



PONTIFICIA UNIVERSIDAD CATOLICA DE CHILE  
ESCUELA DE INGENIERIA

# **INVESTIGACIÓN, ANÁLISIS Y MODELACIÓN APLICADA EN EL ESTUDIO DE INGENIERÍA FINANCIERA EN “VALUE INVESTING”**

**ANDRÉS HOLTHEUER LE FOULON**

Trabajo de Título para optar al título de Ingeniero Civil de  
Industrias, Diploma en Ingeniería Ambiental.

Profesor Guía:  
**TOMÁS REYES**

Santiago de Chile, 2020



PONTIFICIA UNIVERSIDAD CATOLICA DE CHILE  
ESCUELA DE INGENIERIA  
Departamento de Ingeniería Industrial y de Sistemas

# **INVESTIGACIÓN, ANÁLISIS Y MODELACIÓN APLICADA EN EL ESTUDIO DE INGENIERÍA FINANCIERA EN “VALUE INVESTING”**

**ANDRÉS HOLTHEUER LE FOULON**

Trabajo de Título presentada a la Comisión integrada por los profesores:

**TOMÁS REYES**

**IGNACIO VARGAS**

**ALVARO CHACÓN**

Para completar las exigencias del título de  
Ingeniero Civil de Industrias, Diploma en Ingeniería Ambiental

Santiago de Chile, 2020

A mi familia, por su apoyo incondicional.

## **AGRADECIMIENTOS**

Quiero darle mis más sinceros agradecimientos a mi profesor guía, Tomás Reyes, que accedió a orientarme en esta investigación y me apoyó con sus consejos y experiencia. También deseo agradecer a Claudia Santibáñez por su confianza, recomendaciones y compañía en el laboratorio. Por último, agradecer al profesor Gustavo Angulo por su apoyo en la confección de uno de los problemas de optimización.

## INDICE GENERAL

	Pág.
DEDICATORIA .....	ii
AGRADECIMIENTOS .....	iii
INDICE DE TABLAS .....	vi
INDICE DE FIGURAS .....	vii
RESUMEN.....	viii
ABSTRACT .....	ix
I. Descripción .....	1
II. METODOLOGÍA Y ANÁLISIS.....	3
2.1 Datos .....	3
2.1.1 Revisión Bibliográfica .....	3
2.1.2 Extracción de Datos .....	5
2.1.3 Estrategia Datos Faltantes.....	5
2.2 Modelo de Selección de Acciones .....	8
2.2.1 Red Neuronal Artificial .....	8
2.2.2 Entrenamiento .....	10
2.2.3 Separación Base de Datos .....	12
2.2.4 Validación Cruzada .....	13
2.3 Evaluación de Resultados .....	15
2.3.1 Análisis Financiero .....	15
2.3.2 Evaluación Graham .....	16
III. RESULTADOS .....	19
3.1 Resultados Validación Cruzada .....	19
3.2 Resultados Testeo.....	20

IV. CONCLUSIÓN .....	25
4.1 Conclusiones del Trabajo .....	25
4.2 Conclusiones Competencias.....	27
BIBLIOGRAFIA .....	30
A N E X O S .....	34
Anexo A: Inteligencia Artificial.....	35
Anexo B: Aplicaciones de IA en el Sector Financiero .....	38
B.1 Detección de Fraude.....	38
B.2 Evaluación Crediticia.....	39
B.3 Minería de Datos .....	39
B.4 <i>Chatbots</i> .....	40
B.5 <i>Trading</i> Algorítmico .....	41
B.6 Gestión de Patrimonio.....	42
Anexo C: Conceptos relevantes <i>value investing</i> .....	44
Anexo D: Parámetros Base de datos .....	46
Anexo E: Datos Perdidos .....	53
Anexo F: Modelación Eliminación Óptima .....	54
Anexo G: Función Objetivo Red Neuronal.....	56
Anexo H: Ejemplo Separación Validación Cruzada.....	57
Anexo I: Lógica Portafolio .....	58
Anexo J: Criterios Graham Actualizados.....	60
Anexo K: Resultados Validación .....	61

## INDICE DE TABLAS

	Pág.
Tabla 3.1: Resultados Validación Cruzada, periodo de tenencia 3 años .....	19
Tabla 3.2: Resultados Validación Cruzada, periodo de tenencia 5 años .....	19
Tabla 3.3: Resultados testeo, periodo de tenencia 3 años .....	21
Tabla 3.4: Resultados testeo, periodo de tenencia 5 años .....	21
Tabla 3.5: Disminuciones porcentuales entre la validación y el testeo .....	22
Tabla D.1: Parámetros Base de Datos .....	46
Tabla J.1: Percentiles Criterios Graham para el 1990 .....	60
Tabla J.2: Criterios Actualizados .....	60
Tabla K.1: Posiciones configuraciones modelos, periodo de tenencia de 3 años .....	61
Tabla K.2: Posiciones configuraciones modelos, periodo de tenencia de 5 años .....	64

## INDICE DE FIGURAS

	Pág.
Figura 2.1: Red Neuronal .....	10
Figura 2.2: Ejemplo separación de validación y entrenamiento .....	15
Figura 3.1: Rentabilidad mensual acumulada, periodo de tenencia de 3 años.....	24
Figura 3.2: Rentabilidad mensual acumulada, periodo de tenencia de 5 años.....	24
Figura A.1: Subáreas de la IA .....	37



## RESUMEN

Este Trabajo de Título tiene como objetivo el desarrollo de un modelo predictivo basado en Inteligencia Artificial, que genere automáticamente un portafolio de acciones que supere el índice de mercado, en base a los parámetros financieros claves según la filosofía de inversión *Value Investing*.

Este trabajo se puede dividir en tres procesos. En primer lugar, se trabajó en la recolección de datos, que incluye la revisión bibliográfica de la filosofía de inversión en valor, la extracción de los parámetros claves del software financiero *Bloomberg* y el manejo de los datos faltantes de la matriz obtenida. Luego, se construyó el modelo de selección de acciones (MSA) ocupando métodos de *Machine Learning*. Específicamente se generó una red neuronal artificial, la cual fue entrenada, validada y testeada para clasificar las empresas sobresalientes de las no sobresalientes. Por último, se realizaron evaluaciones financieras, midiendo la rentabilidad y el *alfa* del portafolio versus el S&P 500. Además, se evaluó si es que las empresas que selecciona el MSA cumplen ciertos criterios propuestos por Benjamin Graham, máximo exponente de la inversión en valor.

Los resultados mostraron qué, si bien los portafolios creados obtuvieron bajos resultados de clasificación, generalmente obtienen resultados positivos de rentabilidad frente al índice de mercado. Desafortunadamente, no existe suficiente evidencia estadística para asegurar que estos resultados no se dieron por chance. Además, los portafolios tienden a cumplir, en promedio, menos criterios Graham que las empresas no seleccionadas; pero no existe evidencia estadística para afirmar que este último cumple más criterios que el portafolio.

En conclusión, no se recomienda invertir en el portafolio que genera el MSA por sobre el índice de mercado; ya que las diferencias obtenidas no son lo suficientemente robustas para poder pensar que se mantendrán en el tiempo.

## **ABSTRACT**

This project aims to develop a predictive model based on Artificial Intelligence that automatically generates a portfolio of shares that exceeds the market index, based on the key financial parameters according to the investment philosophy Value Investing.

This work can be divided into three processes. First, the student worked on the collection of data, which includes the bibliographic review of the philosophy of investment in value, the extraction of the key parameters of the Bloomberg financial software and the management of the missing data of the matrix obtained. Then, the stock selection model was built using Machine Learning methods. Specifically, an artificial neural network was generated, which was trained, validated and tested to classify outstanding companies from non-outstanding companies. Finally, financial evaluations were carried out, measuring the profitability and the alpha of the portfolio versus the S&P 500. Besides, it was evaluated whether the companies selected by the stock selection model meet certain criteria proposed by Benjamin Graham, the greatest exponent of the value investing philosophy.

The results showed that, although the created portfolios obtained low classification results, they generally obtained positive results of profitability against the market index. Unfortunately, there is not enough statistical evidence to ensure that these results were not given by chance. Also, portfolios tend to meet, on average, less Graham criteria than unselected companies; but there is no statistical evidence to affirm that the latter meet more criteria than the portfolio.

In conclusion, it is not recommended to invest in the portfolio generated by the stock selection model over the market index; since the differences obtained are not robust enough to be able to think that they will remain in time.

## I. DESCRIPCIÓN

En este trabajo de título se busca la automatización de la selección de una cartera de acciones con algoritmos de Inteligencia Artificial. Se trabajó específicamente con las empresas miembros del índice S&P 500, extrayendo su información financiera del *software* Bloomberg, siguiendo la filosofía de inversión *Value Investing*. Luego se generó un modelo que, con métodos de *Machine Learning*, busca elegir las empresas que van a superar al mercado considerando un periodo de tenencia de 3 o 5 años. Por último, se busca observar si es que el portafolio creado automáticamente obtuvo mejores resultados que el mercado, ajustado al riesgo, y si es que efectivamente este sigue una estrategia parecida a lo propuesto en la inversión en valor.

Entonces, el objetivo principal del proyecto es el desarrollo de un modelo predictivo que seleccione las mejores acciones a invertir en base a los parámetros financieros claves según la filosofía de inversión *Value Investing*. Para realizar este trabajo de título se contó con 16 semanas, desde el 14 de octubre de 2019 hasta el 31 de enero del 2020. El alcance del proyecto se dividió en tres fases: Datos, Modelo de Selección de Acciones y Evaluación de Resultados.

En la fase de Datos se tuvo tres tareas principales, la revisión bibliográfica, extracción de datos y la estrategia para manejar los datos faltantes. Se realizó una revisión bibliográfica de *Value Investing* para entender los parámetros que iban a ser fundamentales para el modelo. Luego, la extracción de datos se realizó a través de los complementos Bloomberg disponibles en Excel. En tercer lugar, se realizó una estrategia para tratar los datos faltantes de manera óptima. En la fase del Modelo de Selección de Acciones (MSA), se ocuparon redes neuronales artificiales para programar el algoritmo que genera un portafolio con baja intervención humana. Por último, en la fase de Evaluación de Rentabilidad se buscó comparar los resultados de los portafolios creados automáticamente con los resultados que obtuvo el mercado para años que el MSA no ha visto y observar, específicamente, si es que los portafolios obtuvieron un *alfa* positivo con significancia estadística. Donde el *alfa* es el parámetro que representa la ganancia o pérdida que obtiene un portafolio con referencia a un índice de mercado, de la teoría del CAPM (Reyes, 2018). También, se desea evaluar

si es que las empresas que selecciona el MSA cumplen ciertos criterios propuestos por Benjamin Graham, el máximo exponente de la inversión en valor.

Este trabajo tiene como limitación fundamental la sencillez de la estrategia de inversión que se sigue. Se ocupa como estrategia el mantener en tenencia el portafolio por 3 o 5 años, sin rebalanceos, sin ventas cuando los precios están muy altos y sin compras para aprovechar nuevas oportunidades de ocasión.

Se decidió utilizar métodos de *Machine Learning*, o aprendizaje automático en español, debido a que este tiene la particularidad de alimentarse de una gran cantidad de información y “aprender” las correlaciones que podrían separar a las empresas sobresalientes (y por lo tanto en las que se desea invertir) de las que no. El *Machine Learning* es capaz de manejar una cantidad de información gigantesca que es imposible para un ser humano, por lo que se podría pensar que el algoritmo será capaz de aprender patrones que los seres humanos no han sido capaces de identificar. En el Anexo A y B se presenta más información sobre la inteligencia artificial y como esta revoluciona el mercado financiero, respectivamente.

Entonces, como resultados esperados se desea obtener una estrategia de generación de portafolios automática que posean un *alfa* positivo para los años de testeo, es decir, superen al mercado. Al mismo tiempo, se busca cumplir con las siguientes competencias del perfil de egreso: aplicar diversos métodos de análisis de datos para la comprensión de los fenómenos abordados; realizar análisis financieros y de costos aplicados a los requerimientos de instituciones privadas y públicas; y aplicar los fundamentos de la ingeniería e investigación operacional, al modelamiento y resolución de problemas complejos.

## **II. METODOLOGÍA Y ANÁLISIS**

En este capítulo se presentará la metodología y también como los resultados van a ser evaluados para determinar su relevancia. Se verán las tres fases anteriormente descritas: Datos, Modelo de Selección de Acciones y Evaluación de Resultados.

### **2.1 Datos**

En esta sección se describe las tres tareas principales que se realizaron en cuanto a la base de datos: la revisión bibliográfica, extracción de datos y la generación de una estrategia para manejar los datos faltantes.

#### **2.1.1 Revisión Bibliográfica**

Como objetivo central del trabajo de título se planteó generar un modelo de selección de acciones según la filosofía de inversión de *Value Investing* o inversión en valor. Esta filosofía recomienda invertir en las oportunidades generadas a partir de las discrepancias entre el valor intrínseco y el precio de mercado de una acción (Chee, Sloan & Uysal, 2013).

Benjamin Graham es reconocido ampliamente como el fundador de la inversión en valor, al escribir los libros “*Security Analysis*” y “*The Intelligent Investor*” (el primero junto con David Dodd). Profesor de la universidad de Columbia, es reconocido como una de las primeras personas en hacer análisis fundamentales técnicos en el mercado de valores. (Chee et al., 2013). Graham y Dodd argumentan que los factores especulativos hacen que los precios del mercado se desvíen de los valores intrínsecos y que, en el largo plazo, existe una tendencia inherente a que las disparidades se corrijan por sí mismas. (Graham, 1949; Chee et al., 2013).

La mayor parte del tiempo, el mercado determina los precios de las acciones con una destacable brillantez. Millones de compradores y de vendedores regateando precios son capaces de conseguir unas valoraciones empresariales destacablemente exactas, es decir, que son lo suficientemente parecidas a su valor intrínseco. Sin embargo, en ocasiones el resultado es un precio tremendamente desajustado (Graham, 1949). Las fluctuaciones de precio ofrecen la oportunidad de

comprar con inteligencia cuando los precios se reducen mucho en comparación a su valor intrínseco, y de vender en el caso contrario. En las demás ocasiones lo mejor que se puede hacer es olvidarse del mercado de valores y prestar atención a su rentabilidad por dividendos y a los resultados de explotación de las empresas (Graham, 1949).

Ahora bien, se debe conocer el valor intrínseco de esta acción para poder saber si esta está sobre o subvaluada según su precio de mercado; pero este es un concepto difícil de definir explícitamente. Graham y Dodd escribieron la siguiente guía para entender qué es el valor intrínseco:

Es un gran error imaginar que el valor intrínseco es tan definido y determinable como el precio de mercado. (...) El valor intrínseco de un negocio está determinado por su poder de ganancias. Pero la frase “poder de ganancias” debe implicar una expectativa bastante segura de ciertos resultados futuros. No es suficiente saber qué promediaron las ganancias pasadas, o incluso si revelan una línea separada de crecimiento o declive. Debe haber motivos fundados para creer que este promedio o esta tendencia es una guía confiable para el futuro (como se citó en Chee et al., 2013, p. 2).

Graham asegura que es muy difícil tener estos “motivos fundados” para predecir las ganancias futuras, por lo que propone un enfoque distinto para tratar el problema de evaluar el valor intrínseco; el margen de seguridad. El margen de seguridad se basa en no pagar un precio muy elevado en relación con las utilidades obtenidas. La función de este es, en esencia, la de hacer innecesaria una estimación exacta del futuro. Si el margen es grande, es más que suficiente para suponer que los futuros beneficios no quedarán por debajo de los del pasado, de modo que el inversor pueda sentirse protegido frente a las vicisitudes del tiempo. Entonces, para obtener un buen margen de seguridad, Graham recomienda seleccionar empresas de buena calidad que estén subvaloradas. Donde una empresa de buena calidad es aquella de gran tamaño que muestra una buena situación financiera y estabilidad en cuanto a las utilidades y dividendos (Graham, 1949).

Se seleccionó esta filosofía de inversión para obtener los parámetros relevantes con los que se alimentó el modelo de selección de acciones debido a que es una de las filosofías que más ha perdurado en el tiempo y que sus seguidores

mencionan que logran superar el mercado sostenidamente en el tiempo (Chee et al., 2013). En el Anexo C se detallan los conceptos más importantes encontrados en el libro “*The Intelligent Investor*”, en los cuales se basó la obtención de parámetros. Siguiendo estos conceptos, se llegó a obtener 234 parámetros relacionados; incluyendo los *alfas* proyectivos que se utilizaron para separar a las empresas sobresalientes de las que no, para los periodos de tenencia de 3 y 5 años. Los parámetros se enlistan en el Anexo D.

### **2.1.2 Extracción de Datos**

Como se mencionó anteriormente, los datos fueron extraídos del Terminal de Bloomberg. Bloomberg es un software proporcionado por el proveedor de datos financieros Bloomberg L.P. que permite a los profesionales del sector de servicios financieros, y otras industrias, acceder al servicio *Bloomberg Professional*. Este reúne datos en tiempo real de cada mercado, noticias de última hora, investigación en profundidad, análisis de gran alcance, herramientas de comunicación y capacidades de ejecución de primer nivel, en una solución totalmente integrada (“La Terminal Bloomberg”, s.f.). La extracción fue realizada a través del complemento de Bloomberg para Excel, lo que permite exportar datos del software a una hoja del programa de Microsoft Office.

