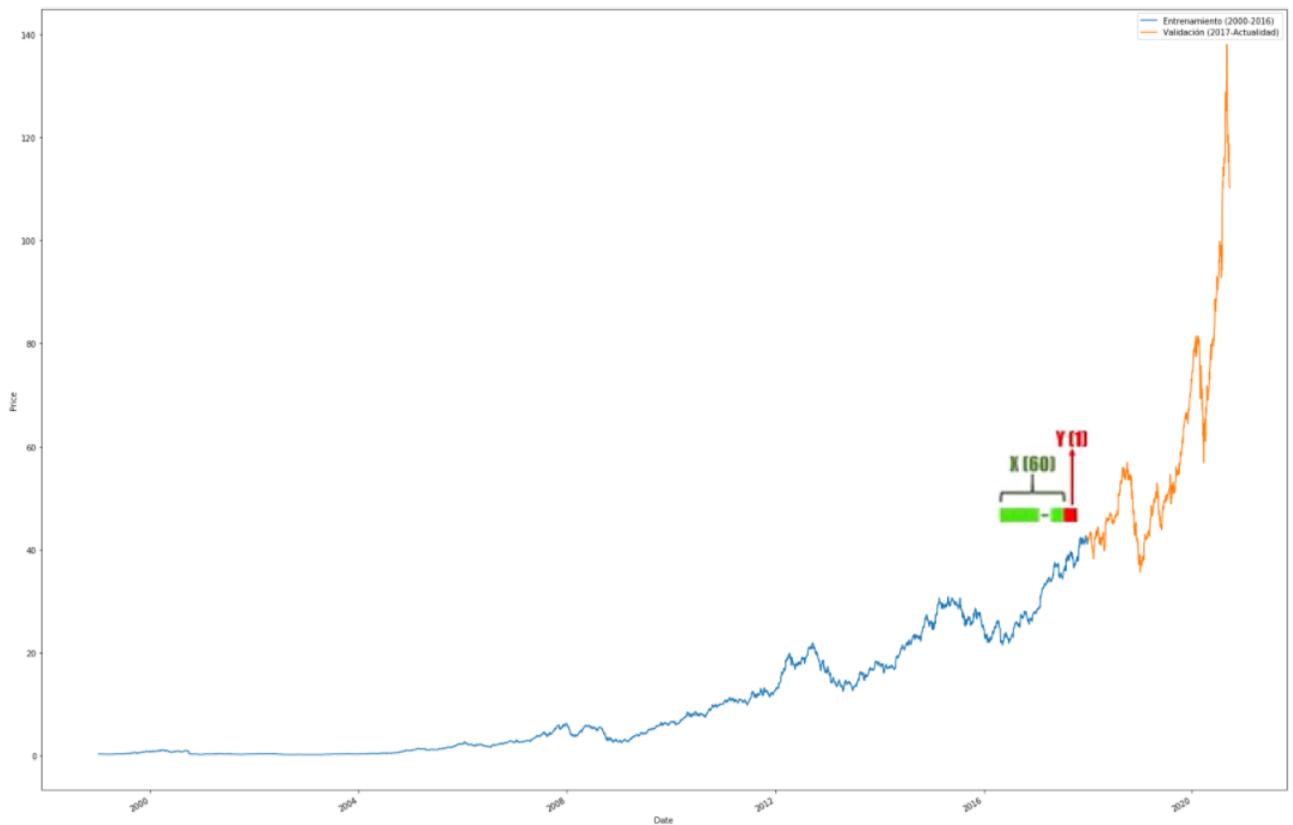


Figura 4.1: Representación de valores obtenidos para entrenamiento (Elaboración Propia)

El resultado para el modelo 1 no es aceptable debido a que es evidente que hace falta disminuir la correlación entre la curva real y la curva de predicción. Sin embargo, es un primer acercamiento ya que permite observar que las redes LSTM son útiles para predicciones con series de tiempo. Para presentar y analizar la predicción de forma más intuitiva, se muestra la comparación del valor real del mercado y la predicción de este método durante un período de tiempo largo en la figura 4.2. Al hacer un acercamiento a la imagen se visualizan algunos períodos específicos más cortos, en los que el mercado se está comportando con diferentes tendencias. Por ejemplo, si se analiza en la última sección (escala 500), se puede interpretar que la predicción trata de acercarse a la línea real. En este caso, la curva no tiene muchos movimientos por lo que se podría interpretar que la predicción se acerca cuando no existe demasiada fluctuación en el mercado. La predicción trata de seguir la línea de tendencia de la curva real, sin embargo, todavía hay una diferencia significativa en la predicción y el valor real debido a que se está pronosticando con una red neuronal LSTM básica, la cual tiene como variable de entrada únicamente los datos históricos, además no hace un preprocesamiento de información. Por otro lado, únicamente se incluyen los datos de los de cierre de cada día, es decir, no se incluye la información intradía.

