

Predicción del precio de acciones de la empresa Telefónica mediante técnicas de minería de datos

Trabajo Final de Máster
Máster en Matemática Computacional
Universidad Jaume I
Facultad de Matemáticas



Phuong Anh Pham
Email: al413620@uji.es
Castellón, Octubre de 2022

Predicción del precio de acciones de la
empresa Telefónica mediante técnicas de
minería de datos

Phuong Anh Pham

dirigido por

Pablo Gregori

Departamento de Matemáticas
Universidad Jaime I de Castellón

AGRADECIMIENTOS

“El mejor profeta del futuro es el pasado”

Lord Byron

Gracias a todos aquellos que confiaron en mí desde que comencé este Máster en Matemática Computacional. A pesar de los obstáculos que uno pueda encontrarse en el camino siempre es adecuado pensar que todo esfuerzo tiene su recompensa. Esa recompensa la veo hoy plasmada en este trabajo que refleja los años de trabajo y dedicación para completarlo. A mi tutor, Pablo Gregori, por despertar en mí la curiosidad por los datos en su asignatura y por apoyarme en todo momento para conseguir llevar a cabo este trabajo.

RESUMEN

Los modelos predictivos permiten al usuario estar preparado frente a posibles eventos futuros, aportándole estimaciones de lo que sucederá. De esta forma, cuando hablamos de modelos predictivos para el mercado bursátil nos referimos a herramientas que pronostiquen el movimiento futuro de los elementos del mercado. La predicción del precio de acciones mediante técnicas de minería de datos es un campo de estudio relativamente reciente y que presenta por tanto, a día de hoy, numerosas incógnitas.

En este proyecto se desarrolla una herramienta de análisis del precio de acciones, basada en las técnicas mencionadas y programada a partir del software estadístico R utilizando el R Markdown para facilitar el trabajo. Mediante esta herramienta se realiza la construcción de diversos modelos estadísticos y se valora la predicción.

Para desarrollar los modelos se usan variables principales los precios de las compañías que están en el IBEX 35, así como una serie de indicadores técnicos relativos al precio. Todos los datos necesarios para el análisis se pueden conseguir de forma sencilla a través de internet.

Los modelos de predicción se construyen por medio de técnicas de minería de datos como CART, Random Forest, Redes Neuronales, Máquinas de Soporte Vectorial y Regresión Lineal.

Se realizan varios experimentos para evaluar aspectos de los modelos. Se escoge la acción de la compañía Telefónica para este proyecto. Las técnicas basadas en el aprendizaje automático tratan de crear programas y rutinas capaces de generalizar comportamientos a partir de una información suministrada en forma de ejemplos de experiencia pasada.

En conclusión, con este trabajo se obtienen diferentes modelos predictivos que presumiblemente detectarán el movimiento futuro del precio de las acciones de Telefónica.

TABLAS DE CONTENIDOS

1. INTRODUCCIÓN.....	1
1.1. Definición.....	1
1.2. Justificación.....	1
1.3. Estructura de la memoria.....	1
2. OBJETIVOS	3
2.1. Generales	3
2.2. Específicos.....	3
2.3. Metodología De Investigación	3
3. ANTECEDENTES.....	5
3.1. Introducción a los mercados financieros	5
3.2. Tipos de análisis existentes en el mercado de valores.....	5
3.2.1. La inversión basada en el análisis gráfico	5
3.2.2. El volumen.....	6
3.2.3. La teoría de Dow	7
3.2.4. Teoría de Wyckoff.....	9
3.2.5. La teoría de la onda Elliot.....	10
3.2.6. El análisis fundamental	10
3.2.7. Máxima pérdida 5% - Stop Loss.....	16
3.2.8. Análisis técnico	16
4. TÉCNICAS DE MINERÍA DE DATOS DISPONIBLE	18
4.1. Regresión logística.....	18
4.2. Árboles de clasificación y regresión.....	19
4.2.1. Árboles de regresión	20
4.2.2. Árbol de clasificación	21
4.3. Random Forest (RF).....	23
4.4. Redes neuronales (NNET)	25
4.4.1. Redes neuronales de una sola capa	25
4.4.2. Redes Neuronales Multicapa.....	27
4.5. Máquinas de soporte vectorial (SVM).....	29
4.5.1. Clasificador de soporte vectorial	30

4.5.2.	Máquinas de soporte vectorial (SVM)	31
4.5.3.	Relación con la regresión logística.....	33
4.6.	Resultados: Trading del sistema.....	34
5.	APLICACIÓN	35
5.1.	Variables utilizadas en el análisis	35
5.2.	Programación del sistema de análisis	38
6.	RESULTADOS Y DISCUSIÓN	40
6.1.	Selección de variables a través del Random Forest.....	40
6.1.1.	Modelo estadístico Max(5)	41
6.1.2.	Modelo estadístico Max(10)	42
6.1.3.	Modelo estadístico Max(15)	43
6.2.	Aplicar las técnicas de minería de datos	44
6.2.1.	Regresión Lineal	44
6.2.1.1.	Modelo estadístico Max(5)	44
6.2.1.2.	Modelo estadístico Max(10).....	46
6.2.1.3.	Modelo estadístico Max(15).....	48
6.2.1.4.	Resumen el caso de los modelos de Regresión Lineal.....	51
6.2.2.	Redes Neuronales.....	51
6.2.3.	Máquinas de soporte vectorial (SVM).....	53
7.	EXPERIMENTO COMPUTACIONAL DE SIMULACIÓN.....	56
7.1.	Regresión Lineal - Max(10) 90% Training.....	56
7.2.	Redes neuronales - Modelo estadístico Max(10) con 80% training.....	57
7.3.	Redes neuronales - Modelo estadístico Max(15) con 80% training.....	57
7.4.	Máquinas de soporte vectorial (SVM) - Max(10) con 90% training	58
7.5.	Máquinas de soporte vectorial (SVM) - Max(15) con 90% training	58
8.	CONCLUSIONES.....	59
9.	LÍNERAS FUTURAS.....	60
10.	BIBLIOGRAFÍA	61
11.	ANEXO I. R MARKDOWN	
12.	ANEXO II. SIMULACIÓN.....	

TABLAS DE FIGURAS

Figura 1. El análisis técnico se basa en gráfico (fuente: https://es.investing.com).....	6
Figura 2. El gráfico de tipo velas de las acciones de Telefónica	6
Figura 3. El análisis fundamental se basa en los estados financieros de la empresa.....	12
Figura 4. El Análisis técnico se basa en las cotizaciones pasadas de la acción.....	17
Figura 5: Estructura jerárquica de un árbol. a) Una división, b) Dos divisiones, c) 26 divisiones [11].....	19
Figure 6. Medidas de impureza de nodo para la clasificación de dos clases, en función de la proporción p en la clase 2. La entropía cruzada se ha escalado para pasar $(0.5, 0.5)$. [4]	22
Figura 7. Bagging, Random Forest y Gradient Boosting aplicados a los datos de spam. Para impulsar, se utilizaron árboles de 5 nodos, y el número de árboles se eligió mediante validación cruzada de 10 veces (2500 árboles=). Cada “paso” en la figura corresponde a un cambio en una sola clasificación errónea (Fuente: [4]).....	24
Figura 8. Red neuronal con una sola capa oculta. [1].....	26
Figura 9. Funciones de activación.....	27
Figura 10. Diagrama de red neuronal con dos capas ocultas y múltiples salidas [1]	28
Figura 11. Izquierda: se aplica una SVM con un núcleo polinomial de grado 3. Derecha: se aplica una SVM con un kernel radial. [1]	32
Figura 12. Se comparan las funciones de pérdida de SVM y regresión logística, en función de $y_i\beta_0 + \beta_1x_{i1} + \dots + \beta_px_{ip}$. Cuando $y_i\beta_0 + \beta_1x_{i1} + \dots + \beta_px_{ip}$ es mayor que 1, entonces la pérdida SVM es cero, ya que esto corresponde a una observación que está en el lado correcto del margen. En general, las dos funciones de pérdida tienen un comportamiento bastante similar. [1]	34
Figura 13. Diagrama de bloques del funcionamiento del sistema de análisis	38
Figura 14. Análisis general de importancia de variables explicativas mediante Random Forest.....	40
Figura 15. Análisis particular de importancia de variables explicativas mediante Random Forest para el modelo Max(5).....	41
Figura 16. Análisis particular de importancia de variables explicativas mediante Random Forest para el modelo Max(10).....	42
Figura 17. Análisis particular de importancia de variables explicativas mediante Random Forest para el modelo Max(15) (Fuente: Elaboración propia).....	43

Figura 18. Gráfico de los precios de TEF reales (línea negra) y la predicción (línea roja) del modelo 80% training.....	45
Figura 19. Gráfico de los precios de TEF reales (línea negra) y la predicción (línea roja) del modelo 90% training.....	46
Figura 20. Gráfico de los precios de TEF reales (línea negra) y la predicción (línea roja) del modelo 80% training del Modelo Max(10)	47
Figura 21. Gráfico de los precios de TEF reales (línea negra) y la predicción (línea roja) del modelo 90% training.....	48
Figura 22. Gráfico de los precios de TEF reales (línea negra) y la predicción (línea roja) del modelo 90% training.....	50
Figura 23. Gráfico de los precios de TEF reales (línea negra) y la predicción (línea roja) del modelo 90% training.....	51

1. INTRODUCCIÓN

1.1. Definición

La minería de datos o exploración de datos (es la etapa de análisis de "Knowledge Discovery in Databases" o KDD) es un campo de la estadística y las ciencias de la computación referido al proceso que intenta descubrir patrones en grandes volúmenes de conjuntos de datos. Utiliza los métodos de la inteligencia artificial, aprendizaje automático, estadística y sistemas de bases de datos. El objetivo general del proceso de minería de datos consiste en extraer información de un conjunto de datos y transformarla en una estructura comprensible para su uso posterior. Además de la etapa de análisis en bruto, supone aspectos de gestión de datos y de bases de datos, de procesamiento de datos, del modelo y de las consideraciones de inferencia, de métricas de intereses, de consideraciones de la teoría de la complejidad computacional, de post-procesamiento de las estructuras descubiertas, de la visualización y de la actualización en línea.

En este trabajo de fin de máster se pretende dar un nuevo impulso al trillado campo del análisis de los precios de una acción en el mercado bursátil aplicando una serie de técnicas novedosas de minería de datos, que han ido desarrollando y que cada vez están siendo probadas en más ámbitos profesionales.

El propósito principal del proyecto es comprobar si estas herramientas tienen realmente aplicación a este campo y si aportan nuevas facilidades y ventajas con respecto a las técnicas ya existentes, como el análisis técnico, basado en series temporales y el análisis fundamental.

1.2. Justificación

Conseguir predecir el movimiento del precio de una acción en el mercado tiene la dificultad extrema, debido a los numerosos factores que intervienen en el mismo, por la ley de oferta y demanda, por parte de los agentes humanos que intervienen en ellos, cada uno con sus propias iniciativas, miedos, estrategias, creencias y condicionantes.

A lo largo de tiempo, se van aplicando las técnicas más modernas y complejas debido a la cantidad enorme de capital que se mueven en estos mercados, y no obstante nunca se ha conseguido llegar a la herramienta perfecta de predicción, ni se estima que se logrará nunca. Por eso, hay que ir mejorando y actualizando las técnicas nuevas.

1.3. Estructura de la memoria

Este documento se divide en varias partes para facilitar la lectura, la estructura propia de este del trabajo es:

- En primer lugar, se desarrolla una introducción del proyecto y la problemática que abordar, se justifica la elección.
- En segundo lugar, se plantean los objetivos generales y específicos del proyecto para que el lector tenga una idea clara del proyecto.
- En tercer lugar, se desarrollan los antecedentes. Se presentan las técnicas de los análisis originales.
- En cuarto lugar, se presentan las técnicas de minería de datos.
- En quinto lugar, se presentan los resultados y discusión del proyecto, donde se puede valorar los avances y los desarrollos obtenidos en el proyecto.
- En sexto lugar, se explicitan las conclusiones a las que se ha llegado durante la realización del proyecto en vista de los resultados.
- En séptimo lugar, se plantean las futuras líneas de investigación que han ido surgiendo como consecuencia del desarrollo del proyecto.
- En octavo lugar, se presentan los recursos bibliográficos utilizados en el trabajo.
- Por último, Se adjunta los anexos del código R Markdown.

2. OBJETIVOS

2.1. Generales

El objeto principal de este proyecto es construir un modelo de predicción del precio de acciones a través de distintas técnicas de minería de datos.

Se intentará poner en práctica y construir el sistema a partir de herramientas gratuitas y accesibles para cualquier individuo con acceso a internet.

2.2. Específicos

Se destacan los siguientes objetivos específicos:

- Comparación, entre las técnicas de minería de datos aplicadas, se comparan para apreciar cuáles de ellas se adaptan de mejor forma de este tipo de análisis y por tanto tienen mayor probabilidad de éxito.
- Comparación de las variables explicativas seleccionadas para saber cuáles de ellas pueden resultar útiles al proceso de predicción y cuáles de ellas no aportan nada o resultan redundantes.
- Establecimiento de posibles relaciones entre compañías a través del análisis de sus respectivos precios. Mediante la inclusión entre las variables explicativas del precio de una empresa, los indicadores de otras compañías se pueden provocar el descubrimiento de patrones similares que pueden llevar a relaciones entre compañías.
- Evaluar el desempeño de las técnicas seleccionadas mediante la comparación de la precisión y los errores de los resultados.

2.3. Metodología De Investigación

Fase	Proceso	Métodos, Herramientas Y Procedimientos
Revisión del estado del arte	Recolección de información	Llevar a cabo una investigación por medio de fuentes bibliográficas sobre los métodos de pronóstico desarrollados hasta el momento, realizando un listado de los mismos con sus principales características, objetivos y resultados
	Revisar qué métodos son aplicables al mercado de valores	Investigar por medio de fuentes bibliográficas, cuáles de los métodos del listado son aplicables al pronóstico del mercado de valores, reduciendo el listado anterior y por medio de un diagrama de Pareto identificar la frecuencia con la que se aplican.

	Identificar cuáles de estos métodos están relacionados	De acuerdo con los conocimientos adquiridos durante la carrera, determinar cuáles de estos métodos están más relacionados con la Ingeniería industrial, reduciendo el listado anterior.
	Seleccionar los métodos	Seleccionar como mínimo tres métodos usados para la proyección del mercado de valores del listado de técnicas de pronóstico.
Recrear los modelos de predicción	Apropiarse de los métodos seleccionados	Investigar y desarrollar los conocimientos necesarios para la correcta aplicación de cada uno de los métodos.
	Seleccionar las acciones de la IBEX para trabajar	Analizar el mercado y seleccionar los de las acciones más activas en este mercado.
	Analizar el comportamiento de las acciones seleccionadas	Mediante gráficos, análisis estadístico y el uso de software, determinar las características y el comportamiento de cada una de las acciones seleccionadas.
	Aplicar las técnicas seleccionadas a las acciones correspondientes	Desarrollar los modelos para cada uno de los métodos seleccionados, generando la proyección para cada una de las acciones elegidas.
Evaluación y comparación de los resultados	Evaluar el nivel de certeza de cada uno de los métodos en cada una de las acciones	Comparar los datos obtenidos mediante cada uno de los métodos contra los valores reales de las acciones en él.
	Determinar el método más adecuado para la proyección de cada acción	Analizar los resultados de cada uno de los métodos de pronóstico, su certeza y ajuste al comportamiento real de cada una de las acciones, determinando cuál de los métodos se adapta mejor a la naturaleza de cierta acción y a qué se debe dicha adaptación.

[15]

3. ANTECEDENTES

3.1. Introducción a los mercados financieros

El Mercado Financiero es el lugar, mecanismo o sistema en el cual se compran y venden cualquier activo financiero. Los mercados financieros pueden funcionar sin contacto físico, a través de teléfono, fax, ordenador. También hay mercados financieros que sí tienen contacto físico. La finalidad del mercado financiero es poner en contacto oferentes y demandantes de fondos, y determinar los precios justos de los diferentes activos financieros.

Las ventajas que tienen los inversores gracias a la existencia de los mercados financieros son la búsqueda rápida del activo financiero que se adecue a nuestra voluntad de invertir, y además, esa inversión tiene un precio justo lo cual impide que nos puedan timar. El precio se determina del precio de oferta y demanda.

Otra finalidad de los mercados financieros es que los costes de transacción sea el menor posible. Pero debemos insistir en que la finalidad principal es determinar el precio justo del activo financiero. Ello dependerá de las características del mercado financiero. Cuanto más se acerque un mercado financiero al ideal de mercado financiero perfecto, el precio del activo estará más ajustado a su precio justo.

3.2. Tipos de análisis existentes en el mercado de valores

3.2.1. *La inversión basada en el análisis gráfico*

Se trata de estudiar la estructura del mercado, debido a que habitualmente el cambio de una tendencia en los precios está precedido de un fortalecimiento o debilitamiento de ésta. Se hace sobre la base de cuatro aspectos: el precio, el tiempo, el volumen y la amplitud. El precio de un título o de un grupo de ellos, refleja el nivel de cambio en la actitud del inversor hacia él o ellos. El tiempo, es la segunda dimensión, cuanto más se tarde en pasar de una tendencia alcista a una bajista, mayor será el cambio en el precio. El volumen mide la intensidad del cambio de actitud de los inversores. Un cambio de precio acompañado de pocas transacciones implica un escaso entusiasmo. La última dimensión es la amplitud, cuando la actitud es favorable a la mayoría de las acciones indicaría una recuperación general del mercado. Los indicadores suelen medir sólo dos de estas dimensiones, normalmente el precio y el tiempo, aunque también es posible relacionar la amplitud y el tiempo. [2]



Figura 1. El análisis técnico se basa en gráfico (fuente: <https://es.investing.com>)



Figura 2. El gráfico de tipo velas de las acciones de Telefónica (fuente: <https://es.investing.com>)

3.2.2. El volumen

El volumen es el número de transacciones que tienen lugar en el mercado. Existe un principio reconocido que se enuncia de la siguiente forma: el volumen es una confirmación de la acción del precio, acompaña a la tendencia. Pero, además, su estudio nos proporciona señales de reversión de la tendencia del precio, de ahí que pueda ser tomado como un indicador de la extensión de la tendencia.

