



**UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL**

**FACULTAD DE INGENIERÍA
CARRERA DE INGENIERÍA EN SISTEMAS
COMPUTACIONALES**

TEMA:

**Modelo predictivo para el comportamiento de las acciones que
cotizan en la bolsa de valores de Nueva York, utilizando minería de
datos**

AUTOR:

Santistevan Pineda, Stephano Paul

**Trabajo de titulación previo a la obtención del título de
INGENIERO EN SISTEMAS COMPUTACIONALES**

TUTOR:

Ing. Miranda Rodríguez, Marcos Xavier, Mgs.

Guayaquil, Ecuador

8 de marzo del 2019



**UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL**

**FACULTAD DE INGENIERÍA
CARRERA DE INGENIERÍA EN SISTEMAS
COMPUTACIONALES**

CERTIFICACIÓN

Certificamos que el presente trabajo de titulación **Modelo predictivo para el comportamiento de las acciones que cotizan en la bolsa de valores de Nueva York, utilizando minería de datos.**, fue realizado en su totalidad por **Santistevan Pineda Stephano Paul** como requerimiento para la obtención del Título de **Ingeniero en Sistemas Computacionales**.

TUTOR

Ing. Marcos Xavier Miranda Rodríguez, Mgs.

DIRECTORA (e) DE LA CARRERA

Ing. Ana Isabel Camacho Coronel, Mgs.

Guayaquil, a los 8 días del mes de marzo del año 2019



**UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL**

**FACULTAD DE INGENIERÍA
CARRERA DE INGENIERÍA EN SISTEMAS
COMPUTACIONALES**

DECLARACIÓN DE RESPONSABILIDAD

Yo, Santistevan Pineda Stephano Paul

DECLARO QUE:

El Trabajo de Titulación **Modelo predictivo para el comportamiento de las acciones que cotizan en la bolsa de valores de Nueva York, utilizando minería de datos** previo a la obtención del Título de **Ingeniero en Sistemas Computacionales**, ha sido desarrollado respetando derechos intelectuales de terceros conforme las citas que constan en el documento, cuyas fuentes se incorporan en las referencias o bibliografías. Consecuentemente este trabajo es de mi total autoría.

En virtud de esta declaración, me responsabilizo del contenido, veracidad y alcance del Trabajo de Titulación referido.

Guayaquil, a los 8 días del mes de marzo del año 2019

EL AUTOR

Santistevan Pineda Stephano Paul



UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL

FACULTAD DE INGENIERÍA
CARRERA DE INGENIERÍA EN SISTEMAS
COMPUTACIONALES

AUTORIZACIÓN

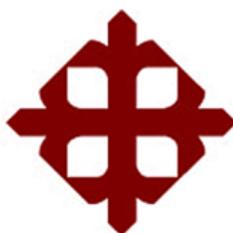
Yo, **Santistevan Pineda Stephano Paul**

Autorizo a la Universidad Católica de Santiago de Guayaquil a la **publicación** en la biblioteca de la institución del Trabajo de Titulación **Modelo predictivo para el comportamiento de las acciones que cotizan en la bolsa de valores de Nueva York, utilizando minería de datos**, cuyo contenido, ideas y criterios son de mi exclusiva responsabilidad y total autoría.

Guayaquil, a los 8 días del mes de marzo del año 2019

EL AUTOR

Santistevan Pineda Stephano Paul



UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL

FACULTAD DE INGENIERÍA
CARRERA DE INGENIERÍA EN SISTEMAS
COMPUTACIONALES

REPORTE DE URKUND

URKUND	
Documento	SANTISTEVAN PINEDA-URKUND_1.docx (D48288401)
Presentado	2019-02-24 23:18 (-05:00)
Presentado por	MARCOS XAVIER MIRANDA RODRIGUEZ (marcos.miranda@cu.ucsg.edu.ec)
Recibido	marcos.miranda.ucsg@analysis.orkund.com
Mensaje	<input type="checkbox"/> Mostrar el mensaje completo
	<div style="background-color: #4CAF50; color: white; padding: 2px;">0%</div> de estas 63 páginas, se componen de texto presente en 0 fuentes.

DEDICATORIA

Dedico este proyecto a Dios Todopoderoso, mi creador, mi roca, mi fortaleza, mi fuente de inspiración, sabiduría, conocimiento y comprensión. Él ha sido la fuente de mi fuerza a lo largo de este proyecto y en sus alas solo me he elevado. A mis padres y hermano, que son mi fortaleza, mi empuje de todos los días, sin el apoyo de ellos, esto no hubiera sido posible. Y a mi papi Jaime.

AGRADECIMIENTO

A Dios por darnos la capacidad de avanzar incluso cuando caemos.

A el Ing. Marco Xavier Miranda Rodríguez por toda la ayuda brindada a lo largo del presente proyecto de titulación.



**UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL**

**FACULTAD DE INGENIERÍA
CARRERA DE INGENIERÍA EN SISTEMAS
COMPUTACIONALES**

TRIBUNAL DE SUSTENTACIÓN

**Ing. Ana Isabel Camacho Coronel, Mgs.
DIRECTORA (e) DE LA CARRERA**

**Ing. Galo Enrique Cornejo Gómez, Mgs.
COORDINADOR DEL ÁREA O DOCENTE DE LA CARRERA**

**Ing. Jorge Salvador Pesantes Méndez, Mgs.
OPONENTE**

ÍNDICE GENERAL

RESUMEN	XII
INTRODUCCIÓN	2
CAPÍTULO I EL PROBLEMA.....	4
1.1 Planteamiento del Problema	4
1.2 Hipótesis y/o Preguntas de Investigación	5
1.2.1 Hipótesis.....	5
1.2.2 Pregunta de la investigación.....	5
1.3 Objetivos de la Investigación	5
1.3.1 Objetivo General	6
1.3.2 Objetivos Específicos.....	6
1.4 Justificación y Alcance	6
CAPÍTULO II MARCO TEÓRICO, CONCEPTUAL Y LEGAL	8
2.1 Elementos Teóricos	8
2.1.1 Mercado de Valores	8
2.1.2 Minería de Datos	17
2.2 Conceptos	27
2.3 Base Legal	29
2.4 Ámbito de Aplicación.....	32
CAPÍTULO III METODOLOGÍA Y RESULTADOS	35
3.1 Metodología de la Investigación.....	35
3.2 Metodología de Desarrollo	36
3.3 Análisis de Resultados.....	41
CAPÍTULO IV Modelo predictivo.....	45
4.1 Objetivo	45
4.2 Descripción del modelo	45
4.3 Requerimientos para su implementación	54
4.4 Resultados esperados del modelo	55
CONCLUSIONES	56
RECOMENDACIONES.....	60
REFERENCIAS BIBLIOGRÁFICAS	62
APÉNDICES	65

ÍNDICE DE TABLAS

Tabla 1. Datos de entrada valores apertura, cierre y volumen (JNJ).....	47
Tabla 2. Datos de salida (1 día en el futuro) apertura, cierre y volumen. (JNJ).....	47
Tabla 3. Secuencia de entrada (longitud 4 días), apertura, cierre y volumen. (JNJ).	48
Tabla 4. Secuencia de salida (longitud 4 días)apertura, cierre y volumen.(JNJ).....	48
Tabla 5. Requerimientos para la implementación del aplicativo.....	54
Tabla 6. Resultados de indicadores matemáticos para el Modelo LSTM.	59
Tabla 7. Resultados de indicadores matemáticos para el Modelo ARIMA.....	59

ÍNDICE DE FIGURAS

Figura 1. Gráfica de Barra y Velas Japonesas.	12
Figura 2. Ilustración de vela cuerpo negra y vela blanca.....	13
Figura 3. Ilustración de vela cuerpo negra y vela blanca.....	14
Figura 4. Ilustración de Hombre colgado y Martillo.	15
Figura 5. Ilustración de Estrella Fugaz y Martillo Invertido..	16
Figura 6. Ejemplo de una red neuronal.....	24
Figura 7. Procesos de Minería de Datos.	36
Figura 8. Ilustración de Análisis de precisión de error de los índices.	38
Figura 9. Gráfica comparativa Valores reales y Valores predichos de las acciones..	38
Figura 10. Metodología desarrollo modelo predictivo para RNN de tipo LSTM.	40
Figura 11. Gráfica de velas de Precio de cierre y apertura de (CL).	41
Figura 12. Gráfica de velas de Precio de cierre y apertura de (RST).	41
Figura 13. Gráfica de velas de Precio de cierre y apertura de FEDEX (FDX).....	42
Figura 14. Gráfica de velas de Precio de cierre y apertura de (JNJ).....	42
Figura 15. Esquema del modelo predictivo propuesto.	46
Figura 16. Definición de la RNN de tipo LSTM, para la presente propuesta.	50
Figura 17. Pasos para el uso del aplicativo de la presente propuesta.	53
Figura 18. Grafica comparativa indicadores ARIMA y LSTM.....	60

ÍNDICE DE APÉNDICES

Apéndice A. Formato del archivo .csv requerido por el modelo.....	65
Apéndice B. Gráfica precio de cierre Predichos y Precio de cierre Reales, de Colgate Palmolive (CL).....	66
Apéndice C. Gráfica indicadores de la corrida del modelo para Colgate Palmolive (CL).....	67
Apéndice D. Tabla de precios de Cierre predichos y Reales para Colgate Palmolive (CL).....	68
Apéndice E. Gráfica precio de cierre Predichos y Precio de cierre Reales, de FEDEX (FDX).....	72
Apéndice F. Gráfica indicadores de la corrida del modelo para FEDEX (FDX).....	73
Apéndice G. Tabla de precios de Cierre predichos y Reales para FEDEX (FDX) ..	74
Apéndice H. Gráfica precio de cierre Predichos y Precio de cierre Reales, de Rosetta Stone (RST).....	78
Apéndice I. Gráfica indicadores de la corrida del modelo para Rosetta Stone (RST).	79
Apéndice J. Tabla de precios de Cierre predichos y Reales para Rosetta Stone (RST).	80
Apéndice K. Gráfica precio de cierre Predichos y Precio de cierre Reales, de Jhonson & Jhonson (JNJ).....	84
Apéndice L. Gráfica indicadores de la corrida del modelo para Jhonson & Jhonson (JNJ).....	85
Apéndice M. Tabla de precios de Cierre predichos y Reales para Jhonson & Jhonson (JNJ).....	86
Apéndice N. Resultados de corrida de modelo ARIMA para Colgate Palmolive (CL).	90
Apéndice O. Resultados de corrida de modelo ARIMA para Rosetta Stone (RST).	91
Apéndice P. Resultados de corrida de modelo ARIMA para FEDEX (FDX).....	92
Apéndice Q. Resultados de corrida de modelo ARIMA para Jhonson & Jhonson (JNJ).	93
Apéndice R. Manual de Usuario.....	94
Apéndice S. Indicadores de Predicciones del modelo Colgate Palmolive (CL).....	97

Apéndice T. Indicadores de Predicciones del modelo Fedex (FDX).	98
Apéndice U. Indicadores de Predicciones del modelo Rosetta Stone (RST).	99
Apéndice V. Indicadores de Predicciones del modelo Jhonson & Jhonson (JNJ). .	100
Apéndice W. Representación gráfica del modelo predictivo definitivo (Red Neuronal).	101
Apéndice X. Representación grafica de los modelos predictivos anteriores al definitivo (Red Neuronal).....	102

RESUMEN

El presente trabajo “**Modelo predictivo para el comportamiento de las acciones que cotizan en la bolsa de valores de Nueva York, utilizando minería de datos**” trata de brindar una herramienta capaz de poder predecir el comportamiento del mercado de valores de la ciudad de Nueva York, mediante la aplicación de la minería de datos, haciendo uso de la data histórica del mercado de valores, teniendo en cuenta este tipo de predicciones, tradicionalmente se ha utilizado el modelo ARIMA para este tipo de predicciones, sin embargo en la actualidad la minería de datos se da de la mano con otra ciencia, tal como las Redes Neuronales, en el caso específico de las series temporales, se plantea el uso de las LSTM que son un tipo de Redes Neuronales Recurrentes.

El mercado de valores presenta una gran volatilidad, lo que a su vez significa que posee un alto grado de riesgo, sin embargo dada la globalización de los mercados, la cantidad de personas que invierten en el mercado de valores va en ascenso, por lo cual se vuelve en una necesidad la existencia de una herramienta que sea capaz de predecir el comportamiento del mercado de valores con un alto grado de certidumbre, pudiendo ser de gran utilidad al momento de realizar transacciones de compra y venta de valores. Para lo cual se tomó como herramienta la minería de datos, la cual es capaz de hacer uso de toda esa gran cantidad de data, proveniente del comportamiento histórico del mercado de valores. De acuerdo a estudios recientes las redes de tipo LSTM, son capaces de poder predecir con un alto grado de certidumbre el comportamiento del mercado de valores, destacando que las Redes Neuronales es una de las técnicas de minería de datos, que hoy en día se encuentra en constante desarrollo, presentándose como una alternativa a las técnicas ya tradicionales, sin embargo al hacer uso de las Redes Neuronales, como modelo, se debe de tener muy en cuenta el valor de la data con la que se entrena la red neural, en gran parte la precisión de los resultados dependerá de la data de entrenamiento del modelo.

Palabras Claves: MINERÍA DE DATOS; MERCADO DE VALORES; BOLSA DE VALORES; REDES NEURONALES; LSTM; SERIES TEMPORALES.

INTRODUCCIÓN

Las acciones son consideradas una gran oportunidad para invertir, el invertir en acciones de una manera correcta y dinámica puede dejar grandes márgenes de utilidad, pero, así como los márgenes de utilidad también pueden llegar a ser altos, los niveles de riesgos que el inversionista tiene que saber sobrellevar también son altos. El mercado de acciones. El mercado de acciones está conformado por los demandantes que tratan de comprar las acciones para después poder venderlas a un precio mayor al que fueron adquiridas, es decir realizan predicciones del valor futuro de las mismas teniendo en cuenta que este deberá de ser mayor al que fueron compradas, para de esta manera poder obtener réditos financieros, esto es si todo fuera de acuerdo a lo planificado, en caso de no ser así deberán de vender las acción a un precio igual al que fueron compradas o tratar de limitar su pérdida hasta cierto porcentaje.

La problemática de la presente investigación es que al día de hoy existen herramientas que permiten hasta cierto punto pueden llegar predecir el comportamiento de acciones que cotizan en la bolsa de valores, sin embargo, la mayoría de estas herramientas son de paga o propiedad exclusiva de las firmas financieras que actúan en el sector bursátil. El poder contar con este tipo de herramientas sería de gran utilidad para los inversionistas, aportando al momento de toma de decisiones, decidiendo si es conveniente comprar o vender una acción, facilitando la tarea de los nuevos inversionistas que incursionan en el mercado bursátil, cantidad que cada día van en aumento dadas las facilidades que han traído los brókers en línea dando apertura a las masas, todo esto gracias a la globalización de los mercados.

Para lo cual en la presente propuesta se plantea el desarrollo de un modelo predictivo el cual sea capaz de predecir el comportamiento del mercado de valores de la ciudad de Nueva York, con un alto grado de certidumbre, capaz de alcanzar e inclusive mejorar los resultados obtenidos por los modelos empleados actualmente, mediante el uso de las nuevas tendencias en el mundo tecnológico. Siendo el modelo propuesto una respuesta a la problemática planteada, brindando una herramienta para

la toma de decisiones de compra o venta de valores en la bolsa de valores de la ciudad de Nueva York.

El resultado de la investigación se presenta en la siguientes páginas, con la siguiente estructura: En el Capítulo I haciendo la presentación de la problemática, la justificación, el alcance y los objetivos; En el Capítulo II, se presenta el marco teórico, conceptos, la legalidad y el ámbito de la propuesta; En el Capítulo III, se presenta tanto la metodología de investigación como la metodología de desarrollo; En el Capítulo IV, la propuesta tecnológica per se; Por ultimo serán presentadas las conclusiones y recomendaciones de la presente investigación.

CAPÍTULO I

EL PROBLEMA

En el presente capítulo se describirá el problema contra el que se enfrentan los inversores, se plantearán los objetivos que cubrirá el presente trabajo, así como el alcance del mismo y la propuesta que trata de dar una solución al problema observado.

1.1 Planteamiento del Problema

Las acciones son una gran oportunidad de inversión para poder obtener ganancias de una manera rápida y de forma legal. El mercado de valores es volátil, lo que quiere decir que existe un alto riesgo, pero en caso de que se tomen las decisiones correctas los beneficios serían muy representativos. Las acciones acreditan a su propietario derechos sobre la compañía; son parte de la misma, es decir que cuando se posee una acción sería como ser dueño de una parte de la misma, el valor de una empresa puede aumentar o disminuir dependiendo de diversos factores, lo cual afecta de igual manera el precio de las acciones de la misma; para poder ganar dinero habría que comprar acciones a un precio determinado y venderlas cuando estas se coticen en un valor más alto al que fueron adquiridas.

El mercado de acciones está influenciado por predicciones que tratan de determinar el valor futuro de las acciones de una empresa, una predicción correcta disminuiría los riesgos que corre el inversionista a la hora de adquirir o vender acciones, las ganancias del inversionista serían maximizadas ya que podría conocer con certeza o con un mayor grado de precisión cuando le sería conveniente adquirir o vender acciones.

Hoy en día, existen herramientas que permiten realizar predicciones del comportamiento de las acciones de las distintas empresas dentro del mercado bursátil, la gran mayoría de las herramientas que están disponibles en el mercado son de paga, o requieren un alto nivel de conocimiento para poder ser utilizadas, muchas otras de estas herramientas son desarrolladas por las firmas corredoras de bolsa y son de propiedad exclusiva de éstas, por lo cual un inversionista que decida incursionar en el

mercado de acciones, corre en gran desventaja, a diferencia de los que ya tiene mucha más experiencia o cuentan con estas herramientas predictivas.

Las distintas técnicas de minería de datos permiten, con ayuda de la data histórica de la cotización de las acciones y con ayuda datos matemáticos, poder establecer un modelo que podría ser de gran utilidad predecir el comportamiento de las acciones en el futuro y así ayudar a los inversionistas a tomar decisiones correctas en el momento correcto.

1.2 Hipótesis y/o Preguntas de Investigación

Por medio de este estudio se busca corroborar o descartar la hipótesis y poder responder la pregunta de investigación.

1.2.1 Hipótesis

Un modelo predictivo, generado con minería de datos, permitirá predecir el comportamiento las acciones de la bolsa valores de Nueva York.

Variable independiente: Modelo predictivo con base en minería de datos.

Variable dependiente: Identificación de patrones en el comportamiento las acciones de la bolsa valores de Nueva York.

1.2.2 Pregunta de la investigación

¿Las técnicas de minería de datos permitirán generar un modelo predictivo para el comportamiento las acciones de la bolsa valores de Nueva York?

1.3 Objetivos de la Investigación

Los objetivos a alcanzar mediante esta investigación y que responden a la problemática planteada, son los siguientes:

1.3.1 Objetivo General

Elaborar un modelo que permita predecir el comportamiento de las acciones que cotizan en la bolsa de valores de Nueva York, utilizando minería de datos, para la toma de decisiones de compra o venta de acciones.

1.3.2 Objetivos Específicos

- Delimitar y recopilar información de las empresas a estudiar, que cotizan en la bolsa de valores de la ciudad de Nueva York, para la extracción de datos e información.
- Analizar la información recopilada del comportamiento de las acciones del mercado bursátil en la ciudad de Nueva York, para el diseño del modelo predictivo.
- Elaborar un modelo predictivo utilizando minería de datos para la toma de decisiones de compra o venta de acciones.
- Evaluar las predicciones generadas por el modelo, para determinar su exactitud.

1.4 Justificación y Alcance

El presente estudio tiene como objetivo brindar una herramienta a inversionistas que quieran incursionar en el mundo de las acciones, los cuales quizás no cuenten con una gran experiencia, en lo que al mercado bursátil se refiere, para poder tomar decisiones de una manera acertada y en el tiempo correcto, en el dinámico mundo de la bolsa de valores.

Los resultados del presente trabajo proporcionan un modelo, el más indicado para poder realizar estimaciones o predicción del comportamiento de las acciones, de esta manera poder minimizar los riesgos que corran los inversionistas, brindando un apoyo a la hora de la toma de decisiones.

Este trabajo de titulación se enmarca en las líneas de investigación y desarrollo de nuevos productos de la Facultad de Ingeniería de la UCSG y de la carrera Ingeniería en Sistemas Computacionales.

El cumplimiento de los objetivos propuestos se refleja en el alcance establecido en este proyecto:

Se utilizan técnicas de minería de datos, con el propósito de poder ser utilizadas sobre una vista minable, la cual será creada a partir de la data histórica de la cotización de las acciones. De esta data se podrán obtener patrones, tendencias y regularidades, los cuales permitirán poder realizar predicciones o estimaciones, las cuales serán analizadas para poder obtener el grado de precisión y confiabilidad de las mismas.

CAPÍTULO II

MARCO TEÓRICO, CONCEPTUAL Y LEGAL

En las últimas dos décadas grandes cambios se han dado en los mercados financieros, entre los cuales están: el desarrollo de poderosas herramientas de comunicación y las facilidades comerciales que han logrado ampliar al ámbito de los inversores. Las predicciones de los valores de las acciones es un importante tema financiero, que ha cautivado la atención de los investigadores a lo largo de los años. El hecho de poder tomar la información histórica fundamental, que es pública, y que esta data tenga la capacidad de establecer una relación para poder predecir el rendimiento futuro de las acciones, resulta muy cautivante. Para poder extraer este tipo de relación de la data que tengamos, existen técnicas de Data Mining permiten extraer este tipo de conocimiento de la data.

Todo lo relacionado con el comportamiento de acciones y su posibilidad de predicción de su valor, algunas conceptualizaciones y términos de sustento legal, así como el contexto general que se analiza, se ha incluido en este capítulo.

2.1 Elementos Teóricos relativos al mercado de valores y su efectividad con base en la tecnología

Para abordar la información relacionada con el uso de la minería de datos para realizar predicciones, concretamente en el comportamiento del mercado de valores, procede analizar algunas teorías y principios relativos a estos campos de estudio; por ello, en este apartado se hace una revisión del ámbito de la bolsa de valores, conceptos, clasificaciones, etc., para luego pasar a analizar aspectos de la tecnología y la técnica de minería de datos con sus correspondientes usos.

2.1.1 Mercado de Valores

Para el entendimiento del complejo mundo bursátil, se debe empezar por la definición del término inversión, la inversión es una manera de adquirir activos para conseguir beneficios en forma de ingresos, los cuales se puedan predecir de forma razonable (dividendos, intereses, rentas) (Malkiel, 2007). Una inversión también

puede ser definida como el instrumento en el cual se depositan fondos con la expectativa de que se generen ingresos o cuyo valor se mantenga o incremente (Gitman & Joehnk, 2009). El termino inversión abre las puertas al mundo bursátil, un mundo que está plagado de inversionistas, unos que venden, otros que compran, ambos son parte importante de este entorno, la inversión no es más sino la capacidad de adquirir un activo, el cual en un futuro ya sea éste a corto, mediano o largo plazo, con el fin de obtener ingresos. Teniendo en cuenta la definición del término inversión, el éxito de la inversión depende de la capacidad que se tenga para poder predecir el futuro.

Históricamente la necesidad de poder predecir con cierto grado de exactitud el comportamiento del mercado bursátil, no es algo nuevo, esta necesidad se remonta algunos siglos atrás desde la apertura del primer mercado bursátil, la apertura de la primera bolsa, la de Ámsterdam en 1602, desde esa época los analistas se han visto en la necesidad de desarrollar distintas teorías y análisis, los cuales serán descritos a continuación.

La **teoría de la base sólida**, presenta un elemento fundamental, el valor de anclaje, el cual viene determinado por un análisis de las condiciones actuales y las perspectivas del futuro, todo instrumento de inversión posee un fuerte anclaje al valor intrínseco. Para poder determinar el valor futuro de las inversiones, con el fin de identificar el momento más adecuado tanto para comprar o para vender alguna inversión, los analistas han definido dos principales métodos: análisis técnico y el análisis fundamental.

El *análisis técnico*, le permite al analista crear e identificar gráficos de títulos, este análisis estudia el pasado en búsqueda de señales que permitan predecir cambios en el futuro, en base a los acontecimientos en el pasado existe una gran probabilidad de predecir el comportamiento de los inversionistas en el futuro. El análisis técnico se basa en dos principios fundamentales, el primero indica que toda la información de una empresa, se encuentra ya reflejada en sus cotizaciones del pasado, el segundo principio menciona que las inversiones tienden a seguir un rumbo, si se encuentran al alza van a seguir al alza, si se encuentran en caída van a seguir en caída (Malkiel, 2007; Garcia Roldan, 2015). El análisis técnico plantea el gran valor que tiene el estudio de acontecimientos pasado de las cotizaciones de una empresa, para poder estimar o

predecir cómo será el comportamiento de la misma en un futuro, este principio es de gran importancia en el presente proyecto, que trata de realizar predicciones del comportamiento de las cotizaciones en la bolsa de valores, un buen punto de partida es un análisis del comportamiento pasado.

El *análisis fundamental*, por otro lado, remarca la diferencia entre el valor presente y el valor verdadero de la inversión, para determinar a ciencia cierta el valor verdadero de una inversión, se deben realizar estimaciones de los futuros beneficios y dividendos de la empresa, haciéndose base en cuatro determinantes: tasa de crecimiento esperada, expectativa de reparto de dividendos, el grado de riesgo y el interés en el mercado. Mientras que el valor presente es el valor actual de todo el flujo de efectivo que el inversor espera recibir de las inversiones (Malkiel, 2007). El análisis fundamental se hace base en las teorías financieras o económicas, para poder determinar las fuerzas de oferta y demanda. En el análisis fundamental se realiza la revisión de indicadores macroeconómicos, los mercados de valores y las condiciones políticas (Morales Sierra, 2011). De acuerdo al análisis fundamental una inversión posee dos valores: el valor verdadero y el valor presente, estos no deben ser confundidos y se los debe manejar como valores completamente independientes. El valor verdadero de la inversión es lo que representara dicha inversión en cuanto a los beneficios y dividendos de la empresa, mientras que el valor presente es lo que espera recibir financieramente hablando, el actual dueño de dicha inversión.

Al igual que las metodologías anteriormente descritas, existen también sistemas técnicos que permiten **predecir el comportamiento futuro del mercado**, entre los cuales están:

El *sistema de filtros*, el principio fundamental de este sistema hoy en día es utilizado en lo que se conoce como: limitar las pérdidas, la teoría menciona que una cotización que ha alcanzado su punto más bajo acto seguido sube, se puede decir esta cotización encuentra a la alza; Una acción que ha alcanzado su cotización más alta, acto seguido baja, quiere decir que la cotización está a la baja (Malkiel, 2007). Teniendo en cuenta este sistema, se puede tener una predicción del comportamiento a futuro en caso de que una cotización alcance su punto máximo o mínimo y acto seguido se dé un cambio en la tendencia, lo más probable es que ese cambio en la tendencia

continúe, ayudando al inversionista en la toma de decisiones en cuanto a compra o venta.

La *teoría de Dow*, esta teoría es considerada el origen del análisis técnico, formulada por Charles H. Dow ayuda a detectar señales de compra-venta a partir de información proporcionada por el mismo mercado (De Miguel Casas, 2016). menciona que cuando el mercado toca techo y hay un cambio de tendencia a la baja, el punto en el cual el mercado toco techo es conocido como área de resistencia, en caso de que el mercado vuelva a estar al alza, pero su techo actual sea inferior al techo anterior, indica que el mercado está a la baja, en caso que alcance el área de resistencia anterior y logre subir un poco más, la área de resistencia pasa a ser área de soporte, y es muy probable el que siga con la tendencia al alza (Malkiel, 2007).

Al igual que con la teoría de los filtros, la teoría de Dow permite tener un claro panorama de las posibles conductas que tome el mercado, basándose en el comportamiento pasado, al igual que la teoría de los filtros claramente esta teoría es de gran utilidad para la toma de decisiones.

El *sistema de precio-volumen*, menciona que cuando la cotización de una acción sube con un alto volumen de contratación o en su defecto en aumento, se puede esperar que la cotización siga en alza. Mientras que cuando la cotización cae con un gran volumen de contratación, indica una presión de venta, provocando una señal a los demás inversionistas de vender (Malkiel, 2007). El sistema de precio-volumen es de gran importancia, dado que permite tener un panorama actual del mercado, de lo que realmente está pasando y de lo que es muy probable que suceda en un futuro. En el caso de que el mercado suba su cotización y un número elevado de inversionistas estén interesados en comprar, podrá generar una demanda insatisfecha, haciendo que el mercado incremente aún más su cotización, en el caso que la cotización caiga, con un gran volumen de inversores comprando, provocara que la cotización del mercado caiga aún más, haciendo que la cantidad de inversionistas a la venta incrementen, provocando una caída en el mercado.

La *lectura de gráficos* es uno de los sistemas de mayor importancia, que ayuda a la hora de elaborar producción o estimaciones en cuanto al comportamiento de

acciones se refiere, siendo este de gran importancia requiere un estudio a profundidad. El sistema de lectura de gráficos revela los movimientos en el precio de las acciones, en muchos casos esos movimientos generan figuras de cambio que los analistas pueden interpretar y sugerir a los inversionistas para la toma de decisiones.

El chartismo consiste en el estudio de los gráficos, el chartista utiliza los gráficos para su análisis, basándose únicamente en el precio, las cotizaciones de los precios de un gráfico se mueven por tendencias, existe una tendencia cuando durante un periodo de tiempo las cotizaciones de un activo tienden a seguir una dirección, el poder determinar la tendencia es uno de los aspectos más importantes para poder invertir (Gallofré, 2014).

En el entorno bursátil el gráfico que mayormente se utiliza es el de *velas japonesas*, debido a su fácil interpretación y la gran cantidad de información de datos de alta significancia que se puede representar, este tipo de gráfico permite observar claramente la constante presión que existe entre los compradores y vendedores.