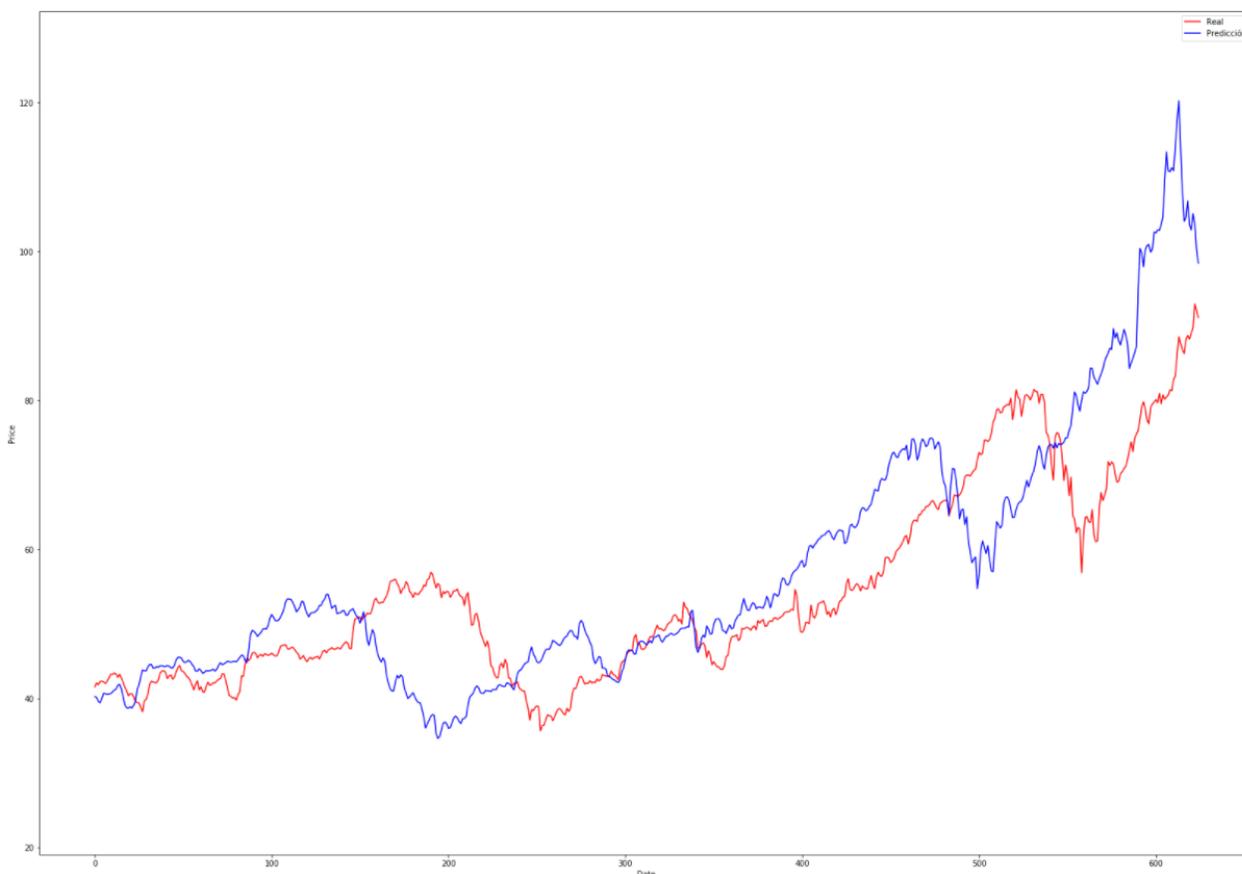


Figura 4.2: Curva de Predicción del Modelo 1
(Elaboración Propia)