Las formas de medir el volumen, resaltando que dentro de ellas existen variantes, son las siguientes:

“Upside/Downside Volume”

“Big-block activity”

“Most active stocks”

3.2.3. La teoría de Dow

[8] Dow sostuvo que los indicadores de la Bolsa reflejaban las decisiones combinadas de los miles de inversores que acababan de actuar en los mercados financieros, pero que además, indirectamente representaban toda la información disponible de cualquier índole, desde la económica hasta la política o social, incluso las condiciones atmosféricas. Sostuvo que el mercado reflejaba las decisiones de los inversores, algo que hoy es aceptado, pero que entonces supuso algo verdaderamente revolucionario, frente a tanto “adivino” que atribuía los movimientos del mercado a causas tan peregrinas como falsas.

En la práctica la Teoría de Dow se suele resumir en quince puntos [8]:

- Primero: Manipulación. Los mercados financieros están expuestos a la manipulación por parte de poderes, entre los que cabe incluir desde inversores hasta políticos, pero esa manipulación sólo tiene consecuencias sobre los resultados del día a día, o sobre la tendencia secundaria, pero nunca sobre la tendencia principal. La tendencia principal de un mercado no puede ser manipulada por nada ni por nadie.
- Segundo: Influencias. Todas las influencias, sean de la naturaleza que sean, quedan reflejadas en los índices de cierre del mercado. Por ejemplo, si un inversor toma una decisión basándose en una información confidencial o privilegiada, el índice de la bolsa lo refleja inmediatamente.
- Tercero: Fiabilidad. La teoría de Dow no es infalible. Estos tres primeros puntos son los axiomas, si no se aceptan no se puede seguir adelante con el resto
- Cuarto: Tendencias. Todo movimiento que existe en un mercado financiero se puede descomponer en tres tendencias: principal, secundaria, y correcciones del mercado.
- Quinto: Tendencia principal. Es el movimiento más amplio que existe en un mercado, y su duración es desde uno a varios, no se puede determinar la duración porque depende de la percepción que tienen los inversores de la economía real. Determinar la dirección de esta tendencia, si es bajista o alcista, es el factor esencial para tener éxito en la inversión. No existe, ni es probable que llegue a existir, ningún método para determinar la duración o la magnitud del movimiento de la tendencia principal.
- Sexto: Tendencia principal bajista. Es un movimiento descendente continuo que tiene lugar cuando ni los especuladores, ni los inversores crean demanda, debido a coyunturas económicas adversas. Se ve interrumpido por tendencias secundarias de signo contrario, finalizando una vez el mercado ha descontado todas las circunstancias adversas. Existen tres fases en una tendencia bajista: Los inversores dejan de crear demanda, se detiene la subida de precios. La reducción del volumen de negocios y de los beneficios, hace que los inversores comiencen a vender y que la demanda no sea suficiente. Como consecuencia caen las cotizaciones, Los inversores al observar la caída de las cotizaciones venden sin tener en cuenta el verdadero valor de los títulos, especialmente lo hacen los que necesitan liquidez. En los mercados se suele decir: que existe un precio de compra, un precio de venta y otro para vender rápidamente. La caída de las cotizaciones es ahora vertiginosa.

- Séptimo: Tendencia principal alcista. La tendencia principal alcista es un movimiento ascendente continuo que tiene lugar cuando se crea una demanda por parte de los inversores, movimiento que se ve interrumpido por tendencias secundarias de signo contrario. Existen tres fases en la tendencia principal alcista: Retomo de la confianza en el futuro de la economía. los precios de los valores comienzan a responder al incremento de los beneficios empresariales. La especulación crece y la inflación dispara los precios; el precio de los valores se adelanta a las expectativas infundadas.
- Octavo: Tendencias secundarias. Son reacciones de signo contrario a la tendencia principal, que pueden durar semanas o meses, intervalo durante el cual el precio puede oscilar entre un treinta y tres y un sesenta y seis por ciento en relación al precio inicial después de producida la última tendencia secundaria.
- Noveno: Fluctuaciones diarias. Son las correcciones que sufre el precio durante una sesión bursátil.
- Décimo: Confirmación. Se refiere a lo ya mencionado sobre los índices
- Undécimo: Determinación de la tendencia. Una sucesión de tendencias secundarias en donde el punto más alto se sitúe por encima del punto más alto alcanzado en la última tendencia secundaria, y en donde el punto más bajo este por encima del alcanzado en la anterior tendencia secundaria, es una clara indicación de una tendencia principal alcista. La inversa, es igualmente cierta para una tendencia principal bajista. Las tendencias secundarias también se componen de correcciones diarias.
- Duodécimo: Una línea horizontal es un movimiento que se tiene que dar en dos índices a la vez, con una duración de varias semanas y sin que llegue a superar una estrecha banda de un cinco por ciento. Cuando la línea horizontal contiene frecuentemente pequeñas alzas, será señal del inicio de una futura tendencia alcista, siendo válido lo contrario para las tendencias a la baja.
- Decimotercero: Volumen. Es de aplicación lo dicho con anterioridad.
- Decimocuarto: Dobles máximos y dobles mínimos. Estos, no son indicio suficiente para determinar la sucesión de una tendencia.
- Decimoquinto: Valores independientes. Cualquier valor suele moverse con la media del indicador del mercado en el que cotiza, pero puede ocurrir que alguno no lo haga.

Actualización de la Teoría de Dow

Se han realizado para adaptar la teoría de Dow a los tiempos actuales [8]. Estas consisten en reducir la duración asignada a cada una de las tendencias: la tendencia principal que inicialmente se consideraba que podía durar cuatro años, ahora se ha reducido a un máximo de un año. La secundaria, cuya duración se estimaba en meses, ahora se ha reducido a semanas o incluso días. Mientras que para las correcciones de mercado, suele seguir utilizándose el día a día, o las que ocurren dentro de la misma jornada. También se argumenta que los índices utilizados, ya mencionados con anterioridad, no son representativos de la economía norteamericana y que deberían ser sustituidos por otros. No se han propuesto, al menos no existe acuerdo, sobre otros que los sustituyan.

3.2.4. Teoría de Wyckoff

Según Wyckoff [8], las técnicas de predicción sobre el comportamiento del mercado no forman parte de una ciencia exacta, porque los valores no tienen una base objetiva, sino que se forman en las mentes de las personas. En consecuencia, ningún pronóstico que se haga de forma puramente mecánica o matemática, basándose en el análisis de gráficos, es suficiente. Debe de estar presidido por el sentido común, que comienza por la comprensión, no por la aplicación, de los ocho puntos denominados por su autor las verdaderas reglas del juego.

- *Las reglas del juego*

Primera: Estas ocho reglas son aplicables a cualquier mercado financiero, sin importar su naturaleza o el tipo de valores negociados.

Segunda: El precio de un valor, de un conjunto de valores o de todos los valores negociados en un mercado, fluctúa debido a un factor único: la ley de la oferta y la demanda.

Tercera: Cuando la demanda de un valor excede a la oferta, su precio sube; cuando la oferta de un valor excede a la demanda, su precio baja.

Cuarta: En cada momento es posible descubrir si un valor determinado o todos los valores negociados en un mercado tienen más oferta o demanda.

Quinta: En un mercado actúan dos tipos de inversores, también denominados fuerzas primarias, que determinan la relación existente entre oferta y demanda. Estos inversores son: los grandes, que actúan de forma independiente o formando grupos; y, los inversores llamados pequeños, que actúan como un único inversor, también llamado inversor múltiple.

Sexta: En un mercado de valores, las negociaciones pueden darse entre dos grandes inversores, entre un gran inversor y un inversor múltiple, o dentro del ámbito del inversor múltiple.

Séptima: Los inversores están sometidos a todo tipo de influencias por toda la información que reciben.

Octava: Cuando un gran inversor compra, o vende, ejerce una presión al alza o a la baja, ésta es una señal de que otros acontecimientos están a punto de suceder en relación con algún valor o con el conjunto del mercado.

La lógica del método está en la idea de que es posible pronosticar el comportamiento futuro estudiando su movimiento presente, ya que éste refleja los planes y propósitos de quienes lo dominan, a través de la única ley que rige el comportamiento del mercado, la ley básica de la oferta y la demanda. Esto se consigue relacionando la evolución del precio de un valor con el volumen contratado. Lo importante es comprender que no todas las transacciones son significativas; es necesario descubrir cuáles lo son. Para poder hacerlo Wyckoff afirma que es necesario adquirir tres tipos de habilidades: en primer lugar las de un

detective, para descubrir las fuerzas que se ocultan detrás de las fluctuaciones de los precios y del volumen contratado; además es necesario ser un psicólogo, para determinar las motivaciones que mueven a los inversores a actuar de esa forma; y finalmente será necesaria la estrategia de un general para planificar actuaciones que permitan aprovechar las oportunidades descubiertas. La relación del precio de un valor con el número de títulos negociados, o volumen contratado del mismo, nos permite saber si en el futuro va a predominar la oferta o la demanda, si el precio de mercado del valor va a subir o bajar y qué ocurrirá con la tendencia del mercado. Es necesario tener en cuenta que no todas las transacciones del mercado serán significativas; es necesario descubrir cuáles lo son.[8]

3.2.5. La teoría de la onda Elliot

[8] Elliot relaciona el precio de los valores con el tiempo gráficamente por medio de dos ejes, uno horizontal donde mide el tiempo; y otro vertical, donde representa el precio de los valores. Cada dato es registrado en su momento y es unido mediante una línea que comienza donde acaba la anterior. A estos trazos continuos los denomina ondas, que pueden ser ascendentes o descendentes, y todas forman parte de otras más grandes y son divisibles en varias más pequeñas. Cuanto más se prolongan en el tiempo mayor categoría se les debe asignar. Los movimientos del mercado se pueden descomponer en dos clases de ondas impulsoras y correctoras. A su vez cada onda impulsora, denominada de esta forma porque el último dato está situado a mayor altura que el primero, se descompone en cinco ondas de la siguiente forma: la primera sube, la segunda baja, la tercera sube, la cuarta baja y la quinta sube. La onda correctora se divide en tres: la sexta onda baja, la séptima sube y la octava baja, quedando a menor altura que el punto de partida. Son conocidas como A, B y C. La segunda corrige a la primera, la cuarta a la tercera, la séptima a la sexta, quedando la quinta y la octava sin corrección alguna.

La nomenclatura utilizada por Elliot para nombrar a las ondas es confusa, es sólo orientativa porque no tienen una duración exacta, de menor a mayor serían las siguientes: onda minuete, onda minuta, onda menor, onda intermedia, onda principal, ciclo, super-ciclo y gran super-ciclo. Basarse en una o en otra depende del plazo del pronóstico que se quiere realizar.

3.2.6. El análisis fundamental

En 1923, durante un curso celebrado en la Universidad de Columbia en Nueva York, titulado “invesimenis” e impartido por Benjamin Graham, nace el llamado Análisis Fundamental que revolucionó el mundo del análisis bursátil y sigue teniendo una gran importancia en nuestros días. Introdujo, el concepto de economía en ese tipo de estudios. El tipo de enfoque dado a esta clase de estudios es el típico de un economista de rama general.

Desde los años cuarenta a los sesenta, el análisis fundamental, forma en que se denominó a este enfoque, se impuso en todos los ámbitos de la gestión de las inversiones hasta llegar a convertirse en la ciencia de la inversión por excelencia.

La década de los setenta aportó grandes innovaciones al campo de la inversión. La diversificación enunciada por Markowitz culmina en modelos como el desarrollado por Sharpe: el CAPM [8]. Aunque este tipo de novedades no sustituyen al análisis fundamental, ni tampoco lo contradicen, de hecho son combinables con él; el CAPM se puede utilizar para la etapa de valoración. Se desplaza el interés de los profesionales, más hacia la gestión que hacia el análisis.

El análisis fundamental se podría definir de acuerdo con los dos objetivos enunciados, cómo el estudio de toda la información disponible en el mercado sobre una determinada compañía, con la finalidad de obtener su verdadero valor, y formular una recomendación sobre su compra o venta. Se trata de anticipar el comportamiento futuro de una compañía en el mercado. La información analizada se conoce con el nombre de “fundamentos de la compañía”.

Para que el análisis tenga alguna utilidad hay que partir de una hipótesis básica: el mercado no es eficiente a corto plazo, aunque sí a largo. De no ser así, no sería posible adelantarse al mercado. Hoy, el precio de una acción no refleja el verdadero valor de la compañía, pero tenderá hacia éste en un futuro más o menos cercano. De lo que se trata es de descubrir acciones minusvaloradas o sobrevaloradas, en base a determinada información todavía no disponible en el mercado.

Uno de los objetivos es adelantarse al mercado y el analista se podrá adelantar siempre que conozca el verdadero valor del título en el instante futuro. Eso implica calcular el valor intrínseco, o precio objetivo de la acción. Por un lado, está el valor bursátil de la acción, que es la cotización que alcanza la acción en el mercado y que depende de la concurrencia de la oferta y la demanda, determinadas a su vez por una multitud de factores (dividendos, perspectivas futuras de beneficios, riesgo de las acciones, modas, pánicos, preferencias de los inversores, etcétera). Por otra parte, está el valor contable de la acción, que es el que se desprende de los libros de contabilidad y que se puede obtener dividiendo los recursos propios de la sociedad entre su capital social.

Finalmente tenemos el valor teórico o intrínseco (2), que es el que debería tener según la opinión de un inversor, después de haber analizado con detenimiento todos los datos disponibles o relevantes en relación con la sociedad de que se trate. El analista fundamental trata de encontrar discrepancias entre el valor intrínseco y el valor bursátil de la acción, pues supone que a medio o largo plazo dicho valor de mercado tenderá a acercarse a su valor intrínseco, a su valor económico.[2]

(En el prólogo, de Dodd, a la quinta edición de su Security Analysis. Éste reconoce que en el mundo se utilizan como sinónimos los términos “valor” y “enfoque de Graham y Dodd”.)

El valor intrínseco, central o económico, varía con el tiempo y tiene que ser revisado. En la práctica las estimaciones se hacen a un año vista. Aunque fluctúa, su volatilidad es menor que la que corresponde al valor en dinero o precio de mercado.

CONCEPTO	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016 estimado
Ingresos	3.764.419	4.008.777	4.389.547	4.771.432	4.869.776	4.029.588	3.684.170	3.634.087	4.368.866	4.456.243
EBITDA	532.093	607.582	746.905	931.951	1.224.900	830.701	1.199.746	758.586	850.558	748.922
Depreciación	158.389	184.893	212.274	237.500	252.370	170.548	168.328	144.494	165.802	132.642
EBIT	373.704	422.689	534.631	694.451	972.530	660.153	1.031.418	614.092	684.756	616.280
Resultado financiero	-147.297	-162.475	-198.739	-327.761	-440.627	-386.767	-523.883	-303.107	-477.715	-396.503
Beneficio neto	140.450	150.686	165.629	195.630	223.306	1.005.512	270.380	23.222	55.632	11.126
Flujo de caja	-105.181	-273.090	-51.014	-238.584	438.209	-441.875	-430.399	-272.454	-12.360	-759.448
Valor en libros	8,01	9,05	12,27	14,01	15,60	25,10	26,95	24,96	26,38	26,22
Ganancias por acción	1,61	1,74	1,84	1,96	2,24	10,13	2,73	0,23	0,42	0,084
Cotización	24,11	7,6	22	26	21,33	25,5	31	19,5	5,3	5,34
Número de acciones	87.236	86.601	90.016	99.811	99.690	99.261	99.040	100.965	132.457	132.457
Capitalización	2.103.261	658.169	1.980.347	2.595.092	2.126.392	2.531.151	3.070.249	1.968.822	702.023	707.321
Deuda neta	2.457.321	3.033.170	3.791.945	4.781.918	5.459.652	4.601.992	5.789.554	5.980.993	4.386.551	3.862.500
Deuda l/p	2.587.282	3.119.753	3.392.960	4.929.381	5.408.996	4.734.684	5.715.633	5.256.369	4.723.169	4.115.200
Deuda c/p	517.881	416.750	1.087.626	1.219.142	1.213.649	650.327	960.604	1.512.533	761.252	330.400
Caja	647.842	503.333	688.641	1.366.605	1.162.993	783.019	886.683	787.909	1.097.870	583.100
Enterprise Value	4.560.582	3.691.339	5.772.292	7.377.010	7.586.044	7.133.143	8.859.803	7.949.815	5.088.574	4.569.821

Figura 3. El análisis fundamental se basa en los estados financieros de la empresa (Fuente: CNMV)

Las etapas del análisis fundamental

1. *Primero:* las proyecciones económicas.

El analista comienza tratando de realizar proyecciones económicas sobre el comportamiento de la economía en el corto plazo, de uno a dos años. Esto permite realizar un pronóstico sobre la posible evolución de los mercados financieros. Determinadas variables económicas (por ejemplo, el índice de producción industrial) normalmente deparan alguna señal sobre los posibles cambios de tendencia en los ciclos económicos y son particularmente importantes en la etapa inicial de proyecciones.