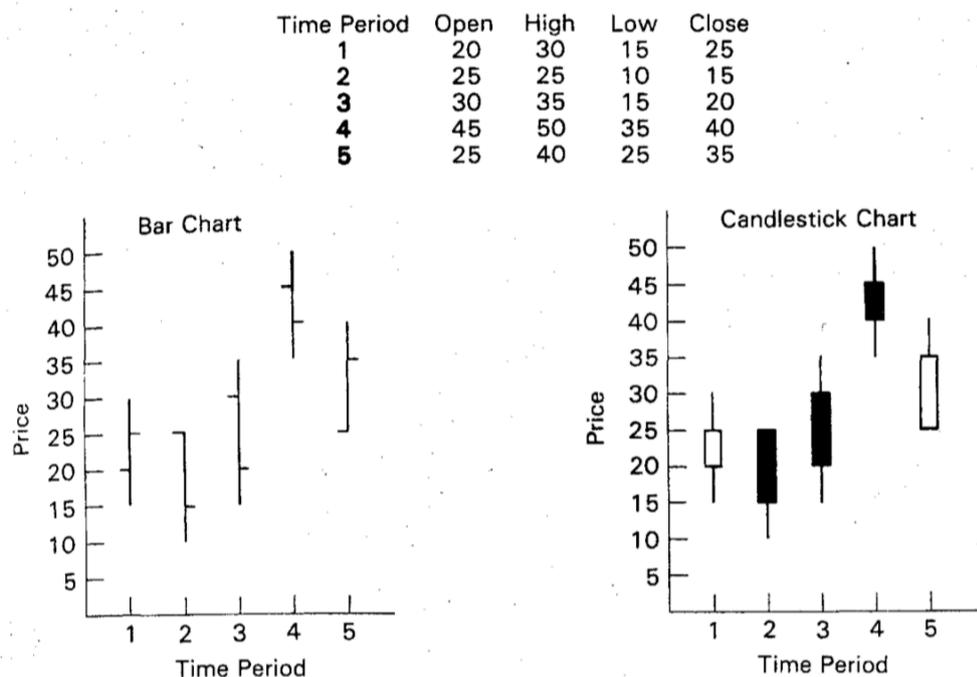


Figura 1. Gráfica de Barra y Velas Japonesas *Nota:* Tomado de Nison (1991). Japanese Candlestick Charting Techniques: Investment Techniques of the far East (p. 21). Nueva York.

En los gráficos de velas japonesas, las velas son representaciones del precio, en caso de que vela tenga un color rojo o negro quiere decir que el precio cayó en ese momento, en el caso de que el cuerpo de la vela tenga un color blanco o verde quiere decir que el precio en ese momento subió. Así como el color de la vela tiene su significado, el cuerpo de la vela también tiene su significado, la longitud del cuerpo de la vela es la diferencia entre el precio de apertura y el precio del cierre (Serrano Ruiz, 2013). El gráfico de velas japonesas permite conocer a profundidad, que es lo que está sucediendo detrás de las cifras de las cotizaciones del mercado, conocer el comportamiento de los inversionistas tanto de los que compran como de los que vende, en el caso del tamaño de las velas, en caso de que el cuerpo de la vela sea grande indica que existe un gran movimiento detrás de ese mercado, en caso de que el cuerpo de la vela sea pequeña indica que el movimiento en ese mercado tuvo muy poca actividad, así como el color permite conocer de manera rápida si el mercado se encuentra en alza o en caída.



Figura 2. Ilustración de vela cuerpo negra y vela blanca. *Nota:* Tomado de Nison (1991). Japanese Candlestick Charting Techniques: Investment Techniques of the far East (p. 24). Nueva York.

En el que caso que de la vela no posea cuerpo, indica que existió un equilibrio entre los que compran y los que venden, a estas velas se las conoce como Doji (Serrano Ruiz, 2013) . Ante la existencia de una vela Doji resulta complejo la toma de decisiones ya que en cualquier momento la tendencia podría inclinarse tanto al alza como a la baja, en el caso de venir de una tendencia alcista se observa un Doji, lo probable es que se presente una tendencia a la caída. El Doji refleja, la indecisión del mercado, la cual a una escala mayor refleja la acción lateral del mercado, un Doji que aparece en

la parte madura de una tendencia alcista o de una bajista puede mostrar unas altas probabilidades de cambio en la tendencia del mercado (Nison, 1996).

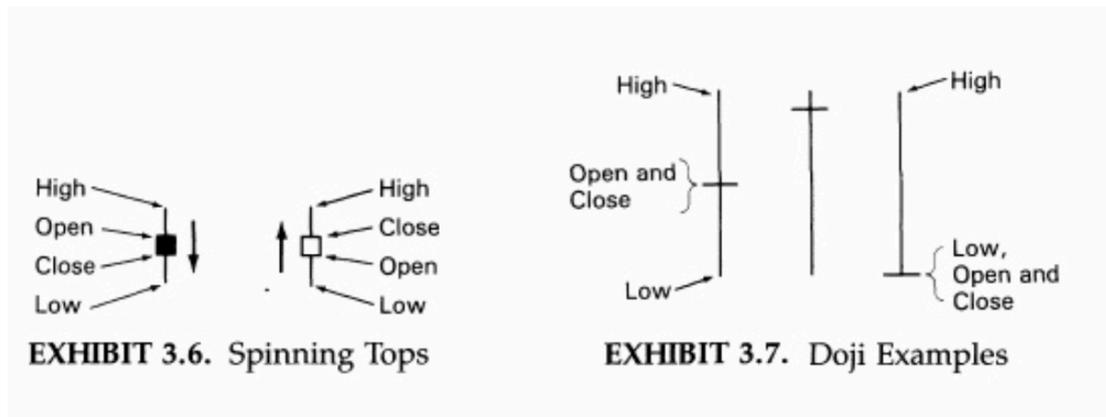


Figura 3. Ilustración de vela cuerpo negra y vela blanca. *Nota:* Tomado de Nison (1991). Japanese Candlestick Charting Techniques: Investment Techniques of the far East (p. 24). Nueva York.

Las formas de las velas tienen su propio significado y pueden ser utilizados a manera de indicadores de cambio de una tendencia. La vela con forma de martillo y la vela con forma de hombre colgado, estos tipos de vela poseen un cuerpo inferior largo, una sombra superior corta o ausente, la vela con forma de martillo tiene su cuerpo verde o blanco, mientras la vela con forma de hombre colgado tiene su cuerpo de color rojo o negro (Serrano Ruiz, 2013). La vela en forma de martillo es entendida como un signo alcista, el termino martillo se deriva de que un mínimo esta tan sólido que no se puede romper, para que el martillo tenga sentido debe de encontrarse después de un giro a la baja. La vela con forma de hombre colgado es una señal de cambio, la cual se produce durante una subida, esta es una señal de venta, caso contrario con el martillo la cual es una señal de compra (Nison, 1996).

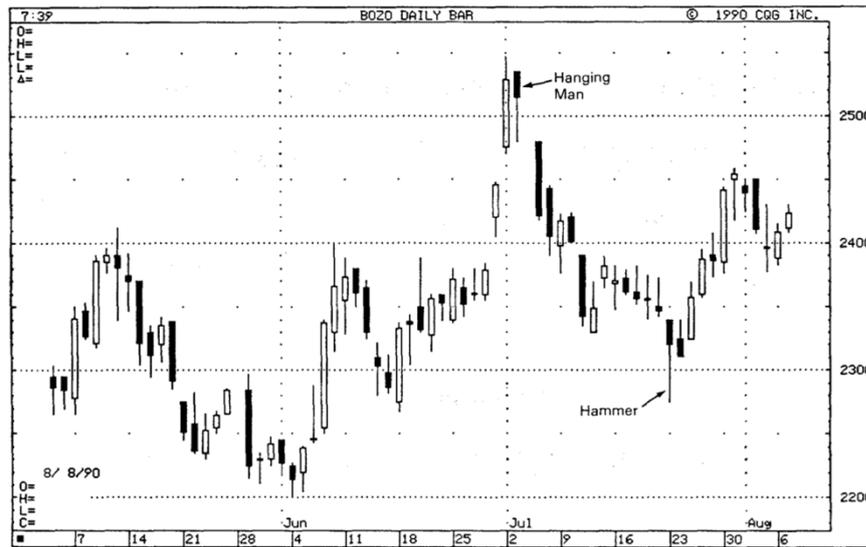


Figura 4. Ilustración de Hombre colgado y Martillo. *Nota:* Tomado de Nison (1991). Japanese Candlestick Charting Techniques: Investment Techniques of the far East (p. 30). Nueva York.

La vela con forma de martillo invertido y la vela con forma de estrella fugaz, estos tipos de vela posee un cuerpo pequeño, con una sombra superior larga y una sombra inferior corta o ausente, la única diferencia es en el color del cuerpo la vela estrella fugaz posee un cuerpo color rojo o negro, mientras que la vela de martillo invertido posee un cuerpo color blanco o verde (Serrano Ruiz, 2013).

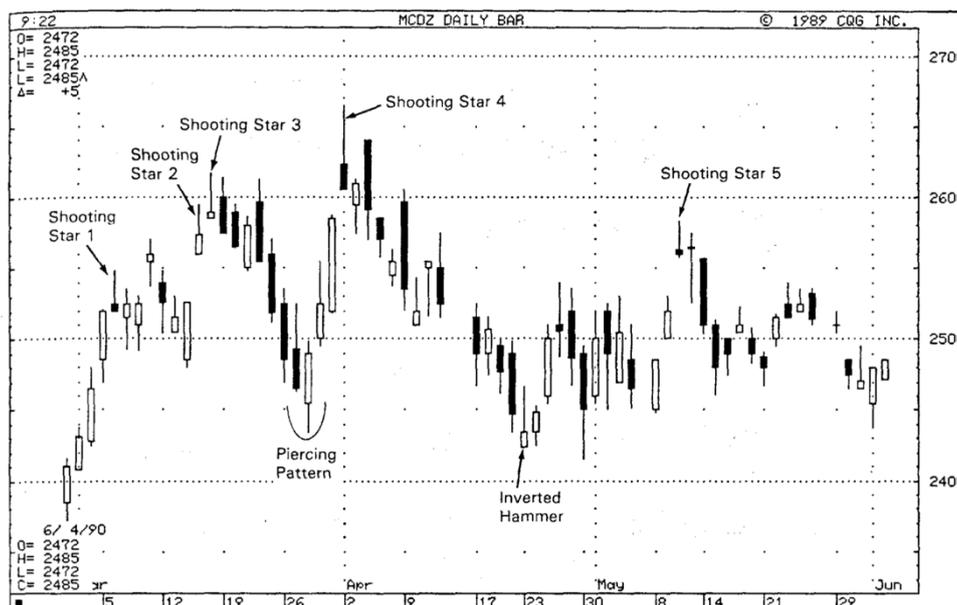


Figura 5. Ilustración de Estrella Fugaz y Martillo Invertido. Nota: Tomado de Nison (1991). Japanese Candlestick Charting Techniques: Investment Techniques of the far East (p. 77). Nueva York.

Los distintos tipo de velas permiten a los analistas conocer el comportamiento de las tendencias, prever el cambio o la continuidad en la tendencia, tal es el caso de la vela martillo, en caso de presentarse al final de una tendencia bajista, indica que el alza de la cotización el mercado está cerca, en el caso de la vela con forma de hombre colgado, en caso de presentarse luego de una tendencia alcista, es muy probable que se acerque un descenso en los precios, en el caso de la vela con forma de martillo invertido, en caso de que se presente luego de una caída en los precios, existe una gran probabilidad que los precios tengan un alza, por lo contrario la vela con forma de estrella fugaz, en el caso de presentarse luego de una tendencia al alza, lo más probable es que los precios sufran una caída.

Por consiguiente, todas estas formas de vela sirven a manera de indicador, pudiendo predecir con cierto grado de seguridad o riesgo o un cambio en la tendencia, brindando al inversionista una herramienta sólida en la toma de decisiones.

En el caso particular de los inversionistas, el *riesgo*, es un elemento financiero que se debe de tener muy en cuenta a la hora de tomar una decisión en cuanto a inversión se refiere, la American Heritage Dictionary define al riesgo como la

posibilidad de sufrir daño o pérdida, financieramente hablando el riesgo es la posibilidad de que no se materialice el rendimiento esperado de los valores y de que dichos valores en un futuro puedan llegar a precios inferiores al que fueron adquiridos. También se suele definir al riesgo como la varianza o la desviación típica. Es decir, una inversión cuyos rendimientos de un año a otro varíen notablemente, se puede decir que es una inversión arriesgada. En caso de que los rendimientos no varíen, se dice que no es arriesgado. En caso de que los rendimientos no se desvíen mucho, se considera un riesgo medio. Es importante definir el termino riesgo ya que un inversor que está dispuesta a correr con un alto grado de riesgo espera que sus beneficios en caso de existir sean proporcionales a el nivel de riesgo que asume, influyendo de manera directa en las cotizaciones del mercado.

Los inversores pueden ser clasificados de acuerdo al nivel de riesgo que estos están dispuestos a asumir. Los inversores conservadores, son cuyo perfil de riesgo es cero. El inversor moderado es aquel que está dispuesto a asumir riesgos, sabiendo que sin riesgo no hay ganancia. El inversor agresivo es aquel que utiliza recursos propios así como recursos provenientes del préstamo, utilizando lo que se conoce como apalancamiento, este tipo de inversor agresivo pueden entrar minutos o segundos al mercado, pueden también realizar operaciones intradiarias, o en pocos días realizan sus transacciones (Serrano Ruiz, 2015). Es importante tener conocimiento de los distintos tipos de inversores que existen en el mercado ya que son estos los que determinan los precios y las conductas del mercado, hoy en día esta es apogeo lo que se conoce como trading, este tipo de inversor está dentro de los agresivos, se mantiene en el mercado por muy poco tiempo realizando operaciones de pocos días hasta intradiarias, todo esto es posibles gracias a la existencia de nuevas tecnologías que al final del día afecta la manera como los inversores trabajan así como influye en el comportamiento del mercado en general.

2.1.2 Minería de Datos

Dados los avances tecnológicos y las nuevas tendencias en el mundo informático y los contantes esfuerzos por parte de los analistas para poder predecir de manera acertada el comportamiento futuro del mercado de valores, tomando como base la data histórica de su comportamiento. La revolución digital ha hecho posible

que la información digitalizada sea fácil de capturar, procesar, almacenar, distribuir, y transmitir, a la hora de realizar predicciones sobre el comportamiento del mercado bursátil, se hallara un inmenso volumen de data, lo verdaderamente importante estará en darle un sentido a toda esta data, en la capacidad que se tenga para poder extraer información útil que pueda ayudar en la toma de decisiones, exploración, la comprensión de algún fenómeno, es ahí donde entra la minería de datos (MD).

Antes de entrar a hablar más a profundidad sobre minería de datos, es necesario dejar bien el claro el término **descubrimiento de conocimiento en base de datos** (KDD, por sus siglas en inglés de Knowledge Discovery in Databases), acuñado en 1989, hace referencia a todo el proceso de extracción de conocimiento a partir de una base de datos y marca un cambio de paradigma, en cual lo realmente importante es el conocimiento útil que se pueda obtener a partir de los datos (Riquelme, Ruiz, & Gilbert, 2006). KDD es la convergencia del aprendizaje automático, la estadística, el reconocimiento de patrones, la inteligencia artificial, las bases de datos, la visualización de datos, los sistemas de apoyo de toma de decisiones, la recuperación de información y otros muchos campos (Beltran Martinez, 2015).

Los términos MD y KDD son erróneamente confundidos como sinónimos, cabe aclarar que la MD (minería de datos) es un paso más dentro de KDD. A parte de MD el KDD cuenta con otros pasos de gran importancia estos son:

1. **Comprender el dominio de aplicación.** - En este paso se incluye la obtención de conocimiento previo y los objetivos de la aplicación.
2. **Extraer la base de datos objetivo.** - En este paso es necesaria la obtención de los datos, realizar la evaluación de estos y realizar un análisis exploratorio de los mismos, con el fin de familiarizarse con ellos.
3. **Preparar datos.** - En este paso se realiza la transformación, reducción e integración de los datos. En este paso se prepara la información con la finalidad de mejorar la calidad de esta, a la vez que se reduce el tiempo que tomara el algoritmo para el aprendizaje.
4. **Minería de datos.** - Esta es la etapa más importante del KDD. Esta etapa está conformada por las siguientes funciones: clasificación, regresión, clustering, resumen, recuperación imágenes, extracción de reglas.

5. **Interpretación.** - En esta etapa se trata de buscarle una explicación a los patrones descubiertos, así como a la posibilidad de visualizarlos.
6. **Uso del conocimiento descubierto.** - Durante esta fase se realizan pruebas con el modelo, es decir se pone en práctica el modelo obtenido.

Por otro lado, la MD es la construcción de un modelo que, al ser ajustado a ciertos datos, proporciona un conocimiento. En la MD se puede distinguir dos pasos principales, el primero es la elección del modelo, el segundo es el ajuste del modelo a los datos (Riquelme et al., 2006). La disciplina de la MD estudia métodos y algoritmos los cuales permiten la extracción automática de información sintetizada, permitiendo el descubrimiento de relaciones escondidas en la gran cantidad de datos, la información obtenida a partir de la minería de datos posee capacidad predictiva, lo cual facilita el análisis de los datos de forma eficiente (Beltran Martinez, 2015).

El objetivo principal de la MD es el de predecir salidas y revelar relaciones en los datos, para lograr esos objetivos, se hace uso de herramientas automáticas, estas herramientas hacen uso de complejos algoritmos con el propósito de descubrir patrones, que no pueden ser detectados a simple vista.

La minería de datos es una tecnología que está compuesta por varias etapas, las cuales integran varias áreas, por lo cual durante el desarrollo de un proyecto de minería de datos se utilizan diferentes software en cada una de estas etapas: estadísticos, visualización de datos o inteligencia artificial (Vallejos, 2006).

En los inicios, la primera ciencia encargada de extraer información de los datos, era la estadística, utilizando metodologías netamente matemáticas, con el avanzar del tiempo y el desarrollo tecnológico, el ordenador juega un papel clave, dando espacio al nacimiento del concepto que hoy se conoce como Machine Learning o aprendizaje automático. Dado el incremento significativo de los datos y la estructuración de los mismos, aparece concepto de minería de datos.

Dada la complejidad que posee la MD, está compuesta por una serie de tareas, estas son:

- **Agrupamiento.** - Es una tarea en donde se busca identificar un conjunto de categorías que permita describir los datos.
- **Clasificación.** - Durante esta tarea se busca adquirir una función que sea capaz de mapear un elemento de datos a una de varias clases predefinidas.
- **Regresión.** - En esta tarea se busca adquirir una función que mapee un elemento de dato a una variable de predicción de valor real.
- **Modelado de dependencias.** - En esta tarea se busca encontrar un modelo el cual pueda describir las dependencias significantes entre variables.
- **Detección de desviaciones, casos extremos o anomalías.** - Esta tarea permite detectar los cambios más significativos en los datos con respecto a valores pasados o considerados normales.
- **Condensación.** - Esta tarea consisten en encontrar un método que permita hallar una descripción simplificada de un subconjunto de datos.

El proceso de MD comienza con la identificación de los datos, saber qué datos son los que se necesita, donde se los puede encontrar y como se los puede conseguir. Una vez obtenidos los datos, deben ser colocados dentro de una base de datos con el formato adecuado, paso seguido se debe realizar una selección de los datos esenciales y eliminación de los innecesarios. Antes de poder realizar el análisis de los datos por medio de las técnicas de Data Mining, se debe tener claro el propósito de la investigación, herramientas a utilizar y como se va a proceder. Una vez que se ha puesto en práctica las herramientas, se debe saber interpretar los resultados o patrones obtenidos, para saber cuáles son significativos y como se puede extraer únicamente los resultados que sean útiles. Los datos útiles que han sido extraídos, deben ser analizados para determinar qué acciones o decisiones deben ser ejecutadas (Riquelme et al., 2006). El *proceso* de MD pasa por los siguientes estados:

- **Procesado de datos.** - En la mayoría de los casos no es posible aplicar una técnica de minería de datos sobre los datos brutos, por lo cual se vuelve necesario el preprocesado, en el cual se filtran los datos, se eliminan los

valores incorrectos, no válidos, desconocidos; también se puede obtener una muestra de los datos o en su defecto reducir el número de valores posibles, mediante el redondeo o el clustering.

- **Selección de características.** - Aun cuando los datos han sido pre procesados, la cantidad de datos es abundante, por lo cual resulta necesario reducir el tamaño de los datos eligiendo las variables más influyentes en el problema, de esta manera la calidad del modelo no se ve sacrificada.
- **Algoritmo de aprendizaje.** - Mediante las técnicas de minería de datos, se obtiene un modelo de conocimiento, el cual representa los patrones de comportamiento observados en los valores de las variables del problema o en las relaciones de asociación entre las variables.
- **Evaluación y validación.** - Con el modelo ya obtenido, se debe de realizar la validación del mismo, comprobando de las predicciones o resultados obtenidos son válidos y satisfactorios. En caso de que se hayan obtenidos varios modelos, es necesario que sean comparados en búsqueda de aquel que mejor se ajuste al problema.

Los *métodos* de MD tienen como metas primarias la predicción de datos desconocidos y la descripción de patrones. Entre los principales métodos de minería de datos están:

- **Agrupamiento.** - Este método permite la identificación de grupos en los cuales los elementos poseen similitud entre sí.
- **Asociación.** - Este método establece las posibles relaciones o correlaciones entre distintas acciones o sucesos que en un principio parecen ser independientes, pero la aparición de ciertos sucesos puede generar la aparición de otros.
- **Secuenciamiento.** - Este método permite identificar como a lo largo del tiempo la ocurrencia de ciertas acciones provoca la ocurrencia de otras acciones, es muy parecida a la asociación, pero en este caso el tiempo juega un rol muy importante.

- **Reconocimiento de patrones.** - Este método permite la identificación de problemas e incidencias y de las posibles soluciones, haciendo uso de técnicas de redes neuronales y algoritmos matemáticos.
- **Forecasting.** - Este método establece el futuro más probable dependiendo de la evolución pasada y presente. Esta tiene un uso fundamental en el tratamiento de series temporales, se hacen fuerte en el uso de información histórica.
- **Simulación.** - Permite la generación de múltiples escenarios o posibilidades sujetas a unas reglas o esquemas, permitiendo analizar todas las posibles variaciones o alternativas de una decisión o situación.
- **Optimización.** - Esta metodología es capaz de resolver el problema de minimización y maximización.
- **Clasificación.** - Este método permite asignar a un elemento la pertenencia a un grupo o clase.

La aplicación de algoritmos de minería de datos requiere que se apliquen previamente una serie de actividades para la preparación de los datos, en muchos casos los datos provienen de diversas fuentes, no se encuentran con el formato debido o poseen ruido. Las *técnicas* más comunes de MD son:

- **Métodos Estadísticos.** - La estadística es la ciencia que tradicionalmente se ha utilizado para el tratamiento de grandes volúmenes de datos numéricos, esta metodología posee una amplia gama de modelos de análisis. Entre los modelos estadísticos más utilizados están: ANOVA, ji cuadrado, componentes principales, análisis de clusters, análisis discriminante, regresión lineal, regresión logística.
- **Métodos Basados en Árboles de Decisión.** - Utilizado para el descubrimiento de reglas y relaciones mediante la ruptura y subdivisión sistemática de la información del conjunto de datos.
- **Reglas de Asociación.** - Extraen información por coincidencias, permite descubrir correlaciones o co-ocurrencias en los sucesos de la base de datos, estas se establecen a manera de condiciones.

- **Redes Neuronales.** - Son una manera de analizar la información, capaces de detectar patrones y características dentro de los datos. Se comportan de forma parecida al cerebro, son capaces de aprender de las experiencias y del pasado, dicho conocimiento es aplicado en la resolución de problemas, una vez que las redes neuronales han sido entrenadas son capaces de hacer previsiones.
- **Algoritmos Genéticos.** - Estos algoritmos hacen uso de las técnicas biológicas de reproducción para utilizarlas en problemas de tipo búsqueda y optimización. Se comienza con una población de partida, la cual se va optimizando para que se ajuste a la solución del problema. Esta herramienta es utilizada en las primeras etapas de data mining.
- **Lógica Difusa.** - Esta técnica surge de la necesidad de modelizar la realidad, para poder evitar el determinismo o la exactitud. Permite y trata la existencia de barreras difusas o suaves entre los distintos grupos.
- **Series Temporales.** - Es el estudio de una variable a través del tiempo, a partir de este conocimiento y bajo el supuesto que no se producirán cambios estructurales, se podrán realizar predicciones. Por lo general se basan en el estudio de la serie en ciclos, tendencias y estacionalidades. Esta metodología puede ser combinada con otros métodos, dando lugar que la serie no solamente pueda ser explicada en función al tiempo sino en función a otras variables más.

Es importante hacer hincapié en las *redes neuronales artificiales (ANN*, por sus siglas en ingles), dado que es una de las áreas en las cuales hoy por hoy refleja más avances y está en constante evolución, dado que gran parte de la comunidad del software se ha volcado a desarrollar proyectos en esta área en específico, con la gran ventaja de que muchos de estos proyectos son de código libre, tal es el caso de Keras, la cual es una biblioteca de redes neuronales de código abierto escrita en Python . Una RNA es un modelo matemático el cual se encuentra compuesto por un gran número de elementos procesales organizados en niveles. De las RNA se puede destacar las siguientes ventajas: aprendizaje adaptivo, auto-organización, tolerancia a fallos, operación en tiempo real y que son de fácil inserción en tecnologías ya existentes.

Una RNA está constituida por neuronas las cuales se encuentra interconectadas. Los datos ingresan por una capa de entrada, pasan por lo que se denomina como capa oculta, de manera general la capa oculta a su vez está conformada por más capas y finalmente los datos salen por la capa de salida (Matich, 2001).

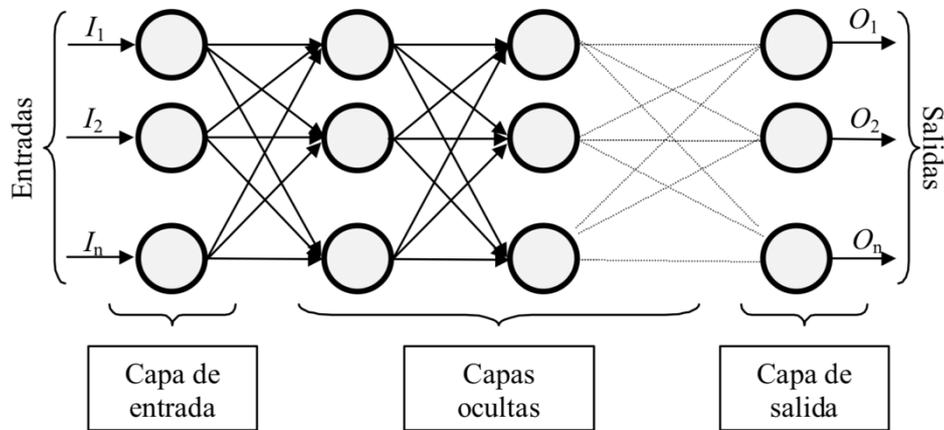


Figura 6. Ejemplo de una red neuronal. *Nota:* Tomado de Matich (2001). Redes Neuronales: Conceptos Básicos y Aplicaciones. (p. 12). Rosario.

Una red neuronal funciona de la siguiente manera: Para cada vector de entrada, este es introducido en la red, de manera que se copia cada uno de los valores que conforman este vector en la célula de entrada correspondiente (neurona), cada una de estas células procesa este valor y genera una salida, dicha salida es transmitida a través de las conexiones entre células para poder llegar a la célula destino, este valor se convierte en valor de entrada para esta célula, una vez que la entrada se ha dispersado a lo largo de toda la red se producirá un vector resultante, conformado por cada uno de los valores de salida de las células de salida (Isasi Viñuela & Galván León, 2004).

Existen dos tipos de problemas básicos que pueden ser resueltos por una RNA, problemas de clasificación y problemas de regresión. Los problemas de clasificación consisten en que crear un procedimiento en el cual se le presenta a la RNA un nuevo caso con atributos o cualidades, los cuales son los datos de entrada, y se lo clasifique

dentro de una serie de clases ya predefinidas. En el caso de los problemas de regresión, se trata de obtener un número en función de los atributos de entrada de la red, dicho en otras palabras encontrar un función continua de ciertas variables, un caso muy claro de este tipo de problemas es realizar una predicción (Palma Ménendez & Marin Morales, 2008).

La parte más importante de una RNA es el aprendizaje, el esquema de aprendizaje de la red neural es lo que determina qué tipo de problemas será capaz de resolver, dado que las RNA aprenden a partir de ejemplos, de esta manera la capacidad de una red neural de poder resolver un problema esta sumamente ligada a la calidad y tipo de ejemplos que le sean proporcionadas en la etapa de aprendizaje o entrenamiento. Por lo cual el aprendizaje debe ser: Significativo, debe de haber el número suficiente de ejemplos, caso contrario la red no será capaz de adaptar sus pesos de manera eficaz. Representativo, los componentes del conjunto de aprendizaje deben ser diversos, en caso de que a la RNA se le proporcionen muchos más ejemplos de un tipo que de otro, harán que la RNA se especialice en este tipo de ejemplos (Isasi Viñuela & Galván León, 2004).

Un tipo de RNA que son comúnmente utilizadas para el caso de predicción de series temporales son las Redes Neuronales Recurrentes (RNN, por sus siglas en ingles). Los problemas de predicción de series temporales, pueden abordarse con estructuras dinámicas de redes recurrentes, las cuales se caracterizan que la propia arquitectura de la red está diseñada para el procesamiento de información temporal, permitiendo que la salida de la red no solo dependa de la entrada, sino de estados anteriores de la propia red. Las RNN, se caracterizan por que crean bucles en las neuronas de la red mediante el uso de conexiones recurrentes, estas conexiones de tipo cíclicas las vuelven más robustas a la hora de modelas secuencias de datos continuos, permitiendo que aparezcan dentro de la red: conexiones de una neurona a ella misma, conexiones entre neuronas de la misma capa o conexiones de una capa a la capa anterior. En las RNN, la activación de una neurona con conexiones recurrentes, no depende solamente de las activaciones de las neuronas de la capa anterior, también depende de la activación de cualquier neurona conectado a ella e incluso de su propia activación (Isasi Viñuela & Galván León, 2004). Las RNN son las que cuyo

funcionamiento más se acerca a la forma de pensar del ser humano, estas redes habilitan conexiones entre todas las neuronas, lo que permite añadir el concepto de temporalidad, dotándolas de memoria, las RNN han podido solventar deficiencias de las redes tradicionales, haciéndolas uno de los pilares de lo que se conoce como Deep Learning.

En particular las LSTM (Long Short-Term Memory), son un tipo de RNN que fueron diseñadas específicamente para modelar secuencias temporales con mayor precisión que las RNN convencionales, dado que a medida que el espacio entre la información requerida y la tarea actual crece, a las RNN les toma más tiempo poder aprender a conectar dicha información, volviéndose incapaces de completar el aprendizaje, aparecen las LSTM, las cuales son capaces de solventar esta problemática dado que incorporan dentro de su arquitectura memoria a corto y largo plazo (González Muñiz, 2018).