El resultado para el modelo 2 mejoro significativamente, si observamos en este caso el precio de la acción de Apple (APPL) recientemente está incrementando, por lo que el modelo

predice un precio al alza para el día siguiente, lo cual es un resultado razonable. Al comparar este modelo con el modelo anterior lo primero que se observa es que la correlación que existe entre la curva real y la curva de predicción es más cercana. Lo que ayudó en este caso es que se incluyan capas adicionales para la red neuronal, es decir, ya no es una red LSTM solamente. Además, se incrementó el número de neuronas para que el aprendizaje fuese más profundo. Se puede notar un efecto en este caso y es que la predicción es más cercana cuando existe una caída brusca del precio y luego vuelve a una tendencia alcista (escala 50). Este patrón es interesante debido a que sería de utilidad a la hora de invertir saber cuándo el mercado marca una nueva tendencia para tomar decisiones. Por ejemplo, si el inversionista es un trader que le gusta operar con operaciones en corto, tener este tipo de predicciones le puede ser útil.

Las ventas en corto se producen cuando la persona que realiza la inversión cree que el precio del valor bajará en el futuro. De esta manera, si el precio de la acción realmente baja después de abrir la operación, el operador debería comprar a un precio menor que el precio por el que las vendió, ganando la diferencia. Al cerrar la operación comprando las acciones, el operador puede devolver las acciones al intermediario que se las prestó y consigue un beneficio, ya que la cantidad que paga por comprar dichas acciones es menor que la cantidad que recibió al venderlas.