2. *Segundo:* el análisis de “ratios” de la empresa y del sector.

El siguiente paso consiste en examinar de una forma específica la empresa que se pretende analizar, así como el sector económico (no bursátil) al que pertenece. El

examen se basa en el análisis de los “ratios” de la empresa, o relaciones entre dos o más variables de las finanzas de dicha empresa. Una vez elaborados se procede a su evaluación, para lo cual suelen compararse tanto con los resultados obtenidos por otras empresas de tamaño similar en el mismo sector o en el conjunto de la industria, y con los resultados obtenidos por la propia empresa en ejercicios anteriores, con el objeto de identificar la tendencia de los “ratios” de la empresa, sus beneficios y su crecimiento. Engloba el estudio de la empresa y de la industria

3. *Tercero:* el análisis de los beneficios y de los dividendos

A continuación, se procede a analizar los beneficios y dividendos repartidos por la sociedad, y a realizar predicciones sobre su evolución futura. Los métodos de pronóstico son diversos y no son consustanciales al enfoque fundamental, y pueden abarcar desde una simple extrapolación de datos pasados utilizando técnicas de regresión, hasta complejos modelos multifactoriales.

4. *Cuarto:* los pronósticos basados en el múltiplo precio/beneficios.

Con el objeto de proyectar la evolución del precio de la acción, el analista que ya cuenta con una proyección de los beneficios futuros debe de realizar una estimación de cómo el mercado va a pagar esas proyecciones futuras . Es decir; de qué forma va el mercado a capitalizar los beneficios a través de la tasa de rendimiento exigida. Alternativamente el analista puede estimar el múltiplo precio/beneficio del mercado para vincularlo a los últimos beneficios por acción. Es esencial resaltar que, en este punto, la tarea del analista es pronosticar la tasa de descuento del mercado, no su propio punto de vista sobre la misma.

5. *Quinto:* especificación del modelo de valoración apropiado.

Con todos estos datos debe construir un modelo de valoración. Los modelos de valoración tampoco son consustanciales con el enfoque fundamental y pueden ser simples o incluir criterios del modelo de valoración del precio de los activos (CAPM). Para poder realizar esto último deberá utilizar el estudio del mercado de valores y el del sector bursátil. Aquellas acciones cuyo valor sea mayor que el precio de mercado actual es susceptible de ser compradas, mientras que las que tengan un valor menor que el actual precio en dinero deberán ser vendidas. Es importante recordar, que los partidarios del enfoque fundamental consideran que el mercado acabará reconociendo el verdadero valor de la acción.[6]

Pasemos ahora a examinar los dos distintos enfoques que se suelen atribuir a este tipo de análisis:

El primer enfoque, también conocido como análisis “top-down”, consiste en ir de lo más global, a lo más específico. Desde un estudio macroeconómico llegaríamos al estudio de la empresa. El entorno global, va a determinar la posible evolución del sistema, especialmente

qué sectores saldrán favorecidos y cuáles perjudicados. Del estudio del sector, obtendremos la forma en la que se compite, la forma en que se va a distribuir esa renta que procede del sistema. Del estudio de la empresa, la forma en que esos fondos, esa demanda, llegará a convertirse en beneficios. Luego este tipo de enfoque comprende tres pasos: un primer paso, es el análisis de las economías nacionales; un segundo paso, es el análisis de los sectores de la economía; el tercer paso, consiste en analizar los fundamentos de las compañías individuales.

- El análisis de las economías nacionales es un primer nivel que comprende el estudio de los países más interesantes desde el punto de vista del inversor. Los criterios para seguir serán el Producto Interior Bruto o las políticas económicas de cada uno. En este sentido, políticas fiscales, son elementos clave para la elección de los países donde invertir. Un crecimiento por encima de la media también sería un buen indicador. No hay que olvidar que vivimos en una economía global y que la existencia de áreas deprimidas debe de ser tenida en cuenta, en el sentido de que podría ocurrir que sus condiciones se trasladaran a otros países y áreas geográficas.
- El siguiente paso, consiste en considerar los sectores dentro de las economías nacionales. Evidentemente, y teniendo en cuenta la influencia de las distintas variables macroeconómicas de un país, existen sectores con más atractivo que otros en cada momento del tiempo. Se buscan aquellos con más crecimiento de ingresos y márgenes, es decir, sectores sanos, con potencial de crecimiento.
- El último paso, sería la elección de las compañías individuales más atractivas dentro de los sectores analizados.
 - i. Uno de los criterios a tener en cuenta es la posición competitiva de la compañía. Está íntimamente relacionado con las fortalezas y debilidades de las compañías y habría que tener en cuenta elementos como la cuota de mercado, la calidad del producto, o barreras existentes para otros competidores.
 - ii. La rentabilidad del negocio a nivel sectorial tendría mucha importancia. Habría que realizar un análisis similar al anterior, pero teniendo en cuenta otras consideraciones como costes, gestión del circulante y rendimiento de las inversiones pasadas.
 - iii. El grado de apalancamiento es otro factor importante a estudiar. Definimos el apalancamiento operativo como la sensibilidad que tienen los beneficios de las mismas al nivel de actividad. Una compañía con muchos costes fijos tiene un apalancamiento operativo alto, Una compañía con muchos costes variables tiene un apalancamiento operativo bajo y está mejor preparada para una bajada en el nivel de actividad, pero no se ve tan beneficiada cuando el ciclo cambia y es bueno. El grado de apalancamiento hace referencia a la forma en que financia su activo, sus inversiones. Una compañía está muy apalancada financieramente hablando, cuando su deuda financiera es alta, en relación a sus recursos propios. Una muy apalancada será muy sensible a los tipos de interés.

Teniendo en cuenta todo lo anterior, el contexto macroeconómico, las tendencias sectoriales y la posición competitiva de cada compañía dentro de su sector, podremos

predecir cuál será la evolución de los resultados de las compañías analizadas, su capacidad de crecimiento y de rentabilizar adecuadamente sus inversiones. Podremos hacer previsiones sobre la creación futura de valor. Así terminaría el estudio.

Otro posible enfoque es el llamado “bottom-up”. En él, no se analizan empresas, sino valores. Un valor tiene una doble dimensión, por un lado, se trata de una empresa y de su comportamiento como tal, y por otro, se hace referencia a su comportamiento en la Bolsa. Lo que se busca son buenas oportunidades de inversión. Ni la situación económica del país, ni el atractivo del sector son relevantes. Es el concepto de valor el que prevalece. Entendiendo por tal el resultado de la empresa, más su comportamiento bursátil.

Una empresa es el resultado de una evolución pasada, pero también de una creación de valor futura. El balance de las compañías nos da una visión estática de su situación actual, fruto de la evolución anterior. El comportamiento bursátil o la rentabilidad para el accionista, es la variación del precio de la acción. Además, se le debe añadir la rentabilidad obtenida por los dividendos.

A la hora de realizar una recomendación de inversión. Intervienen tres elementos básicos:

- El negocio. Es muy importante analizar y sacar conclusiones sobre el negocio de la compañía. La información histórica nos va a permitir conocer su evolución hasta el momento presente, pero no es lo más importante. Es muy relevante intentar adivinar las tendencias. Los elementos a tener en cuenta son: a) Tendencias macroeconómicas, b) Tendencias sectoriales y regulaciones, c) Gestión de la compañía, d) Estrategia, e) Evolución de resultados.
- Valoración, como veremos a continuación todo nuestro análisis quedará plasmado en una valoración. Existen numerosos métodos, algunos de carácter absoluto y otros de carácter relativo. Esta valoración, se comparará con la cotización de la compañía. El objeto es comparar la infravaloración o sobrevaloración en el precio de mercado con respecto a nuestras estimaciones. Si el título estuviera infravalorado la recomendación sería comprar, mientras que si está sobrevalorado la recomendación sería vender.
- Riesgo. El riesgo de una compañía es el último elemento a tener en cuenta en la recomendación. Existen dos tipos de riesgo: el de negocio o específico, y el de la acción o de mercado. El riesgo de negocio está determinado por variables como el apalancamiento financiero y la capacidad para hacer frente a los intereses anuales, o la predictibilidad de los resultados y dividendos. El riesgo de la acción está determinado por dos variables:
 - Volatilidad de la acción, su comportamiento con respecto al mercado.
 - La liquidez de la acción, es decir, la frecuencia con que cotiza en el mercado y su volumen de contratación diario. Este último factor es muy importante a la hora de intentar vender o comprar títulos.

3.2.7. Máxima pérdida 5% - Stop Loss

El stop loss es una orden que se introduce en el mercado y que se ejecuta automáticamente cuando se traspasa a la baja para las posiciones largas y al alza para las posiciones cortas: representa la pérdida que estamos dispuestos a asumir (y alrededor de no más que eso) cuando es bastante obvio que nos hemos equivocado acerca de la dirección del mercado. Es aconsejable usar siempre el 5 % como límite de pérdida máxima.[2]

3.2.8. Análisis técnico

El análisis técnico es un método de análisis basado en observar la evolución a lo largo de tiempo de los movimientos de las acciones. Según el análisis técnico el precio se forma por la libre concurrencia de la oferta y la demanda. El objetivo fundamental del análisis técnico consiste en intentar predecir el movimiento del precio de las acciones en un futuro a través de datos pasados de precios.

Según J. Murphy [10], las ventajas principales del análisis técnico son:

- Es un análisis relativamente sencillo de realizar en el cuál no se requieren sólidos conocimientos financieros, macroeconómicos, etc.
- El acceso a la información es público.
- Se puede aplicar cualquier valor, índice, materia prima, o cualquier activo financiero.
- Se analiza todo tipo de empresas, tengan buenos fundamentos o no, al contrario que el análisis fundamental.
- Se adapta cualquier horizonte temporal de inversión.
- Es un análisis menos costoso en tiempo y esfuerzo al análisis fundamental.

Como inconveniente principal se destaca [10]:

- No existe ninguna evidencia científica acerca de su validez como elemento de predicción fiable del precio de una acción.
- Existencia de subjetividad elevada en el análisis. Los resultados pueden ser muy distintos en función del analista.
- Dificultad de cuantificar los patrones de comportamiento de los precios.
- Gran facilidad para caer en predisposiciones psicológicas.



Figura 4. El Análisis técnico se basa en las cotizaciones pasadas de la acción
(Fuente: <https://www.binarias.org/blog/analisis-fundamental-microeconomico>)

4. TÉCNICAS DE MINERÍA DE DATOS DISPONIBLE

En esta sección vamos a describir brevemente las herramientas de modelado que se utilizarán para llevar a cabo las tareas de predicción. Las herramientas se han escogido en función de su posibilidad de éxito, de su simplicidad de aplicación mediante R Studio y a criterio del tutor y del autor del proyecto.

4.1. Regresión logística

El modelo de regresión logística surge del deseo de modelar las probabilidades posteriores de las K clases a través de funciones lineales en x , al mismo tiempo que se asegura que suman igual uno y permanecen en $[0, 1]$. El modelo tiene la forma [4]:

$$\log \frac{\Pr(G = 1|X = x)}{\Pr(G = K|X = x)} = \beta_{10} + \beta_1^T x$$

$$\log \frac{\Pr(G = 2|X = x)}{\Pr(G = K|X = x)} = \beta_{20} + \beta_2^T x$$

...

$$\log \frac{\Pr(G = K - 1|X = x)}{\Pr(G = K|X = x)} = \beta_{(K-1)0} + \beta_{K-1}^T x$$

El modelo se especifica en términos de $K - 1$ log-odds o transformaciones logit. Aunque el modelo utiliza la última clase como denominador en las razones de probabilidad, la elección del denominador es arbitraria porque las estimaciones son equivalentes bajo esta elección. Un simple cálculo muestra que [4]:

$$\Pr(G = k|X = x) = \frac{\exp(\beta_{k0} + \beta_k^T x)}{1 + \sum_{l=1}^{K-1} \exp(\beta_{l0} + \beta_l^T x)}, \quad k = 1, \dots, K - 1,$$

$$\Pr(G = K|X = x) = \frac{1}{1 + \sum_{l=1}^{K-1} \exp(\beta_{l0} + \beta_l^T x)}, \quad (4.18)$$

y claramente suman uno. Para enfatizar la dependencia de todo el conjunto de parámetros $\theta = \{\beta_{10}, \beta_1^T, \dots, \beta_{(K-1)0}, \beta_{(K-1)}^T\}$, denotamos las probabilidades $\Pr(G = K|X = x) = p_k(x; \theta)$.

Cuando $K = 2$, este modelo es especialmente simple, ya que solo hay una función lineal. Es ampliamente utilizado en aplicaciones bioestadísticas donde las (dos clases) ocurren con bastante frecuencia.

Ajuste de modelos de regresión logística.

Los modelos de regresión logística generalmente se ajustan por máxima verosimilitud, usando la probabilidad condicional de G dado X, dado que $\Pr(G|X)$ especifica completamente la distribución condicional. La log-verosimilitud para N observaciones es:

$$l(\theta) = \sum_{i=1}^N \log p_{g_i}(x_i; \theta),$$

Donde $p_k(x_i; \theta) = \Pr (G = k|X = x_i; \theta)$. [4]

4.2. Árboles de clasificación y regresión

Primero describimos un método popular para la clasificación y regresión basada en árboles llamado CART (*Classification and Regression Trees*). Los árboles de clasificación y regresión son el algoritmo más popular que forma parte del proceso de minería de datos. [12]

Un árbol se construye a un análisis de un conjunto amplio de datos que va separando la población en subgrupos de individuos que difieren entre sí según un criterio discriminante, basado en la relación entre las características de los individuos de estudios y la variable a predecir.

En la figura 5 se puede ver representados árboles de distinta profundidad:

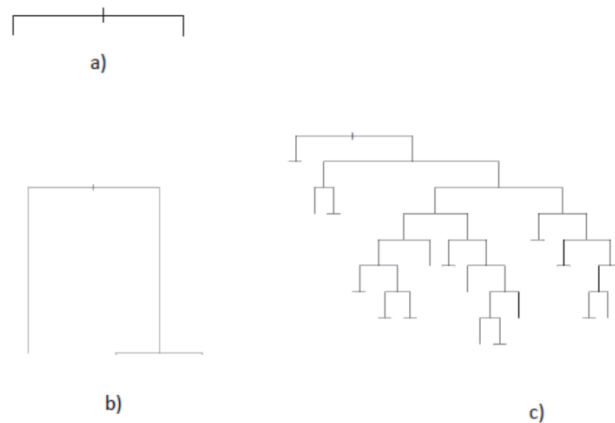


Figura 5: Estructura jerárquica de un árbol. a) Una división, b) Dos divisiones, c) 26 divisiones [11]

Considere un problema de regresión con respuesta continua Y y entradas X1 y X2, cada una tomando valores en el intervalo unitario. El modelo de regresión correspondiente predice Y con un constante c_m en la región R_m es:

$$\hat{f}(X) = \sum_{m=1}^5 c_m I\{(X_1, X_2) \in R_m\}.$$

Este mismo modelo puede ser representado por el árbol binario arriba.

4.2.1. Árboles de regresión

Ahora pasamos a la cuestión de cómo hacer crecer un árbol de regresión. Nuestros datos consisten en p entradas y una respuesta, para cada una de las N observaciones: es (x_i, y_i) para $i = 1, 2, \dots, N$ con $x_i = (x_{i1}, x_{i2}, \dots, x_{ip})$. El algoritmo debe decidir automáticamente las variables de división y los puntos de división, y también qué topología (forma) debe tener el árbol. Supongamos primero que tenemos una partición en M regiones R_1, R_2, \dots, R_M , y modelamos la respuesta como un constante c_m en cada región:

$$f(x) = \sum_{m=1}^M c_m I(x \in R_m).$$

Si adoptamos como nuestro criterio de la minimización de la suma de cuadrados $\sum (y_i - f(x_i))^2$, es fácil ver que el mejor \hat{c}_m es solo el promedio de y_i en la región R_m :

$$\hat{c}_m = \text{avg}(y_i | x_i \in R_m).$$

Ahora, encontrar la mejor partición binaria en términos de suma mínima de cuadrados es generalmente computacionalmente inviable. Por lo tanto, procedemos con un algoritmo. Comenzando con todos los datos, considere una variable de división j y un punto de división s , y defina el par de semiplanos

$$R_1(j, s) = \{X | X_j \leq s\}, \quad R_2(j, s) = \{X | X_j > s\}.$$

Luego buscamos la variable de división j y el punto de división s que resuelven

$$\min_{j,s} \left[\min_{c_1} \sum_{x_i \in R_1(j,s)} (y_i - c_1)^2 + \min_{c_2} \sum_{x_i \in R_2(j,s)} (y_i - c_2)^2 \right].$$

Para cualquier elección j y s , la minimización interna se resuelve mediante

$$\hat{c}_1 = \text{avg}(y_i | x_i \in R_1(j, s)) \quad \text{y} \quad \hat{c}_2 = \text{avg}(y_i | x_i \in R_2(j, s)).$$

Para cada variable de división, la determinación del punto de división s se puede realizar muy rápidamente y, por lo tanto, mediante la exploración de todas las entradas, es posible determinar el mejor par (j, s) . Habiendo encontrado la mejor división, dividimos los datos en las dos regiones resultantes y repetimos el proceso de división en cada una de las dos regiones. Luego, este proceso se repite en todas las regiones resultantes. Entonces, ¿Cómo debemos hacer crecer el árbol? Claramente, un árbol muy grande podría sobreajustar los datos, mientras que un árbol pequeño podría no capturar la estructura importante.