El manejo de los datos para este proyecto en particular se da en series temporales, dada su naturaleza; Una serie temporal está definida como la secuencia ordenada de observaciones cada una de las cuales se encuentra asociada a cierto momento en el tiempo, cabe indicar que el orden en una serie temporal es de suma importancia y que en general las observaciones son dependientes entre si (González Casimiro, 2009). El análisis de series temporales, es asociado al descubrimiento y uso de patrones como: periodicidad, estacionalidad o ciclos, la predicción de valores futuros. El análisis clásico de series temporales comprende el estudio de cuatro componentes básicos, dichos componentes básicos, descritos en la literatura son:

- **Tendencia.** - Componente que refleja la evolución a largo plazo del fenómeno observado.
- **Variación estacional.** - Es el movimiento periódico de corto plazo, componente que refleja la influencia de fenómenos que ocurren cada cierto tiempo, se repiten de manera periódica.
- **Variación cíclica.** - Al igual que la variación estacional existen variaciones que se repiten cada cierto tiempo de mayor longitud.

- **Variación aleatoria.** - Denominada como residuo, no muestra regularidad, obtenida una vez eliminadas las tendencias y las variaciones cíclicas de la serie.

2.2 Conceptos

Entre los términos más utilizados en esta investigación, cabe algunos de gran importancia que sirvieron de base para el diseño del modelo de predicción. A continuación, se hace referencia a algunos de estos términos.

La **bolsa de valores** es una organización privada o pública que permite a sus componentes introducir órdenes y realizar negociaciones de compra y venta de valores. La negociación de los valores en los mercados bursátiles se hace tomando precios conocidos en tiempo real en un entorno seguro para la actividad de los inversionistas. Las transacciones están totalmente reguladas, proporcionando legalidad y seguridad (Serrano Ruiz, 2013). La bolsa dependiendo de sus regulaciones, permite que los mercados bursátiles, permitan que las empresas sean capaces de financiar sus proyectos e inversiones a través de la venta de activos o títulos y que los inversionistas sean capaces de adquirir dichos activos o títulos. La bolsa cumple funciones vitales dentro del sistema financiero de un país: promueve el mercado de valores, propicia eficiente canalización del ahorro nacional hacia entidades que requieran recursos financieros adicionales, propicia al inversionista condiciones de legalidad y seguridad, adopta medidas que le brindan confianza al sistema, lleva el completo registro de todas las negociaciones, la difusión de información (Puentes Pérez, 2015).

Por otro lado, es necesario analizar el término **índice bursátil** que bien puede ser definido como la ponderación matemática de un conjunto de valores, que cotizan en un mismo mercado. Con el índice se pretende medir el crecimiento o la pérdida del mismo sobre los valores que lo componen. Los índices más importantes a nivel mundial son: FTSE 100(Gran Bretaña), DAX 30 (Alemania), CAC 40(Francia), IBEX 35(España), EUROSTOXX 50 (Europa zona euro), STOXX (Europa), Dow Jones (EE.UU.), Nasdaq 100(EE.UU.), Bovespa (Brasil), Merval (Argentina Standard), Poor's 500 (EE.UU.), Nikkei 225(Japón), Shanghai Composite (China).

Por otro lado, y como elemento importante que se maneja en una bolsa de valores, están las **acciones** que se constituyen en la división del capital de una empresa y la persona que poseen acciones de dicha empresa es un accionista. Una acción es un título de propiedad el cual corresponde a una fracción de capital de una empresa, dicho valor es variable, dado que depende de los beneficios de la empresa, las acciones pueden ser emitidas en cualquier momento por las empresas (Puentes Pérez, 2015). Todas tienen el mismo valor individual, aunque la empresa podría aumentar el número de acciones mediante ampliación de capital o podría disminuirlas con amortizaciones de cartera (Serrano Ruiz, 2013).

Otros términos importantes son el soporte y la resistencia. El **soporte** es un nivel de precios dentro de una tendencia bajista, en la cual el precio detiene su caída por un incremento del número de compradores, en casos se puede ocasionar un rebote en los precios a manera de una tendencia alcista (Caicedo Pulido & Gaviria Franco, 2010). Si el precio del valor en el que se opera ha formado una tendencia a la baja, el soporte será la línea imaginaria horizontal con la que choque; mientras que la **resistencia**, que es lo contrario al soporte, en donde el precio del valor encuentra su techo, es decir que el precio choque dos veces contra una línea horizontal imaginaria, que se trazara desde ese momento, la resistencia, en caso de tener una tendencia alcista, se puede llegar a concentrar tal cantidad de ofertantes como para frenar la tendencia alcista (Caicedo Pulido & Gaviria Franco, 2010). La rotura de la resistencia provocará una subida del precio del valor. Cuantos más momentos la línea de resistencia ha sido tocada sin que se rompa, más poderosa y fuerte se vuelve, evitando así la subida del precio del valor. Cuando un valor rompe una resistencia y no hay otra por encima, se encuentra en subida libre (Serrano Ruiz, 2013).

La situación en la que se encuentran los máximos y mínimos crecientes en el tramo alcista, y máximos y mínimos decrecientes en el tramo bajista es la denominada **tendencia**. La tendencia es la dirección del mercado, en la dirección en la cual se mueve el mercado, dado que el mercado no se mueve en línea recta, el mercado por lo general se mueve en zigzag, la dirección en la cual se mueven estos picos y valles es lo que se denomina como tendencia (Murphy, 2000). Es necesario dos puntos para

trazar una tendencia y tres para confirmarla. Al igual que ocurre con el soporte o con la resistencia, cada vez que el precio toca a la tendencia, esta se hace más fuerte.

El **volumen**, por otro lado, es el intercambio de dinero que se produce en el mercado bursátil entre los que compran y los que venden, cuanto más dinero cambia de manos, más sube o baja el precio de los valores. El volumen también puede ser definido como el número de entidades que han sido objeto de operaciones durante un periodo determinado (Murphy, 2000). Las entradas de volumen muchas veces son señal de que algo va a suceder. Existe una vinculación entre el precio y lo que hace el volumen (Serrano Ruiz, 2013).

En lo que respecta a Redes Neuronales es necesario presentar ciertos conceptos, **Epochs** es el número de veces que el modelo es expuesto a la data de entrenamiento, en cada iteración, el optimizador trata de ajustar los pesos haciendo que la función objetivo sea minimizada. El **Batch Size**, es el número de instancias de entrenamiento son observadas antes de que el optimizador realice una actualización de pesos (Gulli & Pal, 2017). Durante la fase de aprendizaje, las conexiones con la siguiente capa pueden ser limitadas a un subconjunto de neuronas, para de esta manera reducir la cantidad de pesos que se actualizarán, a esto se lo conoce como **Dropout**, esta es una técnica para reducir el sobreajuste del modelo, por lo general la capa de Dropout es colocada justo después de una capa que posee una gran cantidad de neuronas adiestrables, permitiendo setear a cero y así excluir la activación de cierto porcentaje de neuronas de las siguiente capa (Zaccone, Karim, & Menshawy, 2017).

2.3 Base Legal

La bolsa de valores de Nueva York se encuentra regulada por la SEC (Comisión de Bolsa Valores), dado que el presente trabajo puede influir sobre las decisiones que los inversionistas puedan tomar, es importante tener en cuenta ciertas normativas:

El Securities Exchange Act de 1934 contiene varias disposiciones relativas a la manipulación del mercado, siendo las más sobresalientes la Sección 9, la Sección 10 (b) y la Sección 15.

Sección 9: declara fuera de la ley aquellas practicas manipulativas en conexión con el comercio de valores listados en bolsa.

Sección 10 (b): Otorga competencias a la SEC para expedir normas que prohíban conductas manipulativas.

Sección 15 (c): Prohíbe conductas fraudulentas y manipulativas de los bróker-dealers.

La Regla 10b-5 prohíbe conductas fraudulentas en relación con la compra o venta de valores. Esta regla dispone lo siguiente:

“Regla 10b-5: Empleo de Dispositivos Manipulativos y Engañosas. Será ilegal para cualquier persona, directa o indirectamente, mediante el uso de cualquier medio o instrumento de comercio interestatal, o de correos o de cualquier facilidad de cualquier bolsa nacional de valores,

a) Emplear cualquier dispositivo, esquema o artificio para defraudar,

b) Hacer cualquier declaración falsa de un hecho material u omitir declarar un hecho material necesario para hacer que las declaraciones hechas, bajo la luz de las circunstancias bajo las cuales fueron hechas, no sean engañosas, o

c) Comprometerse en cualquier acto, practica o curso de negocios que opere como un fraude o engaño frente a cualquier persona, en conexión con la compra o venta de cualquier valor.”

La Regla 10b-5 es una de las más importantes reglas anti-fraude dentro de la regulación del mercado de valores de los Estados Unidos y de su texto, se pueden identificar las características más sobresalientes de la misma:

La Regla aplica a “cualquier valor”, esto es, valores que se transen en forma pública o privada, sujetos a registro o exentos del mismo.

La Regla aplica en relación con la “compra o venta” de cualquier valor. Esto es que no solo puede cometer fraude bajo esta regla el emisor o aquel que venda el valor sino también aquella persona o inversionista que lo compre.

La Regla aplica para “cualquier persona, en conexión” con la compra o venta de cualquier valor. Lo anterior significa que cualquier persona cuya conducta haya afectado la transacción puede ser sujeta a esta regla, a pesar de no estar involucrada directamente o hacer parte de la transacción.

En el literal c) se establece que es ilegal para cualquier persona, directa o indirectamente, comprometerse en cualquier acto, práctica o curso de negocios que opere como un fraude o engaño frente a cualquier persona, en conexión con la compra o venta de cualquier valor. De lo anterior, surgió la regla de “informar o abstenerse”, la cual significa que “cuando un comerciante potencial de valores ha obtenido información material de forma tal que puede verse sujeto al alcance de la Regla 10b-5, éste deberá o abstenerse de transar, o divulgar la información antes de transar con los valores.

Dadas las anteriores normativas cabe indicar que en el presente proyecto lo que se plantea, es sino un modelo, basado netamente en el uso de técnicas de minerías de datos, lo cual quiere decir que las predicciones que el modelo arroje como resultado, deben ser consideradas como predicciones más no como una verdad absoluta, el modelo pretende brindar una herramienta a los inversionistas, cuyas predicciones pueda servir de punto de partida para la toma de decisiones. De acuerdo al literal c de la Regla de 10b-5, el inversionista que cuente con información privilegiada deberá de abstenerse de transar o informar al resto de la comunidad. El modelo de que se plantea en el presente proyecto será diseñado para trabajar con información de conocimiento público, para poder trabajar con cualquier otra información requerirá de una reestructuración en su diseño.

2.4 **Ámbito de Aplicación**

La bolsa de valores de Nueva York es el mayor mercado en el mundo, en volumen monetario y es el que mayor número de empresas adscritas tiene a nivel mundial. Cabe indicar que todo esto va acompañado, de que tiene un volumen de movimiento de alrededor 21 mil millones de dólares.

El mercado de valores de la ciudad de Nueva York se apertura a las 9:30AM, previo a la hora de apertura del mercado de valores, a las 7:30AM, se da lo que se conoce como subasta de apertura (Opening Auction). Durante la subasta de apertura se pueden introducir dos tipos de ordenes: Market on Open (MOO, por sus siglas en ingles) y Limit on Open (LOO, por sus siglas en ingles). La orden Market on Open es una orden de compra la cual hace uso del precio del mercado al momento de la apertura del mismo, por otro lado la Limit on Open es una orden de compra en caso de que el mercado alcance un precio por debajo del actual o una orden de venta cuando el mercado alcance un precio por encima del actual (Smith, 2016).

El mercado de valores de la ciudad de Nueva York se cierra a las 16:00PM, pero a las 15:59PM existe lo que conoce como Subasta de Cierre o Closing Auction, las ordenes que son introducidas al mercado de valores durante la subasta de cierre no pueden ser canceladas, cabe mencionar que el precio de cierre del mercado es determinado a partir de las ordenes generadas durante la Subasta de Cierre.

Todas las ordenes que sean realizadas a partir de la hora de cierre del mercado, se las introducirá en la subasta de apertura del día siguiente, así como estas órdenes en específico se las ejecutará en la apertura del mercado de ese mismo día. Al igual que en la subasta de apertura, existen dos tipos de órdenes, que pueden ser realizadas durante la subasta de apertura: La orden de tipo Market on Close (MOC, por sus siglas en ingles), es en la cual se lanzan órdenes de compra de acciones al precio de cierre del mercado. También está la orden de tipo Limit on Close (LOC, por sus siglas en ingles), muy parecida a la Limit on Open, pero en este caso se da en la hora de cierre del mercado de valores, es en la cual se trata de vender en caso de que el mercado alcance un precio superior o de comprar en caso de que alcance un precio inferior (Smith, 2016).

Los 3 índices más seguidos en la bolsa de valores de Nueva York son:

- 1. Dow Jones:** Compuesto por las 30 compañías industriales más importantes y representativas de Estados Unidos.
- 2. S&P 500:** Es uno de los índices más representativos de la economía real de EEUU por su diversidad. Conformado por 500 grandes empresas las poseen acciones que cotizan en las bolsas NYSE o NASDAQ.
- 3. Nasdaq:** La segunda bolsa de valores electrónica y automatizada más grande de los Estados Unidos. Compuesto por miles de empresas tecnológicas, tales como: Apple, Facebook, Microsoft, etc.

Dados los avances tecnológicos y la gran facilidad del internet han cambiado completamente el esquema de hace algunos años atrás, ya no solo las grande firmas financieras son capaces de realizar transacciones en la Bolsa de Valores, sino también un usuario con una computadora y acceso a internet, dado la existen de un sin número de plataformas de trading online, puede realizar este tipo de transacciones sin tener que salir de la comodidad de su hogar, para muchas de estas personas, conocidas como traders, este tipo de transacciones se convertido en un modo de vida, lo que convierte el hecho de conocer sobre estos mercados y estas operaciones algo cada vez más necesario, ya sea para poder obtener un conocimiento general, ya que lo ocurre en este tipo de mercados bursátiles a la larga influye en nuestra economía diaria, o bien sea que se tome la decisión de invertir o realizar transacciones en estos mercados, como una forma de ingreso.

En este capítulo se recaba información que se considera importante para el desarrollo del presente proyecto, resulta de gran importancia tener claro el funcionamiento de mercado bursátil, saber cuáles son los actores del mismo, bajo que normativas opera dicho mercado, así como el saber cómo los analistas en épocas pasadas realizaban estimaciones o predicciones del mercado, permitiéndoles tomar decisiones de compra o venta de valores, la interpretación de los gráficos de velas que son ampliamente usados en el mercado de valores y poder reconocer los distintos indicadores que se pueden presentar en estos gráficos, en este capítulo también se destaca los avances que trajo el internet en el mercado de valores, volviéndolo mucho más accesible para todos y no solo para unos pocos como fue en un comienzo. En este

capítulo se habla de metodología KDD, la cual es una de las metodologías más usadas cuando nos referimos a la minería de datos, esta metodología permite tener una ruta a seguir garantizando el éxito en un proyecto de minería de datos, así como se ahonda en la minería de datos, la cual es un paso dentro de la KDD, las distintas técnicas y métodos de minería de datos las cuales se debe saber escoger cual utilizar de acuerdo a las necesidades del proyectos. Todos estos conceptos y conocimiento son necesarios para poder escoger la ruta correcta a tomar para el desarrollo eficaz y eficiente del presente proyecto.

CAPÍTULO III

METODOLOGÍA Y RESULTADOS

En este capítulo se presenta la metodología a utilizar para obtener la información la cual permitirá cumplir con el desarrollo de la presente propuesta.

El presente capítulo está conformado por dos apartados, en el primero se presenta la metodología de investigación, en el segundo se presenta la metodología de desarrollo.

3.1 Metodología de la Investigación

El enfoque de la presente investigación es cuantitativo y documental, dado que la obtención de la información que se utilizará en el presente proyecto se encuentra colgada a manera de estadística en la página web del NASDAQ, el cual es uno de mayores mercados de valores del mundo y que negocia más volumen por hora que ninguna otra bolsa a nivel mundial, nació en 1971 por iniciativa de las autoridades norteamericanas con el fin de proporcionar más transparencia a los mercados no regulados (Fuentes Diaz, 2016). Cabe mencionar que la información que se obtuvo de la página web del NASDAQ es oficial, legítima y pública.

A la par de la obtención de la información estadística e histórica que se obtuvo desde el portal oficial del NASDAQ, también se realizó una investigación documental sobre el comportamiento de los mercados bursátiles y de las tendencias que hoy en día se presentan para poder realizar predicciones que aporten significancia a la hora de la toma de decisiones de compra-venta de valores del mercado bursátil.

En cuanto a las técnicas que levantamiento de información, se seleccionó la técnica documental, la investigación documental es una técnica que se orienta en la selección y recopilación de información por medio de la lectura crítica de documentos bibliográficos, la cual tiene como objetivo y el levantamiento de un marco teórico y de datos (Baena Paz, 2014).

3.2 Metodología de Desarrollo

La información obtenida de la página web de la NASDAQ, dado el enfoque cuantitativo de la presente investigación, en la cual se plantea una hipótesis, se determinan las variables que influyen directamente sobre el problema, por consiguiente, se elabora un modelo, el cual cuya exactitud se medirá mediante indicadores matemáticos, generando así una conclusión a partir de la hipótesis propuesta en un principio.

Dado que el presente proyecto es de minería de datos, es necesario seguir una metodología que vaya acorde a la propuesta. Un proceso típico de minería de datos consta de 6 pasos: Selección de los datos, Análisis de las propiedades de los datos, Transformación de los datos de entrada, Selección y aplicación de la Técnica de minería de datos, Extracción de conocimiento, Interpretación y evaluación de datos (Maimon & Rokach, 2010) .

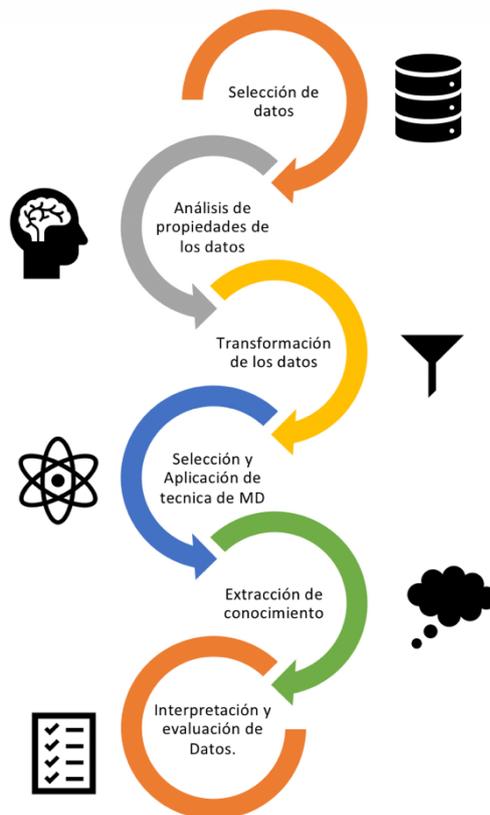


Figura 7. Procesos de Minería de Datos.

En el proceso de selección de la técnica de minería de datos es importante tener en cuenta la naturaleza de los datos con los que se trabajar son series temporales, que son datos que ha sido observados en diferentes instantes de tiempo, por lo cual existen un correlación entre lo que se observa y el tiempo, esto limita el uso de los métodos estadísticos que son utilizados comúnmente en la minería de datos (González Castellanos & Soto Valero, 2013). Las series temporales también puede ser definidas como una colección de observaciones realizadas de manera secuencial en el tiempo (Chatfield, 2003).

El análisis clásico de las series temporales, es conformado por el estudio de cuatro componentes principales sean estos: la tendencia, la variación estacional, la variación cíclica y la variación aleatoria. Es de destacas que no siempre la descomposición de estos factores es el mejor enfoque. Dada la naturaleza de los datos con los que se trabajara en esta propuesta, las técnicas de minería de datos a utilizar en reducen a tres: Redes Neuronales, Regresión Lineal y ARIMA. Uno de los modelos más utilizados para el análisis de series temporales es el modelo autorregresivo integrado móvil (ARIMA, por sus siglas en inglés Autorregresive Integrated Moving Average), el cual utiliza variaciones y regresiones de datos estadísticos con el fin de encontrar patrones para efectuar su predicción (Cowpertwait & Metcalfe, 2009).

El método ARIMA está limitado por los requerimientos de estacionariedad e invertibilidad del modelo estimado, el sistema generador de la serie debe de ser invariante y estables, las diferencias entre la series de tiempo y el modelo ARIMA deben de ser independientes así como presentar un distribución normal (González Castellanos & Soto Valero, 2013). De acuerdo a los estudios realizados el modelo ARIMA no presenta mayor exactitud a la hora de realizar predicciones en cuanto al mercado de valores se refiere, manejando porcentajes de error que fluctúan 16%-30%, en el cual se analizó información recolectada del NSE.com, durante un periodo de cinco años, del 2007 al 2011, utilizando el porcentaje de error como medida, midiendo la diferencia entre la predicción y la data histórica recuperada (Uma Devi, Sundar, & Alli, 2013).

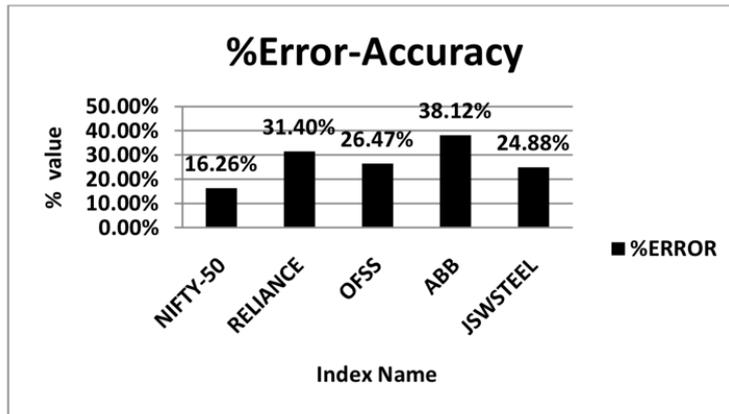


Figura 8. Ilustración de Análisis de precisión de error de los índices. Nota: Tomado de Uma Devi et al. (2013) An Effective Time Series Analysis for Stock Trend Prediction Using ARIMA Model for Nifty Midcap-50 (p. 74). Madurai.

En otro estudio se determina que el proceso para la construcción de un modelo ARIMA para la predicción de valores de mercado financiero tienen un gran potencial en cuanto a la predicción en un corto plazo, mientras que para predicciones a un largo plazo este modelo disminuye su grado de confiabilidad (Adebiyi, Adewumi, & Ayo, 2014).

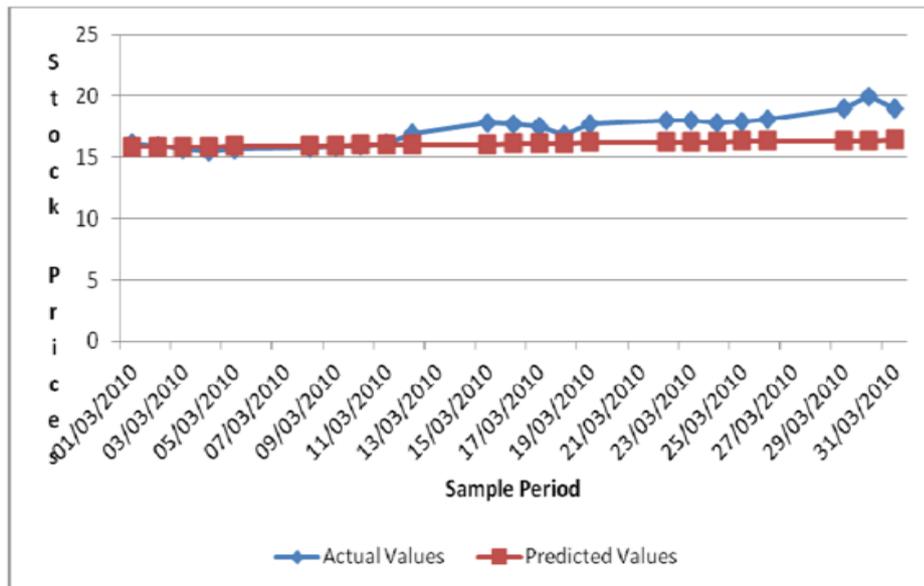


Figura 9. Gráfica comparativa Valores reales y Valores predichos de las acciones. Nota: Tomado de Adebiyi et al. (2014) Stock Price Prediction Using the ARIMA Model (p. 7). Reino Unido.

Teniendo en cuenta los resultados de estudios anteriores, es importante la búsqueda de alternativas alternas a el modelo ARIMA, que hagan uso de las nuevas tecnológicas con las que contamos al día de hoy que puedan ser capaces de mejorar los resultados obtenidos en estudios anteriores que han hecho uso del modelo ARIMA. En un estudio realizado sobre series temporales del consumo de energía eléctrica, se demuestra que resulta viable el uso de Redes Neuronales para la predicción de series temporales, en estudio se hace uso de las Redes Neuronales Recurrentes (RNN, por sus siglas en ingles Recurrent Neural Network), de tipo LSTM (por sus siglas en ingles Long-short term memory).

Las redes neuronales de tipo LSTM son un tipo especial de las redes RNN, las cuales son ampliamente utilizadas en problemas referentes a predicción de series temporales, dado que su diseño permite el almacenamiento de información por largos periodos de tiempo, lo cual es necesario en los casos de realizar estimaciones futuras basándose en registros históricos, estas poseen bloques de memoria, los cuales están conectadas a través de las capas (Gers, Schraudolph, & Schmidhuber, 2002). En el estudio se hace mención que con la ayuda de este tipo de redes neuronales se llega a alcanzar un 93% de precisión. Así como se menciona que la validez de las predicciones alcanzadas con el modelo depende en gran manera de la calidad de los datos proporcionados al modelo (González Avella, Tudurí, & Rullan, 2017).

Es necesario definir una metodología a utilizar cuando se trata del uso de RNN de tipo LSTM, en un estudio realizado para la predicción de precios de acciones de NIFTY 50, se estableció la siguiente metodología, la cual consta de 5 etapas principales (Roondiwala, Patel, & Shraddha, 2015):

- 1. Obtención de la data bruta.** - En esta etapa se concentra en la búsqueda de los datos, en el caso específico de esta propuesta, se enfoca en la búsqueda de la data histórica, para lo cual se usará la web oficial de la NASDAQ (<https://www.nasdaq.com>).
- 2. Preprocesamiento de la data.** - En la etapa del preprocesamiento de la data se incluyen operaciones como: La *discretización de data*, en la cual la data es reducida a solo la de mayor importancia para el modelo, especialmente cuando hablamos de data numérica. Otra operación es la *transformación de*

la data, es decir la normalización de la misma. La *limpieza de data* es otra operación importante, en la cual se llenan espacios vacíos de la data. Por último, la *integración de la data*, la integración de los data files.

3. **Extracción de características.** - En esta etapa se determina que variables o características formaran parte del modelo.
4. **Entrenamiento de la red Neuronal.** - En esta etapa se escoge una porción de la data, la cual servirá para entrenar el modelo, toda la data a utilizar deberá ser dividida en dos, una parte para el entrenamiento del modelo, otra que se utilizará a manera de test, para probar la exactitud o precisión del modelo, es de indicar que la data de entrenamiento deberá ser mayor en cantidad que la data de test. La data para el test de acuerdo a estudios anteriores se considera que sea alrededor del 5%-10% del total de la data a utilizar para el entrenamiento del modelo (Roondiwala et al., 2015).
5. **Generación de resultados.** - En esta etapa el modelo ya entrenado genera un numero de predicciones, las cuales serán comparadas con la data de test, para lo cual se debe de utilizar conceptos matemáticos que permitan determinar la precisión del modelo.

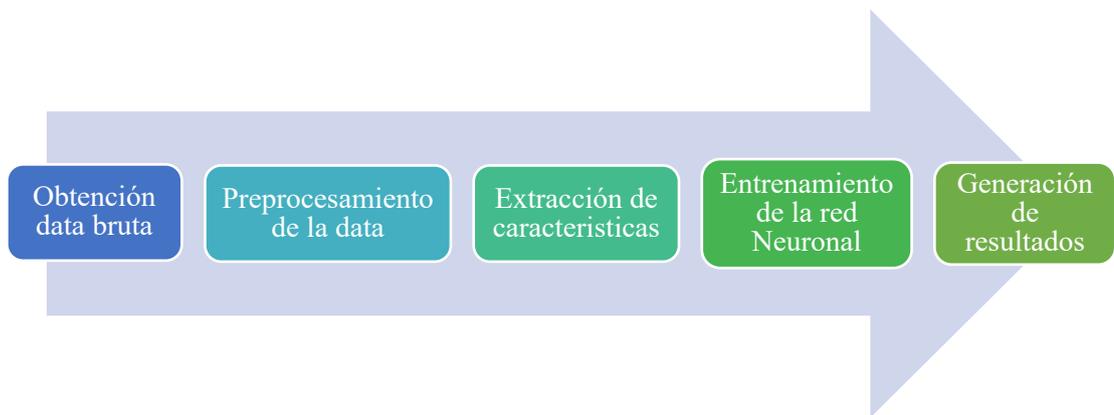


Figura 10. Metodología desarrollo modelo predictivo para RNN de tipo LSTM.