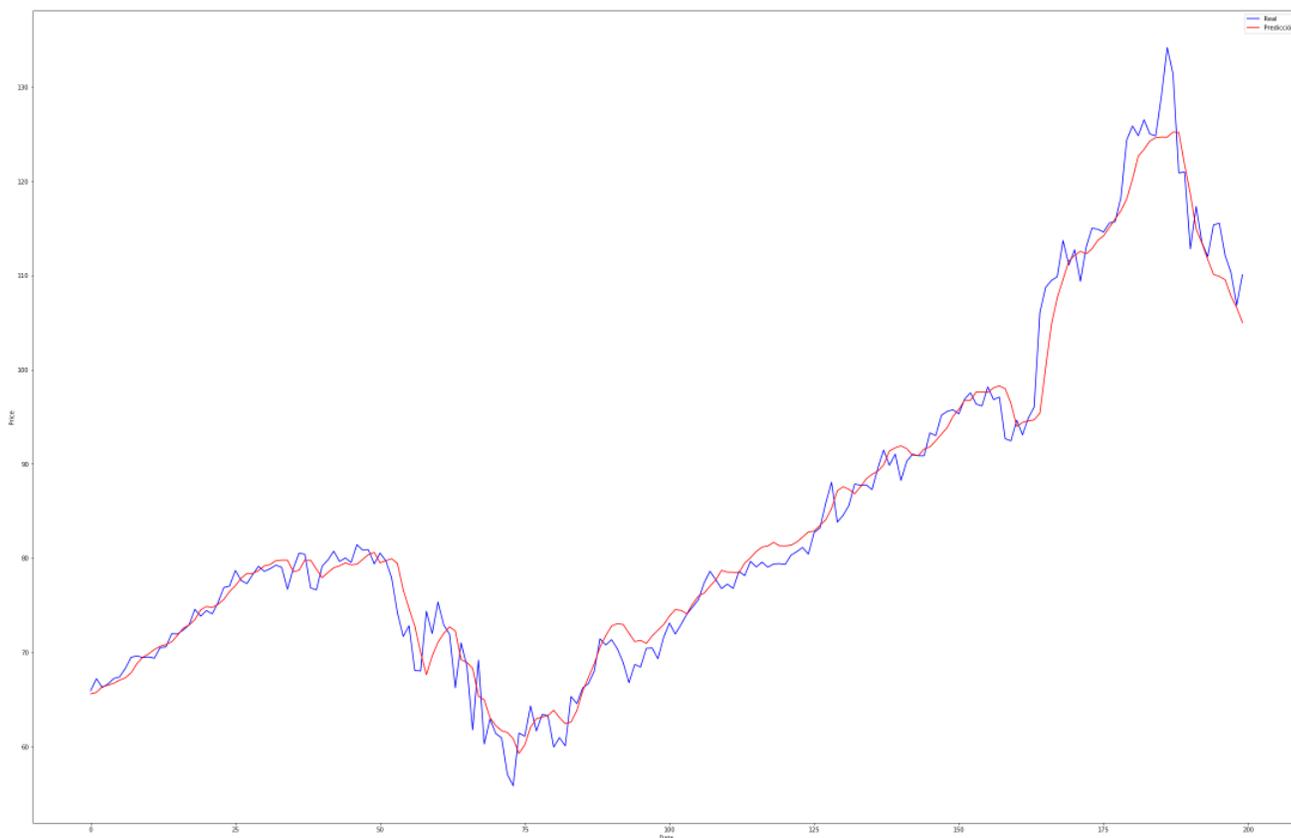


Figura 4.3: Curva de Predicción del Modelo 2
(Elaboración Propia)

En el tercer modelo la predicción del algoritmo mejora y la correlación es pequeña com-

parada con los modelos anteriores. Esto se debe a que se incluyó una nueva entrada que es el indicador SMA el cual ayuda a la precisión del algoritmo, es decir, el modelo es más robusto ya que incluye dos entradas de información como son los datos históricos intradía, es decir toda la fluctuación diaria de la acción. Así mismo se aplica uno de los indicadores técnicos usado por los traders como es el promedio móvil siempre que viene siendo la media no ponderada de los datos n anteriores. Este indicador fácil de implementar marca la diferencia en el algoritmo debido a que permite un tratamiento previo de los datos históricos antes que se ingresen a la red neuronal.

La curva de predicción en este caso es más confiable, pues la diferencia es menor con la curva de precio real y hace más fiable al algoritmo (Figura 4.4). Un punto importante, es que el primer día de cotización en bolsa de las empresas, es decir, cuando las empresas cotizan por primera vez en bolsa, el capital que venden en una oferta pública inicial tiende a estar infravalorado [46], lo que resulta en un aumento del precio el primer día de negociación. Esto afectó la forma en que se escalaron otros valores de volumen en el conjunto de datos al normalizar los datos, por lo que se optó por eliminar los puntos más antiguos de cada conjunto. También se eliminó la fecha ya que el modelo no necesita saber nada sobre cuándo ocurrieron las operaciones, todo lo que necesita son datos de series de tiempo bien ordenados.