El tamaño del árbol es un parámetro de ajuste que rige la complejidad del modelo, y el tamaño óptimo del árbol debe elegirse de forma adaptativa a partir de los datos. Un enfoque sería dividir los nodos del árbol solo si la disminución en la suma de los cuadrados debido a la división supera algún umbral. Sin embargo, esta estrategia es bastante miope, ya que una división aparentemente sin valor podría conducir a una muy buena división debajo de ella.

La estrategia preferida es hacer crecer un árbol grande T_0 , deteniendo el proceso de división solo cuando se alcanza un tamaño mínimo de nodo (por ejemplo, 5). Luego, este árbol grande se poda utilizando *la poda de complejidad de costos (cost-complexity pruning)*, que ahora lo vamos a describir.

Definimos un subárbol $T \subset T_0$ como cualquier árbol que se puede obtener podando T_0 , es decir, colapsando cualquier número de sus nodos internos (no terminales). Indexamos los nodos terminales por m , donde el nodo m representa la región R_m . Sea $|T|$ denote el número de nodos terminales en T . Sea

$$\begin{aligned} N_m &= \# \{x_i \in R_m\}, \\ \hat{c}_m &= \frac{1}{N_m} \sum_{x_i \in R_m} y_i, \\ Q_m(T) &= \frac{1}{N_m} \sum_{x_i \in R_m} (y_i - \hat{c}_m)^2, \end{aligned} \tag{9.15}$$

definimos el criterio de complejidad de costes

$$C_\alpha(T) = \sum_{m=1}^{|T|} N_m Q_m(T) + \alpha |T|.$$

La idea es encontrar, para cada α , el subárbol $T_\alpha \subseteq T_0$ para minimizar $C_\alpha(T)$. El parámetro de ajuste $\alpha \geq 0$ gobierna el equilibrio entre el tamaño del árbol y su bondad de ajuste a los datos. Valores grandes de α dan como resultado árboles T_α más pequeños, y viceversa para los valores más pequeños de α . Como sugiere la notación, con $\alpha = 0$ la solución es el árbol completo T_0 .

Para encontrar T_α usamos *weakest link pruning*: Colapsamos sucesivamente el nodo interno que produce el aumento más pequeño por-nodo crece en $\sum_m N_m Q_m(T)$, y continuamos hasta producir el árbol de un solo nodo (raíz). Esto da una secuencia (finita) de subárboles, y uno puede mostrar que esta secuencia debe contener T_α . [1][4]

4.2.2. Árbol de clasificación

Si el objetivo es un resultado de clasificación tomando valores $1, 2, \dots, K$, los únicos cambios necesarios en el algoritmo del árbol están relacionados con los criterios para dividir nodos y podar el árbol. Para la regresión usamos la medida de error cuadrático de impureza del nodo

$Q_m(T)$ definida en (9.15), pero esto no es adecuado para la clasificación. En un nodo m , que representa una región R_m con N_m observaciones, sea

$$\hat{p}_{mk} = \frac{1}{N_m} \sum_{x_i \in R_m} I(y_i = k),$$

la proporción de observaciones de clase k en el nodo m . Clasificamos las observaciones en el nodo m en clase $k(m) = \arg \max_k \hat{p}_{mk}$, la clase mayoritaria en el nodo m . Las diferentes medidas $Q_m(T)$ de la impureza del nodo incluyen las siguientes:

Error de clasificación errónea:

$$\frac{1}{N_m} \sum_{i \in R_m} I(y_i \neq k(m)) = 1 - \hat{p}_{mk}.$$

Índice Gini:

$$\sum_{k \neq k'} \hat{p}_{mk} \hat{p}_{mk'} = \sum_{k=1}^K \hat{p}_{mk} (1 - \hat{p}_{mk}).$$

Entropía cruzada / desviación:

$$-\sum_{k=1}^K \hat{p}_{mk} \log(\hat{p}_{mk}).$$

Para dos clases, si p es la proporción en la segunda clase, estas tres medidas son: $1 - \max(p, 1 - p)$, $2p(1 - p)$, y, $-p \log p - (1 - p) \log(1 - p)$, respectivamente. Se muestran en la Figura 6. Las tres son similares, pero la entropía cruzada y el índice de Gini son diferenciables y, por lo tanto, más susceptibles de optimización numérica.

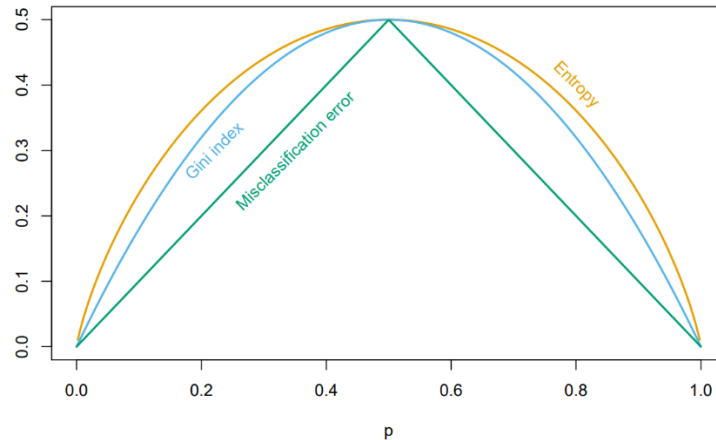


Figure 6. Medidas de impureza de nodo para la clasificación de dos clases, en función de la proporción p en la clase 2. La entropía cruzada se ha escalado para pasar $(0.5, 0.5)$. [4]

Además, la entropía cruzada y el índice de Gini son más sensibles a los cambios en las probabilidades de los nodos que la tasa de error de clasificación.

Por ejemplo, en un problema de dos clases con 400 observaciones en cada clase (denote esto con (400, 400)), supongamos que uno divide los nodos creados (300, 100) y (100, 300), mientras el otro crea nodos (200, 400) y (200, 0). Ambas divisiones producen una tasa de clasificación errónea de 0,25, pero la segunda división produce un nodo puro y probablemente sea preferible. Tanto el índice de Gini como la entropía cruzada son más bajos para la segunda división. Por esta razón, se debe usar el índice de Gini o la entropía cruzada al hacer crecer el árbol. Para guiar la reducción de la complejidad de costos, se puede usar cualquiera de las tres medidas, pero generalmente es *la tasa de error de clasificación*. [4]

4.3. Random Forest (RF)

Bagging o *Bootstrap aggregation* es una técnica para reducir la varianza de una función de predicción estimada. *Bagging* parece funcionar especialmente bien para procedimientos de alta varianza y bajo sesgo como los árboles. Para regresión, simplemente ajustamos el mismo árbol de regresión muchas veces para arrancar versiones muestreadas de los datos de entrenamiento y promediar el resultado.

Random forests – Bosques aleatorios (Breiman, 2001) es una modificación significativa del *Bagging* que construye una gran colección de árboles no correlacionados y luego los promedia.

El uso de Random Forest no se limitará al pronóstico de nuevos datos sino también al análisis de la relación entre la variable respuesta con el resto de las variables, estableciéndose un ranking de importancia de las variables bajo diferentes criterios de valoración.

El algoritmo para Random Forest [4]:

- Para $b = 1$ a B :
 - a) Extraer una muestra Bootstrap Z^* con el tamaño N de los datos de entrenamiento.
 - b) Haz crecer un árbol de bosque aleatorio T_b a los datos de Bootstrap, repitiendo recursivamente los siguientes pasos para cada nodo terminal del árbol, hasta alcanzar el tamaño mínimo de nodo n_{min} .
 - i. Selecciona m variables aleatoriamente de p variables.
 - ii. Elige las mejores variables entre m variables.
 - iii. Divide el nodo en dos nodos secundarios.
- Salida del conjunto de árboles $\{T_b\}_1^B$.

Para hacer una predicción en un nuevo punto x :

$$\text{Regresión: } \hat{f}_{rf}^B = \frac{1}{B} \sum_{b=1}^B T_b(x).$$

Clasificación: Sea $\hat{C}_b(x)$ la predicción de clase del árbol de bosque aleatorio b^{th} . Entonces $\hat{C}_{rf}^B(x) = \text{Voto mayoritario} \{\hat{C}_b(x)\}_1^B$.

Un promedio de B variables aleatorias, cada una con la varianza σ^2 , tiene varianza $\frac{1}{B}\sigma^2$. Con correlación positiva por pares ρ , la varianza del promedio es:

$$\rho\sigma^2 + \frac{1-\rho}{B}\sigma^2$$

A medida que B aumenta, el segundo término desaparece, por lo tanto, el tamaño de la correlación de pares de árboles de árboles embolsados limita los beneficios del promedio. La idea en Random Forest es mejorar la reducción de la varianza del ensacado al reducir la correlación entre los árboles, sin aumentar demasiado la varianza. Esto se logra en el proceso de crecimiento de árboles a través de la selección aleatoria de las variables de entrada.

Normalmente, los valores de m son \sqrt{p} o incluso 1.

Después de que B tales árboles $\{T(x; \theta_b)\}_1^B$ se crezcan, el Random Forest (regresión) predictor es:

$$\hat{f}_{rf}^B(x) = \frac{1}{B} \sum_{b=1}^B T(x; \theta_b).$$

θ_b caracteriza el árbol del bosque aleatorio b^{th} en términos de variables divididas, puntos de corte en cada nodo y valores de nodo terminal. Intuitivamente, la reducción de m reducirá la correlación entre cualquier par de árboles en el conjunto y, por lo tanto, reducirá la varianza del promedio. [4]

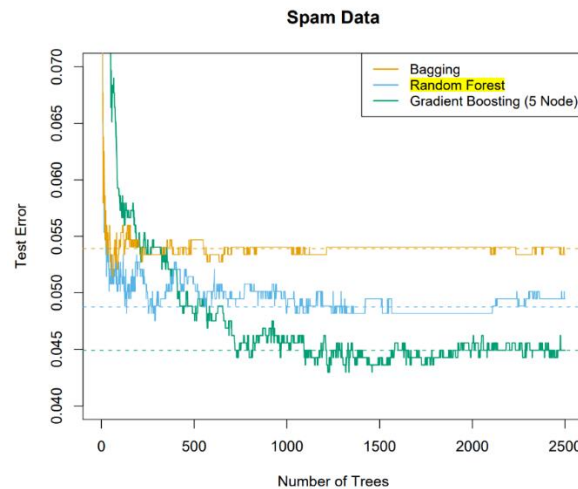


Figura 7. Bagging, Random Forest y Gradient Boosting aplicados a los datos de spam. Para impulsar, se utilizaron árboles de 5 nodos, y el número de árboles se eligió mediante validación cruzada de 10 veces (2500 árboles=). Cada “paso” en la figura corresponde a un cambio en una sola clasificación errónea (Fuente: [4])

Modelo	Ventajas	Desventajas
CART	Mejora los métodos convencionales permitiendo trabajar con múltiples variables predictoras. Fácil de interpretar.	Árboles más inestables porque hay que podarlos. Pequeñas variaciones en los datos crean cambios significativos en el árbol.
Random Forest	La cantidad tan grande de árboles que genera evita el sobreajuste en los datos. Mayor fiabilidad en la predicción al no discriminar variables y elegir las de forma aleatoria con cada árbol. Árboles no podados.	No permite examinar de forma individualiza cada árbol. No es posible interpretar los resultados gráficamente como en el caso de CART.

Tabla 1: Comparativa entre los modelos de pronósticos que utilizan árboles

(Fuente: Elaboración propia)

4.4. Redes neuronales (NNET)

Las Redes Neuronales (*Neural network*) empezaron a tener fama a finales de la década de 1980. Las propiedades de las redes neuronales fueron analizadas por investigadores de aprendizajes automáticos, matemáticos y estadísticos; se mejoraron los algoritmos y se estabilizó la metodología. Luego llegaron las SVM, los Boosting y los bosques aleatorios. Las redes neuronales son frecuentemente utilizadas en el análisis financiero por su habilidad para tratar con problemas altamente no lineales. Están formadas por un conjunto de unidades de computación, llamadas neuronas, que se unen unas a otras creando una red cuya extensión se fija a criterio experto.

4.4.1. Redes neuronales de una sola capa

Una red neuronal toma un vector de entrada de p variables $X = (X_1, X_2, \dots, X_p)$ y construye una función no lineal $f(X)$ para predecir la respuesta Y . Lo que distingue a las redes neuronales de los métodos anteriores es la estructura particular del modelo. La figura 10.1 muestra una red neuronal simple *feed-forward* para modelar una respuesta cuantitativa usando $p = 4$ predictoras. [1]

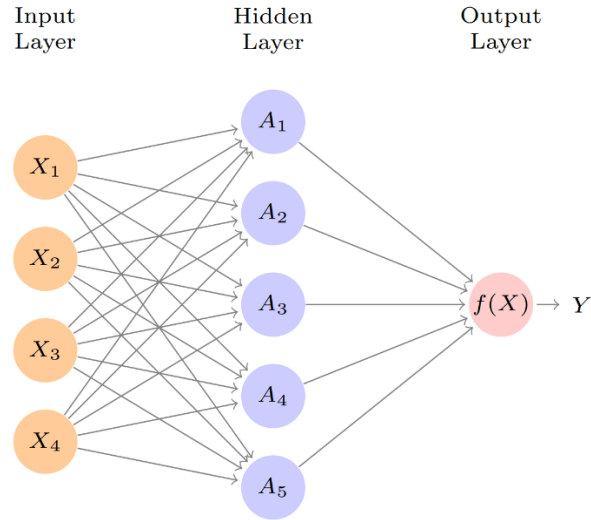


Figura 8. Red neuronal con una sola capa oculta.[1]

En la terminología de las redes neuronales, las cuatro características X_1, \dots, X_4 constituyen las unidades en la capa de entrada. Las flechas indican que cada una de las entradas de la capa de entrada alimenta cada una de las K *unidades ocultas* (podemos elegir K ; aquí elegimos 5). El modelo de red neuronal tiene la forma

$$f(X) = \beta_0 + \sum_{k=1}^K \beta_k h_k(X) = \beta_0 + \sum_{k=1}^K \beta_k g \left(w_{k0} + \sum_{j=1}^p w_{kj} X_j \right). \quad (4.1)$$

Se construye aquí en dos pasos. Primero las K *activaciones* $A_k, k = 1, \dots, K$, en la capa oculta se calculan como funciones de las características de entrada X_1, \dots, X_p ,

$$A_k = h_k(X) = g \left(w_{k0} + \sum_{j=1}^p w_{kj} X_j \right), \quad (4.2)$$

donde $g(z)$ es una función de activación no lineal que se especifica de antemano. Podemos pensar en cada función A_k como una transformación diferente $h_k(X)$ de las características originales. Estas activaciones K de la capa oculta luego alimentan la capa de salida, lo que da como resultado:

$$f(X) = \beta_0 + \sum_{k=1}^K \beta_k A_k,$$

un modelo de regresión lineal en las $K = 5$ activaciones. Todos los parámetros β_0, \dots, β_K y w_{10}, \dots, w_{Kp} deben estimarse a partir de los datos. En los primeros casos de redes neuronales, se favorecía la función de activación sigmoidea

$$g(z) = \frac{e^z}{1 + e^z} = \frac{1}{1 + e^{-z}},$$

La opción preferida en las redes neuronales modernas es ReLU (*rectified linear unit*) función de activación que tiene la forma

$$g(z) = (z)_+ = \begin{cases} 0 & \text{si } z < 0 \\ z & \text{otros caso.} \end{cases}$$

Una activación ReLU se puede calcular y almacenar de manera más eficiente que una activación sigmoidea. Aunque su umbral es cero, debido a que lo aplicamos a una función lineal (4.2), el término constante w_{k0} cambiará este punto de inflexión. En otras palabras, el modelo representado en la Figura 4.1 deriva cinco características nuevas mediante el cálculo de cinco combinaciones lineales diferentes de X , y luego aplasta cada una a través de una función de activación $g(\cdot)$ para transformarla.[1]

El nombre de red neuronal se derivó originalmente de pensar en estas unidades ocultas como análogas a las neuronas en el cerebro: los valores de las activaciones $A_k = h_k(X)$ cercanas a uno están activas, mientras que las cercanas a cero están en silencio.

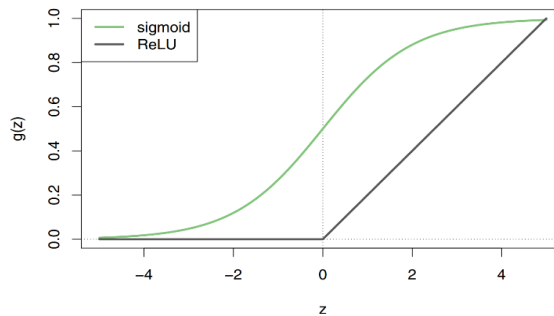


Figura 9. Funciones de activación.

(Fuente: An Introduction to Statistical Learning with Applications in R, Gareth, Daniela, Trevor, Robert)[1]

Ajustar una red neuronal requiere estimar los parámetros desconocidos en (4.1). Para una respuesta cuantitativa, normalmente se utiliza el error cuadrático, de modo que los parámetros se eligen para minimizar:

$$\sum_{i=1}^n (y_i - f(x_i))^2.$$

4.4.2. Redes Neuronales Multicapa

Las redes neuronales modernas suelen tener más de una capa oculta y, a menudo, muchas unidades por capa. En teoría, una sola capa oculta con una gran cantidad de unidades tiene la capacidad de aproximar la mayoría de las funciones. Sin embargo, la tarea de aprendizaje de descubrir una buena solución se hace mucho más fácil con múltiples capas, cada una de tamaño modesto.