3.3 Análisis de Resultados

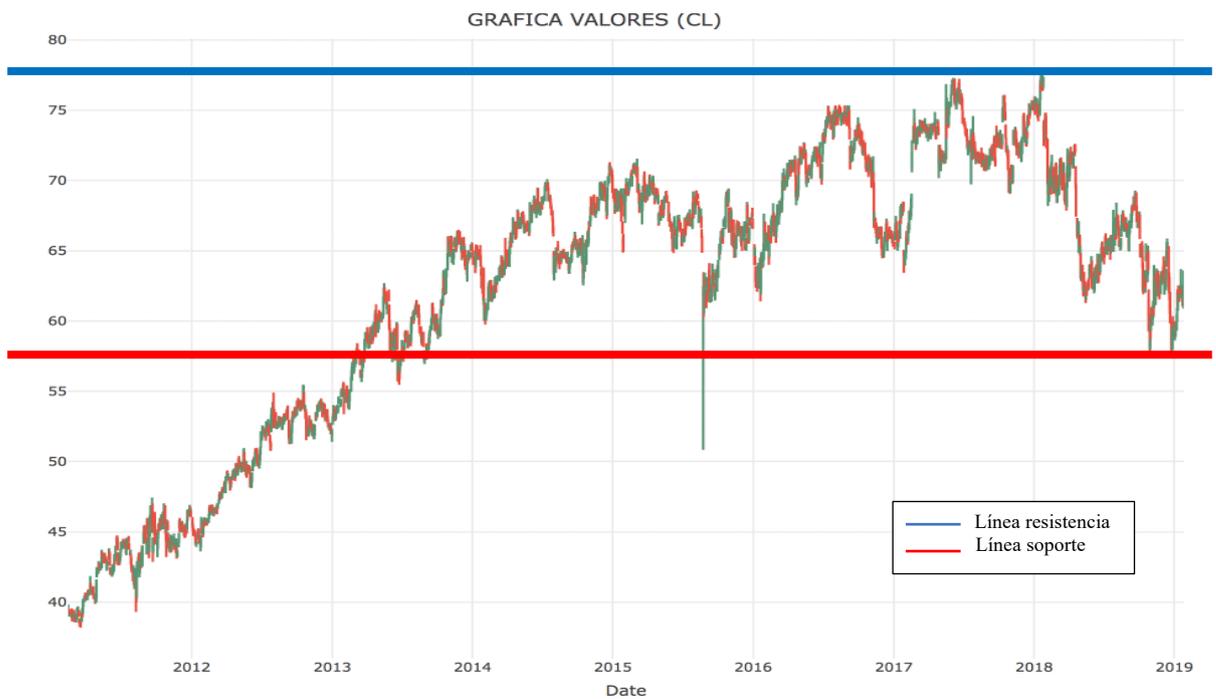


Figura 11. Gráfica de velas de Precio de cierre y apertura de Colgate Palmolive (CL).



Figura 12. Gráfica de velas de Precio de cierre y apertura de Rosetta Stone (RST).



Figura 13. Gráfica de velas de Precio de cierre y apertura de FEDEX (FDX).



Figura 14. Gráfica de velas de Precio de cierre y apertura de Jhonson & Jhonson (JNJ).

Se tomó de la decisión de graficar la data de apertura y cierre del mercado desde Febrero del 2011 hasta el 28 de Enero del 2019, es decir un periodo de ocho años, de cuatro empresas que cotizan en NYSE con el fin de limitar la data de entrenamiento del modelo, es necesario mencionar que la durante la fase de entrenamiento del modelo, la precisión que el modelo logre alcanzar dependerá en gran medida de la data que se entregue al modelo para aprender, sin considerar que mientras seamos capaces de limitar la data sin disminuir la precisión estaremos ahorrando tiempo y esfuerzo de procesamiento, por lo cual resulta necesario identificar el soporte y la resistencia de las distintas gráficas, en especial en el periodo de tiempo del último año.

En el caso de *Jhonson & Jhonson (JNJ)* se puede notar que la línea de soporte se encuentra en un valor de 110, mientras que la línea de resistencia se encuentra alrededor de 150. Por lo tanto, se podrá utilizar un periodo de tiempo que va del 2016 hasta la actualidad. En el caso de *Fedex (FDX)* se puede identificar que la línea de soporte se encuentra en 150, mientras que la resistencia se encuentra alrededor de 275, teniendo en cuenta se podrá utilizar un rango de tiempo que vaya desde inicios del 2014 hasta la actualidad. En el caso de *Rosetta Stone Inc. (RST)*, se puede observar una línea de soporte que se encuentra alrededor de 7 y una línea de resistencia que se encuentra alrededor de 21, en este caso se podría decidir tomar la data desde el 2013 hasta la actualidad. En el caso de *Colgate Palmolive Company (CL)*, se puede observar una línea de soporte que se encuentra alrededor de 57 y una línea de resistencia la cual se encuentra alrededor de 77, por lo cual se podría utilizar un rango de fecha que vaya desde mediados del 2014 hasta la actualidad.

Este análisis permitirá que se tenga un rango de fechas iniciales para poder trabajar sobre el modelo, dependiendo de la precisión que alcance el mismo, se podrá aumentar o reducir el rango de fechas requeridos, siempre y cuando se busque una mejora en el modelo. Es decir, sin importar la empresa sobre la que se desea efectuar la predicción es importante limitar el rango de data con la que se va a entrenar el modelo, ya que esto depende el grado de precisión que el modelo logre alcanzar en cuanto a sus predicciones. La parte más importante de una red neuronal es el aprendizaje como ya se mencionó en el capítulo dos, las redes neuronales basan su

aprendizaje en ejemplos, por lo cual la capacidad que posee la red neuronal para resolver estará dada por el tipo de ejemplos que se le dispone en el proceso de aprendizaje, por tal motivo existen dos factores muy importantes para la data de entrenamiento, esta deber ser, significativa y representativa.

En este capítulo se determinó la metodología que seguirá a lo largo del desarrollo de la propuesta, se toma la decisión de utilizar las Redes Neuronales como una vía alterna a la ya tradicional con el fin de determinar si es posible obtener mejores resultados en comparación a la metodología ARIMA, la cual se utiliza comúnmente en proyectos enfocados en la predicción de series temporales. Una vez que se determinó que el camino a tomar son la Redes Neuronales, se determinó de igual manera el tipo de Redes Neuronales que se podrían utilizar en la presente propuesta y de acuerdo a trabajos anteriores, se determinó que para proyectos en los cuales se busca predecir información a partir de data histórica, siendo esta data series temporales, las Redes Neuronales tipo LSTM son las más adecuadas, ya que cuentan con bloques de memoria los cuales son capaces de almacenar información, en este caso sería la información histórica para de esta manera producir una salida. También se logró establecer un rango de tiempo para las distintas cotizaciones de las distintas acciones, para que sean ingresadas dentro del modelo y de esta manera poder conseguir los resultados con una mayor precisión, dado que de la data con la que se vaya a entrenar el modelo depende la precisión o exactitud que este pueda alcanzar.

CAPÍTULO IV

Modelo predictivo para el comportamiento de las acciones que cotizan en la bolsa de valores de Nueva York.

El modelo predictivo propuesto tiene como fin realizar un pronóstico del comportamiento del mercado de valores, en específico los valores que cotizan en la bolsa de valores de la ciudad de Nueva York. Para lo cual se hizo uso de la minería de datos, la cual cuenta con varias técnicas de acuerdo a las necesidades de cada problema, para el desarrollo de este modelo se optó por las redes neuronales, debido a que es una ciencia que se encuentra en plena vigencia y es utilizada comúnmente en proyectos que utilicen series temporales para realizar predicciones, así como la presente propuesta.

4.1 Objetivo

Realizar pronósticos del comportamiento del mercado de valores de la ciudad de Nueva York, con la finalidad de ser una herramienta para los inversionistas al momento de decisiones de compra o venta en el mercado de valores.

4.2 Descripción del modelo

Módulos y Componentes de la solución

El presente modelo está conformado por dos grandes componentes, la primera es la GUI, la cual es la encargada de interactuar con el usuario, el segundo es el back-end, el cual se encarga de procesar toda la información que el usuario le entrego a la GUI. También la GUI como el back-end de la presente propuesta fueron desarrollado en Python.

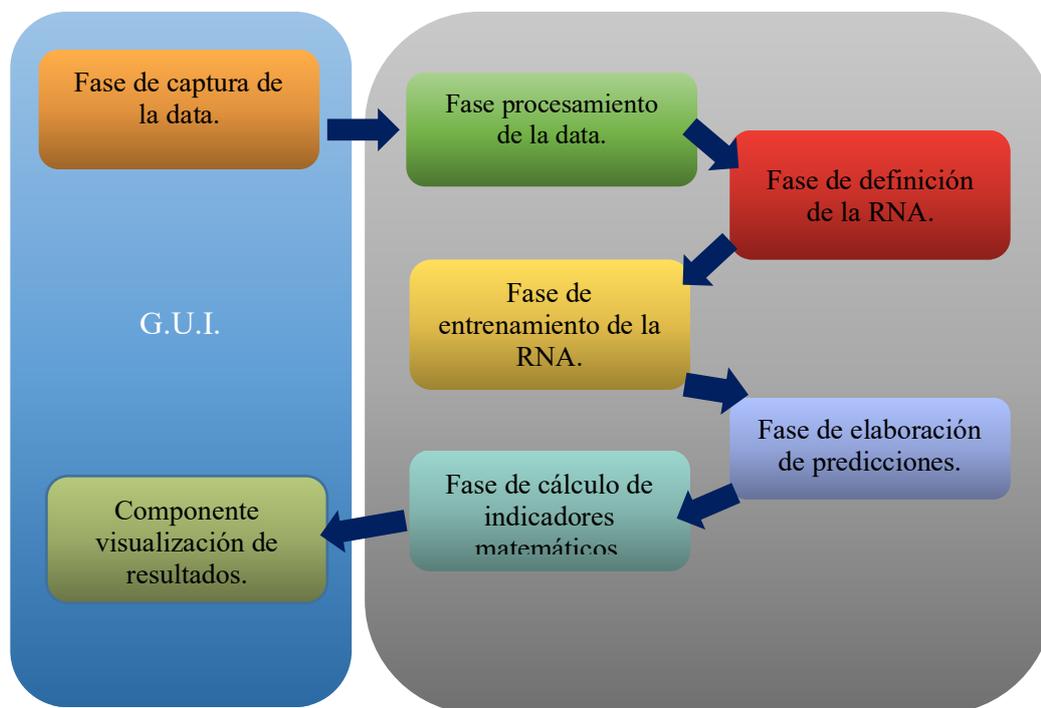


Figura 15. Esquema del modelo predictivo propuesto.

La G.U.I. está conformada por la fase de captura de la data, la cual es la que captura la data que es proporcionada por el usuario, es decir la data histórica de la cotización del valor al cual se quiere predecir el comportamiento, esta primera fase implementa controles los cuales validan que la data que el usuario ingrese al modelo tenga la estructura y forma requerida (ver Apéndice A), para así poder enviarla al back-end. A su vez la G.U.I posee un componente el cual es encargado de mostrarle de manera gráfica al usuario los resultados que arrojó el modelo, este componente se ayuda de una librería de Python llamada Plotly, la cual es capaz de generar gráficas en un entorno web, permitiendo que las gráficas sean interactivas y amigables con el usuario.

Por el otro lado el componente Back-End, está compuesto por varias fases: *Preprocesamiento de la data*, en esta fase la data es preparada para poder ser ingresada al modelo, para lo cual solamente se deja la data que el modelo va a utilizar, en este caso el modelo solamente hace uso de la información histórica del volumen, el precio de cierre y el precio de apertura. Durante la fase del preprocesamiento, es vital que la data sea normalizada, ya que cada una de las características de los datos que fueron ingresados por el usuario lo más probable es que se encuentren dentro de rangos de valores muy distintos, por lo cual para poder alcanzar un mayor grado de eficiencia en

el modelo se opta generalmente por normalizar la data entre 0 y 1, contando con la data normalizada se le facilita al modelo poder realizar cálculos con mayor facilidad y sobre todo que las características que posean valores más altos no dominen sobre las que poseen valores más bajos. Para lo cual se hace uso de la librería Sklearn y en específico de la función `fit_transform` (Sanchez, 2014). Una vez que la data que ha sido ingresada por el usuario ha sido normalizada y se ha dejado solamente las características que se han definido que el modelo utilizara, es necesario que se defina una variable la cual contenga los resultados esperados, para cada dato de entrada, de la siguiente forma:

Tabla 1. *Datos de entrada valores apertura, cierre y volumen. Jhonson & Jhonson (JNJ).*

DATOS ENTRADA (X) JNJ			
Fecha	Apertura	Cierre	Volumen
01/02/2019	128.13	127.75	6,985,739
01/03/2019	128.135	125.72	8,597,698
01/04/2019	127.12	127.83	8,746,471
01/07/2019	127.63	127.01	8,381,321
01/08/2019	128.18	129.96	9,198,666
01/09/2019	129.83	128.93	7,400,359
01/10/2019	127.69	129.71	6,217,193

Tabla 2. *Datos de salida (1 día en el futuro), valores apertura, cierre y volumen. Jhonson & Jhonson (JNJ).*

DATOS SALIDA (Y) JNJ			
Fecha	Apertura	Cierre	Volumen
01/02/2019	128.135	125.72	8,597,698
01/03/2019	127.12	127.83	8,746,471
01/04/2019	127.63	127.01	8,381,321
01/07/2019	128.18	129.96	9,198,666
01/08/2019	129.83	128.93	7,400,359
01/09/2019	127.69	129.71	6,217,193
01/10/2019	N/A	N/A	N/A

Como se puede apreciar, la Tabla 2, posee valores de un día en el futuro con respecto a X (ver Tabla 1), en este caso en específico se optó por que la variable de salida este un día en el futuro, pudiendo ser cualquier número de días en el futuro dependiendo de las necesidades.

Una vez que se cuenta con la data de entrada y salida es necesario que armar secuencias a partir de esta data, de la siguiente manera:

Tabla 3. *Secuencia de entrada (longitud 4 días), valores apertura, cierre y volumen. Jhonson & Jhonson (JNJ).*

SECUENCIA ENTRADA (X) JNJ				
#	Fecha	Apertura	Cierre	Volumen
1	01/02/2019	128.13	127.75	6,985,739
	01/03/2019	128.135	125.72	8,597,698
	01/04/2019	127.12	127.83	8,746,471
	01/07/2019	127.63	127.01	8,381,321
2	01/03/2019	128.135	125.72	8,597,698
	01/04/2019	127.12	127.83	8,746,471
	01/07/2019	127.63	127.01	8,381,321
	01/08/2019	128.18	129.96	9,198,666
3	01/04/2019	127.12	127.83	8,746,471
	01/07/2019	127.63	127.01	8,381,321
	01/08/2019	128.18	129.96	9,198,666
	01/09/2019	129.83	128.93	7,400,359

Tabla 4. *Secuencia de salida (longitud 4 días), valores apertura, cierre y volumen. Jhonson & Jhonson (JNJ).*

SECUENCIA SALIDA (Y) JNJ				
#	Fecha	Apertura	Cierre	Volumen
1	01/07/2019	128.18	129.96	9,198,666
2	01/08/2019	129.83	128.93	7,400,359
3	01/09/2019	127.69	129.71	6,217,193

El armado de las secuencias permite que el entrenamiento del modelo se realice de la mejor manera, la longitud de la secuencia que se designe es la cantidad de datos que se necesite que la red retenga en memoria para proporcionar un resultado (Liu,

Qiu, Chen, Wu, & Huang, 2015). Para la presente propuesta se utilizó secuencias de una longitud de 50 días, esta longitud fue la que mostró mejores resultados. El último paso es dividir el total de las secuencias en dos, se aconseja que un 5-10% sea para la data de comprobación o test, el resto para entrenamiento, en este caso se asignó el último 5% de las secuencias para la data de comprobación.

Para la fase de *definición de la RNA*, se hizo uso de Keras, la cual es una API, escrita en Python, que permite la construcción de RNA de una manera simplificada, con la capacidad de correr sobre Tensorflow, CNTK y Theano. Keras permite un prototipado fácil y rápido de las RNA, tiene soporte para RNN y CNN, es capaz de correr de igual manera tanto en CPU como GPU (Cholllet, 2018). Es necesario definir el tipo de modelo que utilizaremos, para este caso en particular el modelo Sequential, el cual está conformado por un conjunto de capas lineales. La primera capa en este modelo corresponde a la capa de entrada, se debe indicar la dimensión de los datos de entrada, en este caso se hace uso de: Apertura, Cierre y Volumen, siendo la dimensión de tres. A continuación, se define el parámetro correspondiente a la longitud de la secuencia, en este caso es de cincuenta, como se expresó anteriormente. Se procede a añadir la capa Dropout, es un método popular y poderoso para regularizar las redes neuronales, utilizada para reducir el sobreajuste del modelo, introduciendo de manera aleatoria un decimal entre cero y uno, correspondiente a la cantidad de neuronas a dejar, la capa Dropout es necesaria porque elimina constantemente y aleatoriamente unidades en cada lote, obligando a aprender, haciéndolo más robusto evitando que el modelo simplemente memorice los datos de entrenamiento, por lo general para las unidades de entrada se utiliza 0.2 y para las unidades ocultas se utiliza 0.5. Se añade una capa igual a la de entrada, no es necesario especificar la dimensión de entrada, se debe indicar la cantidad de capas ocultas que se requiere en la neurona, 250 proporcionó los mejores resultados en el modelo, seguido de esta capa se añade otra capa Dropout. Finalmente es necesario añadir una capa Dense, indicando la dimensión de la salida, en este caso tres, se debe de indicar la función de activación que utilizará el modelo, en la mayoría de los casos referentes a series temporales se utiliza relu, la cual presentó los mejores resultados.

```

model = Sequential()
model.add(LSTM(input_shape=(None, INPUT_DIM), units=UNROLL_LENGTH,
return_sequences=True))
model.add(Dropout(0.2))
model.add(LSTM(250, return_sequences=False))
model.add(Dropout(0.2))
model.add(Dense(3, activation='relu'))

```

Figura 16. Definición de la RNN de tipo LSTM, para la presente propuesta.

Para la *fase de entrenamiento del modelo*, se le proporciona al modelo las secuencias X de entrenamiento y las secuencias Y de entrenamiento, así como se asigna parámetros, como el número de Epochs, el cual corresponde al número de pasadas que va a realizar el modelo a la data de entrenamiento, las pasadas son de ida y vuelta, de manera aleatoria, para esta propuesta en específico 100 Epochs proporcionan los resultados sobresalientes. También es necesario proporcionar el parámetro correspondiente al Batch Size, el cual corresponde al número de ejemplos que el modelo debe procesar para poder realizar ajustes en sus parámetros internos, en este caso se asignó un Batch Size de 32, comúnmente utilizado en modelos para predicción de series temporales.

La *fase de elaboración de predicciones*, se hace uso de la función predict, a la cual se le proporciona la secuencia X test de entrada, junto el Batch Size, esta función proporciona un resultado, el cual contiene data de un día en el futuro respecto a la secuencia de entrada, conteniendo información del Cierre, Apertura y Volumen. Para poder generar predicciones de más días en futuro lo correspondiente es agregar este dato a la secuencia, de esta manera el modelo podrá generar predicciones a partir de las predicciones anteriores. Es necesario que los resultados obtenidos por el modelo, los cuales se encuentran normalizados, devolverlos a su estado inicial, utilizando la función inverse_transform.

La *fase de cálculo de indicadores matemáticos*, una vez que el modelo proporcionó los resultados es pertinente conocer la precisión y exactitud de estos, mediante la comparación de las salidas obtenidas a partir de la secuencia X test de entrada, con la secuencia Y test de salida. Para lo cual se hace uso de tres indicadores: Error porcentual absoluto Medio (MAPE, por sus siglas en ingles Mean Absolute Percentage Error), permite medir la precisión en porcentaje, con el fin de evaluar el

rendimiento de modelos de pronóstico; Error medio absoluto (MAE, por sus siglas en ingles Mean Absolute Error), su propósito es determinar qué tan cercano es el pronóstico del valor real, se realiza una resta del valor obtenido del real, se obtiene el valor absoluto de esta diferencia y se procede a obtener un promedio; Error cuadrático Medio (RMSE, por sus siglas en ingles Root Mean Square Error), realiza una diferencia entre los valores pronosticados por el modelo y los reales, estas diferencias son elevadas al cuadrado, se obtiene el promedio de todas estas diferencias al cuadrado y al final se obtiene la raíz cuadrada de este promedio (Negrón Baez, 2014).

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2$$

$$RMSE = \sqrt{MSE}$$

$$MAPE = \frac{1}{N} \sum_{i=1}^N \left| \frac{y_i - \hat{y}_i}{y_i} \right|$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i|$$

El *componente de visualización de resultados* es el encargado consolidar la data que se obtuvo del modelo e indicadores, de manera de gráfica para que el usuario la visualice y comprenda de una manera mucho más sencilla e interactiva, haciendo uso de la librería Plotly, que permite la visualización de estos gráficos en un entorno web.

Herramientas Tecnológicas

En la presente propuesta se hace uso de las siguientes herramientas tecnológicas:

Keras, es una API que está escrita en Python, la cual es capaz de ejecutarse sobre Tensorflow, Theano y CNTK. Las principales características de Keras son: es de fácil y rápido prototipado esto dado a su modularidad, minimalismo y extensibilidad; Es capaz de soportar redes neuronales recurrentes y redes neuronales convolucionales; Puede soportar esquemas de conectividad arbitrarios; Corre de igual manera tanto en CPU como en GPU. Por defecto Keras utiliza Tensorflow como motor, sin embargo esto puede ser cambiado a cualquiera de los motores anteriormente mencionados, realizando un cambio en el fichero de configuración (Cortés Antona, 2017).

TensorFlow es una API de código abierto, desarrollada por Google con licencia Apache 2.0, puede ser utilizada por desarrolladores e investigadores que requieren la incorporación de Machine Learning en sus proyectos. Tensorflow es capaz de utilizar tanto la CPU como la GPU. Un punto fuerte de Tensorflow es Tensorboard, el cual permite monitorizar y visualizar, el rendimiento del algoritmo. Tensorflow permite a los desarrolladores construir, entrenar y ejecutar redes neuronales de una manera ágil. Cuando se trabaja en Python, se recomienda el uso de entorno virtuales como Anaconda, esto permitirá mantener las dependencias de los proyectos en distintos lugares del mismo ordenador (Torres, 2016).

Plotly es una API para Python la cual permite representar de manera gráfica la información, lo importante de esta librería es que lo hace bajo un entorno web, el cual es más dinámico e interactivo de cara al usuario, en comparación a que se utilizara Python.

Scikit-Learn es la principal API en cuanto a Machine Learning se refiere, construida a partir de las librerías Numpy y matplotlib, incluye la implementación de varios algoritmos de aprendizaje, esta librería puede ser utilizada para clasificaciones, extracción de características, regresiones, agrupaciones, reducción de dimensiones, selección de modelos o preprocesamiento de la data. Puede ser utilizada para

experimentar con diferentes tipos de algoritmos con solo cambiar unas cuantas líneas de código, es muy utilizada en cuanto a investigaciones académicas se refiere dado a que se encuentra muy bien documentada, es de fácil uso y versátil (Hackeling, 2014).

Proceso

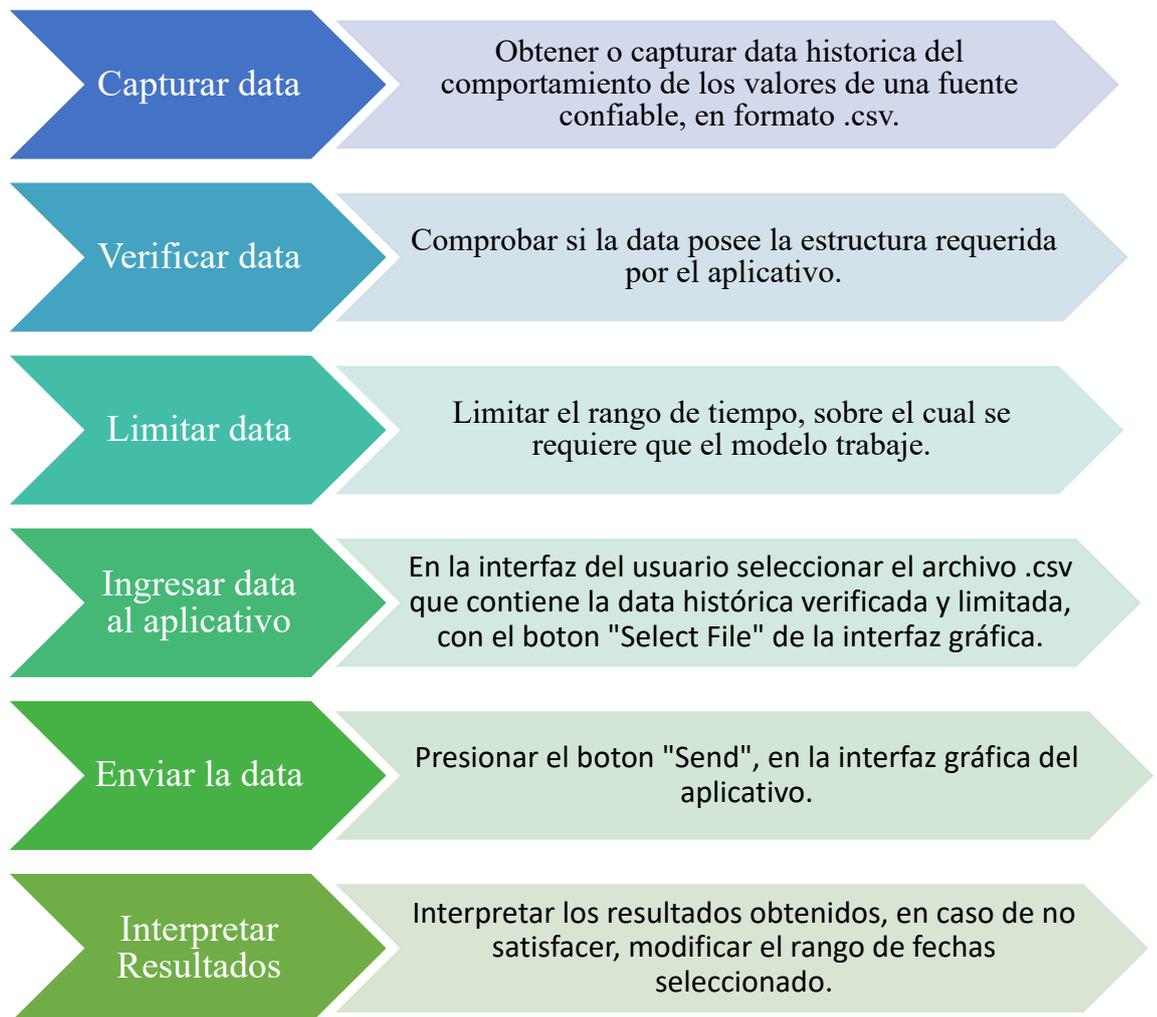


Figura 17. Pasos para el uso del aplicativo de la presente propuesta.

4.3 Requerimientos para su implementación

Tabla 5. *Requerimientos para la implementación del aplicativo.*

REQUERIMIENTOS	
HARDWARE	Requerimientos mínimos del ordenador: <ul style="list-style-type: none">• Procesador: 4 núcleos Core i5.• RAM: 4GB.• Almacenamiento (Disco Duro): 1GB.
SOTWARE	Entorno virtual (Anaconda) con las siguientes librerías: <ul style="list-style-type: none">• Pandas.• Numpy.• Plotly.• Keras.• Sklearn.• Tensorflow. <p>Editor de archivos de texto.</p>

4.4 Resultados esperados del modelo

Mediante la implementación del presente modelo se logrará lo siguiente:

- Predecir el comportamiento de las acciones que cotizan en el mercado de valores de la ciudad de Nueva York.
- Brindar una herramienta gratuita a los inversionistas, para la toma de decisiones de compra y venta de valores.
- Proporcionar una herramienta para la predicción del comportamiento del mercado de valores de la ciudad de Nueva York, la cual haga uso de las nuevas tecnologías.
- Lograr obtener predicciones más precisas, en comparación a las metodologías tradicionales.

CONCLUSIONES

Los mercados de valores los últimos años han sufrido grandes cambios, muchos de estos debidos al internet, y la existencia de brókers digitales, dando apertura de este mercado a las masas. El mercado de valores presenta un comportamiento no lineal, muy frecuentemente se presentan cambios bruscos, en muchos casos impredecibles, los cuales rompen con la tendencia, como consecuencia el tratar de predecir el comportamiento del mercado a simple vista resulta muy complejo y solo con un alto nivel de conocimiento y entrenamiento se podría llegar a realizar una predicción, que posea un alto grado de confiabilidad. La naturaleza propia del mercado de valores y la globalización del mismo, vuelve necesario la existencia de una herramienta que apoye a los inversionistas durante la toma decisiones, sobre todo a los usuarios más inexpertos y a los que se comienzan a adentrar en el mercado de valores.

Mediante el uso de la Minería de Datos y las Redes Neuronales se logró desarrollar un modelo predictivo, una red neuronal de tipo LSTM (ver Apéndice W), el cual es capaz de alcanzar muy bajos porcentuales de error en sus predicciones (ver Tabla 6), pero mucho más importante que eso, es que las predicciones del modelo desarrollado, son capaces de seguir la tendencia y el comportamiento del mercado de valores en la vida real (ver Apéndice B, E, H). Lo que indicaría que el modelo puede ser utilizado como herramienta valida, en la toma de decisiones para la compra o venta de valores, cabe mencionar que esta herramienta debe ser tomada como un punto de partida y debe ser complementada con otro tipo de análisis, tales como: Los Estados de Resultados, los balances generales y los propios encabezados de las noticias.

Es importante mencionar que la obtención del modelo predictivo definitivo requirió de varias corridas, hasta que se llegó a un modelo el cual cuyo grado de precisión presentaba mejoras significativas en comparación al resto, teniendo en cuenta que exista un balance en cuanto al uso del recurso hardware y la precisión de los resultados alcanzados (ver Apéndice X).

Se seleccionó un grupo experimental de 4 de las 2800 empresas que cotizan en la bolsa de valores de la ciudad de Nueva York. Dada la gran cantidad de empresas que cotizan en este mercado de valores y la diversidad de comportamientos que presenta cada una de ellas, se concluyó en la necesidad de entrenar al modelo cada vez que la data le es ingresada, de esta manera se puede asegurar la generalidad del modelo y que la precisión del mismo se mantenga dentro de márgenes confiables.

Mediante el análisis de la data histórica recopilada, la data presentaba las características de una serie temporal, de las series temporales solo ciertas características de todas las capturadas (ver Apéndice A) aportan significancia para el modelo predictivo, siendo estas: fecha, precio de apertura, precio de cierre y volumen. Durante el análisis de las gráficas elaboradas a partir de la data histórica de cierre y apertura, se pudo evidenciar que no toda la data histórica conservaba relación con la data actual (últimos dos años), la limitación de data histórica es una operación que afecta directamente el grado de confiabilidad de las predicciones del modelo. Con el fin de delimitar la data histórica y que esta data guarde relación, se recurrió al trazado de las líneas de soporte y resistencia correspondiente a los dos últimos años (ver Figura 11, 12, 13 y 14), esta brindó los mejores resultados comprobados durante la fase de pruebas. Las predicciones obtenidas por el modelo están directamente relacionadas con la data de entrenamiento que se le proporciona. La reducción de características y el rango de data, ayuda también en la reducción de los tiempos de procesamiento del modelo y el consumo del recurso hardware.