Se obtuvo la información financiera de los miembros del índice S&P 500 desde 1990 hasta 2017. Para cada una de las empresas se obtuvo el valor de los 234 parámetros para el año fiscal correspondiente, es decir, la extracción de la información para un miembro del S&P 500 en un año en particular se hace en el mes que esa empresa específica liberó su reporte anual de ese año. Siguiendo así fielmente el año fiscal de cada empresa para hacer la clasificación apenas se tenga la información. Como resultado se obtuvo una matriz que cuenta con 234 columnas y 13.977 filas, donde cada columna representa un parámetro y cada fila representa una observación de un miembro del S&P 500 para un año específico.

### **2.1.3 Estrategia Datos Faltantes**

La matriz obtenida en el proceso anterior poseía muchos datos faltantes, específicamente el 14,9%. Poseer una fracción de datos perdidos imposibilita el ajuste

de la red neuronal en el MSA, ya que no se puede realizar los procesos del algoritmo. La literatura especializada recomienda dos posibles estrategias para solucionar este problema: imputar la matriz o eliminar los datos faltantes de manera óptima (Roy, 2019; Swalin, 2018). El primero se refiere a predecir los datos faltantes y el segundo a buscar una forma de eliminar las filas y/o columnas con datos faltantes de manera que se maximice el tamaño de la matriz restante. En el Anexo E se explica brevemente la naturaleza de los datos perdidos, la cual indica que estrategia seguir; y se desarrolla las razones de porqué se pueden ocupar ambas metodologías.

Ahora bien, la eliminación de filas y/o columnas disminuye el tamaño de la matriz lo que afecta negativamente a la red neuronal que se desea ajustar con la base de datos, ya que limita los ejemplos que observa la red para establecer las correlaciones. La imputación de la matriz tiene como desventaja su alto costo computacional y el que no asegura mejores resultados, pero tiene la ventaja de no disminuir el tamaño de la muestra (Swalin, 2018).

Entonces, cómo en la literatura no se tiene consenso de cuál método es mejor, sino que depende de la misma base de datos cual ocupar, se decidió generar tres bases de datos distintas: dos con eliminación de datos y una con imputación. En este trabajo se plantea tratar el método de manejo de datos faltantes como un hiper parámetro para la red neuronal, es decir un parámetro que la red ocupa como *input* y por lo tanto no puede estimar. Entonces, a través de un proceso de validación (que será explicado en el punto siguiente), se decide cuál base de datos genera un mejor modelo. A continuación, se explica cómo se generaron las distintas bases de datos, las cuales se llamarán base de datos 1, 2 y 3 según el orden de aparición.

### **2.1.3.1 Base de Datos 1: Eliminación óptima**

Lo recomendado por la literatura es ir eliminando las observaciones, en este caso particular las empresas, que posean uno o más datos faltantes; ya que borrar los parámetros de la base de dato afecta directamente al rendimiento de la red neuronal (Roy, 2019). Ahora bien, si se utiliza este enfoque se obtiene una matriz con sólo 28 filas. Entonces, se decidió crear un modelo de optimización que maximizara el tamaño de la matriz bajo la restricción de que no hubiera ningún dato faltante, teniendo como



variables el poder eliminar filas y columnas de esta. La formulación matemática del problema de optimización se muestra en el Anexo F.

El problema de optimización fue resuelto ocupando el *solver* Gurobi, un *software* que resuelve problemas de optimización matemática, el cual es considerado uno de los más poderosos del mundo (“The Fastest Solver”, s.f.). El modelo fue programado en Python, el cual accede a Gurobi a través de una licencia académica. Este obtuvo como resultado una matriz óptima de 10.311 empresas y 140 parámetros. Los parámetros seleccionados se muestran en el Anexo D.

### **2.1.3.2 Base de Datos 2: Eliminación óptima con parámetros fijos**

La Base de Datos 1 presenta el problema de que se eliminan variables que eran fundamentales para la inversión en valor. Por esta razón se decidió modificar el problema de minimización del Anexo F, agregando restricciones que fijan 8 variables relevantes en la matriz resultante que, según Graham en su libro *The Intelligent Investor*, son las fundamentales que un inversor defensivo debiera evaluar para invertir en una empresa. Estas variables se explicarán con más detalle en la parte de evaluación de resultados. En el Anexo F se muestra el cambio a la formulación matemática del modelo. El modelo también fue programado en Python y resuelto con el *solver* Gurobi. La matriz resultante quedó de un tamaño de 4.976 empresas y 165 parámetros. Estos últimos se muestran en el Anexo D.

### **2.1.3.3 Base de Datos 3: Imputación**

La imputación múltiple se refiere a la predicción de datos perdidos simultáneamente en más de una variable. Existen dos métodos para la solución de este problema, Modelación Conjunta o Imputación Multivariada por Ecuaciones Encadenadas (JM y MICE por sus siglas en inglés) (van Buuren & Groothuis-Oudshoorn, 2011). En este trabajo se utilizó MICE debido a que JM requiere supuestos de que la base de datos distribuye normal multivariante, lo cual es un supuesto bastante fuerte para las 234 variables que presenta la base de datos (y, por esta razón, MICE es el método más ocupado en la literatura) (van Buuren & Groothuis-Oudshoorn, 2011).

MICE realiza la imputación múltiple variable por variable mediante un conjunto de densidades condicionales, una para cada variable incompleta. A partir de una imputación inicial, MICE genera imputaciones iterando sobre las densidades condicionales (van Buuren & Groothuis-Oudshoorn, 2011). Se utilizó el lenguaje de programación R, que se enfoca en el análisis estadístico, en donde se ocupó la librería *mice*. Entonces utilizando este método se obtiene la tercera base de datos, la cual presenta la cantidad de filas y matrices de la base de datos original y que tiene un 14,9% de sus datos estimados.

## **2.2 Modelo de Selección de Acciones**

El modelo de selección de acciones (MSA) fue creado con el objetivo de generar un algoritmo que elija en qué acciones invertir para superar al mercado en el mediano y largo plazo. Esto es un problema de clasificación, ya que se le está pidiendo al modelo contestar a la pregunta si se debe invertir en una empresa en particular con un “sí” o un “no”. Para esto, las redes neuronales han mostrado supremacía debido a su capacidad de generar funciones complejas para encontrar las correlaciones entre una cantidad importante de parámetros (Ng, 2011).

Para decidir cuando una empresa es sobresaliente se observó el *alfa* futuro que presentaría esta, es decir, el *alfa* que la compañía tendrá en el período de tenencia que se desea mantener. Entonces, si el *alfa* pertenece al percentil 66,6 o mayor de la muestra, se le confiere la etiqueta “1”, en caso contrario, la etiqueta “0”. Estos umbrales tomaron los valores de 0,170 y 0,154 para los periodos de tenencia de 3 y 5 años, respectivamente.

A continuación, se explicará el funcionamiento de la red neuronal artificial y los procesos de entrenamiento, validación cruzada y testeo de la red. Este punto ocupa como base el curso *on-line* llamado “*Machine Learning*”, del profesor Andrew Ng, de la universidad de Standford, disponible en Coursera desde el 2011.

### **2.2.1 Red Neuronal Artificial**

La red neuronal artificial es una colección de componentes de procesamiento en paralelo interconectados que imitan la estructura y funcionalidad de

una neurona humana. Estas neuronas imitadas son los elementos de procesamiento de la red neuronal artificial y se les llama nodos o unidades de activación. La capacidad de procesamiento de la red se almacena en los pesos de cada unidad, que se obtienen mediante el proceso de aprendizaje o entrenamiento (Zavadskaya, 2017).

Para entender mejor el funcionamiento de una red neuronal primero se debe entender cómo funciona la unidad de activación. Sea el vector  $x$  los parámetros que recibe esta unidad de activación y  $\theta$  los pesos que le asigna esta unidad a los parámetros que recibió. Entonces, la unidad entrega como resultado un escalar entre  $(0,1)$ , ocupando la función de sigmoid; la cual se muestra a continuación.

$$g(\theta^T x) = \frac{1}{1+e^{-\theta^T x}} \quad (2.1)$$

La red neuronal consiste en capas, como bien se observa en la Figura 2.1. La primera es la capa de *inputs*, la cual es sencillamente los valores de los parámetros de una empresa cualquiera de la base de datos. Las capas intermedias son llamadas capas escondidas, las cuales se componen de una o más unidades de activación. La última capa es la capa de *output* la cual consiste en sólo una unidad de activación que entrega como resultado la función de hipótesis ( $h_{\theta}(x)$ ), es decir, la clasificación. Si esta es mayor o igual a 0,5 el resultado se aproxima a 1 y por ende se considera una empresa sobresaliente y se aconseja invertir en esta; en caso contrario, no se aconseja invertir.

Para poder entender cómo se relacionan las capas entre sí se expondrá un ejemplo con 3 parámetros de *inputs* y una capa escondida con cuatro unidades de activación, como se muestra en la Figura 2.1. Sea  $x$  el vector con los parámetros de *inputs*;  $a_i^{(j)}$  el resultado o “activación” de la unidad  $i$  de la capa  $j$ ; y  $\Theta^{(j)}$  la matriz de pesos que mapean los cálculos de la capa  $j$  a la capa  $j+1$ . Se debe considerar que en cada capa se agrega un parámetro o unidad extra, llamado unidad de sesgo, que siempre toma el valor de 1 y que actúa en la coordenada 0. Entonces los cálculos que se tendrían que hacer para obtener la función de hipótesis se muestran en la Ecuación 2.2; donde la función  $g(\cdot)$  está representada en la Ecuación 2.1.

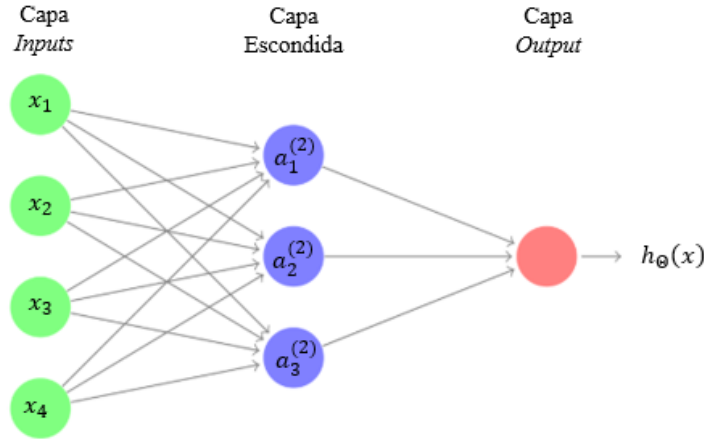


Figura 2.1: Red Neuronal (Ng, 2011)

Al final, lo que hacen las redes neuronales es que cada unidad de activación de una capa recibe los resultados de las unidades de activación de la capa anterior, para luego ocupar la ecuación 2.1 y entregar estos resultados a la capa siguiente. Esto lo hacen todas las unidades de activación de una capa, hasta llegar a la capa de *output* que genera la clasificación.

$$\begin{aligned}
 a_1^{(2)} &= g(\Theta_{10}^{(1)} * x_0 + \Theta_{11}^{(1)} * x_1 + \Theta_{12}^{(1)} * x_2 + \Theta_{13}^{(1)} * x_3 + \Theta_{14}^{(1)} * x_4) \\
 a_2^{(2)} &= g(\Theta_{20}^{(1)} * x_0 + \Theta_{21}^{(1)} * x_1 + \Theta_{22}^{(1)} * x_2 + \Theta_{23}^{(1)} * x_3 + \Theta_{24}^{(1)} * x_4) \\
 a_3^{(2)} &= g(\Theta_{30}^{(1)} * x_0 + \Theta_{31}^{(1)} * x_1 + \Theta_{32}^{(1)} * x_2 + \Theta_{33}^{(1)} * x_3 + \Theta_{34}^{(1)} * x_4) \\
 h_{\Theta}(x) &= a_1^{(3)} = g(\Theta_{10}^{(2)} * a_0^{(2)} + \Theta_{11}^{(2)} * a_1^{(2)} + \Theta_{12}^{(2)} * a_2^{(2)} + \Theta_{13}^{(2)} * a_3^{(2)})
 \end{aligned} \tag{2.2}$$

### 2.2.2 Entrenamiento

Ya entendiendo como la red genera la clasificación, se va a explicar a continuación cómo se calculan los parámetros  $\Theta$ , proceso llamado entrenamiento o aprendizaje de la red. El proceso de entrenamiento consiste en el aprendizaje automático de la matriz  $\Theta$  por parte de la red. De manera general, el entrenamiento se basa en alimentar la red con un set de ejemplos u observaciones, llamado set de entrenamiento, en donde cada ejemplo consiste en los parámetros de *input* de cierta empresa y la clasificación correcta que se le debe asignar a esta; entonces, la red ajusta de manera automática los parámetros  $\Theta$  que minimizan el error de equivocarse en la asignación de las etiquetas de todos los ejemplos del set de entrenamiento, lo cual es

un problema de minimización irrestricto. Este se muestra en la ecuación 2.3, donde  $m$  es la cantidad de ejemplos del set de entrenamiento;  $L$  la cantidad total de capas de la red;  $s_l$  la cantidad de unidades (sin contar la unidad de sesgo) en la capa  $l$ ;  $y^i$  la etiqueta correcta del ejemplo  $i$ . En el Anexo G se detalla más sobre la lógica que sigue esta función.

$$\min_{\Theta} J(\Theta) = \min_{\Theta} \left\{ -\frac{1}{m} \sum_{i=1}^m \left[ y^i * \log(h_{\Theta}(x^i)) - (1 - y^i) * \log(1 - h_{\Theta}(x^i)) \right] + \frac{\lambda}{2} \sum_{l=1}^{L-1} \sum_{i=1}^{s_l} \sum_{j=1}^{s_{l+1}} (\theta_{j,i}^l)^2 \right\} \quad (2.3)$$

El primer sumando representa el costo de clasificación errónea para todos los ejemplos. El segundo sumando de la función objetivo consiste en la regularización. Esta metodología busca combatir el *overfitting*, es decir que la función de hipótesis de la red neuronal se sobreajuste a los datos de entrenamiento, pero que no se generalice bien para predecir datos que la red no ha visto. La otra forma de combatir el *overfitting* es con la eliminación de parámetros, pero esto es un proceso arduo que se hace manualmente y que podría eliminar parámetros relevantes (Ng, 2011). Para que la función objetivo sea cercana a cero al incorporar el segundo sumando, se deben reducir los pesos  $\Theta$  cercanos a cero; y de esta forma se reduce el efecto de parámetros que puedan generar un sobre ajuste al set de entrenamiento y una mala generalización a nuevos ejemplos. El parámetro  $\lambda$  es llamado el parámetro de regularización y determina cuánto se inflan los costos de los parámetros  $\Theta$ . Este es un hiper parámetro, es decir, se debe ajustar a través del proceso de validación

Para poder minimizar la función objetivo, el cual es un problema de minimización irrestricto, se ocupa el método del gradiente. Este método ocupa como dirección de descenso el negativo del gradiente de la función a minimizar, ya que entrega la mejor tasa marginal de disminución de esta (Vera, 2018). Ahora bien, para poder calcular el gradiente de esta función se ocupa el algoritmo *Backpropagation* o propagación para atrás. Este algoritmo permite calcular el gradiente desde la capa de *output* hasta la capa de *input*, es decir, en el sentido contrario en que se calcula la función de hipótesis (Ng, 2011).

Se eligió ocupar una red neuronal para el modelo de selección de acciones debido a la robustez que muestra este algoritmo. Al conectar diversas unidades de las múltiples capas, se puede generar una función de hipótesis que capture correlaciones

muy complejas entre los parámetros, a diferencia de una regresión. Ahora bien, no se optó por una red más compleja, como las utilizadas en *Deep Learning* o aprendizaje profundo, debido a que estos métodos muestran una verdadera ventaja sobre la red neuronal artificial a partir de bases de datos de millones o billones de ejemplos, por lo que no era necesario sobre complejizar el modelo (Ng, 2011).

### 2.2.3 Separación Base de Datos

La base de datos se debe separar en dos subconjuntos, conjunto de entrenamiento y el conjunto de testeo. El conjunto de entrenamiento se divide a su vez en dos sets de datos, el set de entrenamiento y el set de validación cruzada. El set de entrenamiento se utiliza para entrenar la red con el proceso explicado anteriormente, mientras que el set de validación cruzada se utiliza para ajustar los hiper parámetros. Por último, el conjunto de testeo se utiliza para comprobar si es que la red ajustada con los parámetros  $\Theta$  y los hiper parámetros obtiene resultados aceptables con datos que la red no ha visto (Ng, 2011; Cochrane, 2018; Ray, 2018).

Para testear los resultados que el modelo obtendría en el mundo real, se debe ocupar datos que ocurren cronológicamente después de los datos utilizados para entrenar el modelo. Se debe evaluar el rendimiento del algoritmo, el cual se encuentra en un presente y necesita pronosticar un futuro, ocupando años más recientes que el modelo no ha visto en el proceso de entrenamiento. Por lo tanto, para los datos de series de tiempo, se debe retener un subconjunto de datos (dividido temporalmente) que esté reservado exclusivamente para testear el rendimiento del modelo, llamado subconjunto de testeo (Cochrane, 2018). Entonces, para este subconjunto se dejan los últimos tres años con información disponible. Lo que equivale los años fiscales de 2011, 2012 y 2013 para el periodo de tenencia de 5 años; y los años 2013, 2014 2015 para el periodo de tenencia de 3 años. El resto de la base de datos fue dejada para el set de entrenamiento y validación; que se utilizarán para realizar el entrenamiento y la validación cruzada, con el objetivo de elegir la configuración de modelo para cada año de testeo que minimice el error de clasificación.