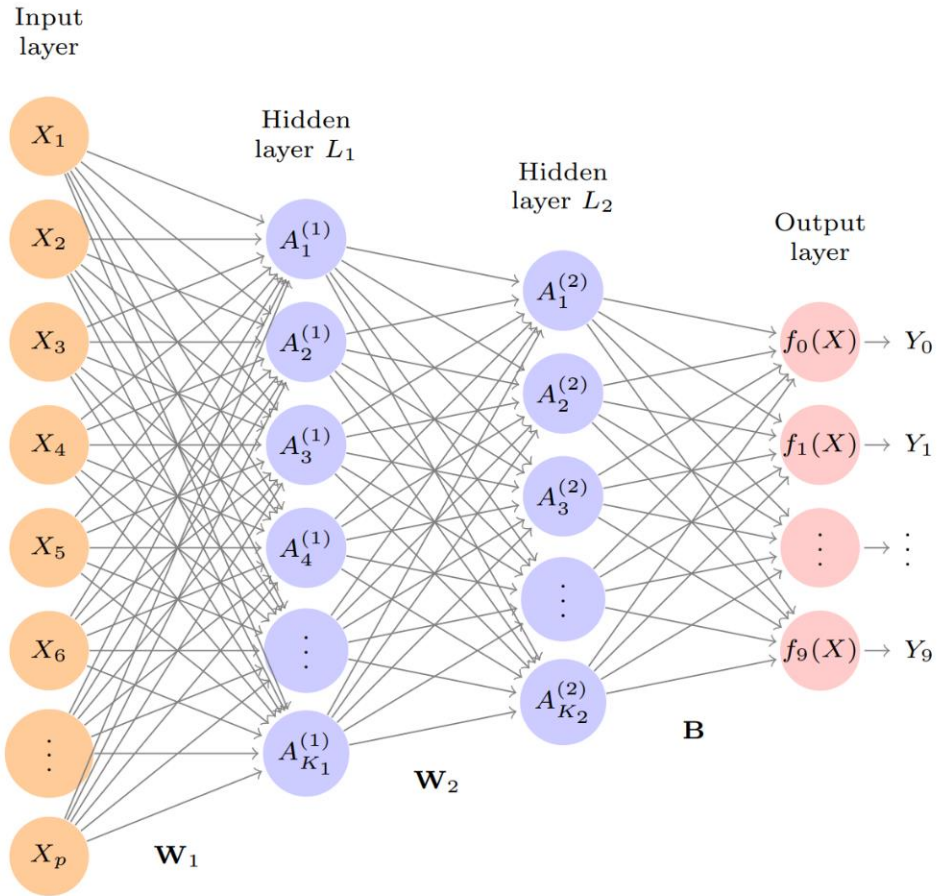


Figura 10. Diagrama de red neuronal con dos capas ocultas y múltiples salidas[1]

Por ejemplo, la capa de entrada tiene $p = 784$ unidades, las dos capas ocultas $K1 = 256$ y $K2 = 128$ unidades respectivamente, y la capa de salida 10 unidades. Esta red tiene 235.146 parámetros

La Figura 10 muestra la red neuronal con dos capas ocultas:

La primera capa oculta según (4.2) tenemos:

$$A_k^{(1)} = h_k^{(1)}(X) = g \left(w_{k0}^{(1)} + \sum_{j=1}^p w_{kj}^{(1)} X_j \right)$$

para $k = 1, \dots, K_1$.

La segunda capa oculta trata las activaciones $A_k^{(1)}$ de la primera capa oculta como entradas y calcula nuevas activaciones:

$$A_l^{(2)} = h_l^{(2)}(X) = g \left(w_{l0}^{(2)} + \sum_{k=1}^{K_1} w_{lk}^{(2)} A_k^{(1)} \right)$$

para $l = 1, \dots, K_2$.

Observe que cada una de las activaciones en la segunda capa $A_l^{(2)} = h_l^{(2)}(X)$ es una función del vector de entrada X .

Ahora llegamos a la capa de salida, donde tenemos diez respuestas en lugar de una. El primer paso es calcular diez modelos lineales diferentes similares a nuestro modelo único (4.1)

$$Z_m = \beta_{m0} + \sum_{l=1}^{K_2} \beta_{ml} h_l^{(2)}(X) = \beta_{m0} + \sum_{l=1}^{K_2} \beta_{ml} A_l^{(2)}$$

para $m = 0, 1, \dots, 9$.

Nos gustaría que nuestras estimaciones representaran probabilidades de clase $f_m(X) = \Pr(Y = m|X)$, al igual que en la regresión logística multinomial. Entonces usamos la función de activación *softmax*:

$$f_m(X) = \Pr(Y = m|X) = \frac{e^{z_m}}{\sum_{l=0}^9 e^{z_l}} \quad (m = 0, 1, \dots, 9)$$

Para entrenar esta red, dado que la respuesta es cualitativa, buscamos estimaciones de coeficientes que minimicen la probabilidad logarítmica multinomial negativa

$$-\sum_{i=1}^n \sum_{m=0}^9 y_{im} \log(f_m(x_i)),$$

también conocida como la entropía cruzada (*cross-entropy*).

Otro aspecto a destacar es que el aprendizaje de la red es de tipo supervisado, ya que durante el entrenamiento se va comparando continuamente el valor de salida de la red con la solución real. La función que define la mejora o empeoramiento de la capacidad predictiva de la red es el error cuadrático medio (MSE). [1][3]

4.5. Máquinas de soporte vectorial (SVM)

Otra técnica que se aplica en este TFM será la de máquinas de soporte vectorial, que se pueden usar para los problemas de clasificación y regresión. La máquina de soporte vectorial (SVM), un enfoque para la clasificación que se desarrolló en la comunidad informática en la década de 1990 y que ha ganado popularidad desde entonces. Este tipo de algoritmos de aprendizaje supervisado fueron desarrollados por Vladimir Vapnik y su equipo en los laboratorios AT&T.

4.5.1. Clasificador de soporte vectorial

El clasificador de vectores de soporte clasifica una observación de prueba según el lado de un hiperplano en el que se encuentra. El hiperplano se elige para separar correctamente la mayoría de las observaciones de entrenamiento en las dos clases, pero puede clasificar erróneamente algunas observaciones. Es la solución al problema de optimización [4]

$$\underset{\beta_0, \beta_1, \dots, \beta_p, \epsilon_1, \dots, \epsilon_n, M}{\text{maximize}} \quad M \quad (4.3)$$

$$\text{subject to} \quad \sum_{j=1}^p \beta_j^2 = 1, \quad (4.4)$$

$$y_i(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip}) \geq M(1 - \epsilon_i), \quad (4.5)$$

$$\epsilon_i \geq 0, \quad \sum_{i=1}^n \epsilon_i \leq C, \quad (4.6)$$

donde C es un parámetro de sintonización no negativo. M es el ancho del margen; buscamos que esta cantidad sea lo más grande posible. $\epsilon_1, \dots, \epsilon_n$ son variables de holgura que permiten que las observaciones individuales estén en el lado equivocado del margen o del hiperplano.

Una vez que hemos resuelto (4.3)–(4.6), clasificamos una observación de prueba x^* como antes, simplemente determinando en qué lado del hiperplano se encuentra. Es decir, clasificamos la observación de prueba según el signo de $f(x^*) = \beta_0 + \beta_1 x_1^* + \dots + \beta_p x_p^*$.

4.5.1.1. Clasificación con límites de decisión no lineales

Por ejemplo, en lugar de ajustar un clasificador de vector de soporte usando p características:

$$X_1, X_2, \dots, X_p,$$

podríamos ajustar un clasificador de vector de soporte usando $2p$ características:

$$X_1, X_1^2, X_2, X_2^2, \dots, X_p, X_p^2.$$

Entonces, (4.3)–(4.6) se convierten en:

$$\underset{\beta_0, \beta_{11}, \beta_{12}, \dots, \beta_{p1}, \beta_{p2}, \epsilon_1, \dots, \epsilon_n, M}{\text{maximize}} \quad M$$

$$\text{subject to} \quad y_i \left(\beta_0 + \sum_{j=1}^p \beta_{j1} x_{ij} + \sum_{j=1}^p \beta_{j2} x_{ij}^2 \right) \geq M(1 - \epsilon_i),$$

$$\sum_{i=1}^n \epsilon_i \leq C, \quad \epsilon_i \geq 0, \quad \sum_{j=1}^p \sum_{k=1}^2 \beta_{jk}^2 = 1.$$

4.5.2. Máquinas de soporte vectorial (SVM)

La máquina de soporte vectorial (SVM) es una extensión del clasificador de soporte vectorial que resulta de ampliar el espacio de características de una manera específica usando *kernels*. [1]

El producto interno de dos observaciones $x_i, x_{i'}$, dado:

$$\langle x_i, x_{i'} \rangle = \sum_{j=1}^p x_{ij} x_{i'j} . \quad (4.7)$$

- El clasificador de soporte vectorial lineal se puede representar como

$$f(x) = \beta_0 + \sum_{i=1}^n \alpha_i \langle x, x_i \rangle \quad (4.8)$$

donde hay n parámetros $\alpha_i, i = 1, \dots, n$, una por observación de entrenamiento.

- Para estimar los parámetros $\alpha_1, \dots, \alpha_n$ y β_0 , necesitamos los $\binom{n}{2}$ productos internos $\langle x, x_{i'} \rangle$ entre todos los pares de observaciones de entrenamiento. $(\binom{n}{2} = \frac{n(n-1)}{2})$

si \mathcal{S} es el conjunto de índices de estos puntos de apoyo, podemos reescribir cualquier función solución de la forma (4.8) como:

$$f(x) = \beta_0 + \sum_{i \in \mathcal{S}} \alpha_i \langle x, x_i \rangle$$

Supongamos ahora que cada vez que aparece el producto interior (4.7) en la representación (4.8), o en un cálculo de la solución para el clasificador de soporte vectorial, lo reemplazamos con una generalización del producto interior de la forma $K(x_i, x_{i'})$, donde K es una función a la que nos referiremos como *kernel*. Un *kernel* es una función que cuantifica la similitud de dos observaciones. Por ejemplo, podríamos simplemente tomar

$$K(x_i, x_{i'}) = \sum_{j=1}^p x_{ij} x_{i'j} , \quad (4.9)$$

lo que simplemente nos devolvería el clasificador de soporte vectorial.

La ecuación 4.9 se conoce como *kernel*/lineal porque el clasificador de soporte vectorial es lineal en las características; el núcleo lineal cuantifica esencialmente la similitud de un par

de observaciones utilizando la correlación de Pearson (estándar). Por ejemplo, reemplazamos cada instancia de $\sum_{j=1}^p x_{ij}x_{i'j}$ con la cantidad:

$$K(x_i, x_{i'}) = \left(1 + \sum_{j=1}^p x_{ij}x_{i'j} \right)^d. \quad (4.10)$$

Esto se conoce como *núcleo polinomial* de grado d , donde d es un número entero positivo.

Cuando el clasificador de soporte vectorial se combina con un núcleo no lineal como (4.10), el clasificador se conoce como *máquina de soporte vectorial*. La función en este caso tiene la forma:

$$f(x) = \beta_0 + \sum_{i \in S} \alpha_i K(x_i, x_i)$$

El panel izquierdo de la figura 11 muestra un ejemplo de SVM con un kernel polinomial. Cuando $d = 1$, entonces la SVM se reduce al clasificador de vector de soporte.

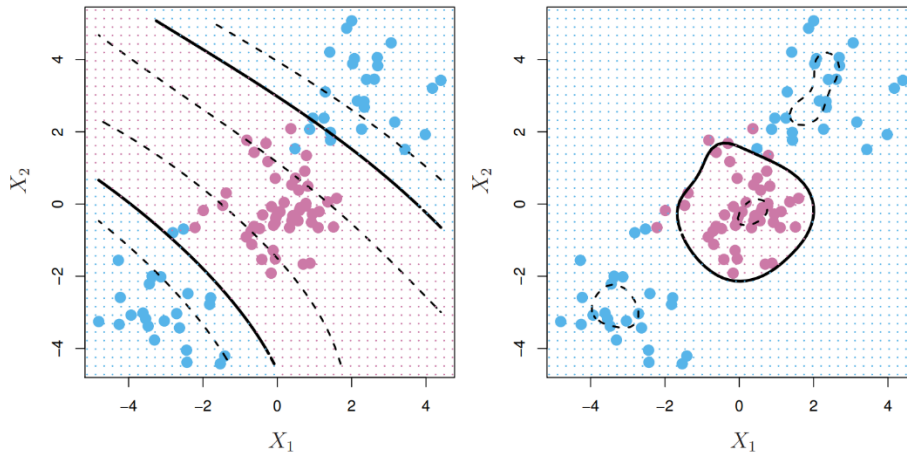


Figura 11. Izquierda: se aplica una SVM con un núcleo polinomial de grado 3. Derecha: se aplica una SVM con un kernel radial. [1]

(Fuente: An Introduction to Statistical Learning with Applications in R, Gareth, Daniela, Trevor, Robert)

Otra opción popular es el kernel radial, que toma la forma

$$K(x_i, x_{i'}) = \exp\left(-\gamma \sum_{j=1}^p (x_{ij} - x_{i'j})^2\right).$$

γ es una constante positiva. A la derecha de la Figura 11, se muestra un ejemplo de una SVM con un núcleo radial con los datos no lineales.

4.5.3. Relación con la regresión logística

Han surgido profundas conexiones entre las SVM y otros métodos estadísticos más clásicos. Resulta que uno puede reescribir el criterio (4.3)–(4.6) para ajustar el clasificador de soporte vectorial $f(X) = \beta_0 + \beta_1 X_1 + \dots + \beta_p X_p$ como:

$$\underset{\beta_0, \beta_1, \dots, \beta_p}{\text{minimize}} \left\{ \sum_{i=1}^n \max[0, 1 - y_i f(x_i)] + \lambda \sum_{j=1}^p \beta_j^2 \right\}, \quad (4.11)$$

donde λ es un parámetro de sintonía no negativo. Cuando λ es grande, entonces β_1, \dots, β_p son pequeños, se toleran más violaciones al margen y resultará un clasificador de baja varianza, pero sesgo alto. Cuando λ es pequeño, se producirán pocas violaciones al margen; esto equivale a un clasificador de alta varianza, pero sesgo bajo.

La función $\lambda \sum_{j=1}^p \beta_j^2$ es el término de penalización, juega un papel similar en el control de la compensación de sesgo-varianza para el clasificador de soporte vectorial.

Ahora (4.11) toma la forma “Pérdida + Penalización”:

$$\underset{\beta_0, \beta_1, \dots, \beta_p}{\text{minimize}} \{L(\mathbf{X}, \mathbf{y}, \beta) + \lambda P(\beta)\}.$$

$L(\mathbf{X}, \mathbf{y}, \beta)$ es alguna función de pérdida que cuantifica la medida en que el modelo, parametrizado por β , se ajusta a los datos (\mathbf{X}, \mathbf{y}) , y $P(\beta)$ es una función de penalización en el vector de parámetros β cuyo efecto está controlado por un parámetro de sintonización no negativo λ . Por ejemplo, la *ridge regression* y el lazo toman esta forma con

$$L(\mathbf{X}, \mathbf{y}, \beta) = \sum_{i=1}^n \left(y_i - \beta_0 - \sum_{j=1}^p x_{ij} \beta_j \right)^2$$

y con $P(\beta) = \sum_{j=1}^p \beta_j^2$ para *ridge regression* y $P(\beta) = \sum_{j=1}^p |\beta_j|$ para el *lasso*.

En el caso de (9.25) la función de pérdida tiene la forma:

$$L(\mathbf{X}, \mathbf{y}, \beta) = \sum_{i=1}^n \max[0, 1 - y_i(\beta_0 + \beta_1 x_{i1} + \dots + \beta_p x_{ip})].$$

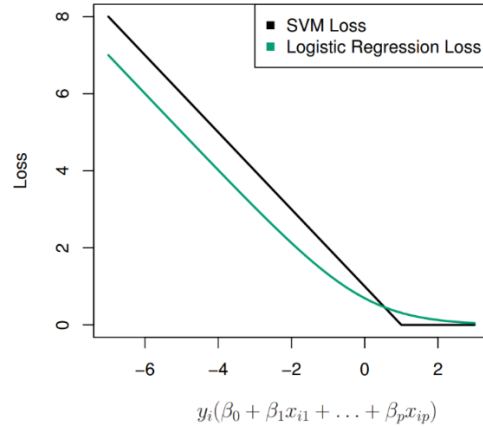


Figura 12. Se comparan las funciones de pérdida de SVM y regresión logística, en función de $y_i(\beta_0 + \beta_1x_{i1} + \dots + \beta_px_{ip})$. Cuando $y_i(\beta_0 + \beta_1x_{i1} + \dots + \beta_px_{ip})$ es mayor que 1, entonces la pérdida SVM es cero, ya que esto corresponde a una observación que está en el lado correcto del margen. En general, las dos funciones de pérdida tienen un comportamiento bastante similar. [1]

4.6. Resultados: Trading del sistema

Una vez llevadas a cabo las predicciones por medio de las técnicas definidas anteriormente es necesario medir el desempeño para poder sacar conclusiones válidas de cara a futuros análisis. Principalmente existen dos formas de evaluar el sistema de predicción de precios: la evaluación estadística y la evaluación económica.

- Evaluación estadística: Este primer tipo de medición se basa en la utilización de las matrices de confusión introducidas anteriormente para poder comparar la precisión y la sensibilidad de los distintos modelos. Aquellos que obtengan un mayor valor de estos parámetros para los datos de test, querrá decir que estadísticamente hablando el modelo es mejor ya que se aproxima en mayor medida a los datos reales en un mayor número de ocasiones.
- Evaluación económica: Esta forma de valorar el sistema es habitual en el caso de sistemas de predicción relacionados con mercados financieros. Su principal argumento radica en que no importa lo bueno que sea el modelo estadístico si luego ese desempeño no se traduce en una ganancia para el inversor que lo utiliza.