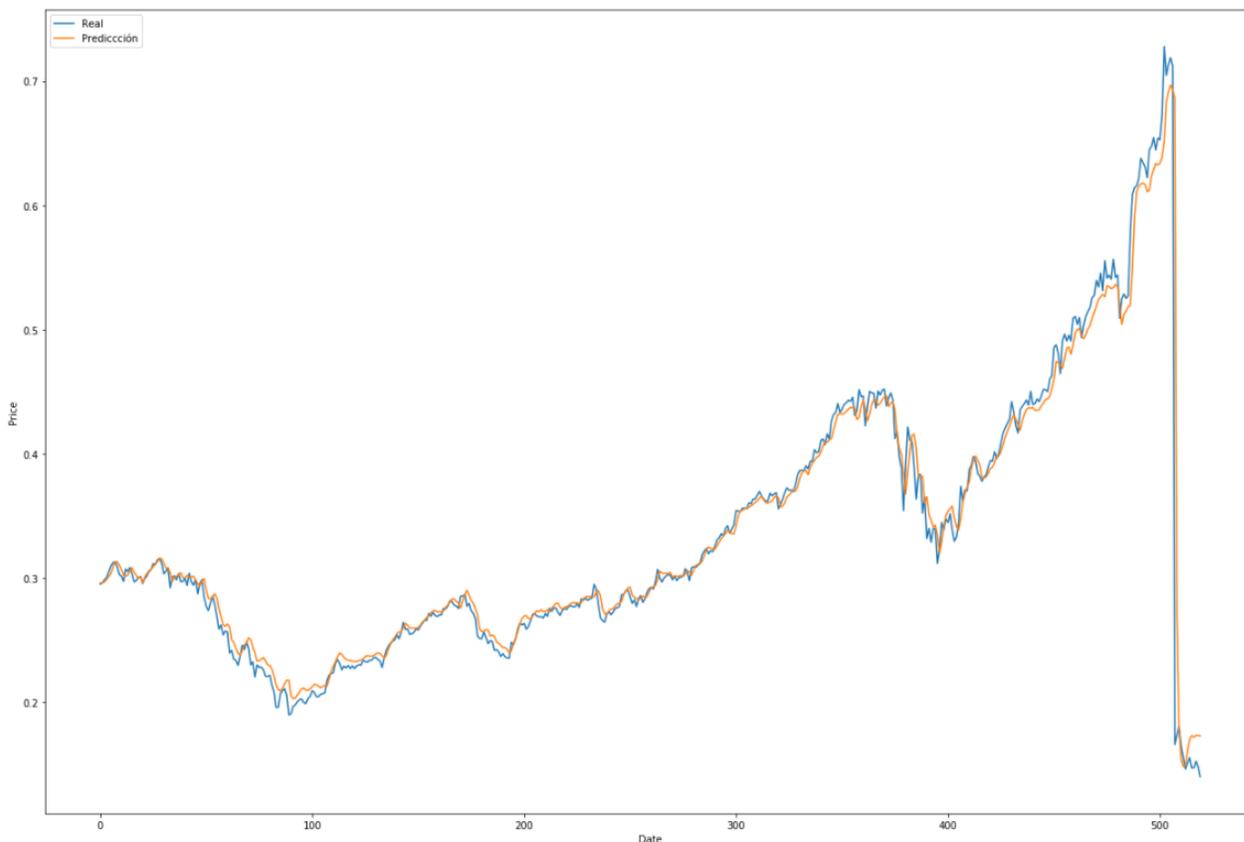


Figura 4.4: Curva de Predicción del Modelo 3
(Elaboración Propia)

Capítulo 5

Conclusiones

Este documento establece un marco de pronóstico para predecir los precios de apertura en el mercado de acciones. Utilizado una red neuronal LSTM la cual tiene como entrada la información histórica de los precios de las acciones, el trabajo consistió en predecir el precio de apertura de las acciones. Los resultados experimentales muestran un acercamiento a los valores reales de una acción debido a que se desarrollaron tres modelos para elegir finalmente cual es el más adecuado. Por lo tanto, el modelo tiene varias posibilidades de aplicación y puede ser competitivo con algunos modelos existentes, principalmente por su facilidad de implementación.

Respecto al tema de noticias del mercado de valores, lo que se puede concluir por el momento es que no es viable incluir esta información como entrada para la predicción debido a que no se cuenta con datos históricos y en este sentido no hace un aporte significativo al algoritmo. Sin embargo, esto tiene una particularidad y es que varias noticias históricas son absorbidas por los precios de los activos financieros en algún momento de su trayectoria. Se requerirán experimentos más exhaustivos en el futuro para comprender mejor cómo el modelo puede ayudar a evaluar con mayor certeza si una noticia es positiva y si esto afecta a la red LSTM. Lo que nos dio mejor resultado en este proyecto es incluir indicadores técnicos.

Ante situaciones de crisis o de expansión crediticia y liquidez, se hace necesario el apoyo de herramientas para la toma de decisiones de inversión. En la pandemia del COVID-19, las inversiones en línea se incrementaron como una opción económica y muchas personas llegaron por ruido del mercado o noticias en redes sociales. Justamente para ese tipo de inversionistas principiantes sin mucho conocimiento del negocio, está pensado este proyecto. En el mercado existen varias herramientas que permiten realizar predicciones con un alto grado de credibilidad que justamente son utilizadas por los traders que operan en la bolsa de valores de manera profesional. Sin embargo, hay varios puntos a tener en cuenta y que de alguna manera permiten mostrar las diferencias con este proyecto. En primer lugar, hay un alto grado de complejidad para utilizar estas herramientas por lo que justamente son utilizadas por traders profesionales lo cual las limita a un sector exclusivo y no a que sean utilizadas por el inversionista común. Lo que se propone con esta herramienta es su fácil uso para cualquier persona que desee invertir de tal manera que con un conocimiento básico, la persona pueda ejecutar su modelo y saber la dirección del mercado.

Otro punto es el costo, pues algunas de las soluciones que existen en el mercado tienen un precio elevado debido a que requieren cierto número de personal para operar e infraestructura especializada (ej.: Bloomberg, Refinitiv) y esto no sería viable para un inversionista pequeño o que está empezando en este negocio. En este caso la solución planteada no tendría un alto costo, si bien el entrenamiento del algoritmo requiere potencia de cómputo también es cierto que con el avance de tecnología se encuentra este tipo de servicios en la nube con costos accesibles. Es por ello que la infraestructura para poner en marcha el desarrollo no representa un costo elevado, además el equipo de desarrollo es bastante reducido.