### 2.2.4 Validación Cruzada

La validación cruzada tiene como objetivo averiguar qué configuraciones de hiper parámetros dan el error mínimo antes de entrenar un modelo final para su posterior testeo (Ng, 2013; “*Cross-Validation strategies for Time Series forecasting*”, 2019). Los hiper parámetros son aquellos que no son ajustados en el proceso de entrenamiento, ya que son un *input* para este proceso. Para las redes neuronales consisten en el parámetro de regularización  $\lambda$ , la cantidad de capas escondidas y las unidades de activación en cada capa escondida; pero, en este proyecto, también se utiliza como hiper parámetro la base de datos a utilizar, como se explicó en el punto 2.1.3. A pesar de esto, la cantidad de unidades de activación fue fijado en 50 en todas las capas escondidas, debido a que Andrew Ng recomienda este número para bases de datos menores a 100.000 ejemplos. El profesor de Standford también recomienda hacer validación cruzada en cuanto a las capas escondidas, aunque resalta que en muchos casos basta sólo con una (Ng, 2011).

Entonces, la validación cruzada genera una cantidad finita de modelos en donde cada uno tiene una configuración única de hiper parámetros, para luego entrenar a todos estos modelos con el set de entrenamiento. Una vez ajustados los parámetros  $\Theta$  de cada modelo, se prueba como clasificarían a un set de ejemplos de empresas que el modelo no ha visto, o sea en el set de validación cruzada. Por último, se elige los hiper parámetros que generen el modelo con el menor error de clasificación en este último set (Ng, 2011).

Cuando hay datos de serie de tiempo, no se puede separar el set de entrenamiento y validación cruzada de manera aleatoria. Por lo que, al igual que en el conjunto de testeo, el set de validación cruzada viene cronológicamente después del set de entrenamiento (Cochrane, 2018). La técnica utilizada para realizar la validación cruzada, llamada encadenamiento hacia adelante, realiza varias iteraciones de entrenamiento y validación. En cada iteración se divide el conjunto de entrenamiento en estos dos sets, a condición de que el set de validación siempre esté por delante del set de entrenamiento cómo se observa en la Figura 2.2. En la primera iteración, se entrena el modelo desde el principio de la base de datos hasta el año  $j$ , para así hacer la validación en el año  $j+1$ . Luego en la siguiente se entrena desde el inicio de la base

de datos hasta el año  $j+1$ , para validar en el año  $j+2$ . Así sucesivamente hasta el último año del conjunto de entrenamiento; respetando así la dependencia temporal (Cochrane, 2018; “*Cross-Validation strategies for Time Series forecasting*”, 2019). En este proyecto se hicieron tres iteraciones de entrenamiento/validación para cada configuración de modelo que será testeado en un año particular, para cada periodo de tenencia. En el Anexo H se muestra un ejemplo para entender mejor esta separación.

Ahora bien, no se utilizó como métrica de error el valor final de la función objetivo debido a que no es relevante para este trabajo que se clasifique bien las etiquetas cero; lo que es verdaderamente importante es que el algoritmo no se equivoque cuando se asigne una etiqueta “1”. Ya que de esta forma las empresas seleccionadas van a superar al mercado, independiente de la cantidad de compañías que conformen ese portafolio.

Sea entonces la precisión del modelo la fracción entre la cantidad de empresas que el algoritmo eligió para invertir y que en realidad tenían la etiqueta “1”, sobre la cantidad total de empresas que el algoritmo eligió para invertir. Sea también la cobertura del modelo la fracción entre la cantidad de empresas que el algoritmo eligió para invertir y que en realidad tenían la etiqueta “1”, sobre la cantidad total de empresas que tenían la etiqueta “1” en ese set. Si se maximiza exclusivamente la precisión, se estaría eligiendo el modelo que justamente es el que menos se equivoca al asignar las etiquetas “1”; pero se tendería a elegir modelos que hacen pocas predicciones, pero seguras, dejando un portafolio muy poco diversificado. En estas situaciones la literatura recomienda ocupar como métrica de error el promedio armónico entre la precisión y la cobertura del modelo, conocido como el *F Score* ( $F_1$ ), ya que de esta manera se busca al mismo tiempo maximizar la precisión ( $P$ ) pero también que el modelo tenga una buena cobertura ( $C$ ) sobre las empresas que tienen los mejores *alfas* (Ng, 2011).

$$F_1 = 2 * \frac{P * C}{P + C} \quad (2.5)$$

Entonces, en cada año de testeo se va a tener la configuración de hiper parámetros que obtuvo el mayor valor promedio de  $F_1$  en la validación cruzada. Para testear el modelo, se entrenó esta configuración de hiper parámetros con el conjunto



de entrenamiento completo, es decir, set de entrenamiento más set de validación. Para estos modelos se calcularon las medidas de rendimiento de la clasificación de: Cantidad de empresas elegidas; Precisión; Cobertura; y  $F Score$ . Además de estas métricas de clasificación, también se hicieron otros tipos de evaluaciones de las empresas seleccionadas; estos serán explicados en el siguiente punto.

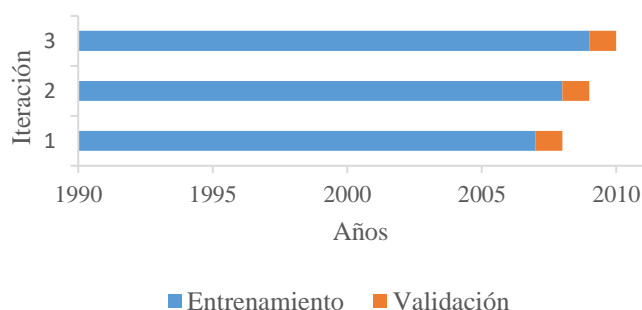


Figura 2.2: Ejemplo separación de validación y entrenamiento

Los procesos descritos anteriormente de funcionamiento, entrenamiento, validación y testeo de la red fueron programados ocupando el *software* libre GNU Octave. Octave es un lenguaje de programación científico orientado a las matemáticas con herramientas integradas de trazado y visualización (“GNU Octave”, s.f.). Se utilizó este lenguaje debido a que Andrew Ng recomienda prototipar las redes neuronales en Octave o Matlab, debido a su buen manejo con vectores y matrices. Se eligió Octave sobre Matlab debido a que este último no es un software libre.

## 2.3 Evaluación de Resultados

A continuación, se harán evaluaciones sobre el conjunto de testeo más allá de las métricas de error propias de los modelos de clasificación.

### 2.3.1 Análisis Financiero

El objetivo de la evaluación financiera es calcular la rentabilidad y el *alfa* que obtuvieron los portafolios elegidos en los diferentes años de testeo y mantenidos por un periodo de tenencia de 3 o 5 años, según corresponda. Los supuestos que se utilizaron para obtener estas métricas fueron que: (i) se puede comprar una cantidad

continúa de acciones; (ii) los datos del reporte anual de una empresa están disponibles en Bloomberg un mes después del final de su respectivo año fiscal; y (iii) no existen costos de transacción ni impuestos.

Para obtener estas métricas, primero se extrae de Bloomberg la rentabilidad mensual de todas las empresas seleccionadas; desde el inicio del año de testeo hasta que se acabe el periodo de tenencia. Estos datos son exportados a un archivo Excel a través de una Macros. Luego, la misma Macros genera una rutina para calcular los pesos del portafolio a inicio de cada mes para así calcular la rentabilidad mensual del portafolio. Con las rentabilidades mensuales del portafolio se puede calcular la rentabilidad total en el periodo de inversión; y también calcular los *alfas* haciendo una regresión lineal entre el premio de mercado y el premio del portafolio. Donde la rentabilidad del mercado se calcula como la rentabilidad mensual del S&P 500 y la rentabilidad del activo libre de riesgo se calcula como la rentabilidad mensual del bono de 10 años de Estados Unidos, ambos datos extraídos de Bloomberg. La regresión lineal se hace a través del complemento de Excel de Análisis de Datos, que otorga como resultado entre otras cosas, el valor del intercepto (*alfa*) y su valor p. El *alfa* obtenido es mensual compuesto mensual, por lo que se transformó a anual compuesto anual.

Esta evaluación de resultados financieros se codificó a través de la herramienta de programación de Excel, Macros, que ocupa el lenguaje *Visual Basic for Applications (VBA)*. Se utilizó esta herramienta debido a que cómo se necesitaba datos de Bloomberg para calcular las métricas, era más fácil ocupar los complementos de Bloomberg en Excel.

Con esta evaluación entonces se obtiene el retorno del portafolio, el retorno del mercado para el mismo periodo, el *alfa*, su intervalo de confianza y su valor p. Para entender la lógica que sigue el portafolio y cómo se calcula los pesos de este, se puede visitar el Anexo I.

### **2.3.2 Evaluación Graham**

La evaluación Graham se refiere a la corroboración de si las empresas elegidas por el MSA cumplen ciertos criterios que Benjamin Graham estipuló para

evaluar acciones en su libro *The Intelligent Investor* (1949). El profesor de la Universidad de Columbia propuso rangos aceptables para los valores de 8 parámetros claves que los inversores defensivos deberían verificar en las empresas que eligen para comprar acciones. Él argumentaba que, si una empresa cumplía estos criterios, lo más probable es que esta sea una compañía que está generando beneficios prolongadamente en el tiempo, que tiene una situación financiera estable y que está siendo subvalorada injustificadamente. A continuación, se detallan los parámetros propuestos por Graham en el 1972; junto a sus rangos aceptables y el objetivo específico que estos representan.

- **Ventas:** Mayor a 100 millones de dólares. Busca filtrar a empresas pequeñas que pueden quedar expuestas a más vicisitudes que la media.
- **Ratio Circulante:** Mayor o igual a 2. Busca una liquidez financiera sólida.
- **Capital de Trabajo y Deuda de Largo Plazo:** Deuda de Largo Plazo debe ser menor al capital de trabajo. También busca estabilidad financiera.
- **Historial Beneficios:** Haber obtenido un cierto beneficio atribuible al capital ordinario en cada uno de los diez últimos ejercicios. Sirve para verificar si es que la empresa ha generado beneficios a lo largo de los años; ya que la probabilidad de que siga generando utilidades en el futuro es alta si es que ha obtenido ganancias constantemente en un periodo lo suficientemente largo.
- **Historial Dividendos:** Haber llevado a cabo pagos de dividendos de manera ininterrumpida por lo menos durante los últimos 20 ejercicios. Mismo argumento que en el punto anterior; pero ahora se busca que los beneficios que ha generado la empresa se les entregue a los inversores.
- **Crecimiento del Beneficio:** Haber conseguido un incremento mínimo de un tercio, por lo menos, en el beneficio por acción durante los tres últimos años. Se busca comprobar si la empresa va a tener capacidad de aumentar los beneficios que va obteniendo.
- **PER:** La cotización actual no debe ser un múltiplo de más de 15 veces las utilidades medias correspondiente a los tres últimos ejercicios. De esta forma se asegura que no se esté pagando un precio muy elevado por la empresa.
- **Precio/Valor Contable:** El precio actual no debiera ser un múltiplo de más de 1,5 veces el valor contable que se haya declarado en la última memoria. Complementa al criterio anterior.

Estos requisitos fueron configurados por Graham para atender a las necesidades de los inversores defensivos. Eliminarán a la inmensa mayoría de las acciones ordinarias como candidatas a la cartera, ya que excluirán a las empresas que sean demasiado pequeñas, se encuentren en condiciones financieras relativamente insatisfactorias, tengan una lacra de déficit en su historial a 10 años, y no tengan un prolongado historial de abono continuo de dividendos (Graham, 1949).

Ahora bien, la última vez que estos criterios fueron revisados por Graham fue en el 1972, por lo que se podría pensar que los rangos propuestos están desactualizados y que pasaron a ser irrelevantes o muy estrictos. Para lidiar con esta limitación se plantea modernizar los rangos propuestos por Graham. Se propone revisar los percentiles que representan estos rangos en los miembros del índice S&P 500 en el 1972 para luego ver los valores de los parámetros que están en esos percentiles en los años de testeo. Por ejemplo, si el ratio circulante igual a 2 está en el percentil 0,8 para los miembros del S&P en el año 1972, entonces se busca este percentil para el año 2011, por lo que se aplica el criterio impuesto por Graham ya actualizado. En el Anexo J se observa los nuevos rangos propuestos debido a la actualización y se explica por qué se evalúan 7 criterios en vez de 8.

Entonces, lo que se busca hacer en esta evaluación es comprobar si es que los portafolios seleccionados para los años de testeo cumplen los criterios de Graham ajustados. Para poder hacer esto se compararon dos grupos, las empresas seleccionadas y las no seleccionadas. Se calculó el promedio de cuantos criterios Graham actualizados cumplen ambos grupos para los años de testeo, para luego hacer una prueba *t de Student* y verificar si es que la diferencia en los promedios observados se dio por chance o por que efectivamente las empresas seleccionadas cumplen más (o menos) criterios que las empresas no seleccionadas, con significancia estadística. En este análisis se va a obtener como resultado el promedio de cuantos criterios Graham ajustados cumple el portafolio y las empresas no seleccionadas y también observar si es que la diferencia entre ellos es significativa a través del valor p de la prueba estadística *t de Student*.

### III. RESULTADOS

#### 3.1 Resultados Validación Cruzada

En esta sección se mostrarán los resultados de la validación cruzada, donde se seleccionó la configuración de hiper parámetros que maximiza el  $F_1$  promedio de los tres años de evaluación para cada año de testeo. Se debe recordar que para el conjunto de testeo se dejan los últimos tres años con información disponible, lo que es equivalente para el periodo de tenencia de 3 años los años fiscales del 2013, 2014, 2015; y para el periodo de tenencia de 5 años los años fiscales del 2011, 2012 y 2013.

En las Tablas 3.1 y 3.2 se muestran la base de datos, cantidad de capas escondidas y parámetro de regularización que maximizaron el  $F_1$  promedio para cada año de testeo para un periodo de tenencia de 3 y 5 años, respectivamente. En el Anexo K se muestra los  $F_1$  promedios de todas las configuraciones evaluadas.

Tabla 3.1: Resultados Validación Cruzada, periodo de tenencia 3 años

Año de Testeo	2013	2014	2015
Base de Datos	2	2	2
Cantidad de Capas Escondidas	1	2	3
Lambda	0,32	0,02	0,02
F Score Promedio	42,7%	41,9%	40,6%
Precisión Promedio	46,9%	41,1%	37,8%
Cobertura Promedio	41,2%	45,7%	45,6%

Tabla 3.2: Resultados Validación Cruzada, periodo de tenencia 5 años

Año de Testeo	2011	2012	2013
Base de Datos	2	2	2
Cantidad de Capas Escondidas	2	2	1
Lambda	0,02	0,02	0,04
F Score Promedio	45,3%	49,4%	49,5%
Precisión Promedio	43,5%	52,9%	61,7%
Cobertura Promedio	50,4%	48,1%	45,3%

En estos resultados sobresale la predominancia absoluta de la base de datos número 2, es decir, la base de datos con eliminación óptima y parámetros fijo, en la cual se fijaron los 8 parámetros relevantes para Graham explicados en el punto 2.3.2. Además, en el Anexo K, se puede observar que la base de datos número 2 ocupa las 23 mejores posiciones en cuanto a  $F_1$  promedio para el periodo de tenencia de 3 años y las 33 mejores posiciones para el de 5 años, si se juntan todos los años de testeo. Que

la base de datos predomine sobre las otras sugiere dos cosas: en primer lugar, que los parámetros de Graham ayudan a la red a ser más efectiva en su selección, ya que la red tiene mejor rendimiento cuando se agregan estos parámetros a pesar del costo de eliminar ejemplos de empresas. En segundo lugar, no es intuitivo el por qué la base de datos número 2 obtiene mejor rendimiento que la base de datos número 3 (la imputación), ya que esta posee más información que las demás al ser la base de datos más grande en cuanto ejemplos y parámetros. El autor insinúa dos posibles razones, quizás el parámetro  $F_1$  favorece a la base de datos número dos, o la imputación no fue correctamente realizada, dificultando el trabajo de la red para clasificar. Esto requiere mayor profundización.

En cuanto la cantidad de capas escondidas hay más dispersión, lo cual realza la importancia de la validación cruzada. En la mayoría de los casos se elige más de una capa escondida, que es lo más común para problemas con pocos datos; por lo que si se hubiera seguido este ejemplo se llegaría a modelos con peores resultados.

Por último, se puede observar que los valores para  $\lambda$  son pequeños, lo cual sugiere que no se da mucha relevancia a la regularización y, por ende, que no tiene importancia al *overfitting*. Por lo tanto, existe una posibilidad que se esté en presencia de *underfitting*, es decir, que la red no se ajuste bien a los datos. Lo cual se observa con los bajos resultados de validación para las métricas de clasificación. Esto era esperable, debido a lo difícil que es separar las empresas sobresalientes de las que no. La decisión de inversión en una empresa es un problema con alta incertidumbre, en donde no existe experto ni algoritmo que haya ganado más que el mercado en todas las empresas que ha decidido invertir. Por lo que era de esperar resultados relativamente deficientes de precisión y cobertura.