5. APLICACIÓN

5.1. Variables utilizadas en el análisis

Como se ha comentado anteriormente, para llevar a cabo la predicción es necesario ajustar la variable deseada, en este caso el indicador de la empresa Telefónica, por medio de una serie de variables explicativas que conforman el modelo de predicción. Este es el primer e importante paso del análisis desarrollado.

Para este caso se han tenido en cuenta, por un lado, variables independientes del precio de Telefónica, como son los indicadores del resto de compañías de las que tenemos datos, y por otro lado, variables dependientes del precio del mismo. A continuación, se presenta una tabla resumen de todas las variables utilizadas. [5] [9]

No	Variable R	Nombre
1	IndTEF	Indicador construido a partir de las cotizaciones de Telefónica S.A.
2	IndTEF_delta_HL	Diferencia entre el precio High y Low de las cotizaciones de Telefónica S.A.
3	IndTEF_lag20	Indicador construido a partir de las cotizaciones de Telefónica S.A. el día t-20
4	IndTEF_lag30	Indicador construido a partir de las cotizaciones de Telefónica S.A. el día t-30
5	IndBBVA_lag5	Indicador construido a partir de las cotizaciones de BBVA S.A. el día t-5
6	IndBBVA_lag8	Indicador construido a partir de las cotizaciones de BBVA S.A. el día t-8
7	IndBBVA_lag13	Indicador construido a partir de las cotizaciones de BBVA S.A. el día t-13
8	IndBBVA_lag20	Indicador construido a partir de las cotizaciones de BBVA S.A. el día t-20
9	IndBBVA_lag30	Indicador construido a partir de las cotizaciones de BBVA S.A. el día t-30
10	IndBKT_lag5	Indicador construido a partir de las cotizaciones de Bankinter S.A. el día t-5
11	IndBKT_lag8	Indicador construido a partir de las cotizaciones de Bankinter S.A. el día t-8
12	IndBKT_lag13	Indicador construido a partir de las cotizaciones de Bankinter S.A. el día t-13
13	IndBKT_lag20	Indicador construido a partir de las cotizaciones de Bankinter S.A. el día t-20
14	IndBKT_lag30	Indicador construido a partir de las cotizaciones de Bankinter S.A. el día t-30
15	IndCABK_lag5	Indicador construido a partir de las cotizaciones de Caixabank S.A. el día t-5
16	IndCABK_lag8	Indicador construido a partir de las cotizaciones de Caixabank S.A. el día t-8
17	IndCABK_lag13	Indicador construido a partir de las cotizaciones de Caixabank S.A. el día t-13
18	IndCABK_lag20	Indicador construido a partir de las cotizaciones de Caixabank S.A. el día t-20
19	IndCABK_lag20	Indicador construido a partir de las cotizaciones de Caixabank S.A. el día t-30
20	IndSAN_lag5	Indicador construido a partir de las cotizaciones de Santander S.A. el día t-5
21	IndSAN_lag8	Indicador construido a partir de las cotizaciones de Santander S.A. el día t-8
22	IndSAN_lag13	Indicador construido a partir de las cotizaciones de Santander S.A. el día t-13
23	IndSAN_lag20	Indicador construido a partir de las cotizaciones de Santander S.A. el día t-20
24	IndSAN_lag30	Indicador construido a partir de las cotizaciones de Santander S.A. el día t-30
25	IndIDR_lag5	Indicador construido a partir de las cotizaciones de Indra S.A. el día t-5
26	IndIDR_lag8	Indicador construido a partir de las cotizaciones de Indra S.A. el día t-8
27	IndIDR_lag13	Indicador construido a partir de las cotizaciones de Indra S.A. el día t-13
28	IndIDR_lag20	Indicador construido a partir de las cotizaciones de Indra S.A. el día t-20
29	IndIDR_lag20	Indicador construido a partir de las cotizaciones de Indra S.A. el día t-30
30	IndAMA_lag5	Indicador construido a partir de las cotizaciones de Amadeus S.A. el día t-5
31	IndAMA_lag8	Indicador construido a partir de las cotizaciones de Amadeus S.A. el día t-8
32	IndAMA_lag13	Indicador construido a partir de las cotizaciones de Amadeus S.A. el día t-13
33	IndAMA_lag20	Indicador construido a partir de las cotizaciones de Amadeus S.A. el día t-20
34	IndAMA_lag30	Indicador construido a partir de las cotizaciones de Amadeus S.A. el día t-30
35	IndELE_lag5	Indicador construido a partir de las cotizaciones de Endesa S.A. el día t-5

36	IndELE_lag8	Indicador construido a partir de las cotizaciones de Endesa S.A. el día t-8
37	IndELE_lag13	Indicador construido a partir de las cotizaciones de Endesa S.A. el día t-13
38	IndELE_lag20	Indicador construido a partir de las cotizaciones de Endesa S.A. el día t-20
39	IndELE_lag30	Indicador construido a partir de las cotizaciones de Endesa S.A. el día t-30
40	IndNTGY_lag5	Indicador construido a partir de las cotizaciones de Naturgy S.A. el día t-5
41	IndNTGY_lag8	Indicador construido a partir de las cotizaciones de Naturgy S.A. el día t-8
42	IndNTGY_lag13	Indicador construido a partir de las cotizaciones de Naturgy S.A. el día t-13
43	IndNTGY_lag20	Indicador construido a partir de las cotizaciones de Naturgy S.A. el día t-20
44	IndNTGY_lag30	Indicador construido a partir de las cotizaciones de Naturgy S.A. el día t-30
45	IndREP_lag5	Indicador construido a partir de las cotizaciones de Repsol S.A. el día t-5
46	IndREP_lag8	Indicador construido a partir de las cotizaciones de Repsol S.A. el día t-8
47	IndREP_lag13	Indicador construido a partir de las cotizaciones de Repsol S.A. el día t-13
48	IndREP_lag20	Indicador construido a partir de las cotizaciones de Repsol S.A. el día t-20
49	IndREP_lag30	Indicador construido a partir de las cotizaciones de Repsol S.A. el día t-30
50	IndITX_lag5	Indicador construido a partir de las cotizaciones de Inditex S.A. el día t-5
51	IndITX_lag8	Indicador construido a partir de las cotizaciones de Inditex S.A. el día t-8
52	IndITX_lag13	Indicador construido a partir de las cotizaciones de Inditex S.A. el día t-13
53	IndITX_lag20	Indicador construido a partir de las cotizaciones de Inditex S.A. el día t-20
54	IndITX_lag30	Indicador construido a partir de las cotizaciones de Inditex S.A. el día t-30
55	IndANA_lag5	Indicador construido a partir de las cotizaciones de Acciona S.A. el día t-5
56	IndANA_lag8	Indicador construido a partir de las cotizaciones de Acciona S.A. el día t-8
57	IndANA_lag13	Indicador construido a partir de las cotizaciones de Acciona S.A. el día t-13
58	IndANA_lag20	Indicador construido a partir de las cotizaciones de Acciona S.A. el día t-20
59	IndANA_lag30	Indicador construido a partir de las cotizaciones de Acciona S.A. el día t-30
60	IndACS_lag5	Indicador construido a partir de las cotizaciones de ACS S.A. el día t-5
61	IndACS_lag8	Indicador construido a partir de las cotizaciones de ACS S.A. el día t-8
62	IndACS_lag13	Indicador construido a partir de las cotizaciones de ACS S.A. el día t-13
63	IndACS_lag20	Indicador construido a partir de las cotizaciones de ACS S.A. el día t-20
64	IndACS_lag30	Indicador construido a partir de las cotizaciones de ACS S.A. el día t-30
65	IndIBE_lag5	Indicador construido a partir de las cotizaciones de Iberdrola S.A. el día t-5
66	IndIBE_lag8	Indicador construido a partir de las cotizaciones de Iberdrola S.A. el día t-8
67	IndIBE_lag13	Indicador construido a partir de las cotizaciones de Iberdrola S.A. el día t-13
68	IndIBE_lag20	Indicador construido a partir de las cotizaciones de Iberdrola S.A. el día t-20
69	IndIBE_lag30	Indicador construido a partir de las cotizaciones de Iberdrola S.A. el día t-30
70	IndFCC_lag5	Indicador construido a partir de las cotizaciones de FCC S.A. el día t-5
71	IndFCC_lag8	Indicador construido a partir de las cotizaciones de FCC S.A. el día t-8
72	IndFCC_lag13	Indicador construido a partir de las cotizaciones de FCC S.A. el día t-13
73	IndFCC_lag20	Indicador construido a partir de las cotizaciones de FCC S.A. el día t-20
74	IndFCC_lag30	Indicador construido a partir de las cotizaciones de FCC S.A. el día t-30
75	IndFER_lag5	Indicador construido a partir de las cotizaciones de Ferrovial S.A. el día t-5
76	IndFER_lag8	Indicador construido a partir de las cotizaciones de Ferrovial S.A. el día t-8
77	IndFER_lag13	Indicador construido a partir de las cotizaciones de Ferrovial S.A. el día t-13
78	IndFER_lag20	Indicador construido a partir de las cotizaciones de Ferrovial S.A. el día t-20
79	IndFER_lag30	Indicador construido a partir de las cotizaciones de Ferrovial S.A. el día t-30
80	IndTL5_lag5	Indicador construido a partir de las cotizaciones de Mediaset S.A. el día t-5
81	IndTL5_lag8	Indicador construido a partir de las cotizaciones de Mediaset S.A. el día t-8
82	IndTL5_lag13	Indicador construido a partir de las cotizaciones de Mediaset S.A. el día t-13
83	IndTL5_lag20	Indicador construido a partir de las cotizaciones de Mediaset S.A. el día t-20
84	IndTL5_lag30	Indicador construido a partir de las cotizaciones de Mediaset S.A. el día t-30
85	IndREDE_lag5	Indicador construido a partir de las cotizaciones de Red Eléctrica S.A. el día t-5
86	IndREDE_lag8	Indicador construido a partir de las cotizaciones de Red Eléctrica S.A. el día t-8
87	IndREDE_lag13	Indicador construido a partir de las cotizaciones de Red Eléctrica S.A. el día t-13

88	IndREDE_lag20	Indicador construido a partir de las cotizaciones de Red Eléctrica S.A. el día t-20
89	IndREDE_lag30	Indicador construido a partir de las cotizaciones de Red Eléctrica S.A. el día t-30
90	IndSCYR_lag5	Indicador construido a partir de las cotizaciones de Sacyr S.A. el día t-5
91	IndSCYR_lag8	Indicador construido a partir de las cotizaciones de Sacyr S.A. el día t-8
92	IndSCYR_lag13	Indicador construido a partir de las cotizaciones de Sacyr S.A. el día t-13
93	IndSCYR_lag20	Indicador construido a partir de las cotizaciones de Sacyr S.A. el día t-20
94	IndSCYR_lag30	Indicador construido a partir de las cotizaciones de Sacyr S.A. el día t-30
95	IndMAP_lag5	Indicador construido a partir de las cotizaciones de Mapfre S.A. el día t-5
96	IndMAP_lag8	Indicador construido a partir de las cotizaciones de Mapfre S.A. el día t-8
97	IndMAP_lag13	Indicador construido a partir de las cotizaciones de Mapfre S.A. el día t-13
98	IndMAP_lag20	Indicador construido a partir de las cotizaciones de Mapfre S.A. el día t-20
99	IndMAP_lag30	Indicador construido a partir de las cotizaciones de Mapfre S.A. el día t-30
100	IndGRLS_lag5	Indicador construido a partir de las cotizaciones de Grifols S.A. el día t-5
101	IndGRLS_lag8	Indicador construido a partir de las cotizaciones de Grifols S.A. el día t-8
102	IndGRLS_lag13	Indicador construido a partir de las cotizaciones de Grifols S.A. el día t-13
103	IndGRLS_lag20	Indicador construido a partir de las cotizaciones de Grifols S.A. el día t-20
104	IndGRLS_lag30	Indicador construido a partir de las cotizaciones de Grifols S.A. el día t-30
105	IndICAG_lag5	Indicador construido a partir de las cotizaciones de IAG S.A. el día t-5
106	IndICAG_lag8	Indicador construido a partir de las cotizaciones de IAG S.A. el día t-8
107	IndICAG_lag13	Indicador construido a partir de las cotizaciones de IAG S.A. el día t-13
108	IndICAG_lag20	Indicador construido a partir de las cotizaciones de IAG S.A. el día t-20
109	IndICAG_lag30	Indicador construido a partir de las cotizaciones de IAG S.A. el día t-30
110	IndVIS_lag5	Indicador construido a partir de las cotizaciones de Viscofan S.A. el día t-5
111	IndVIS_lag8	Indicador construido a partir de las cotizaciones de Viscofan S.A. el día t-8
112	IndVIS_lag13	Indicador construido a partir de las cotizaciones de Viscofan S.A. el día t-13
113	IndVIS_lag20	Indicador construido a partir de las cotizaciones de Viscofan S.A. el día t-20
114	IndVIS_lag30	Indicador construido a partir de las cotizaciones de Viscofan S.A. el día t-30
115	IndACX_lag5	Indicador construido a partir de las cotizaciones de Acerinox S.A. el día t-5
116	IndACX_lag8	Indicador construido a partir de las cotizaciones de Acerinox S.A. el día t-8
117	IndACX_lag13	Indicador construido a partir de las cotizaciones de Acerinox S.A. el día t-13
118	IndACX_lag20	Indicador construido a partir de las cotizaciones de Acerinox S.A. el día t-20
119	IndACX_lag30	Indicador construido a partir de las cotizaciones de Acerinox S.A. el día t-30
120	IndAENA_lag5	Indicador construido a partir de las cotizaciones de Aena S.A. el día t-5
121	IndAENA_lag8	Indicador construido a partir de las cotizaciones de Aena S.A. el día t-8
122	IndAENA_lag13	Indicador construido a partir de las cotizaciones de Aena S.A. el día t-13
123	IndAENA_lag20	Indicador construido a partir de las cotizaciones de Aena S.A. el día t-20
124	IndAENA_lag30	Indicador construido a partir de las cotizaciones de Aena S.A. el día t-30
125	IndMRL_lag5	Indicador construido a partir de las cotizaciones de Merlin Properties Socimi S.A. el día t-5
126	IndMRL_lag8	Indicador construido a partir de las cotizaciones de Merlin Properties Socimi S.A. el día t-8
127	IndMRL_lag13	Indicador construido a partir de las cotizaciones de Merlin Properties Socimi S.A. el día t-13
128	IndMRL_lag20	Indicador construido a partir de las cotizaciones de Merlin Properties Socimi S.A. el día t-20
129	IndMRL_lag30	Indicador construido a partir de las cotizaciones de Merlin Properties Socimi S.A. el día t-30

Tabla 2. Tabla de variables explicativas iniciales a considerar en el análisis

5.2. Programación del sistema de análisis

Una vez llevado a cabo el análisis de las variables implicadas y escogidas aquellas a utilizar, es hora de programar todas las funciones que ayuden a la realización de la simulación determinista. Hay que señalar que gran parte de la carga del proyecto ha consistido en el desarrollo y programación de estas funciones R. Su funcionamiento se explica mediante los dos siguientes diagramas de bloques:

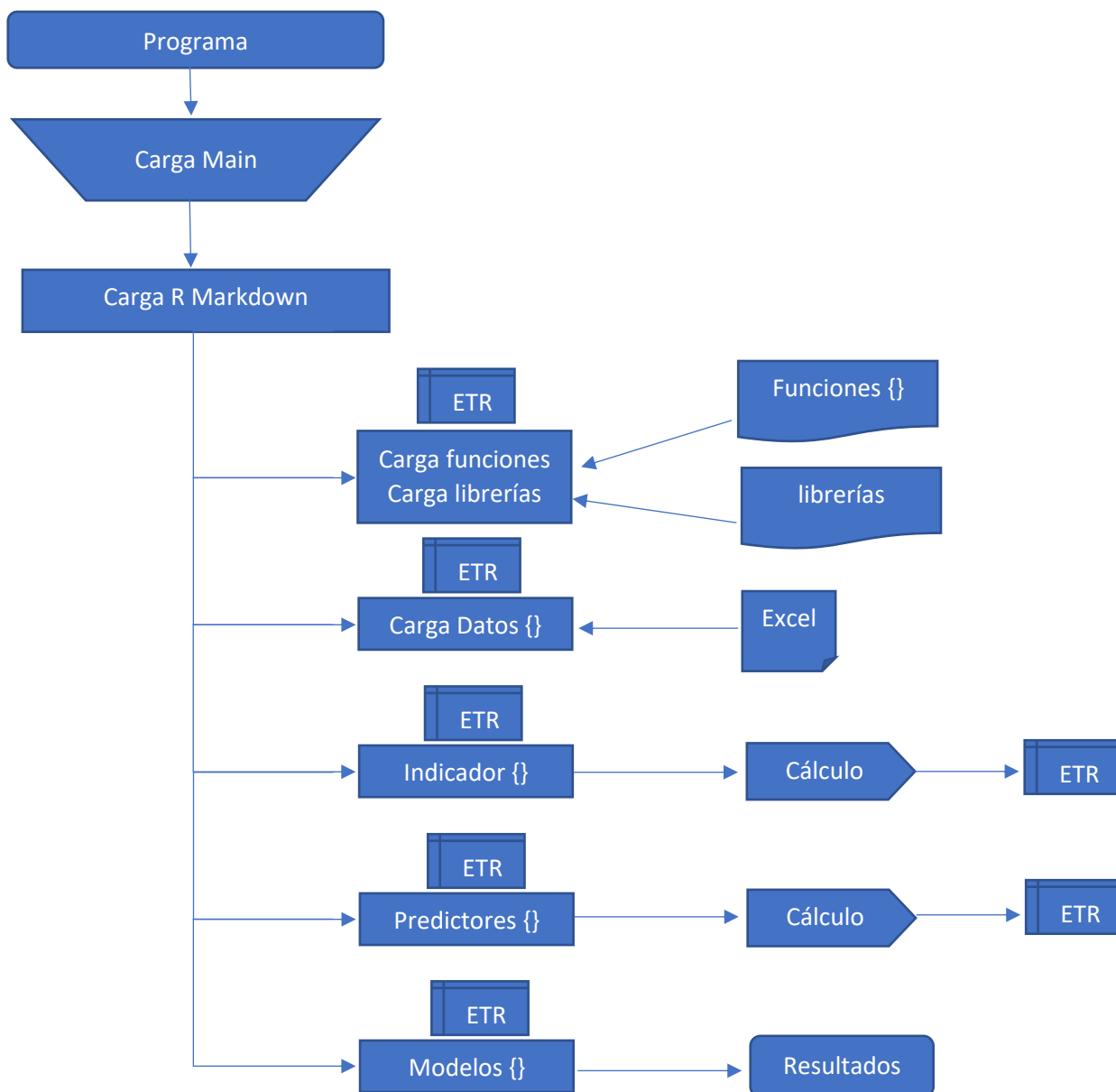


Figura 13. Diagrama de bloques del funcionamiento del sistema de análisis

Explicación el diagrama: Primeramente, se carga la función Main en el espacio de trabajo de R y se ejecuta. Esto debe hacerlo el usuario de forma manual. En segundo lugar, al comenzar a correr esta función carga de forma automática el resto de las funciones, así como las librerías necesarias para el análisis. A continuación, ejecuta la función Carga_Datos() que calcula la variable respuesta de Telefónica, así como del resto de compañías (variables explicativas). Posteriormente entrar en acción la función Predictores() que calcula las variables explicativas relacionadas con el análisis técnico. Por último en la función Modelo() se lleva a cabo un análisis por medio de random forest para conocer y elegir variables explicativas más representativas.