A lo largo de la historia los inversionistas tuvieron herramientas para lograr aciertos en sus decisiones de inversión como indicadores técnicos y noticias del entorno confiando además en los sentimientos que de una u otra manera influyen en la decisión final, es así que para ayudar a unificar los diferentes componentes que se necesitan para un inversionista hoy en día se puede utilizar algoritmos de aprendizaje automático con una serie de variables y que gracias a las herramientas implementadas en los datos masivos se encuentra información disponible en la red para lograr mayor precisión. En este sentido el trabajo futuro tiene varias direcciones ya que este trabajo ha encontrado que un LSTM basado en la economía del comportamiento tiene más aciertos para la predicción de precios que otros métodos, es decir, el impacto de los datos históricos sobre las tendencias de los precios es demasiado singular y es posible que no se pueda pronosticar completamente y con precisión el precio en un día determinado. Se puede pensar en incluir noticias relacionadas con acciones e información básica, para mejorar la estabilidad y precisión del modelo en el caso de un evento importante de más sitios o redes sociales (Facebook, Twitter, etc) u otros factores económicos del momento para dar mayor precisión al algoritmo. Además, se puede pensar en incluir indicadores técnicos más avanzados (osciladores) como nuevas variables para probar la efectividad del modelo.

A futuro se busca seguir perfeccionando el algoritmo de tal manera que pueda ser más confiable para el pequeño inversor. Además con ciertos cambios en las variables de entrada se podría aplicar no sólo para renta variable, sino para otros mercados como puede ser criptomonedas, monedas y algunos derivados financieros. Además se espera publicar la solución en la nube y agregar un front amigable al usuario para que el proyecto sea más atractivo para lo cual se contara con un esfuerzo adicional en desarrollo e inversión. En definitiva hay mucho por aprender y perfeccionar en este mundo de inversiones basado en el mundo del aprendizaje automático y el aprendizaje profundo.

Bibliografía

- [1] F. D. Sciorio, *From Efficient Market to Behavioral Finance*. LAP LAMBERT Academic Publishing, 2018.
- [2] H. Shefrin, *Beyond greed and fear: Understanding behavioral finance and the psychology of investing*. Oxford University Press, 2002.
- [3] B. Mandelbrot and R. L. Hudson, *The Misbehavior of Markets: A Fractal View of Financial Turbulence*. Basic Books, 2006.
- [4] G. Cooper, *The Origin of Financial Crises: Central Banks, Credit Bubbles, and the Efficient Market Fallacy*. Vintage Press, 2008.
- [5] R. J. Shiller, *Irrational exuberance: Revised and expanded third edition*. Princeton university press, 2015.
- [6] O. A. Lamont and R. H. Thaler, “Can the market add and subtract? mispricing in tech stock carve-outs,” *Journal of Political Economy*, vol. 111, no. 2, pp. 227–268, 2003.
- [7] H. K. Baker and J. R. Nofsinger, *Behavioral finance: investors, corporations, and markets*. John Wiley & Sons, 2010, vol. 6.
- [8] K. Tseng, “Behavioral finance, bounded rationality, neuro-finance, and traditional finance,” *Investment Management and Financial Innovations*, no. 3, Iss. 4, pp. 7–18, 2006.
- [9] R. J. Shiller, “From efficient markets theory to behavioral finance,” *Journal of economic perspectives*, vol. 17, no. 1, pp. 83–104, 2003.
- [10] E. F. Fama, “Market efficiency, long-term returns, and behavioral finance,” *Journal of Financial Economics*, vol. 49, no. 3, pp. 283 – 306, 1998. [Online]. Available: [https://doi.org/10.1016/S0304-405X\(98\)00026-9](https://doi.org/10.1016/S0304-405X(98)00026-9)
- [11] S. F. LeRoy and R. D. Porter, “Stock price volatility: tests based on implied variance bounds,” *Econometrica*, vol. 49, no. 97, p. 113, 1981.
- [12] A. Shleifer, *Inefficient markets: an introduction to behavioral finance*. Oxford University Press, 2000.
- [13] J. Fagerström, M. Bång, D. Wilhelms, and M. S. Chew, “Lisep lstm: A machine learning algorithm for early detection of septic shock,” *Scientific reports*, vol. 9, no. 1, pp. 1–8, 2019. [Online]. Available: <https://doi.org/10.1038/s41598-019-51219-4>