### **3.2 Resultados Testeo**

El testeo tiene como objetivo evaluar el algoritmo con años que el modelo no ha utilizado para ajustar los parámetros. En las tablas 3.3 y 3.4 se muestran los resultados obtenidos para los años de testeo de los periodos de tenencia de 3 y 5 años, respectivamente.

Tabla 3.3: Resultados testeo, periodo de tenencia 3 años

<b>Año Testeo</b>	<b>2013</b>	<b>2014</b>	<b>2015</b>
<b>Cantidad Empresas Seleccionadas</b>	9	33	34
<b>F Score</b>	17,1%	15,2%	18,8%
<b>Precision</b>	66,7%	21,2%	26,5%
<b>Recall</b>	9,8%	11,9%	14,5%
<b>Alfa Aca</b>	12,5%	-1,2%	2,5%
<b>Valor P</b>	0,07	0,74	0,45
<b>Retorno Anualizado Portafolio</b>	25,6%	10,9%	8,7%
<b>Retorno Anualizado Mercado</b>	12,5%	12,4%	7,1%
<b>Diferencia</b>	13,13%	-1,56%	1,59%
<b>Promedio Cantidad Criterios Graham Cumplidos ES</b>	4,44	4,55	4,50
<b>Promedio Cantidad Criterios Graham Cumplidos ENS</b>	4,65	4,64	4,61
<b>Valor P</b>	30,6%	66,6%	62,1%

Tabla 3.4: Resultados testeo, periodo de tenencia 5 años

<b>Año Testeo</b>	<b>2011</b>	<b>2012</b>	<b>2013</b>
<b>Cantidad Empresas Seleccionadas</b>	35	54	4
<b>F Score</b>	23,5%	22,0%	3,4%
<b>Precision</b>	28,6%	18,5%	25,0%
<b>Recall</b>	20,0%	27,0%	1,9%
<b>Alfa Aca</b>	3,6%	1,0%	-0,03%
<b>Valor P</b>	0,114	0,540	0,996
<b>Retorno Anualizado Portafolio</b>	15,3%	15,31%	10,14%
<b>Retorno Anualizado Mercado</b>	12,3%	14,25%	11,26%
<b>Diferencia</b>	2,99%	1,05%	-1,11%
<b>Promedio Cantidad Criterios Graham Cumplidos ES</b>	4,47	4,52	5,25
<b>Promedio Cantidad Criterios Graham Cumplidos ENS</b>	4,61	4,59	4,64
<b>Valor P</b>	0,459	0,685	0,307

Observando las Tablas 3.3 y 3.4, se puede observar que se cumple el objetivo de generar portafolios diversificados a través de maximizar el parámetro  $F_1$  para la clase de empresa sobresaliente; excepto para el año 2013 en ambos periodos de tenencia. Lo cual se puede explicar por la baja cantidad de capas escondidas que ocupan los modelos de es año, ambos de sólo una. Es probable que estos modelos más sencillos generen una zona de aceptación menos compleja y, para vencer a los otros modelos, tiendan a aumentar la precisión en desmedro de la cobertura. Lo cual se observa en las Tablas 3.1 y 3.2; los modelos del 2013 son aquellos que obtienen la mayor precisión y menor cobertura en la validación de los años a testear, para ambos periodos de tenencia.

Las métricas de clasificación, en general, sufrieron bajas significativas en comparación a la validación cruzada. El  $F_1$  disminuyó en promedio un 59,1% para el periodo de tenencia de 3 años y un 65,5% para el de 5. Estos cambios se deben, en general, a un menor rendimiento en las dos métricas relevantes, precisión y cobertura. Ahora bien, la disminución de la cobertura es más grande que la disminución en la precisión, como se puede observar en la Tabla 3.5, la cual muestra las disminuciones porcentuales entre la validación y el testeo.

Tabla 3.5: Disminuciones porcentuales entre la validación y el testeo

Periodo de Tenencia	3 Años	5 Años
<b>Disminución F Score Promedio</b>	59,1%	65,5%
<b>Disminución Precisión Promedio</b>	12,1%*	52,9%
<b>Disminución Cobertura Promedio</b>	72,8%	66,7%

\*El bajo cambio en la precisión promedio para el periodo de tenencia de 2013 se debe al buen rendimiento que se obtuvo para el 2013, una precisión de 66,7%. Sin contar este año, el promedio sería de 39,2%.

Estos resultados son una clara señal del *underfitting*, ya que la red mostró un rendimiento deficiente que empeoró en el testeo. La red no se ajusta bien al conjunto de entrenamiento y tampoco es capaz de generalizar bien con datos que no ha visto. Ahora bien, es relevante recordar que la clasificación es entre empresa sobresaliente y empresa no sobresaliente; por lo que existe la posibilidad de que las empresas seleccionadas que no fueron sobresalientes aún sean lo suficientemente buenas como para que el portafolio creado obtenga resultados mejores que el mercado.

En cuanto a las métricas financieras, los portafolios obtuvieron resultados relativamente satisfactorios para los *alfas*. En dos de los tres años de testeo se obtuvo un *alfa* positivo, acompañado de una rentabilidad superior a la del mercado, para ambos periodos de tenencia. Lamentablemente, el valor p no fue menor 0,05 en ninguno de los casos de *alfa* positivo, lo cual se traduce en que los resultados en donde los *alfas* son mayores a cero no son estadísticamente significativos con un intervalo de confianza de 95%. Por lo que existe probabilidad importante que estos resultados se hayan realizado por chance en vez de por el buen manejo de cartera del MSA. Sin embargo, en dos casos se estuvo cerca de tener significancia estadística, para el año 2013 en el periodo de tenencia de 3 años y en el 2011 para el periodo de tenencia de 5 años, con un valor p de 0,067 y 0,114, respectivamente.



Si se observa las diferencias obtenidas entre la rentabilidad anualizada del portafolio y el mercado de las tablas 3.3 y 3.4 se puede advertir que estas no son considerables, tanto cuando es positiva o cuando es negativa; a excepción del año 2013 para el periodo de tenencia de 3 años. Lo más relevante de esta diferencia es observar que, en los años de testeo, las diferencias negativas (donde se perdió contra el mercado) no son grandes; por lo que se podría pensar que los resultados de los portafolios no son tan insatisfactorios. Entonces, los portafolios seleccionados tienden a superar al mercado en los años de testeo, y cuando pierden no son por diferencias muy grandes, pero no existe evidencia estadística para asegurar que estos resultados no se dieron por azar.

En las Figuras 3.1 y 3.2 se observa la rentabilidad acumulada de los portafolios en los tres años de testeo para los periodos de tenencia de 3 y 5 años respectivamente. En estas figuras se observa que, generalmente, los portafolios creados por el modelo de selección de acciones tienen un comportamiento similar al portafolio de mercado. Probablemente, esto se deba a que el portafolio mientras más se diversifique, más similar sea al portafolio de mercado ya que el MSA elige acciones miembros de este último. Además, los portafolios que poseen una baja cantidad de empresas (los portafolios comprados en el 2013 y mantenidos por 3 y 5 años) no siguen un rendimiento similar al mercado; lo cual se podría explicar con el argumento anterior.

Por último, en cuanto a los criterios de Graham se puede observar que en casi todos los escenarios testeados las empresas no seleccionadas cumplen más criterios que las empresas seleccionadas, en promedio. Ahora bien, los valores  $p$  del test *t de Student* son bastantes altos, ninguno es menor a 0,3 (un valor muy alto para nivel de significancia). Por lo que, si bien existe una diferencia, no hay evidencia suficiente para rechazar que los promedios sean iguales. Esto sugiere que las diferencias de estos promedios fueron aleatorias y que el modelo de selección de acciones no elige acciones que cumplan más o menos criterios de Graham. Al final, el algoritmo no sigue una estrategia similar a la que Graham propuso.

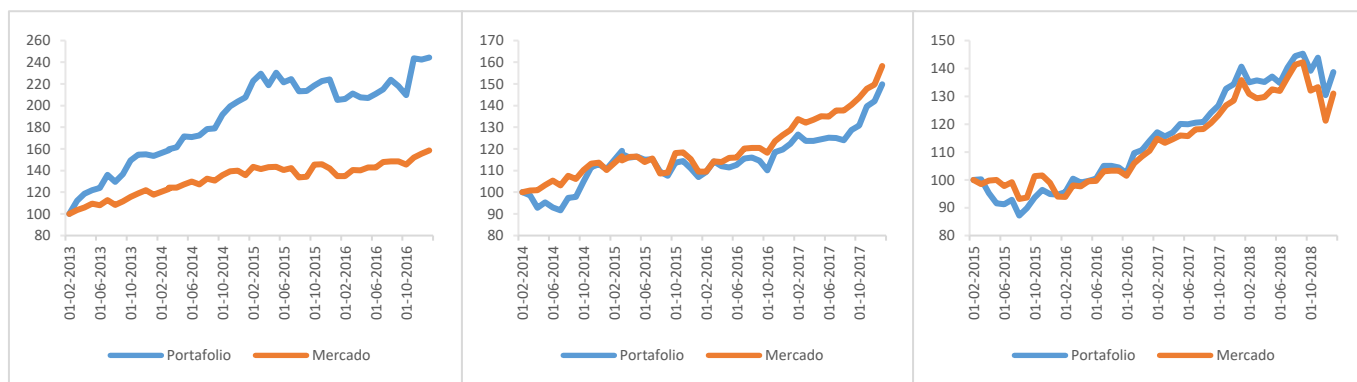


Figura 3.1: Rentabilidad mensual acumulada, periodo de tenencia de 3 años. (De izquierda a derecha: 2013, 2014, 2015)

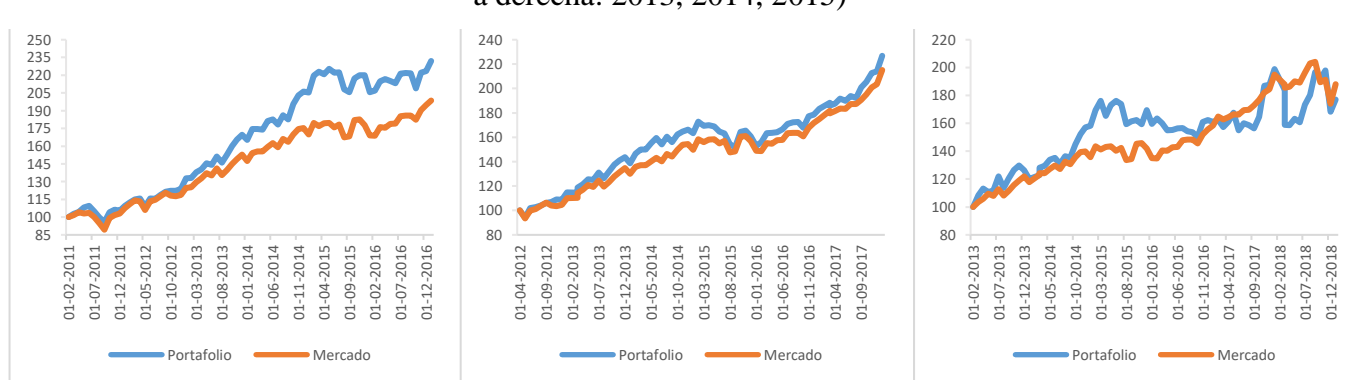


Figura 3.2: Rentabilidad mensual acumulada, periodo de tenencia de 5 años. (De izquierda a derecha: 2011, 2012 y 2013)

Una posible razón de esto es que el modelo intenta seleccionar a las empresas que tendrán *alfas* sobresalientes en los años que se decida mantener la acción en la cartera, las cuales, probablemente, no sean las mismas que Graham recomienda comprar. Como se describió anteriormente, Benjamin Graham propone comprar empresas en las que existan discrepancias entre su precio con su verdadero valor pero que sean financieramente estables, y muy probablemente estas empresas no son las que obtienen *alfas* sobresalientes; si no aquellas que obtienen *alfas* positivos pero no tan grandes.

## IV. CONCLUSIÓN

### 4.1 Conclusiones del Trabajo

Con lo expuesto en el punto anterior, queda en evidencia que la red neuronal tiene *underfitting*, es decir, se ajusta mal tanto al conjunto de entrenamiento como al conjunto de testeo. Esto era un resultado esperable, debido a la dificultad de la clasificación. Andrew Ng menciona en su curso de *Machine Learning* que para saber si los parámetros con los que se alimenta a una red son suficientes para tener una clasificación aceptable, se debe pensar si es que un experto es capaz de hacer una correcta asignación con la información disponible. En las inversiones no existe un experto que logre separar correctamente a las empresas que van a superar al mercado de las que no, por lo que existía una alta probabilidad de que el modelo poseyera *underfitting*.

A pesar de esto, no es necesario que el modelo de selección de acciones sea un excelente clasificador; basta con que obtenga una precisión para la etiqueta “1” (sobresaliente) aceptable para que así los portafolios superen al mercado. Entonces, existe una probabilidad que se pueda obtener mejores resultados de rentabilidad si es que se mejora este problema de subajuste.

Andrew Ng afirma que la mejor forma de corregir el *underfitting* es a través de la incorporación de más parámetros. Entonces, a partir de esto, el estudiante propone dos líneas de acción para mejorar el subajuste. En primer lugar, se podría intentar recolectando más parámetros en el *software* financiero Bloomberg, intentando que estos sean más complejos. En segundo lugar, se podría probar ocupar la base de datos número tres, es decir, la imputada. Esta es la base de datos que más información posee (tanto parámetros como ejemplos) y, si bien no obtuvo los mejores resultados de validación, quizás la red entrenada con más parámetros se generalice mejor a los años de testeo. También, bajo este mismo punto, se podría mejorar la imputación realizada y ocupar algoritmos más sofisticados para esto.

Una tercera línea de acción que se podría realizar, pero que no tiene relación con la cantidad de parámetros, es cambiar la métrica que decide cual modelo es el mejor en la validación; del parámetro  $F_1$  por la precisión del modelo. De esta

manera se estaría enfatizando en la necesidad de sólo elegir empresas sobresalientes, despreciando la necesidad de tener portafolios diversificados (simplificando la diversificación a portafolios más grandes). Estas posibles mejores no fueron testeadas en este trabajo de título debido a la falta de tiempo.

Por otro lado, en la descripción se mencionó las ventajas de utilizar métodos de *Machine Learning* en las inversiones, por lo que se estima necesario compartir algunas conclusiones sobre los aprendizajes obtenidos de trabajar con este tipo de algoritmos. Si bien es verdad que poseen una capacidad importante de manejar mucha información y que automatizan el proceso de selección de acciones para la inversión; también poseen desventajas.

En primer lugar, se tiene un subajuste esperado dado por la dificultad de clasificar las acciones sobresalientes de las no sobresalientes. Como un experto no puede hacerlo bien, es esperado que un algoritmo no lo haga perfecto.

En segundo lugar, se tiene irregularidad en la obtención de los parámetros para las empresas. Bloomberg presenta muchos datos faltantes en la base de datos extraída lo cual obliga a generar estrategias para manejar los datos faltantes, las cuáles no son óptimas.

En tercer lugar, la base de datos extraída y corregida con la eliminación óptima o con la imputación, aún es muy pequeña comparada con las bases de datos que se utilizan en *Machine Learning*. Generalmente, los problemas que se resuelven con aprendizaje automático tienen matrices de *inputs* de billones de datos, mientras que la que se utilizó en este trabajo ronda los 800.000. Se podría hacer un esfuerzo y obtener más parámetros para la inversión en acciones en el largo plazo, pero nunca a este nivel. Las empresas liberan su información financiera trimestralmente, por lo que los datos disponibles aún siguen siendo muy pocos en comparación a, por ejemplo, la cantidad de imágenes con las que se alimenta a un algoritmo que reconoce patentes de los automóviles en la autopista.

Por último, otra desventaja es el tiempo en que estos tipos de algoritmo se demoran en entrenar. Las redes neuronales son algoritmos que ajustan una cantidad considerable de parámetros, lo cual los transforma en un problema de optimización

costoso, computacionalmente hablando. Lo cual dificulta la estrategia de “prueba y error” para ir mejorando el rendimiento de la red.

Con todo lo mencionado en este reporte, ¿se debería ocupar el portafolio creado por el MSA o invertir en el índice de mercado? Si bien el modelo no clasifica con una precisión aceptable a las empresas sobresalientes, el portafolio automatizado igual obtiene un mejor rendimiento que el mercado en la mayoría de los años de testeo; y cuando pierde es por poco. A pesar de esto, no se tiene un nivel de confianza lo suficientemente estricto como para pensar que estos resultados no fueron aleatorios. Además, se debe considerar que en la evaluación de resultados no se está considerando los costos por transacción; que la intuición dice que deberían ser mayores para la compra de acciones individuales que para un fondo que sigue el índice de mercado. Con estos dos argumentos, la decisión tendería a ser invertir en el índice de mercado. Ahora bien, si con las líneas de acción propuestas, se puede replicar con mayor frecuencia las rentabilidades obtenidas para el 2013 en el periodo de tenencia de 3 años, entonces si sería recomendable invertir en el portafolio.