En el proceso de desarrollo del modelo se pudo concluir que las Redes Neuronales, de tipo Redes Neuronales Recurrentes (RNN), se acoplan en gran medida al tratamiento de series temporales del mercado de valores, estas redes poseen un bucle de recurrencia en sus neuronas permitiéndoles mantener la información en memoria, el problema aparece cuando se trabaja con grandes almacenes de datos, los estados más lejanos se van perdiendo, a este problema se lo conoce como: “vanish gradient problem”, para solucionar este problema, existen las LSTM, que son un tipo de RNN que incorpora una célula de memoria, por lo cual se decidió por este tipo de red neuronal en específico. También se pudo concluir que alrededor de la 100 Epochs, el aprendizaje del modelo se estabiliza, el hacer mucho más por encima de las 100

Epochs, no representaba un aumento significativo en la mejora de los resultados, pero si demanda un gasto innecesario en cuanto a hardware se refiere. De igual manera se pudo establecer que para el parámetro del Batch Size, el modelo presentó los mejores resultados al utilizar como parámetro 32, así como el largo de la secuencia de entrenamiento, el modelo presentó los mejores resultados en 50. Fue necesario que asignar al modelo una capa de entrada y de salida de la misma dimensión, de esta manera el modelo podría generar predicciones a partir de sus propias predicciones, mediante un bucle. Las herramientas utilizadas tanto Python como Keras brindan las herramientas necesarias para que el desarrollo del modelo sea ágil y sencillo.

De acuerdo a las predicciones obtenidas por el modelo se puede concluir que el porcentaje de error del modelo para los cuatro casos analizados es relativamente bajo, desde un 1.5% hasta un 3.5% en el cálculo del MAPE, en cuanto para el MAE va desde 0.6 hasta 6 y el RMSE 0.8 hasta 8.21, en estos dos últimos indicadores se debe tener en cuenta que son proporcionales a los precios de los valores (ver Tabla 6). Dado estos resultados se puede concluir que el modelo planteado puede llegar a predecir el comportamiento del mercado de valores de la ciudad de Nueva York con un alto nivel de exactitud, pero como ya se mencionó anteriormente estas predicciones deben ser complementadas con otros análisis, que ya fueron mencionados. Las predicciones del modelo propuesto fueron comparadas con predicciones obtenidas con el modelo ARIMA (ver Apéndice N, O, P y Q), modelo tradicionalmente utilizado para la predicción de series temporales y en precio para el mercado financiero de valores. Los resultados obtenidos fueron muy alentadores, en todos los escenarios el modelo propuesto superó notablemente al modelo ARIMA, es decir el modelo propuesto puede ser considerado como alternativa al modelo tradicionalmente utilizado, proporcionando resultados más precisos (ver Tabla 7) y comparativa (ver Figura 18).

También se pudo notar que mientras más grande es la brecha entre la línea de soporte y la línea de resistencia como se puede observar en el preciso caso de Rosetta Stone (ver Figura 12), predicciones obtenidas a partir del modelo son menos precisas, esto podría ser mejorada con la existencia de mayor cantidad de data histórica relevante para el modelo.

Tabla 6. Resultados de indicadores matemáticos para el Modelo predictivo LSTM.

Nombre	MAPE	MAE	RMSE
Colgate Palmolive (CL)	1.49102%	0.942279	1.37845
Rosetta Stone (RST)	3.50098%	0.608768	0.826862
FEDEX (FDX)	2.86335%	6.03561	8.21595
Jhonson & Jhonson (JNJ)	1.64157%	2.2203	3.14412

Tabla 7. Resultados de indicadores matemáticos para el Modelo predictivo ARIMA.

Nombre	MAPE	MAE	RMSE
Colgate Palmolive (CL)	4.107798%	2.565767	3.103611
Rosetta Stone (RST)	28.799514%	4.5634012	4.608127
FEDEX (FDX)	22.615359%	38.000356	46.023368
Jhonson & Jhonson (JNJ)	5.836885%	7.945114	8.193998

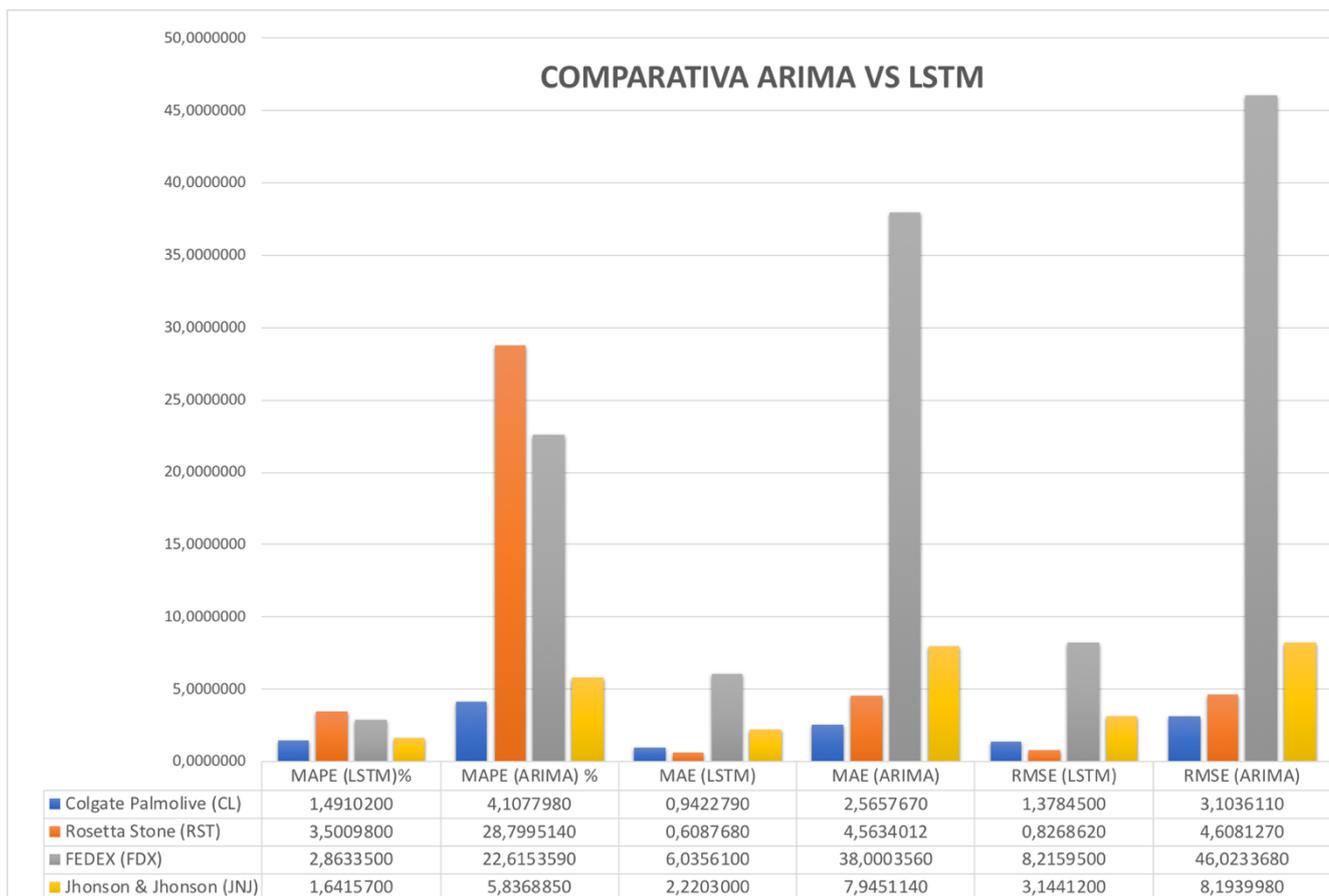


Figura 18. Grafica comparativa indicadores resultado de predicciones modelo ARIMA y LSTM.

Para finalizar se puso a prueba el modelo en la vida real, para lo cual se tomó las predicciones realizadas por el modelo y se las comparo de manera manual con los datos de la vida real (ver Apéndice S, T, U, V), de acuerdo a estos resultados se puede observar un mayor grado de precisión en comparación a un modelo predictivo que utilice ARIMA, el modelo propuesto presenta porcentajes de error que van de un 2% hasta un 5% para los casos estudiados, demostrando que el modelo propuesto puede ser aplicable en la vida real como una alternativa al ya tradicional ARIMA.

RECOMENDACIONES

Se recomienda a manera de complemento de la presente propuesta, determinar la influencia de los encabezados de las noticias en el mercado de valor, en caso de haber relación convendría desarrollar una funcionalidad que permita identificar encabezados de las noticias y un. Modelo que permita predecir los efectos de estos sobre el mercado de valores. También se recomienda el análisis de los Balances Generales y los Estados de Resultado de las empresas y su efecto en el comportamiento del mercado de valores, caso de existir algún tipo de relación lo más conveniente sería el desarrollo de un modelo que permita predecir el efecto de estos sobre el mercado de valores. Dado que una de las practicas más utilizadas al día de hoy son las negociaciones intradiarias, se recomienda la implementación de un módulo que se capaz de poder capturar la data del mercado de valores con mayor frecuencia, es decir poder obtener información cada minuto del comportamiento del mercado de valores, sería lo deseable, cabe mencionar que al día de hoy esta información es de paga, el poder contar con esta información con mayor frecuencia no solo permitirá poder realizar predicciones para negociaciones intradiarias, también aportaría considerablemente para los otros tipos de negociaciones ya que las redes neuronales son hambrientas de datos, siempre y cuando estos datos sean de utilidad.

REFERENCIAS BIBLIOGRÁFICAS

Adebiyi, A. A., Adewumi, A. O., & Ayo, C. K. (2014). Stock Price Prediction Using the ARIMA Model (p. 7). Presentado en International Conference on Computer Modelling and Simulation, Reino Unido.

Baena Paz, G. (2014). *Metodología de la Investigación* (Primera). México.

Beltran Martinez, B. (2015). *Minería de Datos*. México: Benemérita Universidad Autónoma de Puebla.

Caicedo Pulido, A. P., & Gaviria Franco, M. C. (2010). *Guía Básica de análisis técnico aplicado al mercado de divisas Forex*. Universidad de la Sabana, Colombia.

Chatfield, C. (2003). *The Analysis of Time Series: An Introduction* (6a ed.). Boca Ratón: CHAPMAN & HALL/CRC.

Chollet, F. (2018). *Introduction to Keras*. Estados Unidos.

Cortés Antona, C. (2017). *HERRAMIENTAS MODERNAS EN REDES NEURONALES: LA LIBRERÍA KERAS*. Universidad Autónoma de Madrid, Madrid.

Cowpertwait, P., & Metcalfe, A. (2009). *Introductory Time Series with R*. Estados Unidos: Springer US.

De Miguel Casas, E. (2016). *Sistemas de trading basados en análisis técnico*. Universidad de la Rioja, España.

Fuentes Diaz, R. B. (2016, noviembre). Nasdaq: Inversiones para aventureros.

Gallofré, F. (2014). *Manual del Chartismo*. bolsaytrading.com.

Garcia Roldan, T. (2015). *Teoría de Dow y ondas de Elliot: Una aplicación práctica*. Universidad Rey Juan Carlos, Madrid.

Gers, F. A., Schraudolph, N. N., & Schmidhuber, J. (2002). Learning Precise Timing with LSTM Recurrent Networks, 28.

Gitman, L. J., & Joehnk, M. (2009). *Fundamentos de Inversiones* (10a ed.). México: Pearson.

González Castellanos, M., & Soto Valero, C. (2013). *Minería de datos para series temporales*. Universidad Central “Marta Abreu” de las Villas, Santa Clara.

González Avella, J. C., Tudurí, J. M., & Rullan, G. (2017). *Análisis de Series Temporales Usando Redes Neuronales Recurrentes*. España.

González Casimiro, M. P. (2009). *Análisis de Series Temporales*. España: Universidad del País Vasco.

González Muñiz, A. (2018). *Aplicaciones de técnicas de inteligencia artificial basadas en aprendizaje profundo (Deep Learning) al análisis y mejora de la eficiencia de procesos industriales*. Universidad de OVIDE, España.

Gulli, A., & Pal, S. (2017). *Deep Learning with Keras*. Reino Unido: Packt.

Hackeling, G. (2014). *Mastering Machine Learning with Scikit-Learn*. Reino Unido: Packt.

Isasi Viñuela, P., & Galván León, I. M. (2004). *Redes de Neuronas Artificiales. Un enfoque practico*. Madrid: Pearson.

Liu, P., Qiu, X., Chen, X., Wu, S., & Huang, X. (2015). *Multi-Timescale Long Short-Term Memory Neural Network for Modelling Sentences and Documents*. Fudan University, Shanghai.

Maimon, O., & Rokach, L. (2010). *Data Mining and Knowledge Discovery Handbook*. Nueva York: Springer US.

Malkiel, B. G. (2007). *Un paseo aleatorio por Wall Street (Novena)*. Madrid: Alianza.

Matich, D. J. (2001). *Redes Neuronales: Conceptos Básicos y Aplicaciones*. Rosario: Universidad Tecnológica Nacional de Rosario.

Morales Sierra, A. (2011). *Análisis para la operación de mercado Forex en el contexto de un mercado Global*. Pontificia Universidad Javeriana, Bogotá.

Murphy, J. J. (2000). *Análisis Técnico de los Mercados Financieros*. Barcelona: Gestión.

Negrón Baez, P. A. (2014). *Redes Neuronales Sigmoidal con Algoritmo LM para Pronostico de Tendencia del Precio de las Acciones del IPSA* (p. 44). Valparaíso: Pontificia Universidad de Valparaíso.

Nison, S. (1991). *Japanese Candlestick Charting Techniques: Investment Techniques of the far East*. Nueva York: Instituto de Finanzas de Nueva York.

Nison, S. (1996). *Más allá de las velas*. España: GESMOVASA.

Palma Ménendez, J. T., & Marin Morales, R. (2008). *Inteligencia Artificial: Técnicas, métodos y aplicación*. Madrid: Mc Graw Hill.

Puentes Pérez, F. J. (2015). *La bolsa, principios y técnicas de análisis*. Universidad Politécnica de Cartagena, Colombia.

Riquelme, J. C., Ruiz, R., & Gilbert, K. (2006). *Minería de Datos: Conceptos y Tendencias*.

Roondiwala, M., Patel, H., & Shraddha, V. (2015). *Predicting Stock Prices Using LSTM*. Mumbai University, Mumbai.

Sanchez, J. M. (2014). *Análisis de técnicas Machine Learning para la estimación de medidas corporales*. Universidad Autónoma de Barcelona, Barcelona.

Serrano Ruiz, F. (2013). *Escuela de Bolsa (Cuarta)*. Córdoba: Almuzara.

Serrano Ruiz, F. (2015). *Day trading y operativa bursátil (Primera)*. Barcelona: Grupo Planeta.

Smith, J. (2016, 21:00 02/03). Horario de negociación de la Bolsa de Nueva York. Recuperado el 26 de noviembre de 2018, de <https://www.rankia.us/blog/stock-market-dummies/3135840-horario-negociacion-bolsa-nueva-york>

Torres, J. (2016). *Hello World En Tensorflow - Para Iniciarse En La Programación del Deep Learning*. Barcelona: lulu.com.

Uma Devi, B., Sundar, D., & Alli, P. (2013). An Effective Time Series Analysis for Stock Trend Prediction Using ARIMA Model for Nifty Midcap-50, 3.

Vallejos, S. J. (2006). *Minería de Datos*. Universidad Nacional del Nordeste, Argentina.

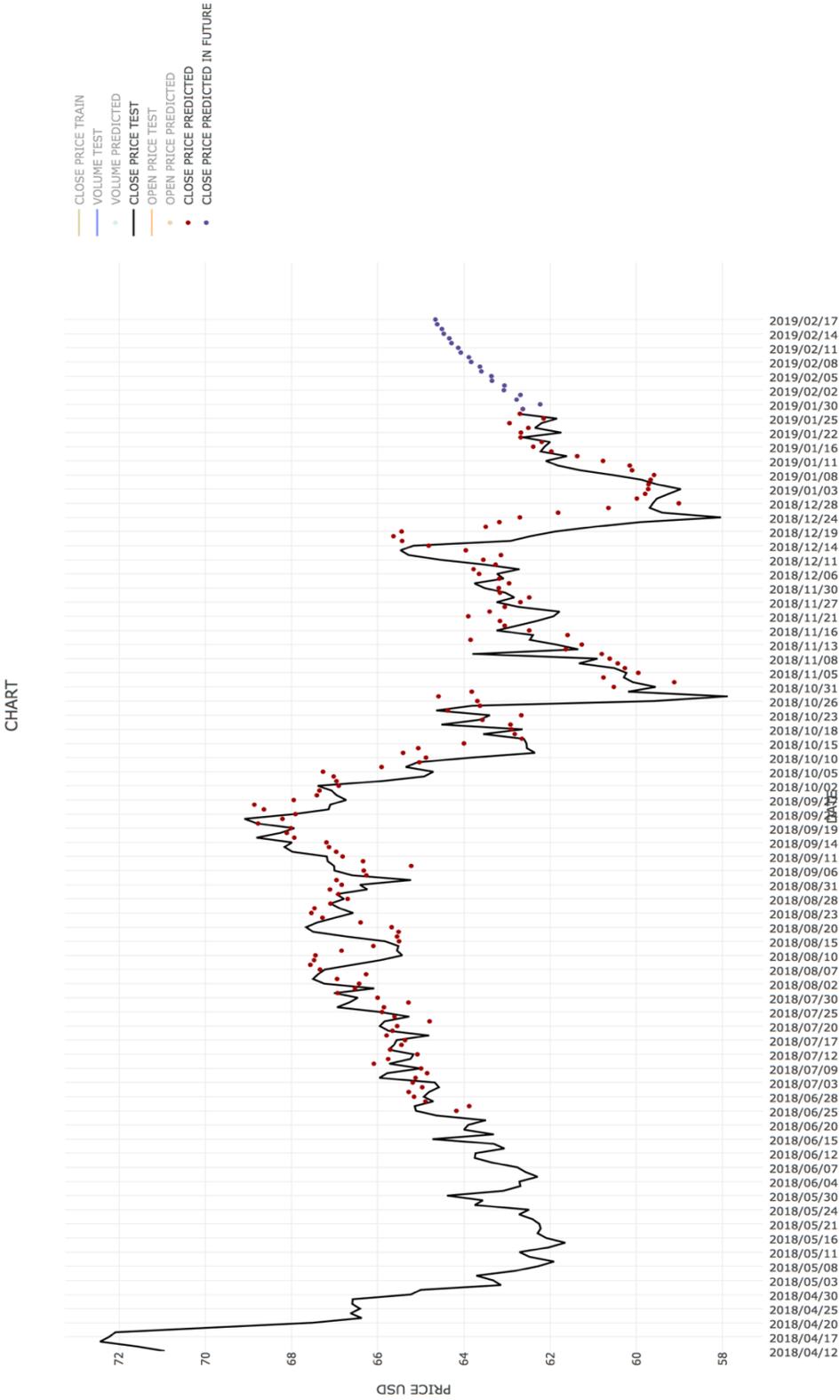
Zaccone, G., Karim, R., & Menshawy, A. (2017). *Deep Learning with TensorFlow*. Birmingham: Packt.

APÉNDICES

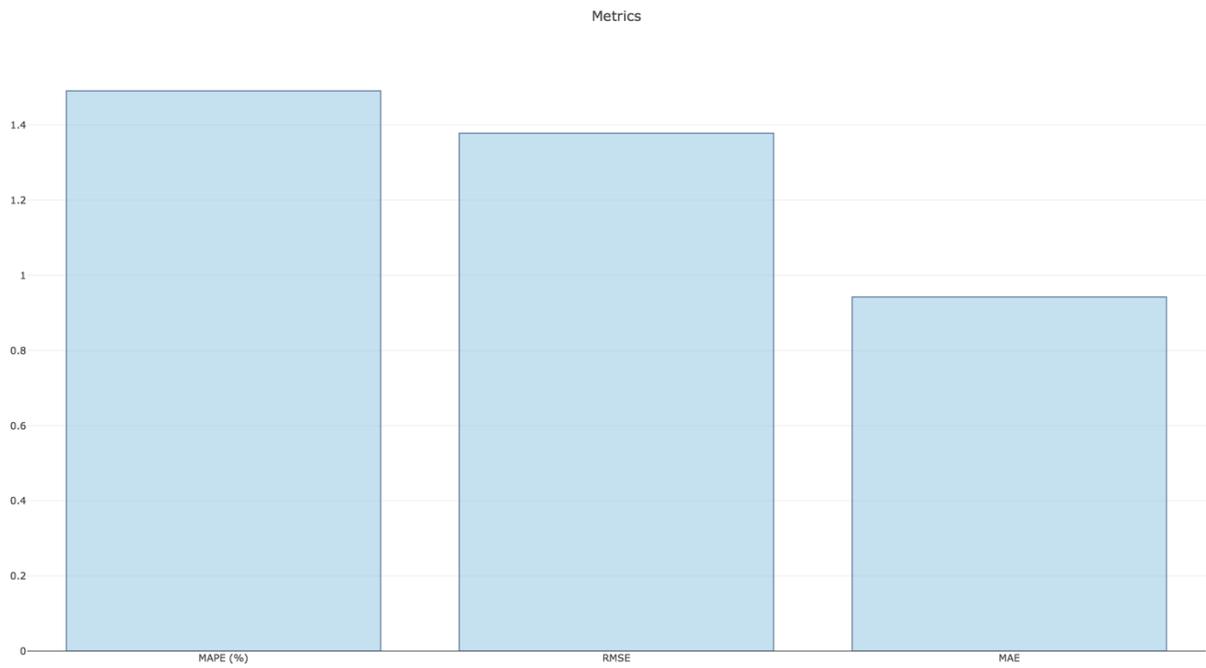
Apéndice A. Formato del archivo .csv requerido por el modelo.

```
"date","close","volume","open","high","low"  
"2019/01/28","62.7400","5752575.0000","61.7100","62.8400","61.6000"  
"2019/01/25","61.8400","8474365.0000","60.9600","63.6000","60.9600"  
"2019/01/24","62.2000","6022838.0000","62.1700","62.2500","61.2700"  
"2019/01/23","62.3500","3881712.0000","61.9300","62.5300","61.6400"  
"2019/01/22","61.7400","5541503.0000","62.1300","62.4500","61.0500"  
"2019/01/18","62.6300","5192768.0000","62.4300","63.7000","62.3300"  
"2019/01/17","62.0000","3544127.0000","62.1200","62.5000","61.9101"  
"2019/01/16","62.1300","3006607.0000","61.8500","62.3300","61.6950"  
"2019/01/15","62.2200","2520310.0000","61.6500","62.5100","61.6000"  
"2019/01/14","61.6100","3887052.0000","61.8100","62.0400","61.3700"  
"2019/01/11","62.1000","2616363.0000","62.2600","62.5400","61.7500"  
"2019/01/10","61.8200","3979359.0000","61.0200","61.8700","60.7200"  
"2019/01/09","61.3000","5218166.0000","61.9700","62.2300","60.9600"  
"2019/01/08","60.5500","5609967.0000","60.0200","60.9500","60.0200"  
"2019/01/07","59.8800","4550793.0000","59.4400","60.4550","59.3100"  
"2019/01/04","59.5200","7016247.0000","59.1100","59.5500","58.8600"  
"2019/01/03","58.9700","4119588.0000","59.1900","59.9850","58.8700"  
"2019/01/02","59.2500","4559379.0000","59.2400","59.5200","58.6200"  
"2018/12/31","59.5200","4271299.0000","59.3400","59.8700","58.9015"  
"2018/12/28","59.6200","5574825.0000","59.8400","60.3500","59.4100"  
"2018/12/27","59.6900","6524824.0000","58.9900","59.6900","57.9700"  
"2018/12/26","59.4000","6514702.0000","58.2000","59.4000","57.5050"  
"2018/12/24","58.0400","3872322.0000","59.9600","60.0000","57.8500"  
"2018/12/21","59.9100","9948337.0000","60.6100","62.0800","59.8900"  
"2018/12/20","60.9800","5387945.0000","61.7400","61.9550","60.5100"
```

Apéndice B. Gráfica precio de cierre Predichos y Precio de cierre Reales, de Colgate Palmolive (CL).



Apéndice C. Gráfica indicadores de la corrida del modelo para Colgate Palmolive (CL).



Apéndice D. Tabla de precios de Cierre predichos y Reales para Colgate Palmolive
(CL).

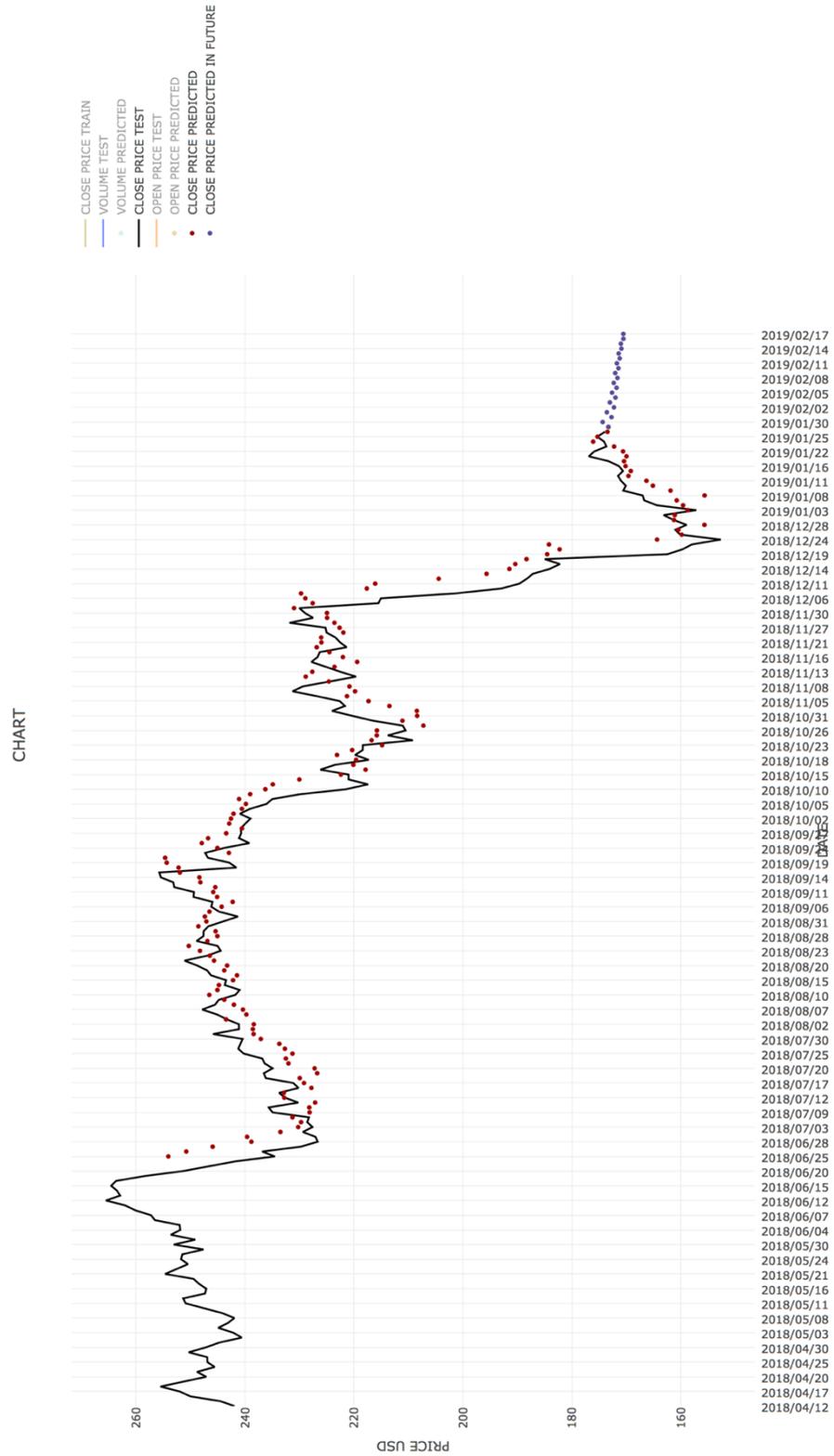
DATE	CLOSE PRICE REAL	CLOSE PRICE PREDICTED
2018/06/25	65,11	64,17332458
2018/06/26	65,14	63,87891388
2018/06/27	64,71	64,8921814
2018/06/28	64,93	65,15663147
2018/06/29	64,81	65,28115082
2018/07/02	64,58	64,9664917
2018/07/03	64,67	65,18605804
2018/07/05	65,96	65,12165833
2018/07/06	65,78	64,85366058
2018/07/09	65	64,99329376
2018/07/10	65,73	66,08869171
2018/07/11	65,24	65,75717163
2018/07/12	65,17	65,07833862
2018/07/13	65,75	65,70955658
2018/07/16	65,62	65,44803619
2018/07/17	65,56	65,36619568
2018/07/18	64,81	65,79388428
2018/07/19	65,75	65,65805054
2018/07/20	65,95	65,54669952
2018/07/23	65,84	64,7967453
2018/07/24	65,27	65,61032104
2018/07/25	65,95	65,89806366
2018/07/26	66,94	65,85496521
2018/07/27	66,66	65,28633881
2018/07/30	66,46	66,00074005
2018/07/31	67,01	66,92823029
2018/08/01	66,08	66,526474
2018/08/02	67,24	66,43000031
2018/08/03	67,5	66,94075775
2018/08/06	67,38	66,26886749
2018/08/07	67,23	67,33757019
2018/08/08	66,59	67,56319427
2018/08/09	65,93	67,47568512
2018/08/10	65,43	67,44265747
2018/08/13	65,55	66,83798218
2018/08/14	65,52	66,09958649
2018/08/15	65,83	65,50392151
2018/08/16	66,72	65,5489502