- [14] M. Rahman, D. Islam, R. J. Mukti, and I. Saha, “A deep learning approach based on convolutional lstm for detecting diabetes,” *Computational Biology and Chemistry*, vol. 88, p. 107329, 2020. [Online]. Available: <https://doi.org/10.1016/j.compbiolchem.2020.107329>
- [15] T. Santosh, D. Ramesh, and D. Reddy, “Lstm based prediction of malaria abundances using big data,” *Computers in Biology and Medicine*, vol. 124, p. 103859, 2020. [Online]. Available: <https://doi.org/10.1016/j.compbiomed.2020.103859>
- [16] H. Jing, Y. Zhang, J. Zhou, W. Zhang, X. Liu, G. Min, and Z. Zhang, “Lstm-based service migration for pervasive cloud computing,” in *2018 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData)*, 2018, pp. 1835–1840. [Online]. Available: [doi:10.1109/Cybermatics.2018.2018.00305](https://doi.org/10.1109/Cybermatics.2018.2018.00305).
- [17] B. M. Ashok Tholusuri, Manish Anumala and G. J. Lakshmi, “Sentiment analysis using lstm,” *International Journal of Engineering and Advanced Technology (IJEAT)*, vol. 8, 2019.
- [18] J. Zhao, D. Zeng, Y. Xiao, L. Che, and M. Wang, “User personality prediction based on topic preference and sentiment analysis using lstm model,” *Pattern Recognition Letters*, vol. 138, pp. 397 – 402, 2020. [Online]. Available: <https://doi.org/10.1016/j.patrec.2020.07.035>
- [19] B. W. Wanjawa and L. Muchemi, “Ann model to predict stock prices at stock exchange markets,” *arXiv preprint arXiv:1502.06434*, 2014.
- [20] R. LaLonde, D. Zhang, and M. Shah, “Fully convolutional deep neural networks for persistent multi-frame multi-object detection in wide area aerial videos,” *arXiv preprint arXiv:1704.02694*, 2017.
- [21] S. A. Hamid and Z. Iqbal, “Using neural networks for forecasting volatility of s&p 500 index futures prices,” *Journal of Business Research*, vol. 57, no. 10, pp. 1116 – 1125, 2004, selected Papers from the third Retail Seminar of the SMA. [Online]. Available: [https://doi.org/10.1016/S0148-2963\(03\)00043-2](https://doi.org/10.1016/S0148-2963(03)00043-2)
- [22] H. Y. Kim and C. H. Won, “Forecasting the volatility of stock price index: A hybrid model integrating lstm with multiple garch-type models,” *Expert Systems with Applications*, vol. 103, pp. 25 – 37, 2018. [Online]. Available: <https://doi.org/10.1016/j.eswa.2018.03.002>
- [23] M. Bildirici and Özgür Ömer Ersin, “Forecasting oil prices: Smooth transition and neural network augmented garch family models,” *Journal of Petroleum Science and Engineering*, vol. 109, pp. 230 – 240, 2013. [Online]. Available: <https://doi.org/10.1016/j.petrol.2013.08.003>
- [24] W. Kristjanpoller and M. C. Minutolo, “Gold price volatility: A forecasting approach using the artificial neural network–garch model,” *Expert Systems with*