## **4.2 Conclusiones Competencias**

Para finalizar, se desea terminar evidenciando el cómo se cumplieron las competencias del perfil de egreso para este trabajo de título. Específicamente de: (i) aplicar diversos métodos de análisis de datos para la comprensión de los fenómenos abordados; (ii) realizar análisis financieros y de costos aplicados a los requerimientos de instituciones privadas y públicas; y (iii) aplicar los fundamentos de la ingeniería e investigación operacional, al modelamiento y resolución de problemas complejos.

Se aplicaron diversos métodos de análisis para la comprensión de los fenómenos abordados. En primer lugar, se calcularon métricas propias de la clasificación para observar el rendimiento de la red neuronal, precisión, cobertura y *F Score*. Estas dan mayor claridad al observar los resultados ya que miden lo relevante para la clasificación de empresas sobresalientes; la precisión mide si es que el MSA clasifica bien la empresa sobresaliente y la cobertura mide la cantidad de empresas clasificadas como sobresalientes del total. Luego, se evaluó financieramente los resultados del portafolio en comparación al mercado. Se calculó la rentabilidad y el *alfa* del portafolio viendo si es que el portafolio le gana tanto en rentabilidad absoluta

y rentabilidad riesgo ajustada; donde también se quiso ver la robustez del *alfa* a través del valor  $p$  que obtuvo en la regresión lineal. Por último, se comprobó si es que el portafolio generado a través del MSA alimentado con parámetros relevantes según la filosofía *Value Investing* cumple los criterios Graham actualizados. Para esto se observó la cantidad de parámetros Graham actualizados que cumplía cada una de las empresas del portafolio y de las empresas no seleccionadas. Se calculó el promedio de cada grupo y se verificó a través de una prueba *t Student* si es que esta diferencia era estadísticamente significativa.

En cuanto a los análisis financieros de instituciones privadas, se cumplió de tres formas. En primer lugar, se cumplió con la revisión bibliográfica de la inversión en valor, donde se aprendieron los conceptos relevantes para decidir cuando una empresa califica como oportunidad de inversión según esta filosofía, y cuáles parámetros son los relevantes para identificar si estos conceptos se cumplen. En segundo lugar, se calculó la rentabilidad del portafolio y también el *alfa* ocupando la fórmula de CAPM y se comparó sus ganancias con el mercado. Se comparó la rentabilidad anualizada del portafolio y la del mercado para ver el rendimiento del primero; pero también se calculó el *alfa* para obtener un parámetro riesgo ajustado. Así se puede saber si es que el portafolio obtiene una rentabilidad mejor que al mercado por tomar mejores decisiones de inversión y no porque tomó más riesgos que el índice. En tercer lugar, se aprendió a ocupar el *software* financiero de Bloomberg. Este *software* es el más utilizado para el estudio de acciones y bonos en el mundo, por lo que tener manejo de este otorga al estudiante herramientas para realizar análisis financieros más complejos.

Por último, se aplicaron los fundamentos de la ingeniería e investigación operacional, al modelar tanto la estrategia de eliminación óptima de datos faltantes y la red neuronal artificial, con el fin de formular un portafolio que supere al mercado. Para los datos faltantes se construyó un problema de optimización que tendría como solución la cantidad óptima de filas y columnas para así maximizar el tamaño de la matriz el cuál se programó en *Python* y se resolvió con *Gurobi*. También se ocupó el paquete *mice*, disponible en el programa estadístico R, para obtener una matriz imputada prediciendo los datos faltantes a través de un conjunto de densidades condicionales. Para el MSA se construyó, entrenó y testeó una red neuronal artificial

que clasificaba las empresas sobresalientes de las que no lo eran. Para la construcción de la red se aplicó álgebra lineal para vectorizar el problema. Para el entrenamiento se debió solucionar un problema de minimización con una cantidad importante de datos, el cual se programó en *Octave* y se solucionó con un optimizador propio de este programa. Para el testeo se ocuparon técnicas estadísticas para verificar la robustez de los datos.

## BIBLIOGRAFIA

Chee, S. Sloan, R & Uysal, A. (2013). A framework for value investing. *Australian Journal of Management*, 38(3), 599-633.

Clifton, C. (s.f.). *Data mining*. Recuperado de:  
<https://www.britannica.com/technology/data-mining>

“Definition of 'Data Mining'”. (s.f.). Recuperado de  
<https://economictimes.indiatimes.com/definition/data-mining>

Cochrane, C. (2018). *Time Series Nested Cross-Validation*. Recuperado de:  
<https://towardsdatascience.com/time-series-nested-cross-validation-76adba623eb9>

“Cross-Validation strategies for Time Series forecasting”. (2019).  
 Recuperado de: <https://hub.packtpub.com/cross-validation-strategies-for-time-series-forecasting-tutorial/>

“GNU Octave”. (s.f). Recuperado de:  
<https://www.gnu.org/software/octave/>

Faggella, D., Azulay D. & Bharadwak, R. (2019). *AI in Banking: Executive Cheat Sheet*. Recuperado de: <https://emerj.com/report>

Graham, B. (1949). *El Inversor Inteligente*. (1a ed.). Barcelona, España: [Deusto].

Kaplan, A. Heanlein, M. (2018). Siri, Siri, in my hand: Who’s the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. *Business Horizons*, 62(1), 15-25.

Kunwar, M. (2019). *ARTIFICIAL INTELLIGENCE IN FINANCE. Understanding how automation and machine learning is transforming the*



*financial industry*. (Tesis maestría inedita). Centria University of Applied Sciences, Kokkola, Finlandia.

“*La Terminal Bloomberg*”. (s.f.). Recuperado de:  
<https://www.bloomberg.com/latam/solucion/la-terminal/>

Leopold H., Krenn W., King, R. & Mateis, C. (2019). *Artificial Intelligence Landscape - An Introduction in Technology Fields & Research Areas*. Recuperado de:  
[https://www.researchgate.net/publication/332471378\\_Artificial\\_Intelligence\\_Landscape\\_-\\_An\\_Introduction\\_in\\_Technology\\_Fields\\_Research\\_Areas](https://www.researchgate.net/publication/332471378_Artificial_Intelligence_Landscape_-_An_Introduction_in_Technology_Fields_Research_Areas)

Machine-learning promises to shake up large swathes of finance. (2017, 17 de Mayo). *The Economist*. Recuperado de:  
<https://www.economist.com/finance-and-economics/2017/05/25/machine-learning-promises-to-shake-up-large-swathes-of-finance>

Ng, A. (2011). *Machine Learning* [Material del aula]. Machine Learning, Stanford University. California, United States of America. Recuperado de: <https://www.coursera.org/learn/machine-learning>

Ray, S. (2018). *Improve Your Model Performance using Cross Validation (in Python and R)*. Recuperado de:  
<https://www.analyticsvidhya.com/blog/2018/05/improve-model-performance-cross-validation-in-python-r/>

Reyes, T. (2018). *Teoría de Portafolios: Definición y relación riesgo-retorno* [Material del aula]. Finanzas, Pontificia Universidad Católica de Chile, Santiago, Chile.

Roy, B. (2019). *All About Missing Data Handling*. Recuperado de:  
<https://towardsdatascience.com/all-about-missing-data-handling-b94b8b5d2184>

Rise of the machines. (2015, 9 de mayo). *The Economist*. Recuperado de: <https://www.economist.com/briefing/2015/05/09/rise-of-the-machines>

Schroer, A. (2019). *AI AND THE BOTTOM LINE: 15 EXAMPLES OF ARTIFICIAL INTELLIGENCE IN FINANCE*. Recuperado de: <https://builtin.com/artificial-intelligence/ai-finance-banking-applications-companies>

Singh, N. (2018). *Artificial Intelligence and It's Sub-Fields*. Recuperado de: <https://medium.com/@neha49712/artificial-intelligence-and-its-sub-fields-a5a63d8263e8>

Stern, M. (2019, 11 de noviembre). McDonald's AI Drive-Thrus May Be Too Smart For Their Own Good. *Forbes*. Recuperado de: <https://www.forbes.com/sites/retailwire/2019/11/11/mcdonalds-ai-drive-thrus-may-be-too-smart-for-their-own-good/#13e9439e6910>

Swalin, A. (2018). *How to Handle Missing Data*. Recuperado de: <https://towardsdatascience.com/how-to-handle-missing-data-8646b18db0d4>

“*The Fastest Solver*”, (s.f.). Recuperado de: <https://www.gurobi.com/es/>

The stockmarket is now run by computers, algorithms and passive managers. (2019, 5 de octubre). *The Economist*. Recuperado de: <https://www.economist.com/briefing/2019/10/05/the-stockmarket-is-now-run-by-computers-algorithms-and-passive-managers>

van Buuren, S. & Groothuis-Oudshoorn, K. (2011). mice: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software*, 45(3).

Vera, J. (2018). *Métodos de Descenso*. [Material del aula]. Métodos de Optimización, Pontificia Universidad Católica de Chile, Santiago, Chile.

Zavadskaya, A. (2017). *Artificial Intelligence in Finance: Forecasting Stock Market Returns Using Artificial Neural Networks*. (Tesis maestría inédita). Hanken School of Economics, Helsinki, Finlandia.

Zheng, X., Zhu M., Li, Q., Chen C. & Tan Y. (2019). Review: FinBrain: when finance meets AI 2.0. *Frontiers of Information Technology & Electronic Engineering*, 914-924.

## **ANEXOS**

## ANEXO A: INTELIGENCIA ARTIFICIAL

La Inteligencia Artificial (IA) es un campo de estudio que tiene como objetivo replicar la inteligencia humana a través de tecnologías artificiales para crear máquinas inteligentes que puedan aprender y mejorar con el tiempo (Kunwar, 2019; Zavadskaya, 2017). La inteligencia artificial puede ser definida entonces como la capacidad de un sistema para interpretar datos externos correctamente, aprender de dichos datos y utilizar esos aprendizajes para lograr objetivos y tareas específicos a través de una adaptación flexible (Kaplan & Haenlein, 2018). Sus capacidades especiales de resolución de problemas pueden aplicarse con éxito y eficacia en una amplia serie de áreas, como, por ejemplo, medicina, finanzas, video juegos, transporte, manufactura, defensa, entre muchísimas otras (Leopold, Krenn, King & Mateis, 2019).

La IA utiliza información externa como datos de entrada para identificar reglas y patrones subyacentes que, recurriendo a enfoques de *machine learning* (ML), describen métodos que ayudan a las computadoras a aprender sin ser programados explícitamente. Estos métodos pueden ser relativamente simples (como la regresión lineal) o increíblemente complejos (como las redes neuronales profundas o *Deep learning*). Ahora bien, la IA es más amplia que el ML, ya que esta también cubre la capacidad de un sistema para percibir datos (por ejemplo, procesamiento de lenguaje natural o reconocimiento de voz / imagen) o para controlar, mover y manipular objetos basadas en la información aprendida; ya sea un robot u otro dispositivo conectado (Kaplan & Haenlein, 2018).

Este es un campo de investigación joven, el cual se ha estudiado durante los últimos 60 años. Alan Turing, matemático y descifrador de códigos de la Segunda Guerra Mundial, es reconocido como uno de los principales individuos en impulsar la idea de las máquinas que “piensan”. Incluso inventó la prueba de Turing, que se usa hoy en día como una guía para decidir la capacidad de “pensar” como humano de la máquina. Sus investigaciones pusieron las ruedas en movimiento y desde entonces los investigadores han estado interesados en la creación de una “máquina pensante” que pueda imitar los procesos del cerebro humano (Kunwar, 2019; Zavadskaya, 2017).

La investigación de IA ha experimentado tres olas distintivas. La primera, impulsada por los descubrimientos de Alan Turing, enfrentaba limitaciones de la

potencia de procesamiento de las computadoras y la capacidad computacional; por lo que no logró avanzar mucho más. La segunda ola apareció en la década de 1980, con el desarrollo de las redes neuronales artificiales que imitaban las funciones del cerebro humano, los sistemas de aprendizaje automático y la capacidad computacional mejorada de las computadoras (Zheng, Zhu, Li, Chen & Tan 2019). Actualmente, el mundo está experimentando una tercera ola de investigación de IA, impulsada por el desarrollo del *Deep learning* que, si bien fue inventado en 1980 por John Hopfield y David Rumelhart, en estos últimos años ha encontrado capacidad computacional para poder desarrollarse en plenitud. El *Deep learning* permite la creación de un sistema neuronal mucho más complejo y una aplicación más eficiente en el mundo real. La diferencia distintiva entre este y los tipos anteriores de IA es la falta de intervención humana en el sistema. Si bien los programadores deben preparar datos de capacitación, entrenar, formular lógica analítica y evaluar los resultados y la precisión del sistema para el ANN, no son necesarios para la capacitación del sistema de aprendizaje profundo, ya que se entrena a sí mismo (Zavadskaya, 2017).

La IA presenta siete subáreas principales: Sistemas Expertos, Robótica, Planificación Automática, Procesamiento del Habla, Visión Artificial, Aprendizaje Automático (*Machine Learning* o ML) y Proceso Natural del Lenguaje (*Natural Language Processing* o NLP) (Kunwar, 2018). En la Figura A.1 se muestra un resumen de estas áreas.

En este informe se trabajó con *Machine Learning*, específicamente en el área supervisada. El ML es un proceso por el cual el computador aprende a realizar una tarea mediante métodos que buscan patrones recurrentes en los datos, reconoce reglas o divide los datos en clases (Leopold et al., 2019). De manera general, a un algoritmo de ML se le debe entregar los parámetros de aprendizaje de muchos ejemplos de entrenamiento, para luego “entrenarlo” con los datos con alguno de los métodos anteriormente descritos, para que así aprenda a realizar la tarea que se le desea enseñar al computador. Por ejemplo, si se desea entrenar a un algoritmo a predecir el precio de una casa, se le debe entregar como parámetros de aprendizaje el tamaño, la ubicación, la cantidad de habitaciones, etc., de muchos ejemplos de entrenamiento, es decir, de muchas casas, para que así el computador aprenda hacer la tarea de predecir el precio de una casa mediante alguno de los métodos descritos.



Figura A.1: Subáreas de la IA (Kuwat, 2018)

En un escenario “supervisado”, la máquina recibe los parámetros de aprendizaje de un ejemplo de entrenamiento sabiendo de antemano cual es el resultado esperado; es decir, los parámetros vienen “etiquetados”. Por ejemplo, si se desea tener un algoritmo de clasificación de spam, entonces en la etapa de entrenamiento el computador deberá conocer cuales mails son spam y cuales no (Ng, 2011). Además, según Andrew Ng, profesor de la Universidad de Standford, para que un algoritmo de ML funcione correctamente un experto debe poder hacer la tarea correctamente, dado los mismos parámetros de aprendizaje (2011). Por ende, la base de datos de entrenamiento debe estar “etiquetada”, lo que resulta un requisito importante para la generación de esta. Pero, además, la base de datos debe ser lo suficientemente grande para que el computador logre aprender los patrones, los datos deben ser apropiados para permitir los efectos de aprendizaje correctos, y los datos no deben contener ningún error; ya que de lo contrario el sistema de IA los aprendería (Leopold et al., 2019).

## ANEXO B: APLICACIONES DE IA EN EL SECTOR FINANCIERO

La IA posee distintas aplicaciones que recorren las áreas de medicina, banco, video juegos, transporte, manufactura, defensa, entre muchísimas otras. Por ejemplo, en la medicina se han creado algoritmos de reconocimiento de imágenes para la detección del cáncer que han alcanzado mayor eficiencia que un grupo de doctores especialistas, a partir del estudio de escáneres del cerebro (Ng, 2011). McDonald's ha gastado cientos de millones de dólares para desarrollar distintos algoritmos de IA para generar sistemas de recomendaciones, intentando de emular a Amazon (Stern, 2019). Al final, la IA tiene como objetivo que las máquinas puedan hacer tareas que antes eran exclusivamente humanas; fenómeno que está cambiando todas las industrias. Y el sector financiero no es una excepción ("*Rise of the machines*", 2015). A continuación, se discute las aplicaciones que actualmente están siendo implementadas en el sector financiero, además de las compañías que las ejecutan.

### B.1 Detección de Fraude

La detección de fraude en el sector financiero se está llevando a cabo mediante la Detección de Anomalías (*Anomaly Detection*), la cual es un enfoque de IA que se ocupa de identificar las desviaciones de la norma dentro de un conjunto de datos. Tiene la particularidad que las desviaciones pueden ser identificadas en tiempo real (Faggella, Azulay & Bharadwak, 2019), buscando identificar fraude, lavado de dinero, robos, incumplimientos contables, entre otros (Zheng et al., 2019).

Empresas que van desde *startups* como Feedzai (para pagos) o Shift Technology (para seguros) hasta gigantes como IBM, están ofreciendo servicios de detección de fraudes. Monzo, una *startup* bancaria británica, creó un modelo lo suficientemente rápido como para evitar que los posibles estafadores completen una transacción, reduciendo la tasa de fraude en sus tarjetas prepagas del 0,85% en junio de 2016 a menos del 0,1% en enero de 2017 ("*Machine-learning promises to shake up large swathes of finance*", 2017). MasterCard introdujo a *Decision Intelligence* a finales de 2016; servicio que utiliza inteligencia artificial para la detección de fraudes. En particular, utiliza algoritmos para identificar patrones de compra normales y anormales para los clientes; como la ubicación de las compras, tiempo, rango de precio típico y otros factores que ayudan a identificar rápidamente el comportamiento



anormal y bloquear usos adicionales de la cuenta hasta nueva aclaración (Zavadskaya, 2017).