2018/08/17	67,5	65,51468658
2018/08/20	67,66	65,67678833
2018/08/21	67,4	66,39671326
2018/08/22	66,99	67,28139496
2018/08/23	66,57	67,53791809
2018/08/24	66,88	67,46670532
2018/08/27	67,09	67,09505463
2018/08/28	66,78	66,69688416
2018/08/29	66,96	66,90869904
2018/08/30	66,24	67,10852051
2018/08/31	66,41	66,83472443
2018/09/04	65,23	66,95329285
2018/09/05	66,59	66,26204681
2018/09/06	67	66,32264709
2018/09/07	67,01	65,22360229
2018/09/10	67,16	66,33930206
2018/09/11	67,18	66,81373596
2018/09/12	67,98	66,9601593
2018/09/13	68,16	67,12981415
2018/09/14	67,99	67,1883316
2018/09/17	68,81	67,93492889
2018/09/18	68,26	68,10906219
2018/09/19	67,93	68,01567078
2018/09/20	68,79	68,77472687
2018/09/21	69,09	68,20838165
2018/09/24	67,94	67,90527344
2018/09/25	67,14	68,63863373
2018/09/26	67,1	68,86116028
2018/09/27	66,73	67,94826508
2018/09/28	66,95	67,40982819
2018/10/01	67,07	67,35058594
2018/10/02	67,39	66,90284729
2018/10/03	65,92	66,95423889
2018/10/04	64,92	67,02100372
2018/10/05	64,71	67,26831818
2018/10/08	65,35	65,90972137
2018/10/09	65,07	65,03578949
2018/10/10	63,83	64,87999725
2018/10/11	62,35	65,41085815
2018/10/12	62,54	65,05812073
2018/10/15	62,55	63,99930573
2018/10/16	62,61	62,65605545
2018/10/17	63,55	62,82065201
2018/10/18	62,64	62,89979553
2018/10/19	64,52	62,92071915

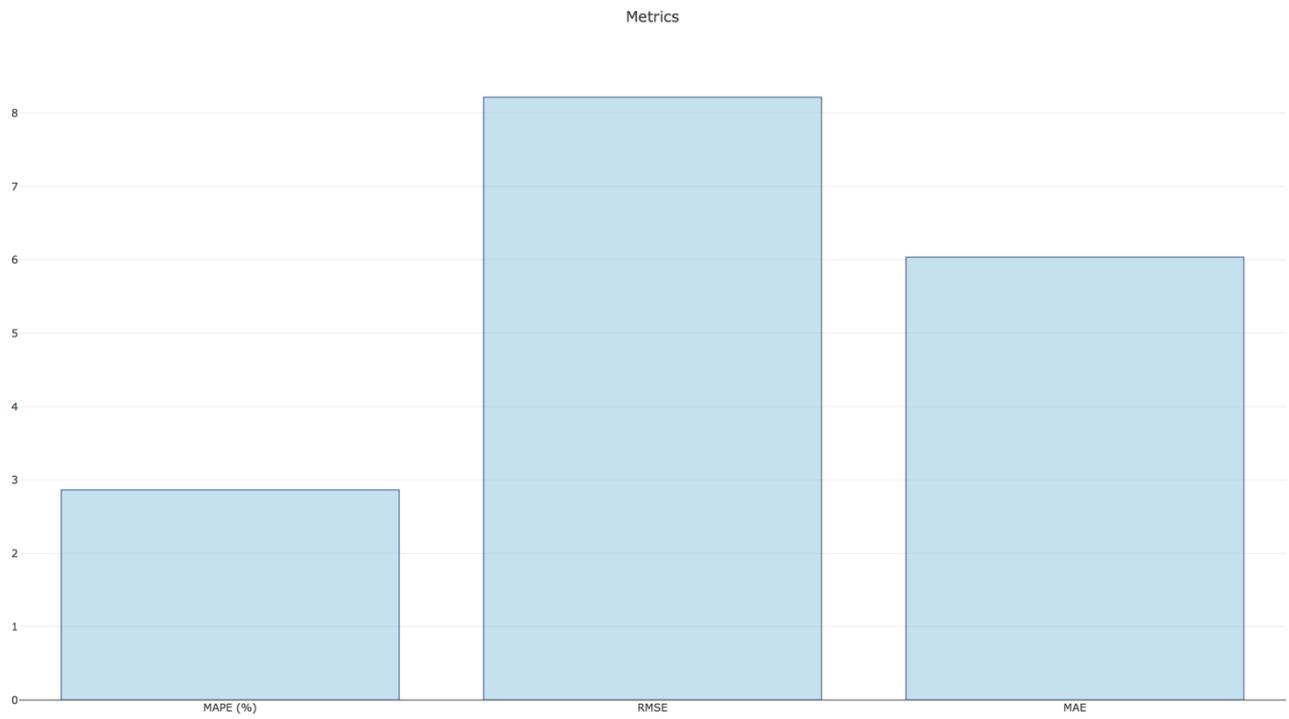
2018/10/22	63,64	63,57052231
2018/10/23	63,4	62,66892624
2018/10/24	64,64	64,37771606
2018/10/25	63,82	63,62778091
2018/10/26	59,58	63,68701553
2018/10/29	57,88	64,58843231
2018/10/30	60,19	63,81982422
2018/10/31	59,55	60,52254486
2018/11/01	60,08	59,12172318
2018/11/02	60,29	60,76337814
2018/11/05	60,23	59,9552269
2018/11/06	60,51	60,2673912
2018/11/07	61,33	60,43107224
2018/11/08	60,9	60,61935425
2018/11/09	63,8	60,80256271
2018/11/12	61,35	61,63759232
2018/11/13	61,86	61,26616669
2018/11/14	62,47	63,84647369
2018/11/15	62,4	61,59563446
2018/11/16	63,24	62,48535919
2018/11/19	62,77	63,05467606
2018/11/20	62,31	63,16298294
2018/11/21	61,92	63,89942169
2018/11/23	61,78	63,40523529
2018/11/26	62,74	63,05300903
2018/11/27	63,24	62,68939209
2018/11/28	62,83	62,4852562
2018/11/29	63,04	63,1633873
2018/11/30	63,52	63,19100952
2018/12/03	63,76	62,95303345
2018/12/04	63,08	63,17763138
2018/12/06	63,23	63,64707947
2018/12/07	62,71	63,77623749
2018/12/10	63,54	63,26544952
2018/12/11	64,55	63,54968262
2018/12/12	65,28	63,14050293
2018/12/13	65,47	63,95616531
2018/12/14	65,17	64,81764984
2018/12/17	62,92	65,43373108
2018/12/18	62,46	65,63352966
2018/12/19	61,88	65,44481659
2018/12/20	60,98	63,49381638
2018/12/21	59,91	63,17706299
2018/12/24	58,04	62,70177841
2018/12/26	59,4	61,81435394

2018/12/27	59,69	60,64826965
2018/12/28	59,62	59,01502609
2018/12/31	59,52	59,9893837
2019/01/02	59,25	59,79640198
2019/01/03	58,97	59,72599792
2019/01/04	59,52	59,7181015
2019/01/07	59,88	59,66991806
2019/01/08	60,55	59,58871841
2019/01/09	61,3	60,10170364
2019/01/10	61,82	60,15354919
2019/01/11	62,1	60,77334595
2019/01/14	61,61	61,37134552
2019/01/15	62,22	61,97336578
2019/01/16	62,13	62,39139175
2019/01/17	62	62,19512939
2019/01/18	62,63	62,68301773
2019/01/22	61,74	62,67789459
2019/01/23	62,35	62,50637817
2019/01/24	62,2	62,94358444
2019/01/25	61,84	62,151577
2019/01/28	62,74	62,70082855
2019/01/29	N/A	62,63338089
2019/01/30	N/A	62,23108595
2019/01/31	N/A	62,7771276
2019/02/01	N/A	62,68853936
2019/02/04	N/A	63,07262688
2019/02/05	N/A	63,05938059
2019/02/06	N/A	63,35020356
2019/02/07	N/A	63,36302066
2019/02/08	N/A	63,59529779
2019/02/11	N/A	63,62778317
2019/02/12	N/A	63,83379168
2019/02/13	N/A	63,88624939
2019/02/14	N/A	64,07291824
2019/02/15	N/A	64,13087526
2019/02/19	N/A	64,28897754
2019/02/13	N/A	64,33883675
2019/02/14	N/A	64,46722757
2019/02/15	N/A	64,50996487
2019/02/16	N/A	64,61858275
2019/02/17	N/A	64,66132882

Apéndice E. Gráfica precio de cierre Predichos y Precio de cierre Reales, de FEDEX (FDX).



Apéndice F. Gráfica indicadores de la corrida del modelo para FEDEX (FDX).



Apéndice G. Tabla de precios de Cierre predichos y Reales para FEDEX (FDX).

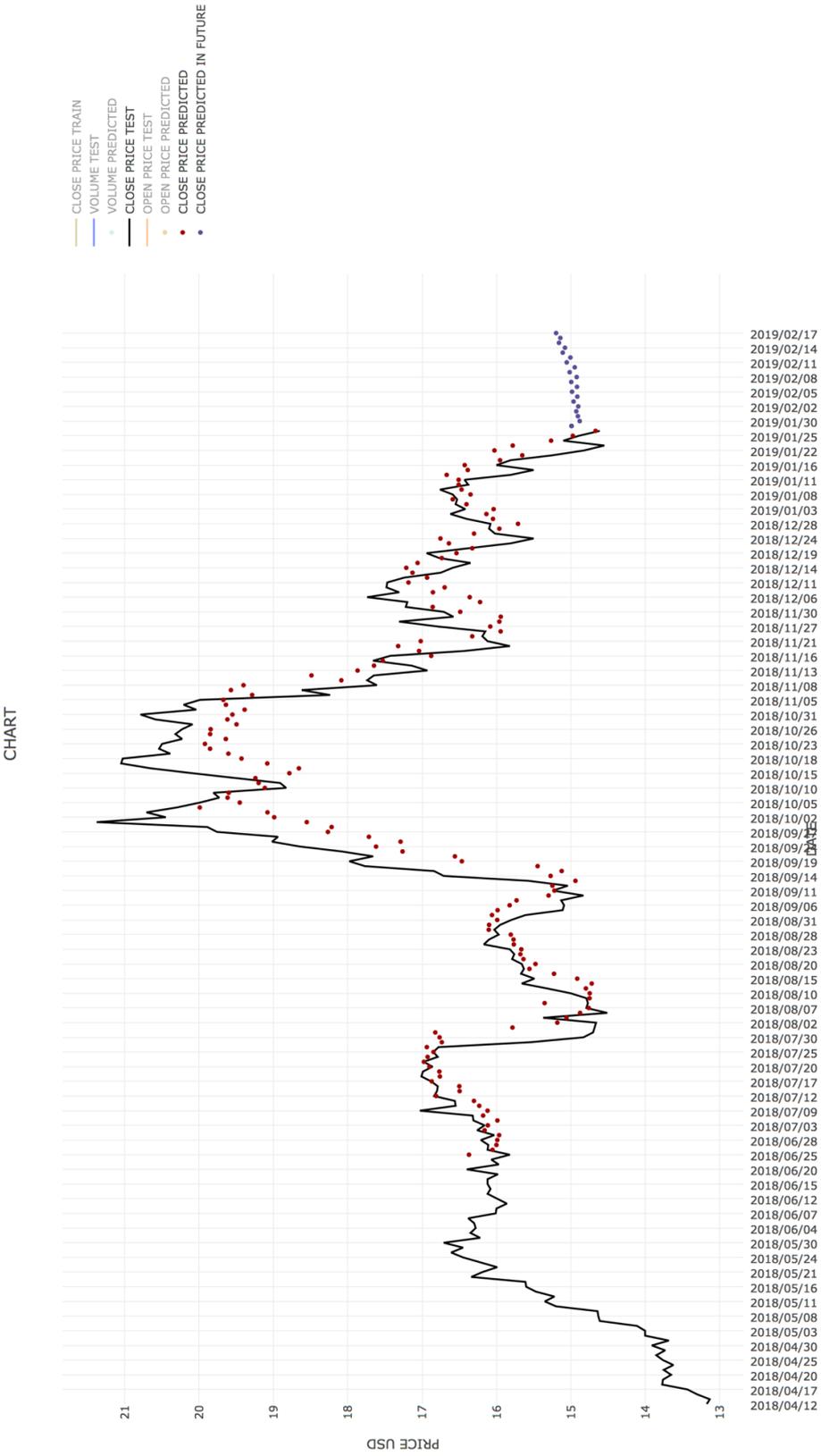
DATE	CLOSE PRICE REAL	CLOSE PRICE PREDICTED
2018/06/25	234,53	254,0751343
2018/06/26	236,85	250,7851257
2018/06/27	229,75	245,937561
2018/06/28	226,67	238,8251495
2018/06/29	227,06	239,6217194
2018/07/02	229,39	233,5113068
2018/07/03	227,61	230,2478943
2018/07/05	228,58	229,7137909
2018/07/06	228,27	231,28862
2018/07/09	234,94	228,1512146
2018/07/10	235,66	228,2098999
2018/07/11	230,22	227,1395569
2018/07/12	232,69	232,8031006
2018/07/13	233,75	232,8724365
2018/07/16	230,18	227,8178253
2018/07/17	231,15	229,1704254
2018/07/18	236,19	229,9489136
2018/07/19	236,55	226,7563934
2018/07/20	234,93	227,2364655
2018/07/23	236,4	232,0273285
2018/07/24	236,83	232,5046539
2018/07/25	240,24	231,2881012
2018/07/26	241,24	232,6856995
2018/07/27	240,84	233,7031708
2018/07/30	240,44	237,1008453
2018/07/31	245,87	238,4280548
2018/08/01	241,1	238,5657501
2018/08/02	241,12	238,3931122
2018/08/03	243,32	243,4736328
2018/08/06	245,13	239,758194
2018/08/07	247,85	240,3813934
2018/08/08	245,49	242,0405426
2018/08/09	244,8	243,815155
2018/08/10	241,76	246,5544739
2018/08/13	240,99	245,0576019
2018/08/14	243,7	244,7915955
2018/08/15	243,48	242,1942291
2018/08/16	246,25	241,4698792
2018/08/17	246,94	243,8195496
2018/08/20	248,79	243,2885742

2018/08/21	251,11	245,687149
2018/08/22	246,94	246,4545441
2018/08/23	244,47	248,2670593
2018/08/24	245,02	250,3147583
2018/08/27	248,9	246,9039764
2018/08/28	247,63	245,0607147
2018/08/29	247,6	245,4026184
2018/08/30	246,74	248,5570984
2018/08/31	243,95	247,0850525
2018/09/04	241,29	247,3729858
2018/09/05	244,76	246,5281525
2018/09/06	246,19	244,2842102
2018/09/07	245,97	242,2604675
2018/09/10	249,45	245,109436
2018/09/11	249,36	245,8262939
2018/09/12	252,98	245,4600983
2018/09/13	253,17	248,19133
2018/09/14	255,44	248,3778076
2018/09/17	255,73	251,9574127
2018/09/18	241,58	252,1967621
2018/09/19	242,88	254,3710632
2018/09/20	246,81	254,6517639
2018/09/21	247,32	242,9541626
2018/09/24	243,84	245,0592651
2018/09/25	239,2	247,9373322
2018/09/26	241,11	246,7796021
2018/09/27	240,7	243,452652
2018/09/28	240,79	240,5713654
2018/10/01	239,93	242,8994141
2018/10/02	239	242,574585
2018/10/03	240,94	242,1080475
2018/10/04	239,07	240,5752563
2018/10/05	236,06	239,8443604
2018/10/08	234,99	241,0974884
2018/10/09	229,75	239,0585175
2018/10/10	221,46	236,2781372
2018/10/11	217,42	234,9111176
2018/10/12	221	230,0340881
2018/10/15	221,02	222,4074097
2018/10/16	226,18	217,8715668
2018/10/17	223,46	220,1133118
2018/10/18	217,29	219,6053467
2018/10/19	219,8	223,1188965
2018/10/22	218,4	220,3306885
2018/10/23	218,38	214,8439941

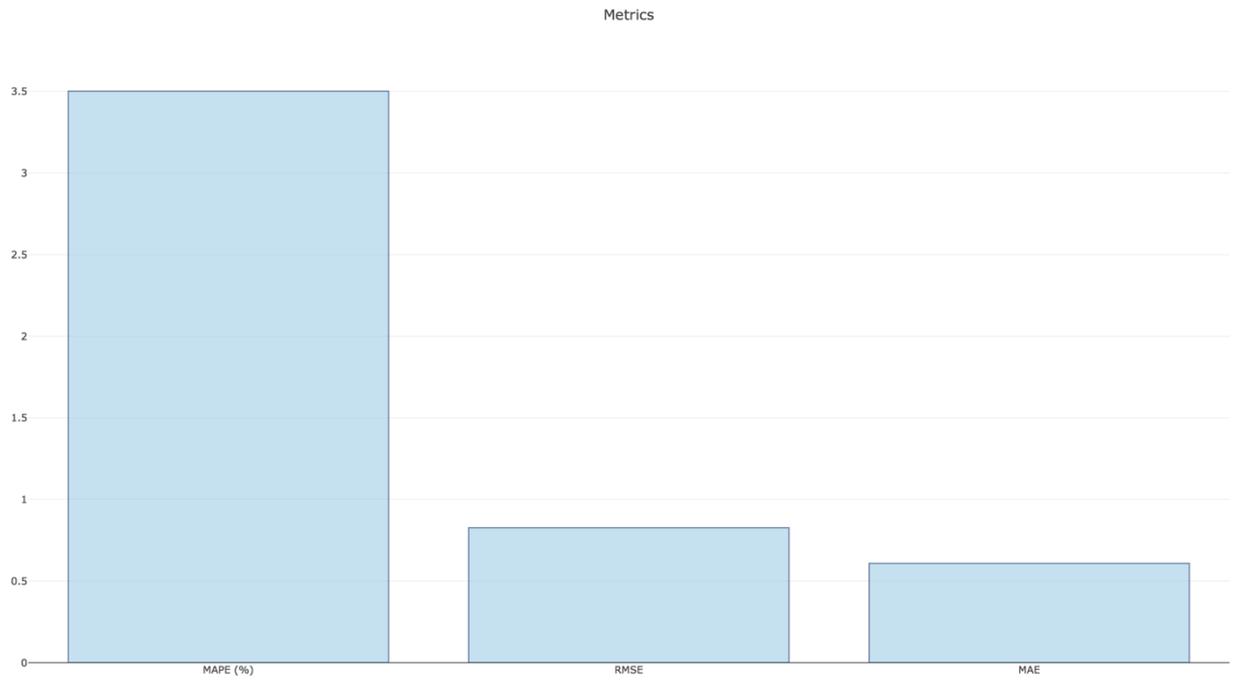
2018/10/24	209,2	216,767807
2018/10/25	213,78	215,8249664
2018/10/26	210,53	215,7957306
2018/10/29	211,04	207,2610779
2018/10/30	216,53	211,1060638
2018/10/31	220,34	208,4117432
2018/11/01	224,04	208,4684296
2018/11/02	221,66	213,496933
2018/11/05	222,61	217,3293457
2018/11/06	226,77	221,2883606
2018/11/07	231,28	219,8152008
2018/11/08	229,4	220,8319855
2018/11/09	224,4	224,6185303
2018/11/12	219,63	228,864502
2018/11/13	222,51	227,6477966
2018/11/14	225,16	223,5757446
2018/11/15	227,81	219,4217377
2018/11/16	226,69	222,0351105
2018/11/19	226,3	224,5170135
2018/11/20	221,34	226,8432922
2018/11/21	222,55	225,9864044
2018/11/23	223,4	226,0260468
2018/11/26	225,04	221,9555664
2018/11/27	225,23	222,6492767
2018/11/28	231,85	223,58815
2018/11/29	227,5	224,9312897
2018/11/30	229	224,9557648
2018/12/03	230,04	231,0038757
2018/12/04	215,52	227,5814667
2018/12/06	215,06	228,9316254
2018/12/07	201,39	229,7209473
2018/12/10	192,93	217,6602173
2018/12/11	189,65	216,1018219
2018/12/12	188,27	204,4473114
2018/12/13	187,18	195,6785889
2018/12/14	184,11	191,4935303
2018/12/17	182,19	190,4008026
2018/12/18	185,01	188,3434906
2018/12/19	162,51	184,5515594
2018/12/20	159,72	182,2646484
2018/12/21	158	184,2226868
2018/12/24	152,7	164,3846893
2018/12/26	160,15	159,8492126
2018/12/27	161,02	160,4719543
2018/12/28	158,98	155,6810303

2018/12/31	161,33	161,2766876
2019/01/02	163,19	161,129776
2019/01/03	157,19	158,7828064
2019/01/04	164,43	159,6039429
2019/01/07	166,73	160,7778625
2019/01/08	167,02	155,653183
2019/01/09	170,59	161,9048462
2019/01/10	170,12	165,1468811
2019/01/11	170,99	166,3316193
2019/01/14	171,55	169,6370239
2019/01/15	170,63	169,1995239
2019/01/16	171,36	170,1564026
2019/01/17	173,27	170,4475708
2019/01/18	176,91	169,9911194
2019/01/22	175,9	170,6253204
2019/01/23	173,68	172,2723389
2019/01/24	174,1	176,1100464
2019/01/25	175,36	175,3140869
2019/01/28	173,96	173,485672
2019/01/29	N/A	173,313797
2019/01/30	N/A	174,370077
2019/01/31	N/A	172,739476
2019/02/01	N/A	173,5953109
2019/02/02	N/A	172,2953588
2019/02/03	N/A	173,0393783
2019/02/04	N/A	172,0144692
2019/02/05	N/A	172,6483265
2019/02/06	N/A	171,8220445
2019/02/07	N/A	172,3410461
2019/02/08	N/A	171,6525153
2019/02/09	N/A	172,0582017
2019/02/10	N/A	171,464694
2019/02/11	N/A	171,7620469
2019/02/12	N/A	171,2291585
2019/02/13	N/A	171,4219345
2019/02/14	N/A	170,9282427
2019/02/15	N/A	171,0248649
2019/02/16	N/A	170,5569603
2019/02/17	N/A	170,5715183

Apéndice H. Gráfica precio de cierre Predichos y Precio de cierre Reales, de Rosetta Stone (RST).



Apéndice I. Gráfica indicadores de la corrida del modelo para Rosetta Stone (RST).



Apéndice J. Tabla de precios de Cierre predichos y Reales para Rosetta Stone (RST).

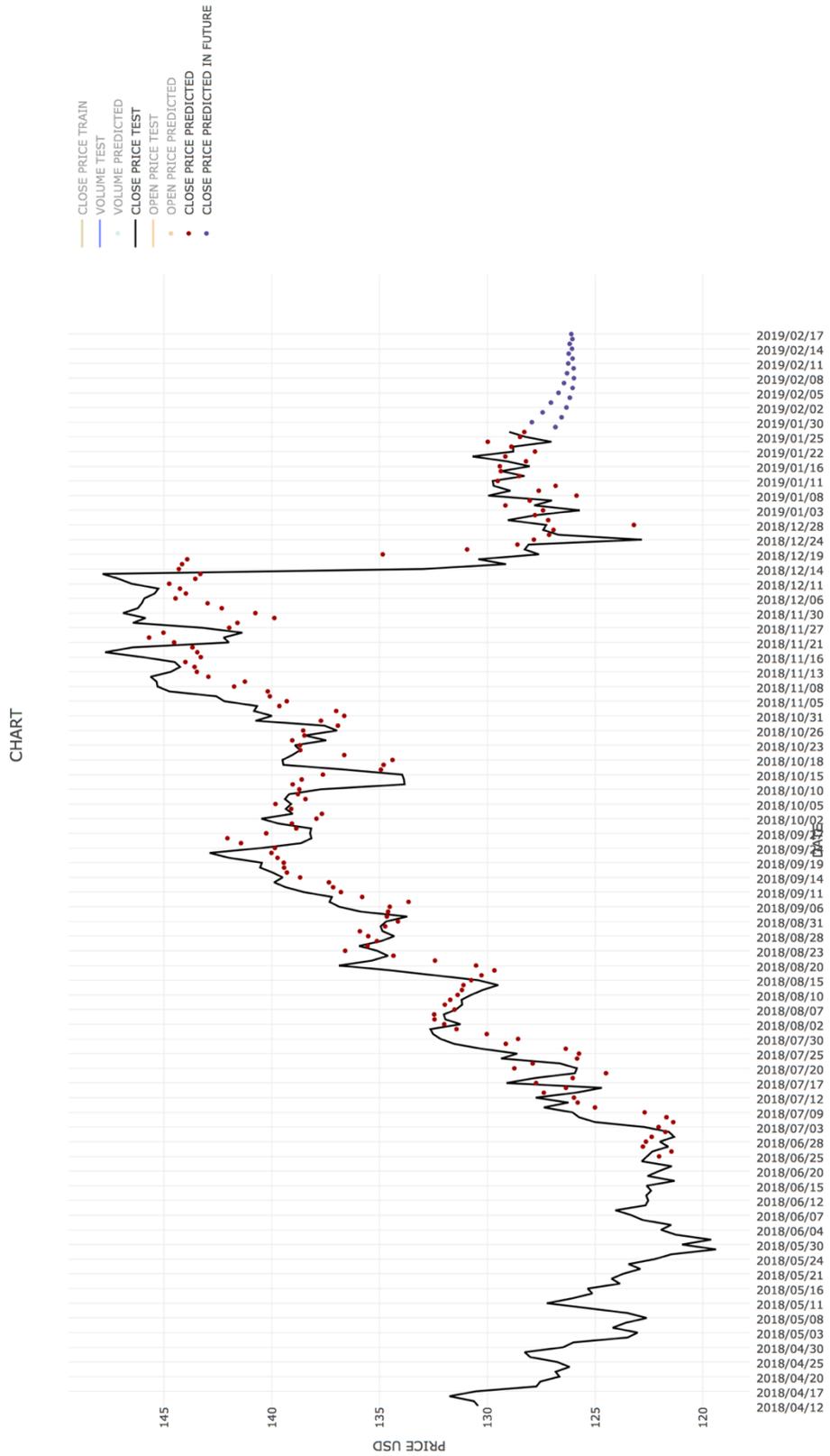
DATE	CLOSE PRICE REAL	CLOSE PRICE PREDICTED
2018/06/25	15,82	16,37018204
2018/06/26	16,12	16,05334854
2018/06/27	16,11	16,00100136
2018/06/28	16,21	15,98917198
2018/06/29	16,03	15,96400166
2018/07/02	16,26	16,16017151
2018/07/03	16,16	16,11661339
2018/07/05	16,31	15,98783112
2018/07/06	16,32	16,1799736
2018/07/09	17,03	16,12051582
2018/07/10	16,55	16,23184013
2018/07/11	16,56	16,30323792
2018/07/12	16,84	16,81195259
2018/07/13	16,8	16,49612808
2018/07/16	16,79	16,50009918
2018/07/17	16,88	16,87106133
2018/07/18	17,01	16,76146507
2018/07/19	16,99	16,76918411
2018/07/20	16,86	16,90307617
2018/07/23	16,98	16,97655106
2018/07/24	16,79	16,92680359
2018/07/25	16,85	16,84740829
2018/07/26	16,78	16,93699074
2018/07/27	15,54	16,73517418
2018/07/30	14,83	16,76537323
2018/07/31	14,7	16,82356453
2018/08/01	14,68	15,78450871
2018/08/02	14,66	15,18134594
2018/08/03	15,37	15,05784798
2018/08/06	14,51	14,87512016
2018/08/07	14,79	14,75829315
2018/08/08	14,77	15,35233116
2018/08/09	14,79	14,74787045
2018/08/10	15	14,74525642
2018/08/13	15,32	14,79690075
2018/08/14	15,66	14,71846104
2018/08/15	15,49	14,91350937
2018/08/16	15,67	15,2259388
2018/08/17	15,63	15,55619717
2018/08/20	15,66	15,47519493

2018/08/21	15,79	15,63681507
2018/08/22	15,76	15,67834282
2018/08/23	15,82	15,66627502
2018/08/24	16,17	15,7664032
2018/08/27	16,1	15,77057743
2018/08/28	15,97	15,80741119
2018/08/29	16,03	16,10404205
2018/08/30	15,95	16,09960938
2018/08/31	15,8	15,99018764
2018/09/04	15,61	16,06056023
2018/09/05	15,11	15,98490906
2018/09/06	15,09	15,82296753
2018/09/07	15,13	15,72968769
2018/09/10	14,83	15,29883194
2018/09/11	15,24	15,22004509
2018/09/12	15,04	15,2478075
2018/09/13	15,57	14,93731213
2018/09/14	16,71	15,27157211
2018/09/17	16,84	15,12268639
2018/09/18	17,77	15,44638443
2018/09/19	17,98	16,46592712
2018/09/20	17,66	16,55899429
2018/09/21	18,08	17,26360512
2018/09/24	18,64	17,62042809
2018/09/25	19,02	17,29063034
2018/09/26	18,94	17,71674919
2018/09/27	19,76	18,26809692
2018/09/28	19,89	18,21834564
2018/10/01	21,38	18,55264091
2018/10/02	20,45	18,99086571
2018/10/03	20,71	19,0812912
2018/10/04	20,3	19,9923439
2018/10/05	19,99	19,45428658
2018/10/08	19,73	19,61805344
2018/10/09	19,81	19,60230637
2018/10/10	18,83	19,11666489
2018/10/11	18,91	19,20059586
2018/10/12	19,53	19,24379539
2018/10/15	20,09	18,78682899
2018/10/16	20,65	18,65862656
2018/10/17	21,05	19,08414078
2018/10/18	21,03	19,43086052
2018/10/19	20,39	19,60734558
2018/10/22	20,54	19,85391617
2018/10/23	20,5	19,92291832

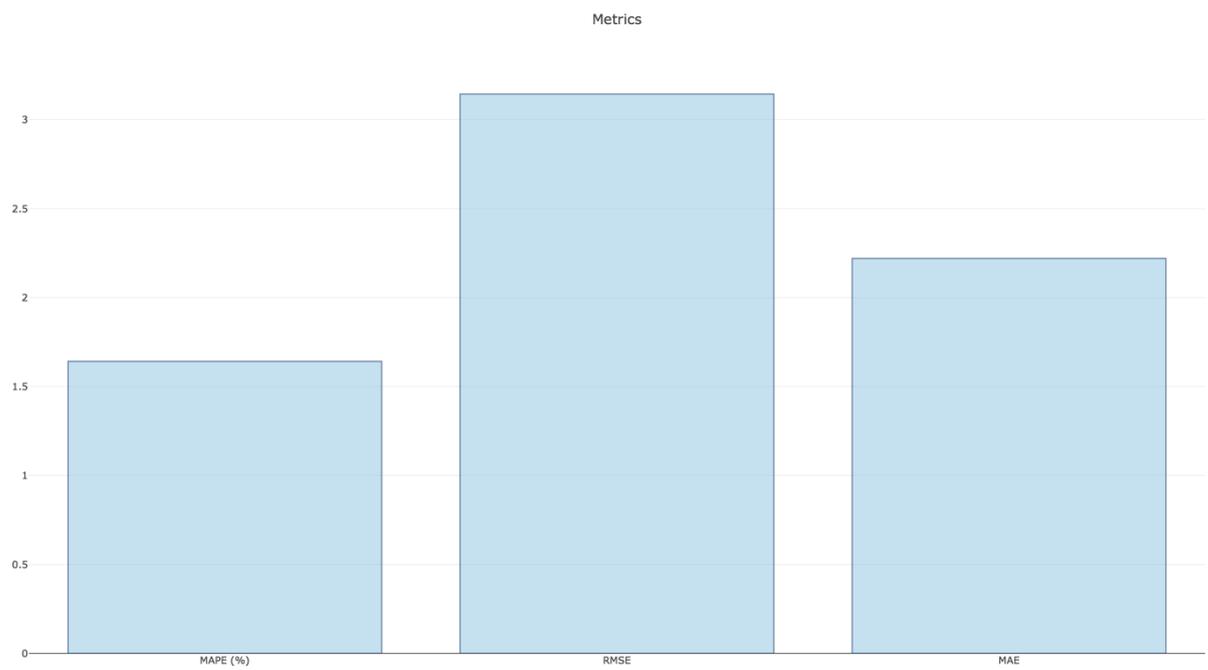
2018/10/24	20,23	19,64178085
2018/10/25	20,32	19,85268974
2018/10/26	20,23	19,84610176
2018/10/29	20,09	19,49801064
2018/10/30	20,59	19,62052917
2018/10/31	20,79	19,55293083
2018/11/01	20,04	19,38902473
2018/11/02	20,21	19,64050674
2018/11/05	19,99	19,67515945
2018/11/06	18,24	19,28769684
2018/11/07	18,62	19,57278633
2018/11/08	17,61	19,40382957
2018/11/09	17,74	18,0871582
2018/11/12	17,65	18,48977852
2018/11/13	16,93	17,86891937
2018/11/14	17,14	17,64805603
2018/11/15	17,66	17,52946663
2018/11/16	17,43	16,87835312
2018/11/19	16,43	17,04154396
2018/11/20	15,82	17,32319641
2018/11/21	16,12	17,01940346
2018/11/23	16,19	16,32647324
2018/11/26	16,15	15,9439888
2018/11/27	16,79	16,08368111
2018/11/28	17,31	15,96140957
2018/11/29	16,58	15,94252777
2018/11/30	16,71	16,4877243
2018/12/03	17,22	16,85812759
2018/12/04	17,2	16,22119331
2018/12/06	17,74	16,36019135
2018/12/07	17,31	16,85521698
2018/12/10	17,48	16,69778442
2018/12/11	17,47	17,18448067
2018/12/12	17,24	16,93329811
2018/12/13	16,75	17,12995911
2018/12/14	16,59	17,21369743
2018/12/17	16,35	17,06152916
2018/12/18	16,77	16,7357769
2018/12/19	16,94	16,53829193
2018/12/20	16,34	16,32609367
2018/12/21	15,81	16,64029694
2018/12/24	15,5	16,7538681
2018/12/26	16,02	16,302351
2018/12/27	16,1	15,96105003
2018/12/28	16,08	15,71055317

2018/12/31	16,4	16,04681778
2019/01/02	16,62	16,13539696
2019/01/03	16,42	16,03858948
2019/01/04	16,55	16,40384293
2019/01/07	16,53	16,58883286
2019/01/08	16,59	16,34898567
2019/01/09	16,76	16,46988678
2019/01/10	16,38	16,50866318
2019/01/11	16,43	16,50895882
2019/01/14	15,81	16,67013359
2019/01/15	15,5	16,38667679
2019/01/16	16	16,42723465
2019/01/17	15,81	15,95193195
2019/01/18	15,25	15,65282822
2019/01/22	14,82	16,0274334
2019/01/23	14,55	15,7814436
2019/01/24	15,1	15,26516151
2019/01/25	14,89	14,97277355
2019/01/28	14,61	14,66684246
2019/01/29	N/A	14,98973083
2019/01/30	N/A	14,88018501
2019/01/31	N/A	14,90479607
2019/02/01	N/A	14,92604453
2019/02/02	N/A	14,89965964
2019/02/03	N/A	14,96247608
2019/02/04	N/A	14,9124949
2019/02/05	N/A	14,9834894
2019/02/06	N/A	14,91665948
2019/02/07	N/A	14,99400098
2019/02/08	N/A	14,92059249
2019/02/09	N/A	15,01525391
2019/02/10	N/A	14,94674224
2019/02/11	N/A	15,05472261
2019/02/12	N/A	15,00538874
2019/02/13	N/A	15,10828901
2019/02/14	N/A	15,07844586
2019/02/15	N/A	15,15881467
2019/02/16	N/A	15,14063198
2019/02/17	N/A	15,19754935

Apéndice K. Gráfica precio de cierre Predichos y Precio de cierre Reales, de Jhonson & Jhonson (JNJ).