- Applications*, vol. 42, no. 20, pp. 7245 – 7251, 2015. [Online]. Available: <https://doi.org/10.1016/j.eswa.2015.04.058>
- [25] Y. Hu, J. Ni, and L. Wen, “A hybrid deep learning approach by integrating lstm-ann networks with garch model for copper price volatility prediction,” *Physica A: Statistical Mechanics and its Applications*, vol. 557, p. 124907, 2020. [Online]. Available: <https://doi.org/10.1016/j.physa.2020.124907>
- [26] Z. Zhao, R. Rao, S. Tu, and J. Shi, “Time-weighted lstm model with redefined labeling for stock trend prediction,” in *2017 IEEE 29th International Conference on Tools with Artificial Intelligence (ICTAI)*. IEEE, 2017, pp. 1210–1217.
- [27] S. Wang and J. Jiang, “Learning natural language inference with lstm,” *arXiv preprint arXiv:1512.08849*, 2015.
- [28] J.-p. Zhao and D.-J. Huang, “Mirror extending and circular spline function for empirical mode decomposition method,” *Journal of Zhejiang University (Science)*, vol. 2, no. 3, pp. 247–252, 2001.
- [29] J. Qiu, B. Wang, and C. Zhou, “Forecasting stock prices with long-short term memory neural network based on attention mechanism,” *PloS one*, vol. 15, no. 1, p. e0227222, 2020. [Online]. Available: <https://doi.org/10.1371/journal.pone.0227222>
- [30] D. M. Nelson, A. C. Pereira, and R. A. de Oliveira, “Stock market’s price movement prediction with lstm neural networks,” in *2017 International joint conference on neural networks (IJCNN)*. IEEE, 2017, pp. 1419–1426.
- [31] Q. Gao, “Stock market forecasting using recurrent neural network,” Ph.D. dissertation, University of Missouri–Columbia, 2016.
- [32] Y. Wang, Y. Liu, M. Wang, and R. Liu, “Lstm model optimization on stock price forecasting,” in *2018 17th International Symposium on Distributed Computing and Applications for Business Engineering and Science (DCABES)*. IEEE, 2018, pp. 173–177.
- [33] L.-J. Cao and F. E. H. Tay, “Support vector machine with adaptive parameters in financial time series forecasting,” *IEEE Transactions on neural networks*, vol. 14, no. 6, pp. 1506–1518, 2003.
- [34] T. Lin, T. Guo, and K. Aberer, “Hybrid neural networks for learning the trend in time series,” in *Proceedings of the twenty-sixth international joint conference on artificial intelligence*, no. CONF, 2017, pp. 2273–2279.
- [35] M. Sangiorgio and F. Dercole, “Robustness of lstm neural networks for multi-step forecasting of chaotic time series,” *Chaos, Solitons & Fractals*, vol. 139, p. 110045, 2020. [Online]. Available: <https://doi.org/10.1016/j.chaos.2020.110045>
- [36] B. Krollner, B. J. Vanstone, and G. R. Finnie, “Financial time series forecasting with machine learning techniques: a survey.” in *Esann*, 2010.

- [37] J. Kumar, R. Goomer, and A. K. Singh, “Long short term memory recurrent neural network (lstm-rnn) based workload forecasting model for cloud datacenters,” *Procedia Computer Science*, vol. 125, pp. 676 – 682, 2018, the 6th International Conference on Smart Computing and Communications. [Online]. Available: <https://doi.org/10.1016/j.procs.2017.12.087>
- [38] C.-J. Lu, T.-S. Lee, and C.-C. Chiu, “Financial time series forecasting using independent component analysis and support vector regression,” *Decision support systems*, vol. 47, no. 2, pp. 115–125, 2009.
- [39] H. Yan, W. Liu, X. Liu, H. Kong, and C. Lv, “Predicting net asset value of investment fund based on bp neural network,” in *2010 international conference on computer application and system modeling (ICCA SM 2010)*, vol. 10. IEEE, 2010, pp. V10–635.
- [40] M. P. Clements, P. H. Franses, and N. R. Swanson, “Forecasting economic and financial time-series with non-linear models,” *International Journal of Forecasting*, vol. 20, no. 2, pp. 169–183, 2004.
- [41] Z. Y. Zhao, M. Xie, and M. West, “Dynamic dependence networks: Financial time series forecasting and portfolio decisions,” *Applied Stochastic Models in Business and Industry*, vol. 32, no. 3, pp. 311–332, 2016.
- [42] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997. [Online]. Available: <https://doi.org/10.1162/neco.1997.9.8.1735>
- [43] S. Nison, *Japanese Candlestick Charting Techniques*. Prentice Hall Press, 2001.
- [44] spacy 101: Everything you need to know. [Online]. Available: <https://spacy.io/usage/spacy-101#pipelines>
- [45] J. J. Murphy, *Technical Analysis of the Financial Markets: A Comprehensive Guide to Trading Methods and Applications*. New York Institute of Finance, 1999.
- [46] A. Ljungqvist, “Ipo underpricing,” in *Handbook of Empirical Corporate Finance*. Elsevier, 2007, pp. 375–422.