## **B.2 Evaluación Crediticia**

La evaluación crediticia es la base de la gestión de riesgos. La información crediticia tradicional generalmente incluye solo atributos financieros sólidos, como crédito, tarjeta de crédito, divisas, préstamos privados y otros datos de transacciones financieras (Zheng et al, 2019). En cambio, la IA ofrece una evaluación más rápida y precisa a menores costos de un posible prestatario y refleja una gama más amplia de variables que conducen a una decisión mejor informada y respaldada por datos. Esto se debe a que la calificación crediticia de IA se puede basar en reglas más complicadas y avanzadas que los esquemas tradicionales (Kunwar, 2018).

ZestFinance, fundada por el ex CEO de Google, Douglas Merrill, (ZestFinance, 2020) es el creador de la plataforma Zest Automated Machine Learning (ZAML), una solución de suscripción basada en inteligencia artificial que ayuda a las empresas a evaluar a los prestatarios con poca información o historial crediticio. La plataforma utiliza miles de datos y proporciona transparencia que otros sistemas de suscripción no pueden, lo que ayuda a evaluar mejor las poblaciones tradicionalmente consideradas “en riesgo” (Schroer, 2019). ZAML tiene como objetivo el encontrar a millones de nuevos prestatarios mediante el uso de Big Data y, también, mantener el sesgo fuera del análisis de crédito (Zavadskaya, 2017). En Zestfinance aseguran que no existe un parámetro que sea decisivo por sí sólo en el momento de hacer un préstamo; por lo que ZAML, ocupando miles de parámetros en conjunto, ayudará a identificar con mayor precisión a los buenos prestatarios.

## **B.3 Minería de Datos**

La minería de datos es el proceso de descubrir patrones y relaciones interesantes y útiles en grandes volúmenes de datos. En este, se combina herramientas de estadística e inteligencia artificial con gestión de bases de datos para así cumplir con el objetivo de extraer datos utilizables de un conjunto más grande de datos sin procesar (Clifton, s.f.; “*Definition of 'Data Mining'*”, s.f.)

En el sector financiero se está ocupando esta tecnología para buscar automáticamente información dentro de las decenas de miles de documentos que se manejan todos los días (Faggella et al., 2019). La máquina impulsada con IA será capaz de leer toda la información relevante que existe sobre cierto *commodity* en solo unos segundos, mientras que los humanos necesitarían muchas horas para esta tarea; y aún no serían capaces de cubrir toda la información que podría estar afectando el rendimiento del *commodity* en particular (Zavadskaya, 2017). Además, puede ser utilizado para sacar información de contratos, facturas y cualquier documento que presente información desestructurada. Generando un aumento importante en la eficiencia.

JPMorgan Chase afirma haber invertido en un proyecto de IA llamado plataforma de inteligencia de contrato (CoiN). El banco afirma que CoiN fue diseñado para analizar documentos legales y extraer cláusulas importantes, una tarea que tradicionalmente realizan los empleados humanos y que a menudo requiere mucho tiempo (Faggella et al., 2019). Otro ejemplo es Dataminr que, fundada en 2009, utiliza un algoritmo para escanear Twitter y otras publicaciones en las plataformas públicas de redes sociales para leer las noticias y enviar alertas accionables a los inversores suscritos antes de que se denuncien las noticias. (Zavadskaya, 2017)

#### **B.4 Chatbots**

La atención de los bancos tradicionales no está siendo suficiente para satisfacer las necesidades de los consumidores con conocimiento digital. Los clientes necesitan de una atención inmediata que los humanos no pueden lograr; pero que sí podría una máquina impulsada con IA. En particular, las grandes empresas del sector financiero tienen una gran base de clientes y, por lo tanto, requieren alternativas de servicio al cliente automatizadas y que brinden servicio las 24 horas, como los *chatbots* (Kunwar, 2018).

Los *chatbots* utilizan IA para generar asesoramiento financiero personalizado y procesamiento del lenguaje natural (NLP por sus siglas en inglés) para proporcionar un servicio al cliente instantáneo y de autoayuda. El NLP es un subcampo de inteligencia artificial que se ocupa para procesar y analizar grandes cantidades de datos que se han capturado en la forma en que los humanos escriben, hablan o

documentan información. Entonces, los *chatbots* tratan de reconocer la intención dentro de los datos de texto así como de responder a los clientes con texto; por lo que pueden contestar preguntas de los clientes o permitirles realizar operaciones de transferencia de fondos desde la ventana de chat.

Bank of America utiliza el *chatbot* llamado Erica. Erica utiliza ML para hacer predicciones sobre el futuro, basadas en datos históricos, para proporcionar orientación financiera a los clientes. Esto implica sugerencias personalizadas sobre cómo un cliente puede ahorrar más dinero en función de sus gastos (Faggella et al., 2019)]. Otro ejemplo es Trim, un asistente para ahorrar dinero que se conecta a las cuentas bancarias de los usuarios y analiza los gastos. La aplicación inteligente puede cancelar suscripciones que desperdician dinero, encontrar mejores opciones para servicios como seguros e incluso negociar pagos (Schroer, 2019). Otra empresa que genera *chatbots* para la atención al cliente es Kasisto, el creador de KAI, una plataforma de IA conversacional. KAI ayuda a los bancos a reducir el volumen de llamadas al *call center*, al proporcionar a los clientes opciones y soluciones de autoservicio. Además, estos *chatbots* brindan a los usuarios recomendaciones calculadas y ayudan con otras decisiones financieras diarias (Schroer, 2019).

### **B.5 Trading Algorítmico**

El *trading* algorítmico es un sistema que realiza un conjunto ordenado de operaciones sistemáticas para generar operaciones automatizadas de compra y venta de *commodities* muy rápidas, basándose en información sobre las condiciones cambiantes del mercado y el nivel de precios. A menudo, las operaciones se realizan tan rápidamente que se les determina “*trading* de alta frecuencia” (Zavadskaya, 2017). Según Deutsche Bank, el 90% de las operaciones de futuros de acciones y el 80% de las operaciones de acciones se ejecutan mediante algoritmos sin ningún aporte humano, a nivel global (“*The stockmarket is now run by computers, algorithms and passive managers*”, 2019).

Las computadoras impulsadas con IA pueden analizar conjuntos de datos grandes y complejos de manera más rápida y eficiente que los humanos, por lo que son capaces de generar operaciones comerciales a velocidades superiores de las que podría hacer una persona. Además, con frecuencia, crea millones de intercambios en

un día sin interferencia humana (Kunwar, 2018; Singh, 2018). Por razones obvias, el ser humano no puede procesar tal nivel de información en un período de tiempo tan corto, por lo tanto, pierde ante el sistema algorítmico (Zavetskaya, 2017).

Muchos fondos de cobertura cuantitativos están empezando a ocupar IA para sus transacciones de alta frecuencia. Castle Ridge Asset Management, una empresa emergente con sede en Toronto, logró un rendimiento promedio anual del 32% desde su fundación en 2013 a finales de 2017. Utiliza un sistema sofisticado de ML, como los utilizados para modelar la biología evolutiva, para tomar decisiones de compraventa (*“Machine-learning promises to shake up large swathes of finance”*, 2017). Walnut Algorithms, finalista en el Premio Fintech 2016 a la Innovación en la gestión de inversiones, utiliza la IA para identificar patrones en los mercados financieros, adaptarse a las condiciones cambiantes del mercado y generar las operaciones de *trading* (Zavetskaya, 2017).

## **B.6 Gestión de Patrimonio**

La IA está siendo utilizada en el sector financiero para establecer estrategias óptimas de inversión. Las máquinas inteligentes han mostrado un notable potencial para proporcionar servicios de gestión de patrimonio a través de perfiles de usuario, análisis predictivo y recomendaciones personalizadas (Zheng et al, 2019) (Zavetskaya, 2017). Por ejemplo, actualmente se puede combinar datos estructurados financieros y datos no estructurados sobre el comportamiento del inversor, para evaluar el estilo de inversión de los clientes, la tolerancia al riesgo, las preferencias de compra y desarrollar carteras optimizadas a bajos costos y altas velocidades. (Zheng et al, 2019)

Principalmente, se han desarrollado dos formas de hacer gestión de patrimonio con IA, la recomendación de productos financieros y la robo-asesoría. El sistema de recomendación tiene como objetivo encontrar productos financieros ya establecidos que mejor se acomoden a las necesidades del cliente. La clave para diseñar un buen sistema de recomendación es caracterizar las preferencias de los clientes y construir modelos de comportamiento personalizados. Los enfoques de recomendación convencionales se han combinado con frecuencia con la teoría de cartera de Markowitz para lograr mejores recomendaciones de inversión (Zheng et al,

2019). En cambio, la tarea principal de Robo-Advisor es crear una cartera optimizada para la gestión de activos utilizando modelos de riesgo, lo que ayuda a disminuir el costo de la construcción de la cartera al tiempo que mejora la calidad. Las compañías automatizadas de gestión de cartera pueden cobrar tarifas más bajas mientras proporcionan al menos resultados igualmente buenos (si no mejores) (Zheng et al, 2019) (Zavetskaya, 2017).

Por ejemplo, Betterment y Schwab Intelligent Portfolios son empresas de gestión de carteras que basan su modelo y análisis con IA para proporcionar una elección de cartera óptima automatizada y un reequilibrio automático, selección de cartera de activos y selecciones de capital (Zavetskaya, 2017). Mientras tanto, Man AHL, un proveedor de fondos cuantitativos de 18.800 millones de dólares ha realizado investigaciones sobre el ML con fines comerciales desde 2009, y lo ha utilizado como una de las técnicas para administrar el dinero de los clientes desde 2014 y también para hacer recomendaciones de sus productos a sus inversionistas (*“Machine-learning promises to shake up large swathes of finance”*, 2017).

## ANEXO C: CONCEPTOS RELEVANTES *VALUE INVESTING*

Los conceptos listados a continuación son aquellos que Benjamin Graham considera relevantes a la hora de seleccionar empresas para hacer un portafolio, en su libro *The Intelligent Investor* (1949)

- **Precio:** Lo principal es que el precio pagado no sea elevado. El inversor debería imponer algún tipo de límite en el precio que está dispuesto a pagar por unas acciones, en relación con los beneficios netos obtenidos a lo largo de, por ejemplo, los siete últimos años. No obstante, esta restricción eliminaría prácticamente a todas las sociedades más fuertes y populares de la cartera. En particular, esto supondría prohibir prácticamente toda la categoría de “acciones de sociedades de crecimiento”.
- **Tamaño:** Graham recomienda invertir en empresas grandes o de “primera categoría” ya que tienden a sobrevivir las bajadas abruptas del mercado y también a las sobre/sub evaluaciones. Para observar si es que una empresa es de tamaño suficiente recomienda mirar las ventas anuales y los activos.
- **Situación Financiera:** Deben poseer una estable situación financiera y una estructura de capital más conservadora, para presentar resiliencia a los cambios abruptos del mercado.
- **Utilidades:** Se busca que la empresa a invertir posea estabilidad en cuanto a las utilidades. Principalmente que no hayan presentado pérdidas en un periodo de tiempo importante; de esta forma es más probable que las utilidades futuras sigan siendo positivas.
- **Dividendos:** Graham declara que un importante factor favorable en la calificación de calidad de la empresa es el historial de pagos de dividendos constantes durante los 10 últimos años o más. Es importante recalcar que cuando se compra una acción, se está comprando la participación en una empresa; entonces si esta tiene un historial de dividendos estables significa que genera utilidad para sus dueños.
- **Perspectivas Generales de Largo Plazo:** El inversor debe intentar de observar que es lo que hace crecer a esta empresa, de donde provienen sus beneficios y de donde van a provenir en el futuro.

- **Sentimiento de Mercado:** Es recomendable las empresas a invertir tengan una falta de interés o impopularidad sostenidas a lo largo del tiempo en la bolsa de valores, porque indica que no estará sobrevalorada.

## ANEXO D: PÁRAMETROS BASE DE DATOS

En la Tabla D.1 se muestra los parámetros de la base de datos original; y también cuáles parámetros quedaron para la base de datos de eliminación óptima (BdD 1) y la base de datos de eliminación óptima con fijación de parámetros (BdD 2). La base de datos imputada tiene la cantidad total de parámetros de la base de datos original.

Tabla D.1: Parámetros Base de Datos

Parámetros	BdD 1	BdD 2
Beta 1 year		
Beta 3 years		
Beta 5 years		
P/E		
P/E (average last 3 years)		
P/(E average last 5 years)		
Net Income		
Net Income average 3 years		
Net Income average 5 years		
Net Income Available to Common		
Net Income Available to Common av 3 years		
Net Income Available to Common av 5 years		
Basic Earnings per Share		
Basic Earnings per Share average 3 years		
Basic Earnings per Share average 5 years		
Normalized Net Income Growth		
Normalized Net Income – 3 years Geometric Growth		
EPS – 1 Yr Growth		
3 years Geometric Growth EPS bef Extraordinary Items		
EPS before XO – 3 year Geometric Growth		
Absence of Profit in 10 years		
EBIT		
EBIT average 3 years		
EBIT average 5 years		
Net Debt to Earnings before Interest Taxes		
Net Debt to Earnings before Interest Taxes average 3 years		
Net Debt to Earnings before Interest Taxes average 5 years		
EBIT Margin		



EBIT Margin average 3 years		
EBIT Margin average 5 years		
EBITDA		
EBITDA average 3 years		
EBITDA average 5 years		
Net Debt/EBITDA		
Net Debt/EBITDA average 3 years		
Net Debt/EBITDA average 5 years		
Price/EBITDA		
Price/EBITDA average 3 years		
Price/EBITDA average 5 years		
EBIT Growth Year over Year		
EBIT Growth 3 years CAGR		
EBITDA Growth Year over Year		
EBITDA Growth 3 years CAGR		
Net Income to Common Margin		
Net Income to Common Margin average 3 years		
Net Income to Common Margin average 5 years		
Revenue		
Revenue average 3 years		
Revenue average 5 years		
Price to Sales Ratio		
Price to Sales Ratio average 3 years		
Price to Sales Ratio average 5 years		
Capital Expenditure to Sales		
Capital Expenditure to Sales average 3 year		
Capital Expenditure to Sales average 5 year		
Revenue Growth Year over Year		
3 years Geometric Growth Net Sales		
Revenue 3 years CAGR		
Profit Margin – 3 year Geometric Growth		
Return on Stock 1 year		
Return on Stock 3 years		
Return on Stock 5 years		
Total Assets		
Total Assets average 3 year		
Total Assets average 5 year		
Dividends per Share		
Dividends per Share average 3 year		

Dividends per Share average 5 years		
Uninterrupted Dividends for 20 years		
Dividend Indicated Yld – Gross		
Gross 1 Year Dividend Growth Rate		
Gross 3 year Growth Rate		
Dividend Gross 3 years Growth Rate		
Dividend Payout Ratio		
Dividend Payout Ratio average 3 year		
Dividend Payout Ratio average 5 years		
Preferred Dividends		
Preferred Dividends average 3 year		
Preferred Dividends average 5 years		
Short and Long Term Debt		
Short and Long Term Debt average 3 years		
Short and Long Term Debt average 5 years		
Net Debt		
Net Debt average 3 years		
Net Debt average 5 years		
Short-Term Debt		
Short-Term Debt average 3 years		
Short-Term Debt average 5 years		
Long Term Debt		
Long Term Debt average 3 years		
Long Term Debt average 5 years		
Current Ratio		
Current Ratio average 3 years		
Current Ratio average 5 years		
Quick Ratio		
Quick Ratio average 3 years		
Quick Ratio average 5 years		
Common Equity to Tot Assets		
Common Equity to Tot Assets average 3 years		
Common Equity to Tot Assets average 5 years		
EBITDA/Interest		
EBITDA/Interest average 3 years		
EBITDA/Interest average 5 years		
EBIT/Interest		
EBIT/Interest average 3 years		
EBIT/Interest average 5 years		

Net Debt/EBITDA		
Net Debt/EBITDA average 3 years		
Net Debt/EBITDA average 5 years		
Total Debt to EBITDA		
Total Debt to EBITDA average 3 years		
Total Debt to EBITDA average 5 years		
Net Debt to Earnings before Interest Taxes		
Net Debt to Earnings before Interest Taxes average 3 years		
Net Debt to Earnings before Interest Taxes average 5 years		
Total Debt to Earnings before Interest Taxes		
Total Debt to Earnings before Interest Taxes average 3 years		
Total Debt to Earnings before Interest Taxes average 5 years		
Working Capital		
Working Capital average 3 years		
Working Capital average 5 years		
Working Capital to Sales		
Net Working Capital Investment		
Net Working Capital Investment average 3 years		
Net Working Capital Investment average 5 years		
P / (Working Capital)		
P / (Working Capital average 3 years)		
P / (Working Capital average 5 years)		
Working Capital – 1 Yr Growth		
Working Capital 3 years Growth		
Total Debt 1 Year Growth		
Total Debt 3 years Growth		
Short Term Debt 1 Year Growth		
Short Term Debt 3 years Growth		
Cash Flow – 1 Yr Growth		
Cash from Operating Act – 3 years Geometric Growth		
Free Cash Flow		
Free Cash Flow average 3 years		
Free Cash Flow average 5 years		
Free Cash Flow Yield		
Free Cash Flow Yield average 3 years		
Free Cash Flow Yield average 5 years		
Free Cash Flow to Total Debt		
Free Cash Flow to Total Debt average 3 years		
Free Cash Flow to Total Debt average 5 years		