Apéndice L. Gráfica indicadores de la corrida del modelo para Jhonson & Jhonson (JNJ).



Apéndice M. Tabla de precios de Cierre predichos y Reales para Jhonson & Jhonson
(JNJ).

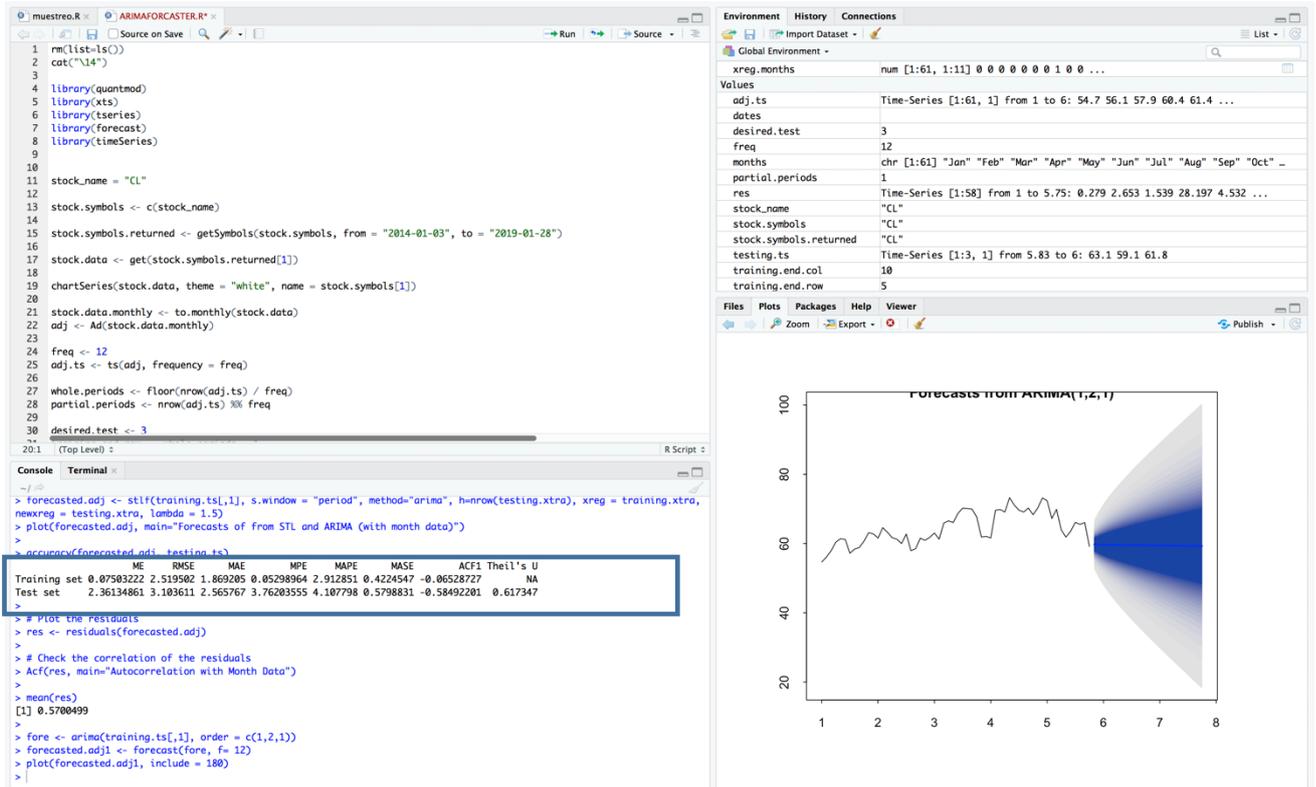
DATE	CLOSE PRICE REAL	CLOSE PRICE PREDICTED
2018/06/25	122,59	122,0369568
2018/06/26	122,35	121,4634552
2018/06/27	121,61	122,7885513
2018/06/28	122	122,6466675
2018/06/29	121,34	122,3818054
2018/07/02	121,58	121,7463379
2018/07/03	122,71	122,0607986
2018/07/05	125	121,3807526
2018/07/06	125,75	121,6926346
2018/07/09	126,05	122,7051468
2018/07/10	127,38	125,0096588
2018/07/11	126,24	125,8128967
2018/07/12	127,76	125,9867477
2018/07/13	125,93	127,3808899
2018/07/16	124,69	126,3576355
2018/07/17	129,11	127,7411804
2018/07/18	127,8	126,0415573
2018/07/19	125,94	124,4995422
2018/07/20	125,85	128,7502289
2018/07/23	126,63	127,8991928
2018/07/24	129,36	125,8397293
2018/07/25	128,62	125,752739
2018/07/26	130,28	126,3664627
2018/07/27	131,55	129,1471252
2018/07/30	132,17	128,577301
2018/07/31	132,52	130,0286407
2018/08/01	132,64	131,432724
2018/08/02	131,25	132,0013275
2018/08/03	131,95	132,4516144
2018/08/06	132,02	132,4681702
2018/08/07	131,46	131,5243073
2018/08/08	131,17	131,9744415
2018/08/09	131,2	131,721817
2018/08/10	130,75	131,3791656
2018/08/13	130,22	131,1800537
2018/08/14	129,5	131,1103821
2018/08/15	130,43	130,7507935
2018/08/16	132,48	130,2728729
2018/08/17	134,47	129,6745605

2018/08/20	136,88	130,5256653
2018/08/21	135,35	132,4302521
2018/08/22	134,61	134,3550262
2018/08/23	135,11	136,5971069
2018/08/24	135,95	135,5595856
2018/08/27	134,92	135,1227722
2018/08/28	134,32	135,5263672
2018/08/29	134,86	135,9139404
2018/08/30	134,95	134,7392731
2018/08/31	134,69	134,1483154
2018/09/04	133,72	134,6617126
2018/09/05	135,86	134,6060333
2018/09/06	136,87	134,5262756
2018/09/07	137,32	133,6557465
2018/09/10	137,21	135,806427
2018/09/11	138,51	136,7932129
2018/09/12	139,36	137,1549835
2018/09/13	139,89	137,3473053
2018/09/14	139,49	138,686264
2018/09/17	139,96	139,2996368
2018/09/18	140,54	139,4276581
2018/09/19	140,46	139,4423065
2018/09/20	141,98	139,7342682
2018/09/21	142,88	140,0130615
2018/09/24	140,47	139,8554535
2018/09/25	138,64	141,4290771
2018/09/26	138,16	142,0576324
2018/09/27	138,22	140,2574768
2018/09/28	138,17	138,865448
2018/10/01	139,69	139,065567
2018/10/02	140,48	137,926651
2018/10/03	139,03	137,6740723
2018/10/04	139,35	139,0986328
2018/10/05	139,1	139,8270569
2018/10/08	139,39	138,4376373
2018/10/09	139,2	138,7846985
2018/10/10	137,73	138,7235565
2018/10/11	133,84	139,0282745
2018/10/12	133,87	138,6104889
2018/10/15	133,95	137,6275787
2018/10/16	136,56	134,9374847
2018/10/17	139,46	134,8111267
2018/10/18	139,5	134,3982849
2018/10/19	139,05	136,6379242
2018/10/22	138,68	138,6763763

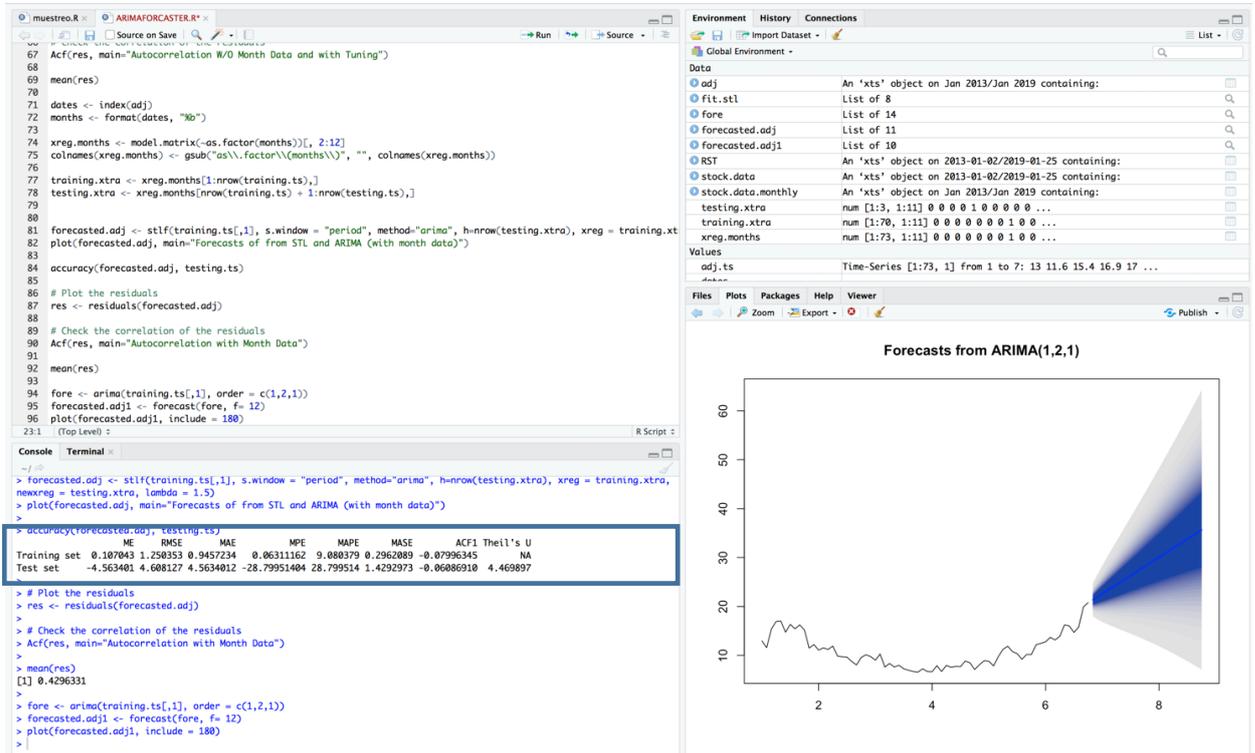
2018/10/23	138,93	138,6925964
2018/10/24	137,47	139,0487518
2018/10/25	138,47	138,4854126
2018/10/26	136,97	138,54245
2018/10/29	137,55	136,9322052
2018/10/30	140,75	137,7177124
2018/10/31	139,99	136,6384125
2018/11/01	140,82	137,0124359
2018/11/02	140,68	139,6463013
2018/11/05	142,2	139,3060303
2018/11/06	142,57	140,0868835
2018/11/07	144,75	140,1866455
2018/11/08	145,3	141,7449951
2018/11/09	145,34	141,2479706
2018/11/12	145,62	142,940094
2018/11/13	144,68	143,4727478
2018/11/14	144,25	143,580246
2018/11/15	144,5	144,0068359
2018/11/16	145,99	143,2908783
2018/11/19	147,73	143,4523163
2018/11/20	146,45	143,6761475
2018/11/21	141,99	144,5277557
2018/11/23	142,23	145,690506
2018/11/26	141,37	145,0241699
2018/11/27	143,22	141,9709015
2018/11/28	146,44	141,5890808
2018/11/29	145,85	139,878891
2018/11/30	146,9	140,7597961
2018/12/03	146,21	142,3188171
2018/12/04	146,03	142,9757996
2018/12/06	145,91	144,4634552
2018/12/07	145,43	143,983078
2018/12/10	145,26	144,2565002
2018/12/11	146,5	144,7569122
2018/12/12	147,1	143,5452728
2018/12/13	147,84	143,3139191
2018/12/14	133	144,3085938
2018/12/17	129,14	144,1614227
2018/12/18	130,42	143,9230804
2018/12/19	127,61	134,8525848
2018/12/20	128,27	130,9401703
2018/12/21	128,09	128,6053162
2018/12/24	122,84	127,8431854
2018/12/26	126,71	127,138504
2018/12/27	127,41	126,9317932

2018/12/28	127,27	123,2063904
2018/12/31	129,05	127,1799011
2019/01/02	127,75	127,7920456
2019/01/03	125,72	127,4195099
2019/01/04	127,83	129,1648712
2019/01/07	127,01	128,0340271
2019/01/08	129,96	125,865181
2019/01/09	128,93	127,616066
2019/01/10	129,71	126,837616
2019/01/11	129,75	129,5211792
2019/01/14	128,28	128,5293884
2019/01/15	129,36	129,3769531
2019/01/16	128,04	129,4236298
2019/01/17	129,09	128,2086487
2019/01/18	130,69	129,1677704
2019/01/22	128,8	127,7909317
2019/01/23	128,8	128,8873901
2019/01/24	127,03	129,9826355
2019/01/25	128,23	128,4856873
2019/01/28	128,99	128,2877045
2019/01/29	N/A	126,8447113
2019/01/30	N/A	127,9362316
2019/01/31	N/A	126,5592514
2019/02/01	N/A	127,4393184
2019/02/02	N/A	126,3330627
2019/02/03	N/A	127,0526485
2019/02/04	N/A	126,1773093
2019/02/05	N/A	126,7008466
2019/02/06	N/A	126,0460752
2019/02/07	N/A	126,4469284
2019/02/08	N/A	125,9868472
2019/02/09	N/A	126,3049398
2019/02/10	N/A	126,0006152
2019/02/11	N/A	126,2495183
2019/02/12	N/A	126,0473243
2019/02/13	N/A	126,2270631
2019/02/14	N/A	126,074067
2019/02/15	N/A	126,1827678
2019/02/16	N/A	126,056249
2019/02/17	N/A	126,1068115

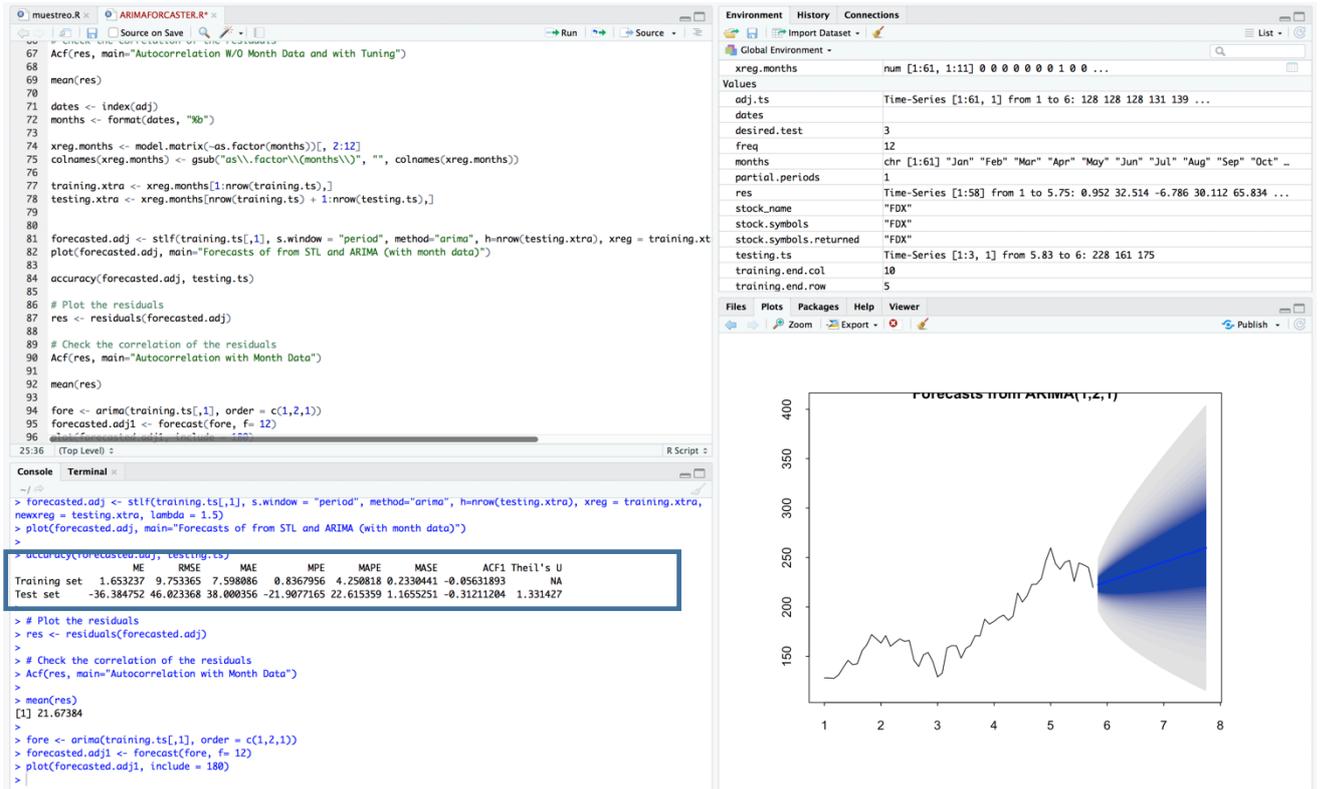
Apéndice N. Resultados de corrida de modelo ARIMA para Colgate Palmolive (CL).



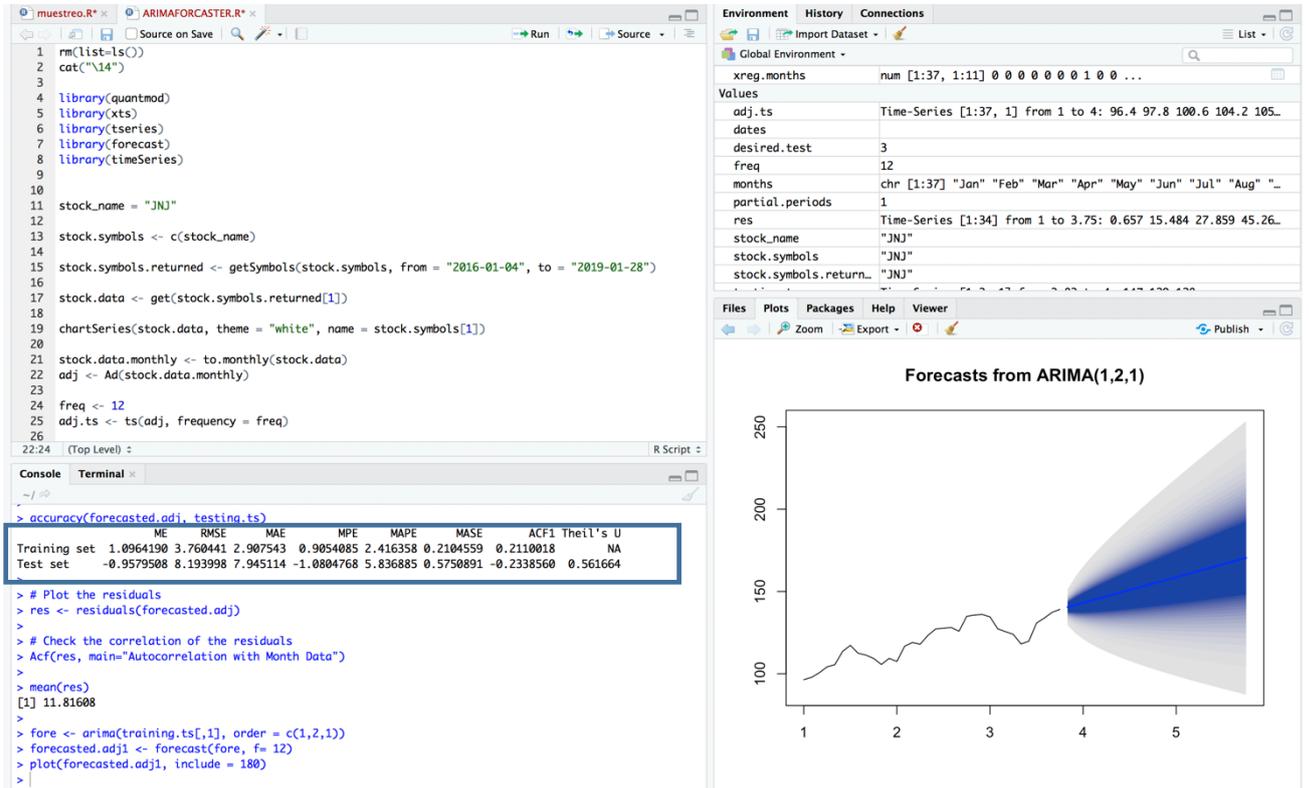
Apéndice O. Resultados de corrida de modelo ARIMA para Rosetta Stone (RST).



Apéndice P. Resultados de corrida de modelo ARIMA para FEDEX (FDX).



Apéndice Q. Resultados de corrida de modelo ARIMA para Jhonson & Jhonson (JNJ).



Apéndice R. Manual de Usuario.

Visit our [historical Nasdaq Official Close Price \(NOCP\)](#) page to view the history.

Get up to 10 years of daily historical stock prices & volumes.
Select the Timeframe: **3 Months**

Results for: 3 Month, From 04-NOV-2018 TO 04-FEB-2019

Date	Open	High	Low	Close / Last	Volume
16:00	65.33	65.43	64.61	65.19	3,612,628
02/04/2019	65.29	65.43	64.61	65.19	3,603,607
02/01/2019	64.9	65.25	64.2	65.05	5,314,498
01/31/2019	63.76	64.85	63.44	64.68	6,182,759
01/30/2019	63.66	64.34	63.46	63.95	6,651,783
01/29/2019	62.7	63.8	62.7	63.69	4,593,131
01/28/2019	61.71	62.84	61.6	62.74	5,752,575

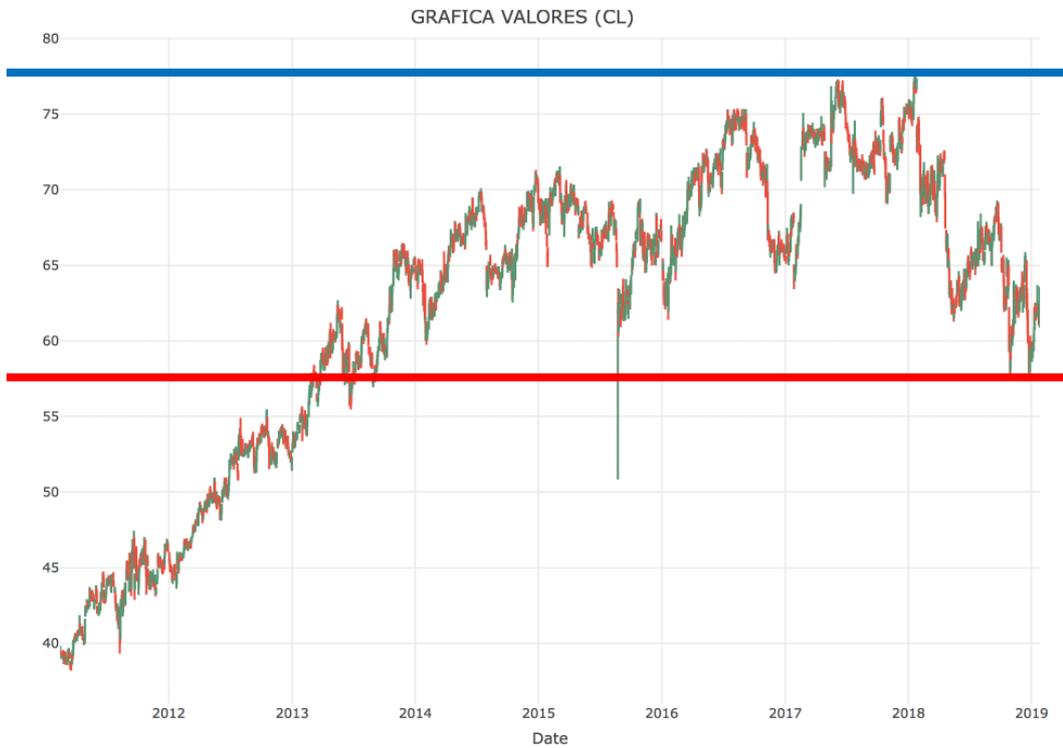
1) Capturar data histórica requerida de una fuente confiable.

```

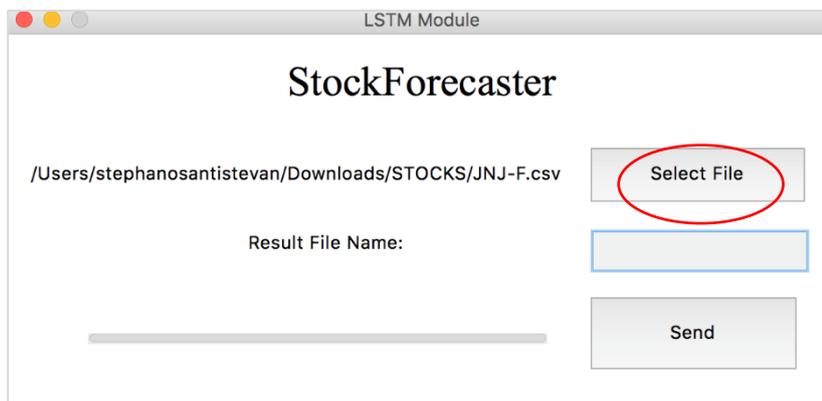
FDX.csv
"date","close","volume","open","high","low"
"2019/01/28","173.9600","1322336.0000","173.5000","174.3300","171.7100"
"2019/01/25","175.3600","1968596.0000","175.8900","177.5600","175.0200"
"2019/01/24","174.1000","1629835.0000","173.3700","174.2800","171.7900"
"2019/01/23","173.6800","1853042.0000","176.8400","177.0800","171.2800"
"2019/01/22","175.9000","2105923.0000","176.3100","176.7200","174.0200"
"2019/01/18","176.9100","3071151.0000","175.1800","179.1900","174.5500"
"2019/01/17","173.2700","1786065.0000","170.9600","175.0000","170.7300"
"2019/01/16","171.3600","2336062.0000","171.1900","174.2295","170.9400"
"2019/01/15","170.6300","2124765.0000","171.4700","172.6500","169.0200"
"2019/01/14","171.5500","2223815.0000","169.8700","173.1000","169.1200"
"2019/01/11","170.9900","2836437.0000","169.4900","171.1400","167.2000"

```

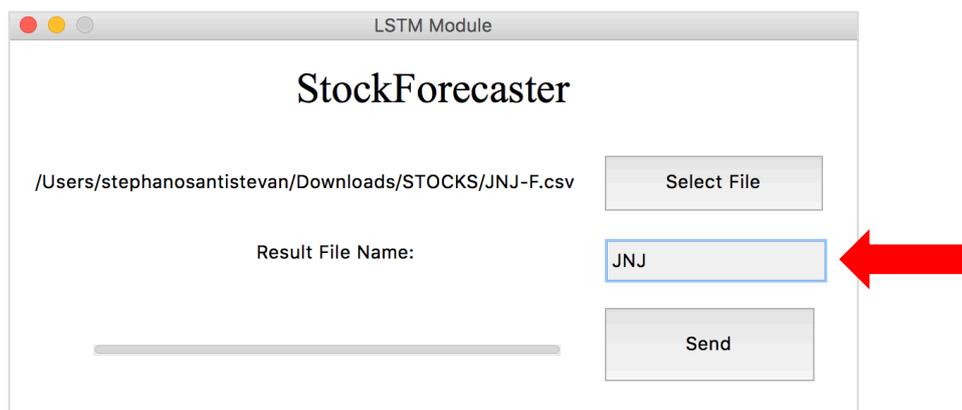
2) Revisar que la data capturada tenga la estructura requerida por el modelo.