6. RESULTADOS Y DISCUSIÓN

6.1. Selección de variables a través del Random Forest

Después de analizar la base de datos generalmente usando el Random Forest, con las 100 variables, tenemos [14]:

Números de árboles	No. de variables probadas en cada división	Media de residuos cuadráticos	% Var explicado
<i>No.trees</i>	<i>var.Split</i>	<i>MSE</i>	<i>%Var.Explain.</i>
500	42	0.8877111	99.82%

Tabla 3. Datos principales del análisis Random Forest para el modelo general

Vemos que el número de árboles construidos para obtener el resultado final es 500. Así mismo, en cada árbol, el número de variables probado para realizar cada separación es 43. El error cuadrático medio sobre 0.8877111. Por último, el nivel de explicación de las variables probadas hasta un 99,82 %.

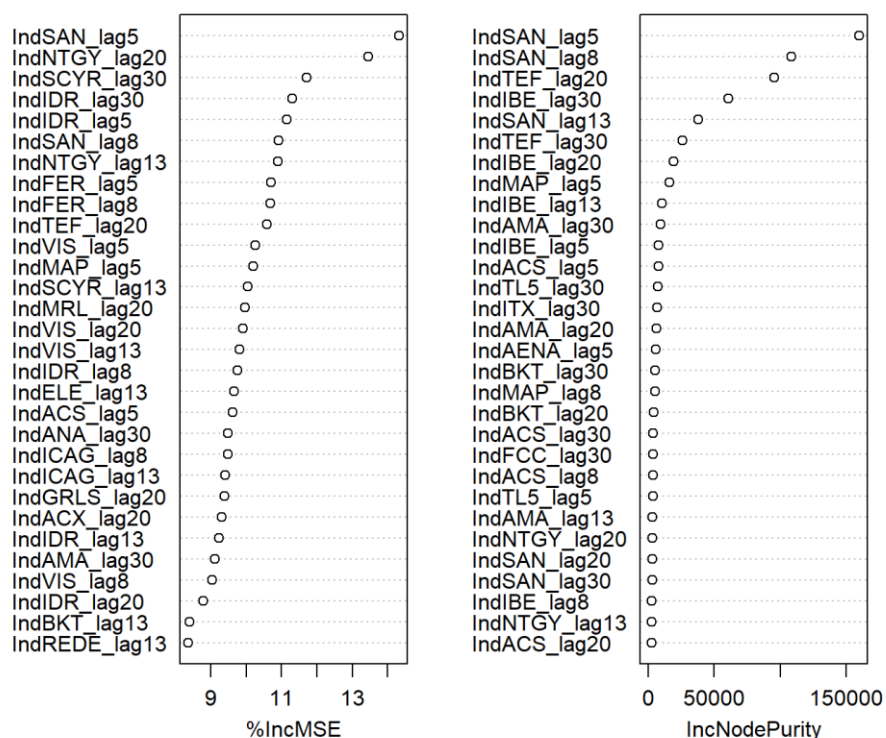


Figura 14. Análisis general de importancia de variables explicativas mediante Random Forest

En la figura 14 puede observarse cómo las variables de mayor importancia, con respecto al indicador de Telefónica son: IndSAN_lag5, IndNTGY_lag20, ...IndREDE_lag13.

Con las 132 variables, se puede complicar mucho para los análisis planteados en este proyecto. Por eso, es necesario reducir el número de variables probando varios modelos para comprobar cuál de ellos se ha comportado mejor. Para ello se experimentan 5 modelos:

- Max(5): Se escogen las 5 variables más relevantes según la figura 14.
- Max(10): Se escogen las 10 variables más relevantes
- Max(15): Se escogen las 15 variables más relevantes

6.1.1. Modelo estadístico Max(5)

Realizamos generalmente las 5 principales variables explicativas usando el análisis de Random Forest, tenemos los siguientes resultados [14]:

Números de árboles	No. de variables probadas en cada división	Media de residuos cuadráticos	% Var explicado
<i>No.trees</i>	<i>var.Split</i>	<i>MSE</i>	<i>%Var.Explain.</i>
500	1	5.865576	98.78%

Tabla 4. Datos principales del análisis Random Forest para el modelo Max(5)

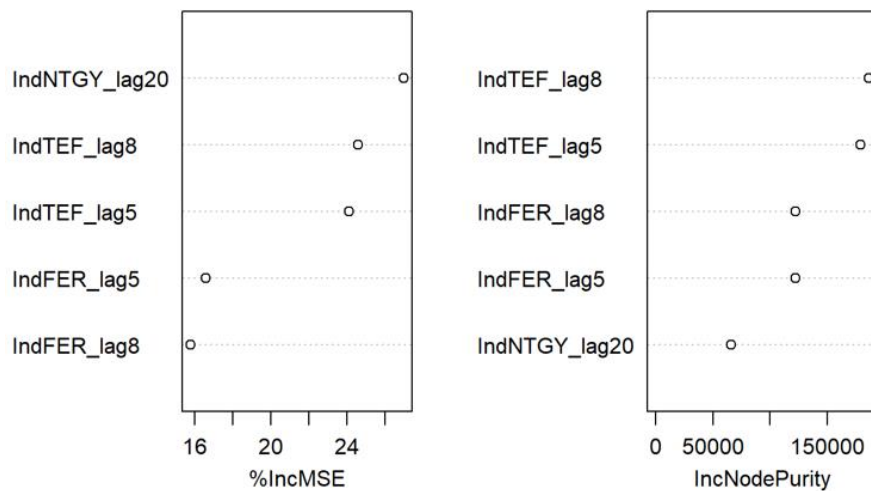


Figura 15. Análisis particular de importancia de variables explicativas mediante Random Forest para el modelo Max(5)

Se ve en la figura 15 que el orden de la importancia de las variables ha cambiado con respecto al caso general en la figura 14. Esto puede ser el descubrimiento de nuevos patrones por parte del algoritmo, que refuerzan el liderazgo de unas variables frente a otras.

En la tabla 4, puede apreciarse cómo el número de árboles empleados para el análisis es el mismo que en caso anterior (500 árboles) mientras que la ramificación se realiza probando únicamente una variable. La media de error cuadrático medio es mucho mayor que el modelo general por la reducción mayor de las variables explicativas.

6.1.2. Modelo estadístico Max(10)

Al usar el análisis de Random Forest con las 10 principales variables explicativas, tenemos los siguientes resultados [14]:

Números de árboles	No. de variables probadas en cada división	Media de residuos cuadráticos	% Var explicado
<i>No.trees</i>	<i>var.Split</i>	<i>MSE</i>	<i>%Var.Explain.</i>
500	3	1.690655	99.65 %

Tabla 5. Datos principales del análisis Random Forest para el modelo Max(10)

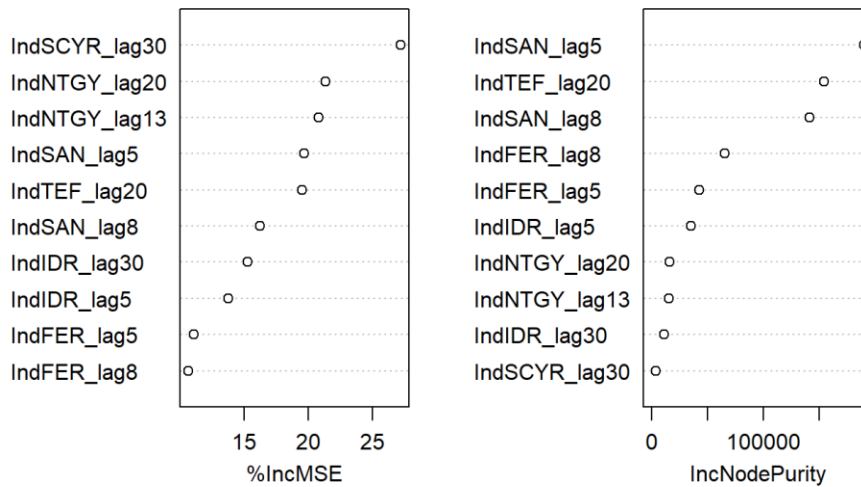


Figura 16. Análisis particular de importancia de variables explicativas mediante Random Forest para el modelo Max(10)
(Fuente: Elaboración propia)

Se puede ver que el orden de la importancia de las variables explicativas ha vuelto a cambiar con respecto a los casos anteriores. Eso ya se ha comentado anteriormente, que al descubrimiento de nuevos patrones en las variables analizadas y unas variables pasan a ostentar mayor importancia para el modelo que otras.

En la tabla 5, el número de árboles es el mismo que en el caso anterior mientras que la ramificación de los árboles se realiza probando tres variables. El error cuadrático medio es mayor que el modelo general pero es mucho mejor que en el caso del modelo Max(5). El porcentaje de varianza explicada es bastante mejor que el modelo anterior.

6.1.3. Modelo estadístico Max(15)

Al usar el análisis de Random Forest con las 15 principales variables explicativas, tenemos los siguientes resultados [14]:

Números de árboles	No. de variables probadas en cada división	Media de residuos cuadráticos	% Var explicado
<i>No.trees</i>	<i>var.Split</i>	<i>MSE</i>	<i>%Var.Explain.</i>
500	5	1.236503	99.74 %

Tabla 6. Datos principales del análisis Random Forest para el modelo Max(15)

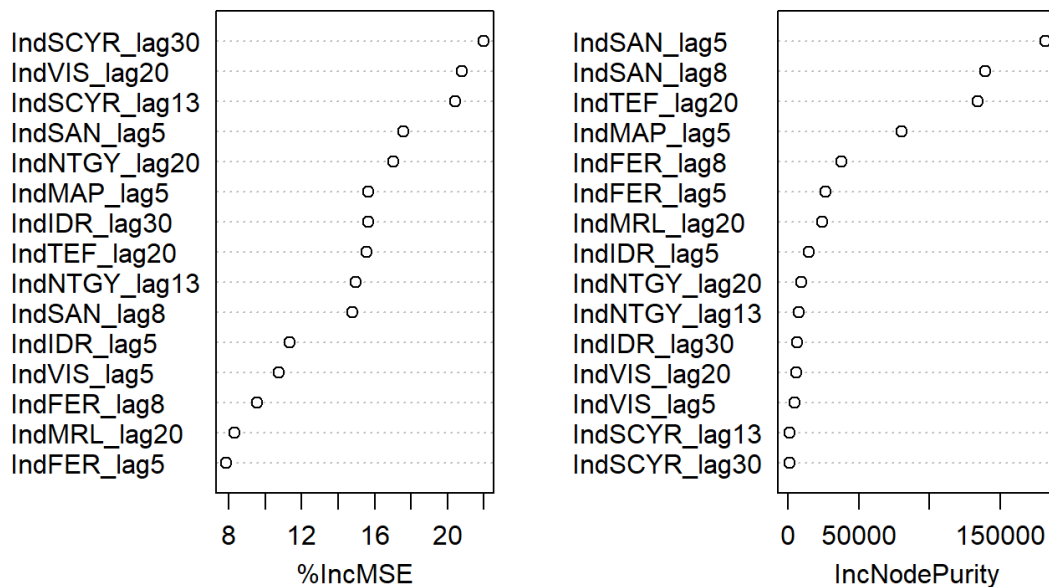


Figura 17. Análisis particular de importancia de variables explicativas mediante Random Forest para el modelo Max(15) (Fuente: Elaboración propia)

Se puede ver que el orden de la importancia de las variables explicativas ha vuelto a cambiar con respecto a los casos anteriores. En la tabla 6, puede apreciarse como el número de árboles empleados para el análisis es el mismo que en caso anterior (500 árboles) mientras que la ramificación se realiza probando 5 variables. El error cuadrático medio es mayor que el modelo general (1.236503vs 0.8877111) pero bastante menor que en el caso del modelo Max(10). El porcentaje de varianza explicada es mejor.

6.2. Aplicar las técnicas de minería de datos

6.2.1. Regresión Lineal

6.2.1.1. Modelo estadístico Max(5)

a. 80% training

Se usa 80% base de datos para el entrenamiento usando la técnica Regresión Lineal [7], tenemos:

```
Call:
lm(formula = IndTEF ~ IndSAN_lag5 + IndNTGY_lag20 + IndSCYR_lag30 +
    IndIDR_lag30 + IndIDR_lag5, data = DB_Training80)

Residuals:
    Min       1Q   Median       3Q      Max
-17.9692  -5.3187   0.1551   4.2916  22.6533

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -3.70117    2.28988  -1.616 0.106307
IndSAN_lag5   1.35601    0.04730  28.665 < 2e-16 ***
IndNTGY_lag20 0.35328    0.09726   3.632 0.000294 ***
IndSCYR_lag30 -7.40386    1.10243  -6.716 2.96e-11 ***
IndIDR_lag30  0.50564    0.29031   1.742 0.081829 .
IndIDR_lag5   1.71720    0.34364   4.997 6.75e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 7.507 on 1120 degrees of freedom
Multiple R-squared:  0.8893, Adjusted R-squared:  0.8888
F-statistic: 1800 on 5 and 1120 DF, p-value: < 2.2e-16
```

El error estándar es 7.507, es la cantidad promedio que la respuesta se desviará de la verdadera línea de regresión. El estadístico $R^2 = 0.8893$ proporciona una medida de ajuste alternativa. Un estadístico R^2 cercano a 1 indica que una gran proporción de la variabilidad en la respuesta se explica por la regresión. Podemos ver en el Figura 18 las dos líneas que reflejan los precios de TEF entre los valores reales y los valores de predicción (con la línea roja).

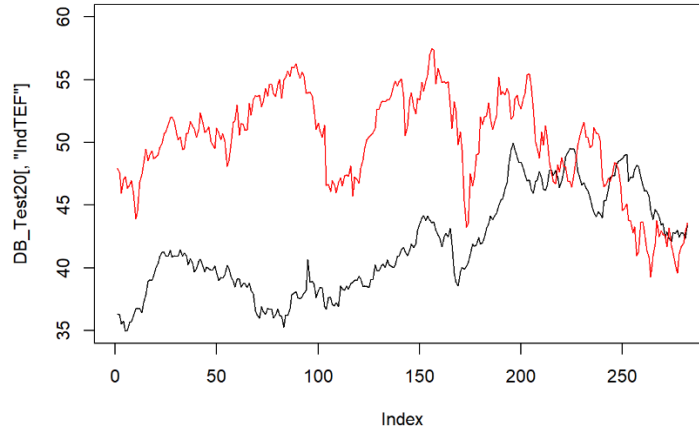


Figura 18. Gráfico de los precios de TEF reales (línea negra) y la predicción (línea roja) del modelo 80% training
 Error medio cuadrático = 103.5175

b. 90% Training

Se usa 90% base de datos para el entrenamiento usando la técnica Regresión Lineal [7].

```
Call:
lm(formula = IndTEF ~ IndSAN_lag5 + IndNTGY_lag20 + IndSCYR_lag30 +
  IndIDR_lag30 + IndIDR_lag5, data = DB_Training90)

Residuals:
    Min       1Q   Median       3Q      Max
-18.9654  -6.6755   0.6273   5.1187  23.9940

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    6.75723    2.19438   3.079 0.002119 **
IndSAN_lag5    1.65627    0.04141  39.998 < 2e-16 ***
IndNTGY_lag20  0.04014    0.09495   0.423 0.672539
IndSCYR_lag30 -8.75708    1.11456  -7.857 8.38e-15 ***
IndIDR_lag30  0.95858    0.27872   3.439 0.000603 ***
IndIDR_lag5   -0.21116    0.31804  -0.664 0.506860
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 7.879 on 1261 degrees of freedom
Multiple R-squared:  0.8788, Adjusted R-squared:  0.8783
F-statistic: 1829 on 5 and 1261 DF, p-value: < 2.2e-16
```

El error estándar es 7.879, es un poco mayor que el caso de 80% training. El estadístico $R^2 = 0.8788$. Podemos ver en el Figura 30 las dos líneas que reflejan los precios de TEF entre los valores reales y los valores de predicción (con la línea roja).