Free Cash Flow 1 Year Growth		
3 years Geometric Growth FCF per Share		
Free Cashflow – 3 years Geometric Growth		
Preferred Equity and Hybrid Capital		
Preferred Equity and Hybrid Capital average 3 years		
Preferred Equity and Hybrid Capital average 5 years		
Historical Market Cap		
Periodic Enterprise Value		
Periodic Enterprise Value average 3 years		
Periodic Enterprise Value average 5 years		
Weighted Average Cost of Cap		
WACC Cost of Equity		
WACC Cost of Debt (After Tax)		
WACC Cost of Preferred		
EV per Share		
EV per Share average 3 years		
EV per Share average 5 years		
Risk Premium		
P/Book		
P/Book average 3 years		
P/Book average 5 years		
P/Tangible Book		
P/Tangible Book		
P/Tangible Book		
Current Shares Outstanding		
Equity Float		
Cash ratio		
Cash ratio average 3 years		
Cash ratio average 5 years		
Long-Term Debt/Equity		
Long-Term Debt/Equity average 3 years		
Long-Term Debt/Equity average 5 years		
Long-Term Debt/Capital		
Long-Term Debt/Capital average 3 years		
Long-Term Debt/Capital average 5 years		
Long-Term Debt/Total Assets		
Long-Term Debt/Total Assets average 3 years		
Long-Term Debt/Total Assets average 5 years		
Total Debt/Equity		

Total Debt/Equity average 3 years		
Total Debt/Equity average 5 years		
Total Debt/Capital		
Total Debt/Capital average 3 years		
Total Debt/Capital average 5 years		
Total Debt/Total Assets		
Total Debt/Total Assets		
Total Debt/Total Assets		
Net Debt/Equity		
Net Debt/Equity average 3 years		
Net Debt/Equity average 5 years		
Net Debt/Capital		
Cash From Investing		
Cash From Investing average 3 years		
Cash From Investing average 5 years		
Cash From Financing		
Cash From Financing average 3 years		
Cash From Financing average 5 years		
Bloomberg Estimation Revenue		
Bloomberg Estimation Earnings Per Share		
Bloomberg Estimation Ebit		
Bloomberg Estimation EBITDA		
Bloomberg Estimation Net Income		
Bloomberg Estimation Target Price		
Bloomberg Estimation Median Target Price		
Bloomberg Estimation High Target Price		
Bloomberg Estimation Low Target Price		
Buys		
Holds		
Sells		
Consensus Rating		
Bloomberg Estimation Est Long Term Growth		
Gain/Loss on Investments		
Gain/Loss on Investments average 3 years		
Gain/Loss on Investments average 5 years		
Marketable Securities & Other Short Term Investments		
Marketable Securities & Other Short Term Investments		
Marketable Securities & Other Short Term Investments		
ARD Purchases of Investments		

GDP CYOY Index		
WGDPU\$ Index		
FDDSGDP Index		
CPI YOY Index		
USURTOT Index		
NHSPSTOT Index		
USTW\$ Index		
USGG10YR Index		
CONCCONF Index		
Alpha 3 years		
Alpha 5 years		

## ANEXO E: DATOS PERDIDOS

En este anexo se explicará el por qué se pierden los datos. La naturaleza de los datos perdidos es relevante debido a que entrega información sobre que estrategia seguir para manejar los datos faltantes (Swalin, 2018; Roy, 2019).

- ***Missing Completely at Random***: En este escenario, la pérdida de datos no tiene nada que ver con la observación que se está estudiando. Es decir, el hecho de que falta un cierto valor no tiene nada que ver con su verdadero valor ni con los valores de otras variables.
- ***Missing at Random***: Significa que la propensión a que falte un punto de datos no está relacionada con los datos faltantes, pero está relacionada con algunos de los datos observados. El hecho de que este dato falte no tiene relación con su verdadero valor, sino que por una razón explicada por otras variables.
- ***Missing not at Random***: En esta situación la falta de un dato se relaciona específicamente con el valor que tomaría esta variable. Por ejemplo, cuando personas con salarios altos no quieren revelar sus ingresos en las encuestas.

Entonces, se decidió trabajar bajo el supuesto de que los datos faltantes de la matriz extraída de Bloomberg son de naturaleza completamente aleatoria, o al menos aleatoria. Se cree que la razón de que estos valores no fueron entregados por Bloomberg no depende del valor en sí de este dato faltante, si no de la forma en que las empresas liberan la información en sus reportes anuales.

En estos escenarios se puede realizar tanto la eliminación de datos, ya que los datos que faltan no sesgan las estimaciones de los coeficientes (Row, 2019), como la imputación. Por lo que queda en libertad del autor cuál de las dos metodologías seguir.

## ANEXO F: MODELACIÓN ELIMINACIÓN ÓPTIMA

En este anexo se muestra la formulación matemática del problema de optimización a resolver para construir las bases de datos 1 y 2; la base de datos de eliminación óptima y la base de datos de eliminación óptima con parámetros fijos. En primer lugar, se mostrará el problema de optimización que generó la base de datos 1, para luego mostrar los cambios que se hicieron al problema para obtener la base de datos 2. Se agradece al profesor Gustavo Angulo por su ayuda en la confección de este problema de optimización.

Sea  $X$  la base de datos original de tamaño  $m \times n$  ( $13.977 \times 234$ ), y  $F = \{1, 2, 3, \dots, m\}$  y  $C = \{1, 2, 3, \dots, n\}$ , la cantidad de filas y columnas que tiene la matriz original, respectivamente. Sea entonces  $D$  la matriz tal que  $D_{ij}$  presenta el valor “0” si  $X_{ij}$  es un dato faltante, y el valor “1” en caso contrario; es decir,  $D$  es una matriz que mapea las posiciones de datos disponibles y de datos perdidos. Entonces sea  $x_i$  una variable binaria que toma el valor de “1” si se decide mantener la fila  $i$  o el valor “0” en caso contrario; y sea  $y_j$  la variable binaria que toma el valor de “1” si se decide mantener la fila  $j$  y “0” en caso contrario. Entonces el  $P1$ ) genera la base de datos 1 al maximizar el tamaño de la matriz, sujeto a la restricción de que se mantiene un valor en la posición  $i, j$  de la matriz resultante sí y sólo sí se mantiene la fila  $i$  y la columna  $j$ .

$$P1) \quad \max \sum_{i \in F} \sum_{j \in C} z_{i,j}$$

$$\begin{aligned} s. a. \quad & z_{ij} = x_i * y_j, & i \in F, j \in C \\ & z_{ij} \leq x_i, & i \in F, j \in C \\ & z_{ij} \leq y_j, & i \in F, j \in C \\ & x_i + y_j \leq D_{i,j} + 1, & i \in F, j \in C \\ & z_{ij}, x_i, y_j \in \{0, 1\}, & i \in F, j \in C \end{aligned}$$

La base de datos 2 se construye a partir de modificar el  $P1$ ) al fijar los 8 parámetros que Graham considera relevantes para la evaluación de acciones y que se detallan en el punto 2.3.2. Entonces, como los parámetros son las columnas de la matriz original, se debe construir restricciones que fijen la variable  $y_k = 1$ , para todo  $k$  igual a la posición en que se encuentran estos parámetros relevantes. Sea entonces



$K$  el conjunto que contiene las posiciones de las columnas de las 9 variables a fijar (son 9 porque un parámetro es la división entre dos). Entonces el  $P2$ ) que genera la base de datos 2 será de la siguiente forma:

$$\begin{aligned}
 P1) \quad & \max \sum_{i \in F} \sum_{j \in C} z_{i,j} \\
 \text{s. a.} \quad & z_{ij} = x_i * y_j, & i \in F, j \in C \\
 & z_{ij} \leq x_i, & i \in F, j \in C \\
 & z_{ij} \leq y_j, & i \in F, j \in C \\
 & x_i + y_j \leq D_{i,j} + 1, & i \in F, j \in C \\
 & y_k = 1, & k \in K \\
 & z_{ij}, x_i, y_j \in \{0, 1\}, & i \in F, j \in C
 \end{aligned}$$

Donde  $K = \{4, 17, 21, 47, 69, 89, 92, 119, 164\}$ .

## ANEXO G: FUNCIÓN OBJETIVO RED NEURONAL

El entrenamiento de la red, o ajuste de los parámetros  $\Theta$ , se resuelve a través de un problema de minimización irrestricto. Para poder calcular la función objetivo a minimizar, se debe primero conocer la asignación correcta del set de entrenamiento; es decir, cuando una empresa es sobresaliente, y por lo tanto lleva la etiqueta “1”, o cuando no lo es, llevando entonces la etiqueta “0”. La variable que guarda las etiquetas se llamará  $y$ .

Entonces, la lógica que debiera seguir la función objetivo es que esta debe ser pequeña cuando la etiqueta fue correctamente asignada y debe ser grande cuando no lo fue. La Ecuación 2.3 muestra el “costo” que tendría en la función objetivo con sólo un ejemplo de entrenamiento, el cual sigue este razonamiento a la perfección. La ecuación 2.3 tiende a cero cuando  $h_{\theta}(x) \approx y$ ; tiende a infinito cuando  $y = 0$  y  $h_{\theta}(x) \rightarrow 1$  o cuando  $y = 1$  y  $h_{\theta}(x) \rightarrow 0$ . Representando perfecto el error de clasificación que se desea minimizar.

$$\text{Costo}(h_{\theta}(x), y) = -y * \log(h_{\theta}(x)) - (1 - y) * \log(1 - h_{\theta}(x)) \quad (\text{H.1})$$

Ahora bien, el costo se debe calcular considerando todos los ejemplos de entrenamiento, por lo que el problema de minimización queda de la forma que se puede ver en la Ecuación 2.4. En donde  $m$  es la cantidad de ejemplos del set de entrenamiento,  $L$  la cantidad total de capas de la red,  $s_l$  la cantidad de unidades (sin contar la unidad de sesgo) en la capa  $l$ .

$$\min_{\Theta} J(\Theta) = \min_{\Theta} \left\{ -\frac{1}{m} \sum_{i=1}^m \left[ y^i * \log(h_{\theta}(x^i)) - (1 - y^i) * \log(1 - h_{\theta}(x^i)) \right] + \frac{\lambda}{2} \sum_{l=1}^{L-1} \sum_{i=1}^{s_l} \sum_{j=1}^{s_{l+1}} (\theta_{j,i}^l)^2 \right\} \quad (\text{H.2})$$

## ANEXO H: EJEMPLO SEPARACIÓN VALIDACIÓN CRUZADA

A modo de ejemplo, se explicará como funcionaría el proceso de entrenamiento y validación cruzada para el año de testeo 2011 para el periodo de tenencia de 5 años. Se desea elegir, en el 2011, las empresas que tendrán un buen rendimiento en un periodo de 5 años, es decir que se quieren mantener hasta el 2016. Para entrenar y validar los parámetros que va a utilizar el algoritmo se separa la base de datos desde 1990 hasta 2010 para el conjunto de entrenamiento y se deja 2011 para el conjunto de testeo. Luego para ajustar los parámetros e hiper parámetros se separa el conjunto de entrenamiento en tres divisiones, dejando para el set de entrenamiento y validación:

- **Iteración 1:** Set de entrenamiento [1990; 2007]. Set de validación 2008.
- **Iteración 2:** Set de entrenamiento [1990; 2008]. Set de validación 2009.
- **Iteración 3:** Set de entrenamiento [1990; 2009]. Set de validación 2010.

Entonces, para cada configuración de modelo con una base de datos  $b$ , una cantidad de capas escondidas  $q$  y un parámetro de regularización  $\lambda$  (hiper parámetros), se entrena y se valida 3 veces, siguiendo la separación del conjunto de entrenamiento anteriormente descrita. Por último, para cada modelo se calcula el promedio de la métrica de error de clasificación de las iteraciones; y se elige la configuración que tenga el menor promedio para seguir a la etapa de testeo. Este proceso se hace para todos los años de testeo de los dos periodos de tenencia.

## ANEXO I: LÓGICA PORTAFOLIO

Para poder obtener las métricas de rentabilidad mensual, rentabilidad anualizada y el *alfa* durante el periodo de tenencia que obtiene el portafolio, se siguieron los pasos que se explicarán a continuación. Primero se extrae de Bloomberg la rentabilidad mensual de todas las empresas seleccionadas; desde el inicio del año de testeo hasta que se acabe el periodo de tenencia. Estos datos son exportados a un archivo Excel a través de una Macros. Luego, la misma Macros genera una rutina para calcular cuánto dinero equivale las acciones de una empresa seleccionada al principio de un mes cualquiera, el cual se detalla en la Ecuación 2.6. Sea entonces  $m^k$  el mes en el que está disponible la información financiera del año fiscal testeado de cierta empresa  $k$ ,  $pt$  el periodo de tenencia para este portafolio y  $r_j^k$  la rentabilidad mensual de la empresa  $k$  en el mes  $j$ ; entonces el valor de dinero que se tendrá a principio de en un mes  $j$  cualquiera ( $v_j$ ) será igual a:

$$v_j^k = \begin{cases} 0 & \text{si } j < m^k \\ h_\phi(x^k) * 100 & \text{si } j = m^k \\ v_{j-1}^k * (1 + r_{j-1}^k) & \text{si } m^k < j \leq m^k + pt * 12 \\ 0 & \text{si } j > m^k + hp * 12 \end{cases} \quad (2.6)$$

La ecuación 2.6 representa la lógica del portafolio. Cuando comienza un mes del año de testeo, se corre el modelo de selección de acciones para todas las empresas que pasaron a tener disponible su información financiera del año fiscal que se está testeando. De esas empresas se seleccionan aquellas que tuvieron un valor de la función de hipótesis de la red ( $h_\phi(x)$ ) mayor a 0,5, y se invierte en estas el valor de la función de hipótesis multiplicado por 100. Entonces se invierte entre 50 y 100 dólares según el valor que haya tenido tal empresa en la clasificación. Luego el valor invertido va a aumentando o disminuyendo según la rentabilidad que se obtiene mes a mes, hasta que se cumpla el periodo de tenencia para esa empresa y se haga efectiva la ganancia o pérdida.

Ya obteniendo cuanto dinero se tiene por acción en el periodo estudiado, se calculan los pesos que tienen en el portafolio las empresas seleccionadas en un mes cualquiera. Por último, se calcula la rentabilidad del portafolio en un mes  $j$  sumando por cada una de las empresas la multiplicación entre el peso del portafolio al inicio de  $j$  por la rentabilidad de tal empresa en el mes  $j$ . Con las rentabilidades mensuales del

portafolio se puede calcular la rentabilidad total en el periodo de inversión; y también calcular los *alfas* haciendo una regresión lineal entre el premio de mercado y el premio del portafolio. Donde la rentabilidad del mercado se calcula como la rentabilidad mensual del S&P 500 y la rentabilidad del activo libre de riesgo se calcula como la rentabilidad mensual del bono de 10 años de Estados Unidos, ambos datos extraídos de Bloomberg. La regresión lineal se hace a través del complemento de Excel de Análisis de Datos, que otorga como resultado entre otras cosas, el valor del intercepto (*alfa*), su intervalo de confianza y su valor p. Considerar que el *alfa* obtenido es mensual compuesto mensual, por lo que hay que transformarlo a anual compuesto anual.

## ANEXO J: CRITERIOS GRAHAM ACTUALIZADOS

Los rangos de Graham fueron modificados por última vez en el 1972, última edición de *The Intelligent Investor* realizada por él. En la Tabla J.1 se muestra los valores críticos de estos rangos y los percentiles que estos representan. Desafortunadamente, los valores de los parámetros para las empresas del S&P 500 para el año 1972 no están disponibles en Bloomberg; por lo que se utilizó el año más actual que la plataforma financiera tuviera disponible; el año 1990.

Tabla J.1: Percentiles Criterios Graham para el 1990

	P/E	Ventas	Ratio Corriente	P/Valor Contable	Aumento Utilidades	WK/DLP
<i>Criterio Graham</i>	<15	>100 MM	>2	<1,5	>33,3%	>1
<i>Valor Crítico</i>	15	100 MM	2	1,5	33,33	1
<i>Percentil Valor Crítico</i>	0,59	0,00	0,77	0,13	0,82	0,61
<i>Empresas Evaluadas</i>	382	455	412	178	12	401

Los criterios de historial de dividendos y utilidades no pueden ser evaluados con percentiles al ser variables binarias; por lo que estos no se actualizaron. Por lo que, para que una empresa cumpla estos criterios, debe tener los 20 y 10 años de estabilidad respectivamente. También es importante agregar que se descartó el parámetro Crecimiento del Beneficio ya que en Bloomberg presentaba el valor de este parámetro sólo para 12 empresas del S&P 500. Por último, se puede observar en la Tabla J.1 que prácticamente todas las empresas del S&P cumplen el criterio de ventas, por lo que se consideró que todas las empresas que evalúa el MSA cumplen esta norma impuesta por Graham ya que son miembros de este índice.