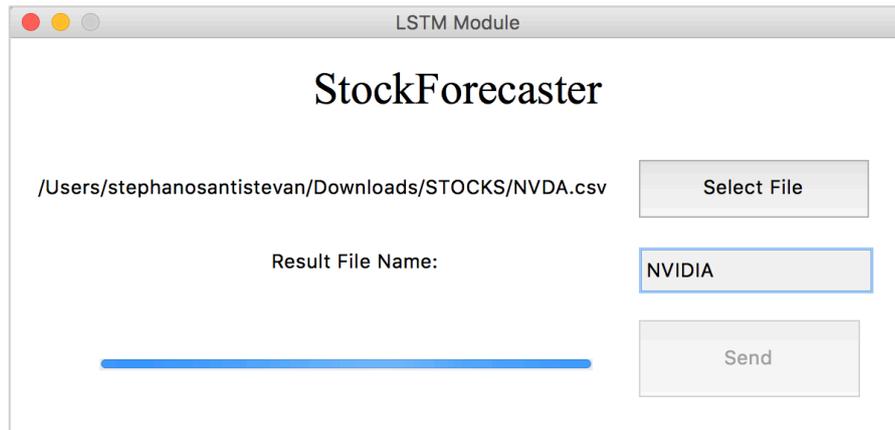
3) Analizar la data captura y delimitar el rango de fechas.



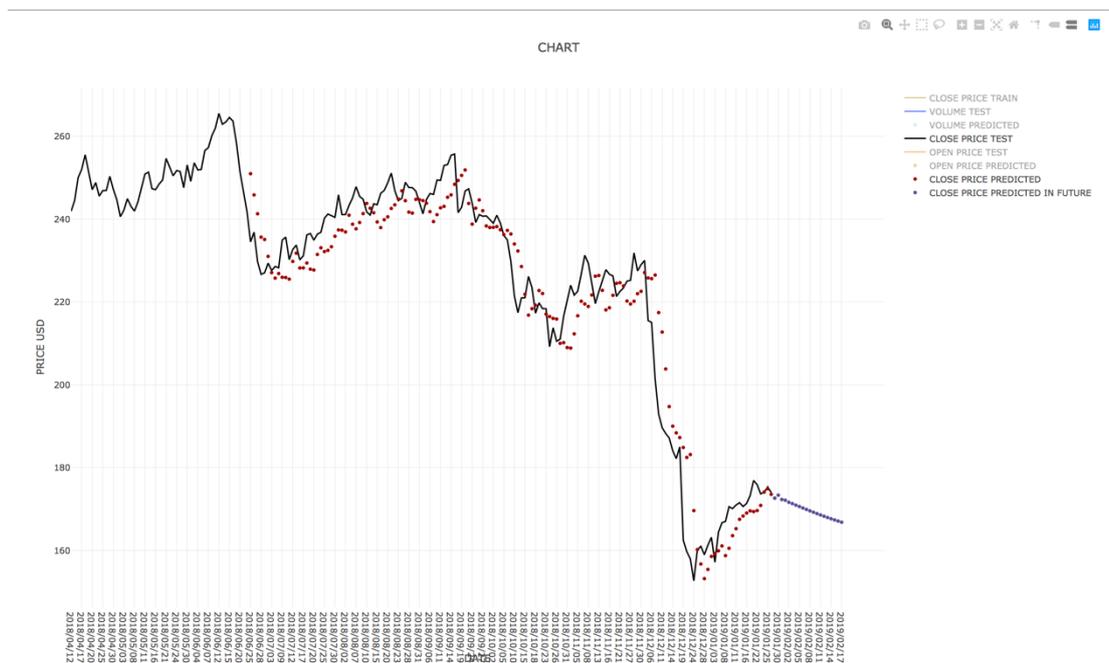
4) Seleccionar el archivo .csv que contenga la data histórica.



5) Ingresar un nombre para los archivos de respaldo.



6) Enviar la información ingresada para ser procesada.



7) Interpretar resultados.

Apéndice S. Indicadores de Predicciones del modelo Colgate Palmolive (CL).

FECHA	REAL	PREDICHO	MAPE	MAE	RMSE
29/1/19	63,69	62,63338089	1,65900316	1,05661911	1,11644394
30/1/19	63,95	62,23108595	2,68790313	1,71891405	2,95466551
31/1/19	64,68	62,7771276	2,94197959	1,9028724	3,62092337
1/2/19	65,05	62,68853936	3,63022389	2,36146064	5,57649635
4/2/19	65,19	63,07262688	3,24800295	2,11737312	4,48326893
5/2/19	64,74	63,05938059	2,59595213	1,68061941	2,8244816
6/2/19	65,33	63,35020356	3,03045529	1,97979644	3,91959394
7/2/19	65,05	63,36302066	2,59335794	1,68697934	2,84589929
8/2/19	65,69	63,59529779	3,18876878	2,09470221	4,38777735
11/2/19	65,79	63,62778317	3,28654329	2,16221683	4,67518162
12/2/19	66,02	63,83379168	3,31143338	2,18620832	4,77950682
13/2/19	65,94	63,88624939	3,11457478	2,05375061	4,21789157
14/2/19	65,45	64,07291824	2,10402102	1,37708176	1,89635417
15/2/19	66,48	64,13087526	3,53358114	2,34912474	5,51838704
19/2/19	66,72	64,28897754	3,64361879	2,43102246	5,9098702
	FINAL		2,97129462	1,9439161	1,97866524

Apéndice T. Indicadores de Predicciones del modelo Fedex (FDX).

FECHA	REAL	PREDICHO	MAPE	MAE	RMSE
29/1/19	173,58	173,313797	0,15336041	0,266203	0,07086404
30/1/19	177,47	174,370077	1,74673072	3,099923	9,60952261
31/1/19	177,57	172,739476	2,72034916	4,830524	23,3339621
1/2/19	179,31	173,5953109	3,18704428	5,7146891	32,6576715
4/2/19	182,73	172,2953588	5,71041493	10,4346412	108,881737
5/2/19	183,73	173,0393783	5,81865874	10,6906217	114,289392
6/2/19	183,81	172,0144692	6,41724106	11,7955308	139,134547
7/2/19	181,37	172,6483265	4,80877405	8,7216735	76,0675886
8/2/19	177,24	171,8220445	3,05684693	5,4179555	29,3542418
11/2/19	181,09	172,3410461	4,8312739	8,7489539	76,5441943
12/2/19	185,45	171,6525153	7,44000253	13,7974847	190,370584
13/2/19	185,62	172,0582017	7,30621609	13,5617983	183,922373
14/2/19	183,92	171,464694	6,77213245	12,455306	155,134648
15/2/19	179,3	171,7620469	4,204101	7,5379531	56,8207369
19/2/19	180,25	171,2291585	5,00462774	9,0208415	81,3755814
	FINAL		4,6118516	8,40627329	9,22882313

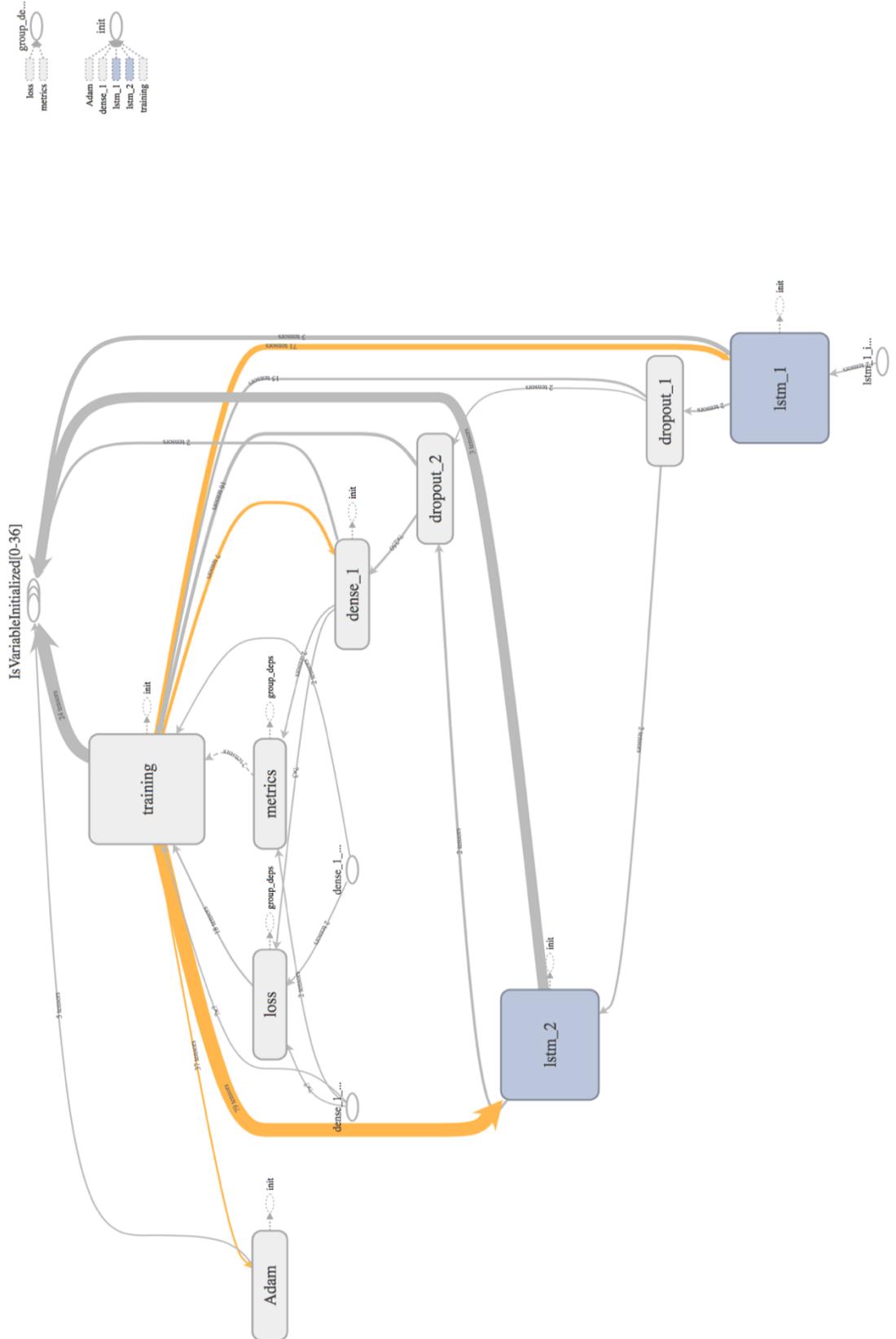
Apéndice U. Indicadores de Predicciones del modelo Rosetta Stone (RST).

FECHA	REAL	PREDICHO	MAPE	MAE	RMSE
29/1/19	14,63	14,98973083	2,45885735	0,35973083	0,12940627
30/1/19	14,83	14,88018501	0,33840196	0,05018501	0,00251854
31/1/19	15,16	14,90479607	1,68340323	0,25520393	0,06512905
1/2/19	15,4	14,92604453	3,07763292	0,47395547	0,22463379
4/2/19	15,59	14,89965964	4,42809724	0,69034036	0,47656981
5/2/19	15,53	14,96247608	3,65437167	0,56752392	0,3220834
6/2/19	15,32	14,9124949	2,65995496	0,4075051	0,16606041
7/2/19	14,91	14,9834894	0,49288665	0,0734894	0,00540069
8/2/19	15,42	14,91665948	3,26420571	0,50334052	0,25335168
11/2/19	15,48	14,99400098	3,13952855	0,48599902	0,23619505
12/2/19	15,84	14,92059249	5,80434034	0,91940751	0,84531017
13/2/19	16,21	15,01525391	7,37042622	1,19474609	1,42741822
14/2/19	16,31	14,94674224	8,35841668	1,36325776	1,85847172
15/2/19	16,26	15,05472261	7,41253007	1,20527739	1,45269359
19/2/19	16,05	15,00538874	6,50848137	1,04461126	1,09121268
	FINAL		2,39727854	0,63963824	0,75526839

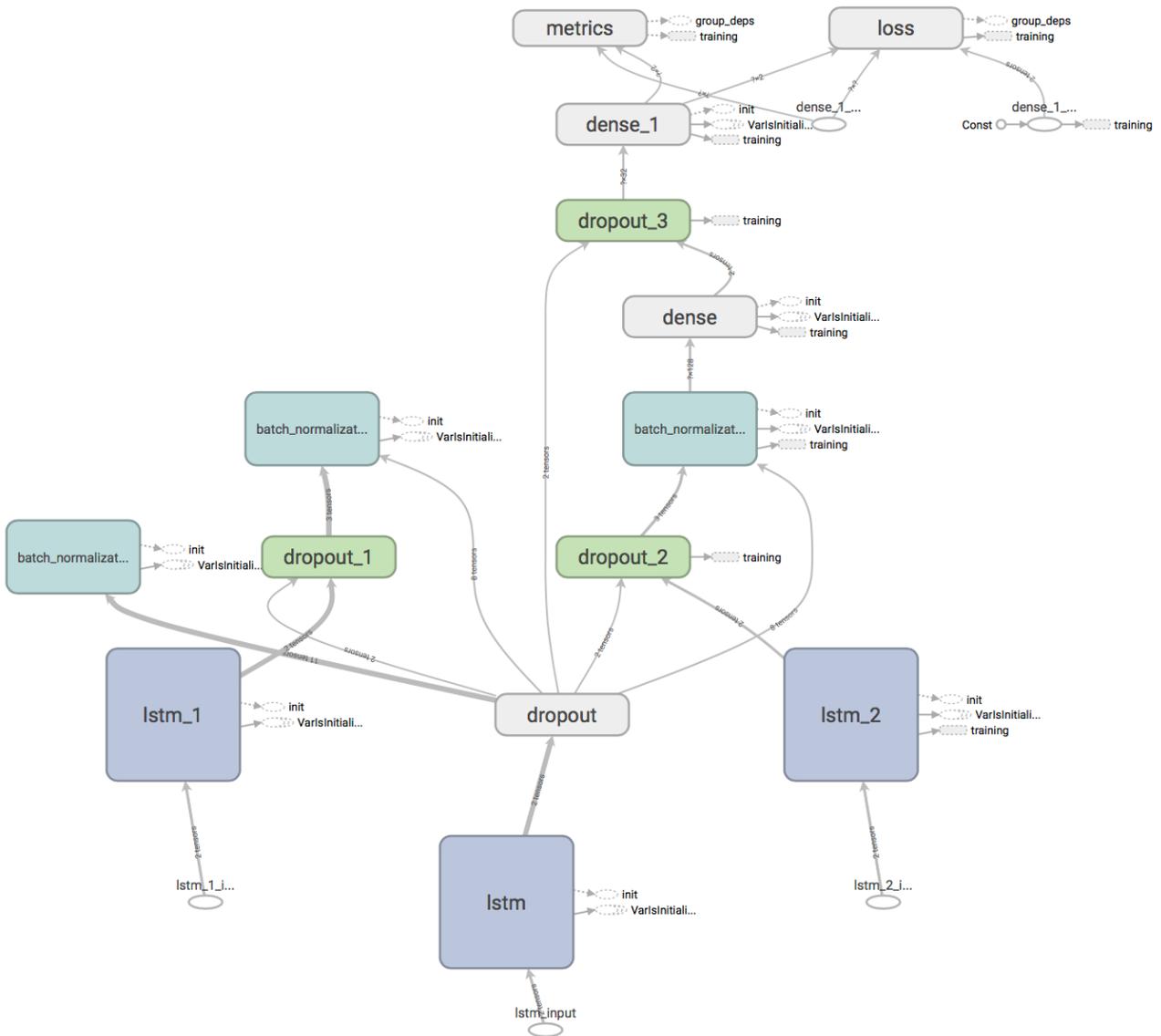
Apéndice V. Indicadores de Predicciones del modelo Jhonson & Jhonson (JNJ).

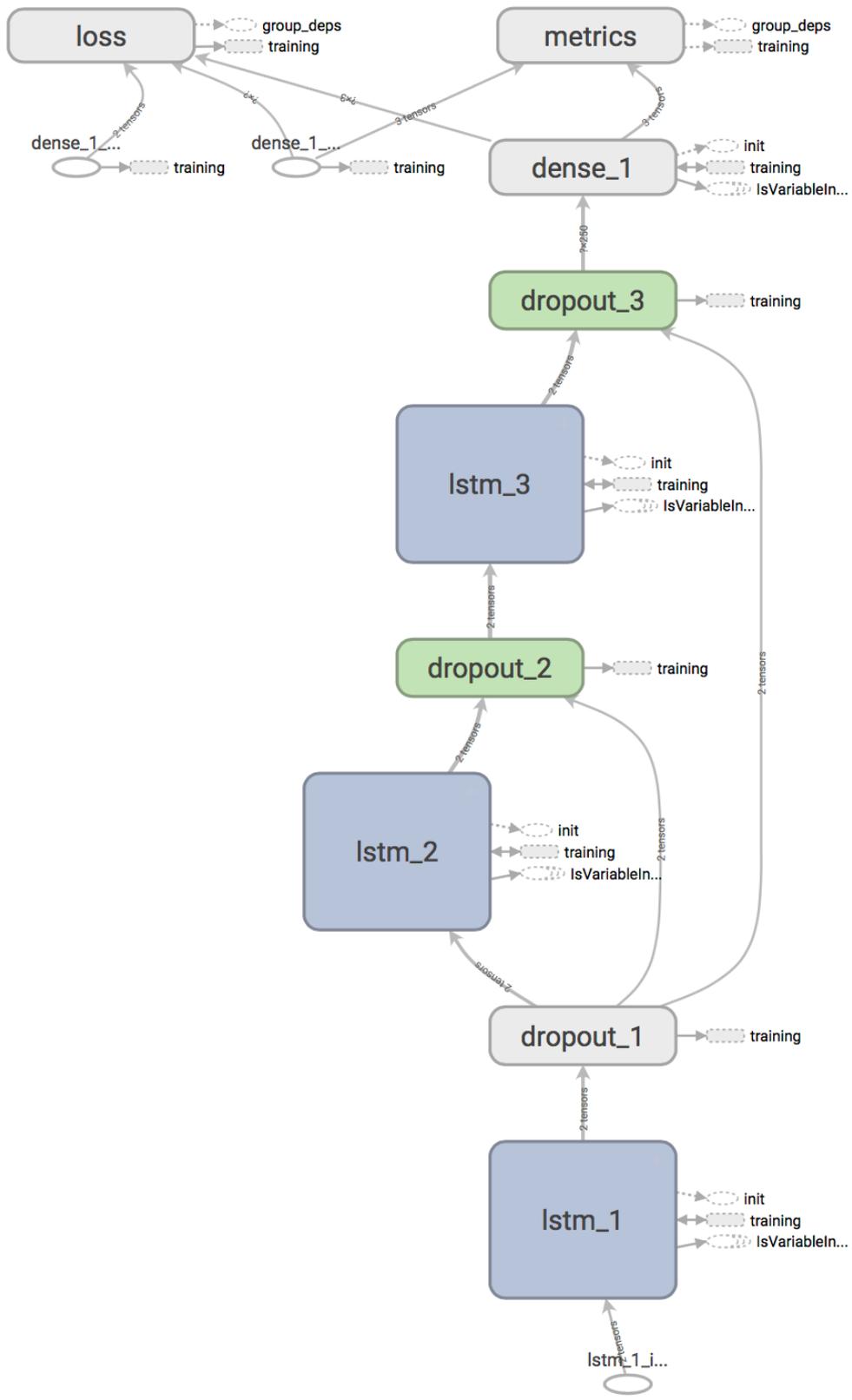
FECHA	REAL	PREDICHO	MAPE	MAE	RMSE
29/1/19	130,33	126,8447113	2,67420295	3,4852887	12,1472373
30/1/19	131,64	127,9362316	2,81355849	3,7037684	13,7179004
31/1/19	133,08	126,5592514	4,89987121	6,5207486	42,5201623
1/2/19	134,2	127,4393184	5,03776572	6,7606816	45,7068157
4/2/19	132,88	126,3330627	4,92695462	6,5469373	42,862388
5/2/19	132,88	127,0526485	4,38542407	5,8273515	33,9580255
6/2/19	133	126,1773093	5,12984263	6,8226907	46,5491084
7/2/19	132,05	126,7008466	4,05085452	5,3491534	28,6134421
8/2/19	132,4	126,0460752	4,79903686	6,3539248	40,3723604
11/2/19	132	126,4469284	4,20687242	5,5530716	30,8366042
12/2/19	134,16	125,9868472	6,09209362	8,1731528	66,8004267
13/2/19	134,45	126,3049398	6,05805891	8,1450602	66,3420057
14/2/19	134,31	126,0006152	6,18672087	8,3093848	69,0458758
15/2/19	136,38	126,2495183	7,42812854	10,1304817	102,626659
19/2/19	135,69	126,0473243	7,10640113	9,6426757	92,9811947
	FINAL		5,05305244	6,75495812	7,00038193

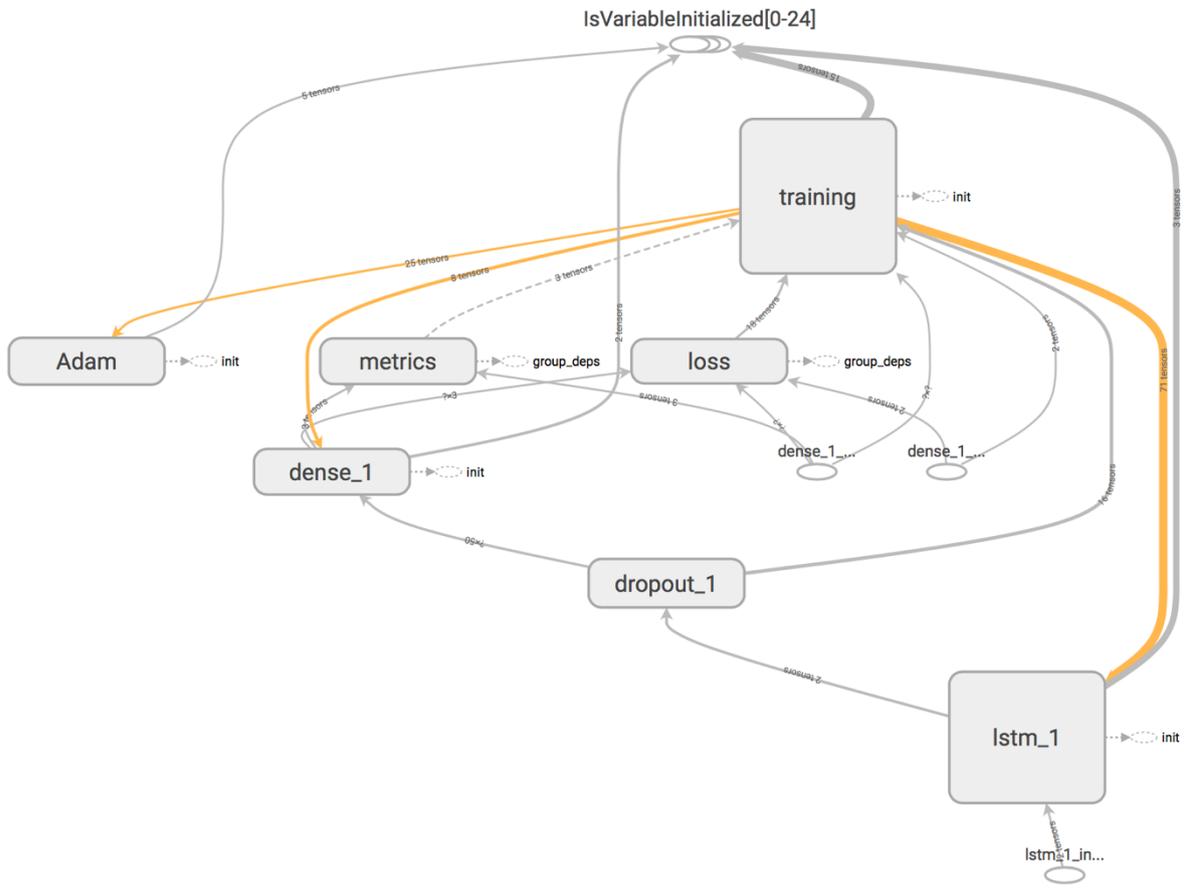
Apéndice W. Representación gráfica del modelo predictivo definitivo (Red Neuronal).



Apéndice X. Representación gráfica de los modelos predictivos anteriores al definitivo (Red Neuronal).









Presidencia
de la República
del Ecuador



Plan Nacional
de Ciencia, Tecnología,
Innovación y Saberes



SENESCYT
Secretaría Nacional de Educación Superior,
Ciencia, Tecnología e Innovación

DECLARACIÓN Y AUTORIZACIÓN

Yo, **Santistevan Pineda, Stephano Paul** con C.C.: # 0924832249, autores del trabajo de titulación: **Modelo predictivo para el comportamiento de las acciones que cotizan en la bolsa de valores de Nueva York, utilizando minería de datos** previo a la obtención del título de **INGENIERO EN SISTEMAS COMPUTACIONALES** en la Universidad Católica de Santiago de Guayaquil.

1.- Declaro tener pleno conocimiento de la obligación que tienen las instituciones de educación superior, de conformidad con el Artículo 144 de la Ley Orgánica de Educación Superior, de entregar a la SENESCYT en formato digital una copia del referido trabajo de graduación para que sea integrado al Sistema Nacional de Información de la Educación Superior del Ecuador para su difusión pública respetando los derechos de autor.

2.- Autorizo a la SENESCYT a tener una copia del referido trabajo de graduación, con el propósito de generar un repositorio que democratice la información, respetando las políticas de propiedad intelectual vigentes.

Guayaquil, 08 de marzo del 2019

Santistevan Pineda, Stephano Paul
C.C: 0924832249

REPOSITORIO NACIONAL EN CIENCIA Y TECNOLOGÍA			
FICHA DE REGISTRO DE TESIS/TRABAJO DE GRADUACIÓN			
TÍTULO Y SUBTÍTULO:	Modelo predictivo para el comportamiento de las acciones que cotizan en la bolsa de valores de Nueva York, utilizando minería de datos		
AUTOR:	Stephano Paul, Santistevan Pineda		
TUTOR:	Ing. Marcos Xavier, Miranda Rodríguez, Mgs.		
INSTITUCIÓN:	Universidad Católica de Santiago de Guayaquil		
FACULTAD:	Ingeniería		
CARRERA	Ingeniería en Sistemas Computacionales		
TÍTULO OBTENIDO:	Ingeniero en Sistemas Computacionales		
FECHA DE PUBLICACIÓN:	08 de marzo del 2019	No. DE PÁGINAS:	120
ÁREAS TEMÁTICAS	Tecnología, sistemas de información		
PALABRAS CLAVE:	Minería de datos; Mercado de valores; Bolsa de valores; Redes neuronales; LSTM; Series temporales.		
RESUMEN:	<p>El presente trabajo “Modelo predictivo para el comportamiento de las acciones que cotizan en la bolsa de valores de Nueva York, utilizando minería de datos” trata de brindar una herramienta capaz de poder predecir el comportamiento del mercado de valores de la ciudad de Nueva York, mediante la aplicación de la minería de datos, haciendo uso de la data histórica del mercado de valores, teniendo en cuenta este tipo de predicciones, tradicionalmente se ha utilizado el modelo ARIMA para este tipo de predicciones, sin embargo en la actualidad la minería de datos se da de la mano de con otra ciencia, tal como las Redes Neuronales, en el caso específico de las series temporales, se plantea el uso de las LSTM que son un tipo de Redes Neuronales Recurrentes.</p> <p>El mercado de valores presenta una gran volatilidad, lo que a su vez significa que posee un alto grado de riesgo, sin embargo dada la globalización de los mercados, la cantidad de personas que invierten en el mercado de valores va en ascenso, por lo cual se vuelve en una necesidad la existencia de una herramienta que sea capaz de predecir el comportamiento del mercado de valores con un alto grado de certidumbre, pudiendo ser de gran utilidad al momento de realizar transacciones de compra y venta de valores. Para lo cual se tomó como herramienta la minería</p>		

de datos, la cual es capaz de hacer uso de toda esa gran cantidad de data, proveniente del comportamiento histórico del mercado de valores. De acuerdo a estudios recientes las redes de tipo LSTM, son capaces de poder predecir con un alto grado de certidumbre el comportamiento del mercado de valores, destacando que las Redes Neuronales es una de las técnicas de minería de datos, que hoy en día se encuentra en constante desarrollo, presentándose como una alternativa a las técnicas ya tradicionales, sin embargo al hacer uso de las Redes Neuronales, como modelo, se debe de tener muy en cuenta el valor de la data con la que se entrena la red neural, en gran parte la precisión de los resultados dependerá de la data de entrenamiento del modelo.

ADJUNTO PDF:	<input checked="" type="checkbox"/> SI	<input type="checkbox"/> NO
CONTACTO CON AUTOR:	Teléfono: +593-0985159457	E-mail: dadsteph@gmail.com
CONTACTO CON LA INSTITUCIÓN:	Nombre: Ing. Edison José Toala Quimí	
	Teléfono: +593-042 20 27 63 / 593-9-90976776	
	E-mail: edison.toala@cu.ucsg.edu.ec	

SECCIÓN PARA USO DE BIBLIOTECA	
Nº. DE REGISTRO (en base a datos):	
Nº. DE CLASIFICACIÓN:	
DIRECCIÓN URL (tesis en la web):	