Entonces con los percentiles de la Tabla J.1 se pueden calcular los criterios actualizados para los años de testeo, que son aquellos que están en el percentil de su respectivo año. Los resultados obtenidos se muestran en la Tabla J.2.

Tabla J.2: Criterios Actualizados

<i>Año testeo</i>	P/E	Ratio Corriente	P/Valor Contable	WK/DLP
<i>2011</i>	16,10	2,33	1,21	0,86
<i>2012</i>	17,18	2,31	1,18	0,88
<i>2013</i>	20,44	2,39	1,16	0,82
<i>2014</i>	21,23	2,26	1,09	0,72
<i>2015</i>	20,85	2,28	1,23	0,53

## ANEXO K: RESULTADOS VALIDACIÓN

Tabla K.1: Posiciones configuraciones modelos, periodo de tenencia de 3 años

P	BdD	CCE	L	AT	F av
1	2	1	0,32	2013	0,427
2	2	2	0,02	2014	0,419
3	2	2	0,32	2013	0,418
4	2	2	0,32	2014	0,412
5	2	3	0,02	2015	0,406
6	2	2	0,16	2013	0,402
7	2	3	0,16	2013	0,402
8	2	2	0,04	2014	0,402
9	2	2	0,16	2014	0,400
10	2	3	0,02	2014	0,397
11	2	3	0,02	2013	0,395
12	2	1	0,16	2014	0,388
13	2	3	0,04	2014	0,388
14	2	2	0,02	2015	0,386
15	2	2	0,02	2013	0,383
16	2	1	0,16	2015	0,383
17	2	3	0,08	2013	0,382
18	2	2	0,08	2015	0,381
19	2	2	0,04	2013	0,381
20	2	2	0,08	2014	0,380
21	2	1	0,02	2015	0,379
22	2	1	0,02	2013	0,379
23	2	1	0,08	2014	0,378
24	3	1	0,08	2013	0,375
25	2	2	0,04	2015	0,374
26	2	3	0,04	2015	0,374
27	1	1	0,04	2014	0,371
28	2	3	0,08	2014	0,366
29	2	1	0,02	2014	0,364
30	3	1	0,08	2014	0,364
31	2	2	0,08	2013	0,362
32	2	1	0,08	2013	0,362
33	2	1	0,04	2015	0,359
34	2	2	0,32	2015	0,358
35	1	1	0,16	2015	0,353
36	1	1	0,02	2013	0,351
37	1	1	0,02	2014	0,347
38	2	1	0,32	2014	0,347
39	2	1	0,04	2014	0,346

P	BdD	CCE	L	AT	F av
40	2	2	0,64	2014	0,344
41	2	2	0,16	2015	0,343
42	1	1	0,02	2015	0,340
43	2	3	0,04	2013	0,337
44	2	2	0,64	2013	0,336
45	2	3	0,16	2014	0,335
46	2	1	0,16	2013	0,333
47	1	1	0,16	2014	0,331
48	2	1	0,08	2015	0,330
49	3	2	0,08	2014	0,329
50	1	1	0,04	2013	0,327
51	1	1	0,64	2015	0,323
52	2	1	0,32	2015	0,321
53	3	1	0,02	2015	0,320
54	1	1	0,08	2014	0,318
55	2	3	0,08	2015	0,316
56	1	1	0,04	2015	0,315
57	3	1	0,02	2014	0,315
58	1	1	0,08	2015	0,311
59	2	3	0,16	2015	0,306
60	2	1	0,64	2013	0,304
61	2	1	0,04	2013	0,304
62	3	2	0,32	2015	0,301
63	2	2	0,64	2015	0,300
64	1	2	0,16	2015	0,295
65	2	3	0,32	2013	0,295
66	3	2	0,02	2015	0,293
67	3	2	0,16	2014	0,292
68	3	2	0,16	2015	0,291
69	3	1	0,02	2013	0,290
70	3	1	0,04	2014	0,287
71	3	1	0,08	2015	0,285
72	3	1	0,32	2013	0,284
73	1	2	1,28	2015	0,283
74	1	1	0,32	2015	0,283
75	1	1	0,64	2014	0,282
76	1	3	0,32	2015	0,280
77	3	2	0,08	2013	0,279
78	1	2	0,16	2014	0,278

P	BdD	CCE	L	AT	F av
79	1	2	0,04	2014	0,277
80	3	1	0,04	2013	0,273
81	3	2	0,08	2015	0,272
82	2	2	1,28	2014	0,270
83	1	1	0,32	2014	0,268
84	1	3	0,16	2014	0,268
85	3	1	0,32	2014	0,268
86	3	2	0,16	2013	0,265
87	1	3	0,08	2015	0,263
88	2	1	0,64	2015	0,263
89	1	3	0,64	2015	0,261
90	1	2	1,28	2014	0,261
91	1	1	0,08	2013	0,259
92	2	1	0,64	2014	0,259
93	2	2	1,28	2013	0,255
94	1	1	0,16	2013	0,251
95	1	3	0,32	2014	0,248
96	3	2	0,32	2014	0,247
97	2	2	1,28	2015	0,247
98	3	1	0,32	2015	0,247
99	3	1	0,04	2015	0,247
100	1	2	0,04	2015	0,244
101	1	3	0,16	2013	0,243
102	1	2	0,08	2015	0,242
103	2	3	0,32	2014	0,241
104	1	2	0,02	2015	0,236
105	3	1	0,64	2015	0,236
106	1	2	0,64	2015	0,235
107	3	2	0,64	2015	0,232
108	1	2	0,08	2014	0,232
109	3	2	0,02	2014	0,232
110	1	3	0,16	2015	0,228
111	2	3	0,32	2015	0,228
112	1	2	0,16	2013	0,228
113	3	2	0,04	2014	0,219
114	3	2	0,04	2015	0,219
115	1	2	0,02	2014	0,217
116	1	3	0,02	2015	0,215
117	3	1	0,16	2015	0,213

P	BdD	CCE	L	AT	F av
118	3	2	0,04	2013	0,212
119	3	2	0,64	2014	0,212
120	1	2	0,32	2014	0,211
121	3	1	0,16	2013	0,209
122	1	3	0,08	2014	0,208
123	1	2	0,32	2015	0,206
124	3	1	0,16	2014	0,205
125	1	2	0,04	2013	0,205
126	1	3	0,64	2014	0,203
127	3	2	0,32	2013	0,199
128	1	3	0,32	2013	0,194
129	1	2	0,32	2013	0,191
130	1	3	0,04	2015	0,188
131	1	2	0,02	2013	0,183
132	1	1	1,28	2015	0,178
133	3	2	0,02	2013	0,177
134	1	2	0,64	2014	0,175
135	1	3	0,02	2014	0,175
136	3	1	0,64	2014	0,172
137	3	2	1,28	2015	0,171
138	1	1	0,32	2013	0,170
139	1	1	0,64	2013	0,168
140	1	2	1,28	2013	0,167
141	3	1	0,64	2013	0,167
142	3	2	0,64	2013	0,159
143	1	3	0,04	2014	0,148
144	3	1	2,56	2015	0,147
145	1	2	0,08	2013	0,147
146	1	3	0,08	2013	0,145
147	3	1	1,28	2015	0,141
148	1	1	1,28	2014	0,134
149	1	2	0,64	2013	0,134
150	3	2	1,28	2014	0,128
151	3	1	1,28	2014	0,124
152	2	2	2,56	2015	0,120
153	2	1	1,28	2014	0,118
154	1	2	2,56	2015	0,116
155	1	3	0,64	2013	0,114
156	2	2	2,56	2014	0,108



P	BdD	CCE	L	AT	F av
157	1	2	2,56	2014	0,103
158	2	1	1,28	2013	0,103
159	3	1	1,28	2013	0,102
160	3	2	1,28	2013	0,100
161	1	3	0,02	2013	0,099
162	1	3	0,04	2013	0,095
163	2	1	1,28	2015	0,094
164	3	1	2,56	2014	0,087
165	2	2	2,56	2013	0,084
166	3	2	2,56	2015	0,084
167	2	3	0,64	2014	0,078
168	2	3	0,64	2015	0,078
169	3	2	2,56	2014	0,072
170	2	1	2,56	2015	0,063
171	1	2	2,56	2013	0,062
172	1	1	2,56	2015	0,059
173	1	3	1,28	2015	0,054
174	3	1	2,56	2013	0,053
175	1	1	1,28	2013	0,048
176	3	2	2,56	2013	0,046
177	2	1	2,56	2014	0,043
178	1	1	2,56	2014	0,037
179	2	1	2,56	2013	0,014
180	1	3	1,28	2013	0,013
181	1	1	2,56	2013	0,000
182	1	3	1,28	2014	0,000
183	1	3	2,56	2013	0,000
184	1	3	2,56	2014	0,000
185	1	3	2,56	2015	0,000
186	2	3	0,64	2013	0,000
187	2	3	1,28	2013	0,000
188	2	3	1,28	2014	0,000
189	2	3	1,28	2015	0,000
190	2	3	2,56	2013	0,000
191	2	3	2,56	2014	0,000
192	2	3	2,56	2015	0,000

Tabla K.2: Posiciones configuraciones modelos, periodo de tenencia de 5 años

P	BdD	CCE	L	AT	F av	P	BdD	CCE	L	AT	F av
1	2	1	0,04	2013	0,495	40	2	2	0,04	2011	0,379
2	2	2	0,02	2012	0,494	41	2	1	0,32	2011	0,379
3	2	2	0,08	2013	0,477	42	2	1	0,32	2012	0,377
4	2	2	0,32	2013	0,470	43	1	1	0,16	2011	0,376
5	2	2	0,04	2013	0,467	44	1	1	0,16	2013	0,374
6	2	1	0,64	2013	0,463	45	2	3	0,02	2011	0,372
7	2	1	0,32	2013	0,454	46	2	1	0,64	2012	0,372
8	2	2	0,02	2011	0,453	47	2	1	0,08	2011	0,365
9	2	3	0,02	2013	0,452	48	2	1	0,04	2011	0,364
10	2	1	0,04	2012	0,448	49	1	1	0,04	2013	0,358
11	2	1	0,16	2013	0,447	50	2	2	1,28	2011	0,342
12	2	2	0,04	2012	0,444	51	2	1	0,02	2011	0,340
13	2	1	0,02	2013	0,443	52	2	2	0,32	2011	0,336
14	2	3	0,08	2013	0,439	53	1	2	0,08	2011	0,331
15	2	2	0,16	2013	0,439	54	2	2	1,28	2012	0,330
16	2	2	0,08	2011	0,435	55	1	2	0,02	2013	0,324
17	2	2	0,02	2013	0,434	56	2	3	0,32	2012	0,324
18	2	2	0,16	2012	0,434	57	2	3	0,08	2012	0,323
19	2	3	0,16	2013	0,433	58	1	1	0,32	2011	0,323
20	2	1	0,08	2013	0,432	59	1	1	0,04	2011	0,321
21	2	2	0,08	2012	0,426	60	3	1	0,02	2011	0,320
22	2	2	0,64	2013	0,425	61	2	1	0,16	2011	0,318
23	2	3	0,02	2012	0,419	62	2	1	1,28	2013	0,316
24	2	3	0,32	2013	0,418	63	1	1	0,08	2013	0,316
25	2	3	0,04	2013	0,414	64	1	2	0,08	2012	0,313
26	2	1	0,02	2012	0,409	65	2	3	0,08	2011	0,311
27	2	1	0,16	2012	0,408	66	1	3	0,04	2011	0,310
28	2	2	0,64	2011	0,404	67	2	3	0,32	2011	0,310
29	2	3	0,04	2012	0,403	68	3	1	0,04	2011	0,307
30	2	2	0,32	2012	0,403	69	1	3	0,32	2011	0,304
31	2	3	0,16	2012	0,402	70	1	1	0,02	2011	0,304
32	2	2	0,64	2012	0,400	71	1	1	0,32	2012	0,303
33	2	3	0,04	2011	0,398	72	1	2	0,32	2011	0,302
34	1	1	0,04	2012	0,396	73	1	2	0,04	2012	0,299
35	2	2	1,28	2013	0,391	74	1	1	0,02	2012	0,297
36	2	3	0,16	2011	0,391	75	2	1	1,28	2012	0,296
37	2	2	0,16	2011	0,391	76	2	1	0,64	2011	0,295
38	2	1	0,08	2012	0,382	77	1	2	0,04	2011	0,295
39	1	1	0,16	2012	0,381	78	3	2	0,64	2013	0,293

P	BdD	CCE	L	AT	F av
79	1	3	0,08	2011	0,292
80	1	2	0,02	2011	0,290
81	3	2	0,02	2011	0,289
82	3	1	0,08	2011	0,288
83	1	1	0,02	2013	0,286
84	1	1	0,32	2013	0,283
85	1	2	0,04	2013	0,281
86	1	2	0,08	2013	0,280
87	3	1	0,08	2013	0,279
88	3	2	0,64	2012	0,279
89	3	1	0,08	2012	0,278
90	1	3	0,16	2011	0,278
91	3	1	0,16	2013	0,278
92	1	1	0,08	2012	0,277
93	3	1	0,04	2012	0,277
94	3	1	0,02	2013	0,274
95	3	2	1,28	2011	0,271
96	3	1	0,02	2012	0,270
97	3	1	0,32	2011	0,269
98	3	2	0,02	2013	0,269
99	2	3	0,64	2013	0,266
100	3	2	0,08	2011	0,265
101	1	3	0,04	2012	0,263
102	2	2	2,56	2013	0,263
103	1	2	1,28	2011	0,262
104	1	2	0,16	2011	0,257
105	1	3	0,08	2012	0,256
106	1	3	0,32	2012	0,256
107	1	2	0,02	2012	0,255
108	1	3	0,02	2011	0,253
109	3	1	0,16	2011	0,249
110	3	1	0,04	2013	0,248
111	3	2	0,02	2012	0,247
112	1	2	0,64	2011	0,245
113	1	1	0,64	2013	0,242
114	3	1	0,32	2013	0,241
115	1	2	0,32	2012	0,236
116	1	2	0,32	2013	0,235
117	3	1	0,64	2012	0,235

P	BdD	CCE	L	AT	F av
118	1	3	0,32	2013	0,233
119	1	1	1,28	2011	0,232
120	1	1	0,64	2012	0,231
121	3	2	0,04	2011	0,230
122	1	3	0,08	2013	0,230
123	1	2	0,16	2013	0,228
124	3	2	0,32	2011	0,228
125	1	3	0,16	2013	0,227
126	3	2	0,08	2013	0,227
127	3	1	0,32	2012	0,224
128	1	1	0,08	2011	0,223
129	1	3	0,16	2012	0,223
130	3	1	1,28	2011	0,221
131	3	2	0,64	2011	0,220
132	1	2	0,16	2012	0,218
133	1	3	0,02	2012	0,215
134	1	3	0,02	2013	0,214
135	1	2	0,64	2012	0,213
136	1	2	0,64	2013	0,212
137	3	1	0,16	2012	0,212
138	1	1	0,64	2011	0,210
139	3	2	0,32	2013	0,209
140	3	1	0,64	2013	0,205
141	1	2	1,28	2012	0,204
142	1	3	0,04	2013	0,201
143	3	2	0,04	2013	0,199
144	1	3	0,64	2011	0,198
145	3	2	1,28	2013	0,198
146	1	2	2,56	2011	0,196
147	1	3	0,64	2012	0,195
148	3	2	0,32	2012	0,193
149	2	2	2,56	2012	0,185
150	2	1	1,28	2011	0,180
151	3	2	0,08	2012	0,177
152	3	2	2,56	2011	0,177
153	3	2	0,16	2012	0,173
154	3	2	0,16	2011	0,173
155	3	2	0,16	2013	0,169
156	3	1	0,64	2011	0,168

P	BdD	CCE	L	AT	F av
157	1	2	1,28	2013	0,168
158	3	2	1,28	2012	0,167
159	1	3	0,64	2013	0,158
160	1	1	1,28	2012	0,157
161	3	1	1,28	2012	0,157
162	3	2	2,56	2012	0,152
163	1	1	1,28	2013	0,152
164	3	2	0,04	2012	0,151
165	2	3	0,64	2012	0,143
166	1	2	2,56	2012	0,142
167	2	2	2,56	2011	0,122
168	3	2	2,56	2013	0,121
169	1	2	2,56	2013	0,100
170	3	1	1,28	2013	0,089
171	3	1	2,56	2011	0,074
172	3	1	2,56	2013	0,069
173	3	1	2,56	2012	0,064
174	1	1	2,56	2013	0,048
175	1	3	1,28	2013	0,046
176	2	1	2,56	2013	0,040
177	1	1	2,56	2011	0,039
178	1	1	2,56	2012	0,028
179	2	1	2,56	2012	0,025
180	2	1	2,56	2011	0,000
181	2	3	0,64	2011	0,000
182	2	3	1,28	2011	0,000
183	2	3	1,28	2012	0,000
184	2	3	1,28	2013	0,000
185	2	3	2,56	2011	0,000
186	2	3	2,56	2012	0,000
187	2	3	2,56	2013	0,000
188	1	3	1,28	2011	0,000
189	1	3	1,28	2012	0,000
190	1	3	2,56	2011	0,000
191	1	3	2,56	2012	0,000
192	1	3	2,56	2013	0,000