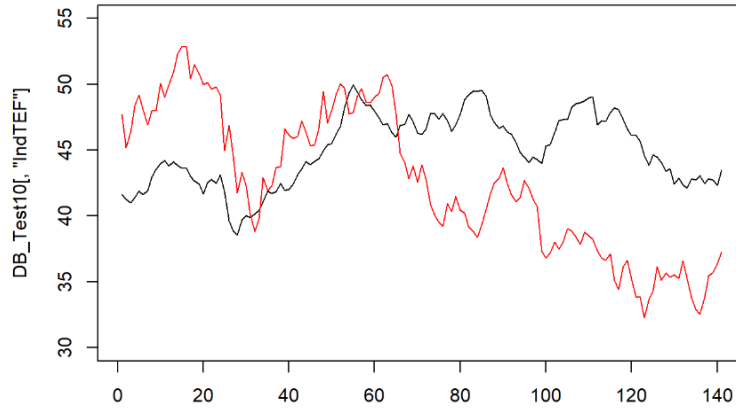


Figura 19. Gráfico de los precios de TEF reales (línea negra) y la predicción (línea roja) del modelo 90% training
 Error medio cuadrático = 47.18739

6.2.1.2. Modelo estadístico Max(10)

a. 80% training

Con las 10 mejores variables, utilizamos un 80% de la base de datos para el entrenamiento usando la técnica Regresión Lineal [7]. Con seguimos los resultados:

```
Call:
lm(formula = IndTEF ~ IndSAN_lag5 + IndNTGY_lag20 + IndSCYR_lag30 +
  IndIDR_lag30 + IndIDR_lag5 + IndSAN_lag8 + IndNTGY_lag13 +
  IndFER_lag5 + IndFER_lag8 + IndTEF_lag20, data = DB_Training80)

Residuals:
    Min       1Q   Median       3Q      Max
-14.9784  -2.4954   0.0006   2.1522  11.4143

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)   13.3963     1.4060   9.528 < 2e-16 ***
IndSAN_lag5    0.8454     0.1018   8.302 2.93e-16 ***
IndNTGY_lag20 -0.9542     0.1512  -6.313 3.95e-10 ***
IndSCYR_lag30 0.3029     0.5593   0.542 0.58814
IndIDR_lag30 -1.6196     0.1591 -10.179 < 2e-16 ***
IndIDR_lag5   0.5512     0.1866   2.953 0.00321 **
IndSAN_lag8   -0.5682     0.1027  -5.534 3.91e-08 ***
IndNTGY_lag13 0.7588     0.1525   4.976 7.52e-07 ***
IndFER_lag5   0.1683     0.2065   0.815 0.41515
IndFER_lag8  -0.4219     0.2073  -2.035 0.04207 *
IndTEF_lag20  0.9005     0.0159  56.623 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.687 on 1115 degrees of freedom
Multiple R-squared:  0.9734, Adjusted R-squared:  0.9732
F-statistic: 4084 on 10 and 1115 DF, p-value: < 2.2e-16
```


El error estándar es 3.687, mucho mejor que el modelo Max(5), y el estadístico $R^2 = 0.9734$ indica que una gran proporción de la variabilidad en la respuesta se explica por la regresión. Podemos ver en el Figura 31 las dos líneas que reflejan los precios de TEF entre los valores reales y los valores de predicción (con la línea roja).

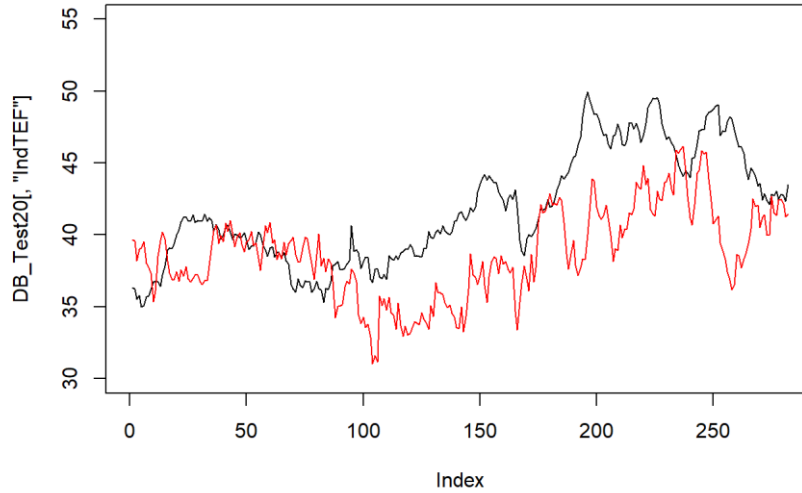


Figura 20. Gráfico de los precios de TEF reales (línea negra) y la predicción (línea roja) del modelo 80% training del Modelo Max(10)
Error medio cuadrático = 21.17588

b. 90% training

Utilizamos un 90% de la base de datos para el entrenamiento, tenemos los resultados:

```
Call:
lm(formula = IndTEF ~ IndSAN_lag5 + IndNTGY_lag20 + IndSCYR_lag30 +
  IndIDR_lag30 + IndIDR_lag5 + IndSAN_lag8 + IndNTGY_lag13 +
  IndFER_lag5 + IndFER_lag8 + IndTEF_lag20, data = DB_Training90)

Residuals:
    Min       1Q   Median       3Q      Max
-15.9009  -2.4399  -0.0755   2.2115  12.0862

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  10.62913   1.23185   8.629 < 2e-16 ***
IndSAN_lag5    0.82019   0.09586   8.556 < 2e-16 ***
IndNTGY_lag20 -0.99622   0.13797  -7.221 8.93e-13 ***
IndSCYR_lag30 -0.16209   0.53007  -0.306  0.7598
IndIDR_lag30  -1.36775   0.13712  -9.975 < 2e-16 ***
IndIDR_lag5    0.83954   0.16339   5.138 3.21e-07 ***
IndSAN_lag8   -0.58417   0.09650  -6.054 1.87e-09 ***
IndNTGY_lag13  0.94318   0.13384   7.047 3.00e-12 ***
IndFER_lag5    0.05124   0.19169   0.267  0.7893
IndFER_lag8   -0.34109   0.19254  -1.772  0.0767 .
IndTEF_lag20  0.86938   0.01362  63.838 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.607 on 1256 degrees of freedom
Multiple R-squared:  0.9747, Adjusted R-squared:  0.9745
F-statistic: 4839 on 10 and 1256 DF, p-value: < 2.2e-16
```

El error estándar es 3.607 y el estadístico $R^2 = 0.9747$, son mejores que el caso anterior. La mayoría de las variables es explicativa. Podemos ver en el Figura 32 las dos líneas que reflejan los precios de TEF entre los valores reales y los valores de predicción (con la línea roja).

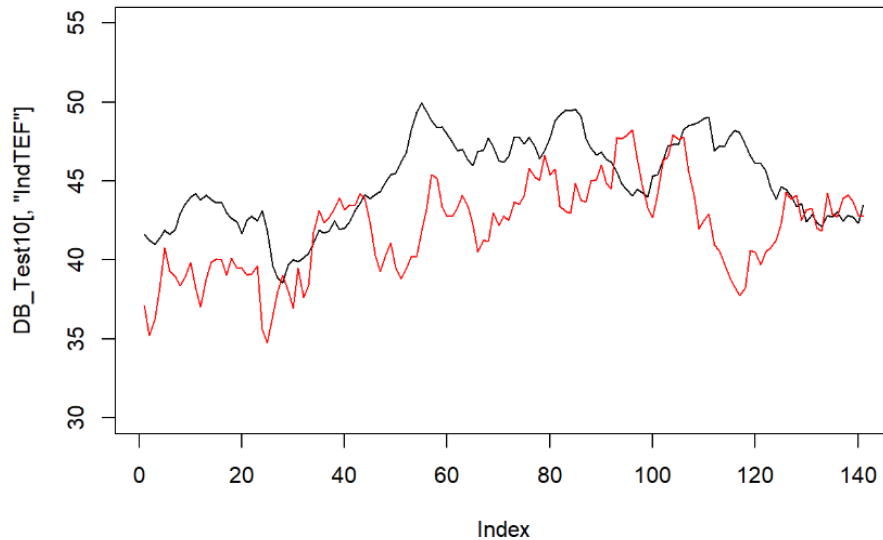


Figura 21. Gráfico de los precios de TEF reales (línea negra) y la predicción (línea roja) del modelo 90% training
Error medio cuadrático = 17.62934

6.2.1.3. Modelo estadístico Max(15)

a. 80% training

Con las 15 mejores variables, utilizamos un 80% de la base de datos para el entrenamiento usando la técnica Regresión Lineal [7]. Con seguimos los resultados:

```

lm(formula = IndTEF ~ IndSAN_lag5 + IndNTGY_lag20 + IndSCYR_lag30 +
  IndIDR_lag30 + IndIDR_lag5 + IndSAN_lag8 + IndNTGY_lag13 +
  IndFER_lag5 + IndFER_lag8 + IndTEF_lag20 + IndVIS_lag5 +
  IndMAP_lag5 + IndSCYR_lag13 + IndMRL_lag20 + IndVIS_lag20,
  data = DB_Training80)

Residuals:
    Min       1Q   Median       3Q      Max
-9.8477 -1.8658  0.0663  1.8532 11.3189

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)   44.78269    2.65119   16.892 < 2e-16 ***
IndSAN_lag5    0.43567    0.08297    5.251 1.81e-07 ***
IndNTGY_lag20 -0.81386    0.12674   -6.422 2.00e-10 ***
IndSCYR_lag30  2.04027    0.56593    3.605 0.000326 ***
IndIDR_lag30  -1.27764    0.14353   -8.902 < 2e-16 ***
IndIDR_lag5    1.16803    0.15736    7.423 2.28e-13 ***
IndSAN_lag8   -0.31690    0.08315   -3.811 0.000146 ***
IndNTGY_lag13  0.55921    0.12514    4.469 8.67e-06 ***
IndFER_lag5   -0.76660    0.16914   -4.532 6.47e-06 ***
IndFER_lag8   -0.14147    0.17185   -0.823 0.410578
IndTEF_lag20  0.52374    0.02041   25.662 < 2e-16 ***
IndVIS_lag5   -0.21071    0.03919   -5.376 9.26e-08 ***
IndMAP_lag5   15.26285    0.79430   19.215 < 2e-16 ***
IndSCYR_lag13 -1.39252    0.69972   -1.990 0.046823 *
IndMRL_lag20  -0.93039    0.12858   -7.236 8.63e-13 ***
IndVIS_lag20  -0.21170    0.03971   -5.331 1.18e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.918 on 1110 degrees of freedom
Multiple R-squared:  0.9834, Adjusted R-squared:  0.9832
F-statistic: 4390 on 15 and 1110 DF,  p-value: < 2.2e-16

```

El error estándar es 2.918, bastante mejor que el modelo Max(10), y el estadístico $R^2 = 0.9834$ indica que una gran proporción de la variabilidad en la respuesta se explica por la regresión. Podemos ver en el Figura 33 las dos líneas que reflejan los precios de TEF entre los valores reales y los valores de predicción (con la línea roja).

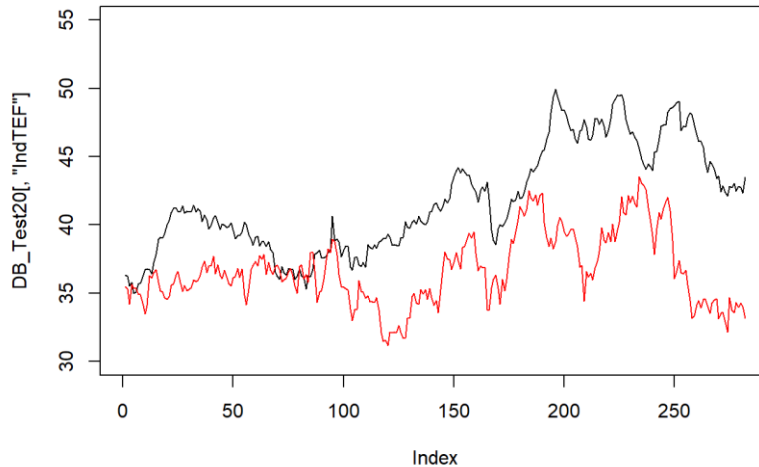


Figura 22. Gráfico de los precios de TEF reales (línea negra) y la predicción (línea roja) del modelo 90% training
Error medio cuadrático = 38.27777

b. 90% training

Utilizamos un 90% de la base de datos para el entrenamiento, tenemos los resultados:

```
lm(formula = IndTEF ~ IndSAN_lag5 + IndNTGY_lag20 + IndSCYR_lag30 +
  IndIDR_lag30 + IndIDR_lag5 + IndSAN_lag8 + IndNTGY_lag13 +
  IndFER_lag5 + IndFER_lag8 + IndTEF_lag20 + IndVIS_lag5 +
  IndMAP_lag5 + IndSCYR_lag13 + IndMRL_lag20 + IndVIS_lag20,
  data = DB_Training90)

Residuals:
    Min       1Q   Median       3Q      Max
-10.1341  -1.7747   0.0177   1.8260  11.1900

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  37.55857    2.50127  15.016 < 2e-16 ***
IndSAN_lag5   0.43834    0.07945   5.517 4.19e-08 ***
IndNTGY_lag20 -0.76473    0.11912  -6.420 1.93e-10 ***
IndSCYR_lag30  2.51123    0.53814   4.666 3.39e-06 ***
IndIDR_lag30 -1.05370    0.12295  -8.571 < 2e-16 ***
IndIDR_lag5   1.36167    0.14264   9.546 < 2e-16 ***
IndSAN_lag8  -0.30572    0.07942  -3.849 0.000124 ***
IndNTGY_lag13  0.77782    0.11175   6.960 5.46e-12 ***
IndFER_lag5  -0.70337    0.15882  -4.429 1.03e-05 ***
IndFER_lag8  -0.09320    0.16196  -0.575 0.565086
IndTEF_lag20  0.52512    0.01980  26.519 < 2e-16 ***
IndVIS_lag5  -0.18063    0.03765  -4.797 1.80e-06 ***
IndMAP_lag5  13.98306    0.75413  18.542 < 2e-16 ***
IndSCYR_lag13 -2.28052    0.65343  -3.490 0.000500 ***
IndMRL_lag20 -1.21108    0.11436 -10.590 < 2e-16 ***
IndVIS_lag20 -0.20914    0.03781  -5.531 3.87e-08 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.912 on 1251 degrees of freedom
Multiple R-squared:  0.9836, Adjusted R-squared:  0.9834
F-statistic: 4993 on 15 and 1251 DF, p-value: < 2.2e-16
```

El error estándar es 2.912 y el estadístico $R^2 = 0.9836$, bastante similares con el caso anterior de 80% training. Podemos ver en el Figura 34 las dos líneas que reflejan los precios de TEF entre los valores reales y los valores de predicción (con la línea roja).

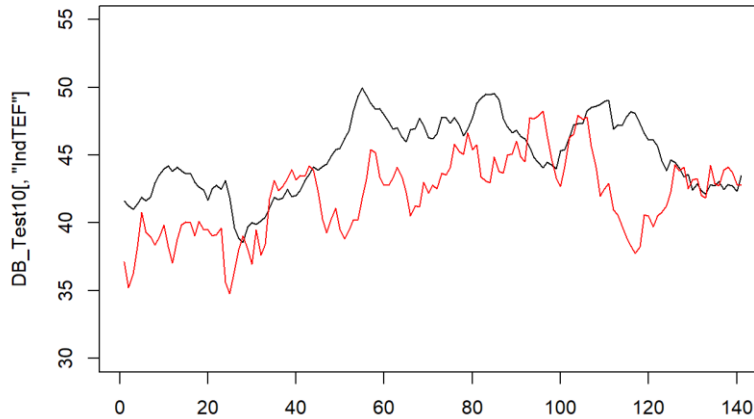


Figura 23. Gráfico de los precios de TEF reales (línea negra) y la predicción (línea roja) del modelo 90% training
Error medio cuadrático = 17.62934

6.2.1.4. Resumen el caso de los modelos de Regresión Lineal

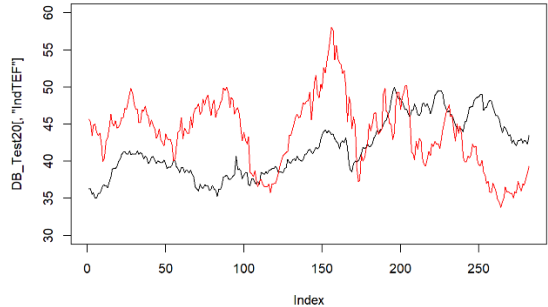
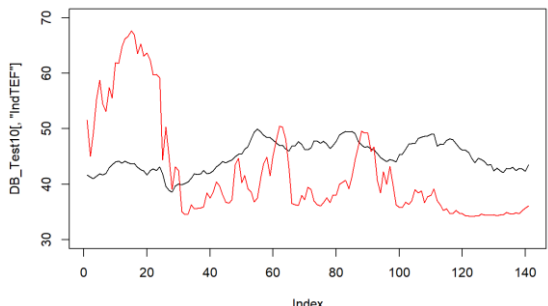
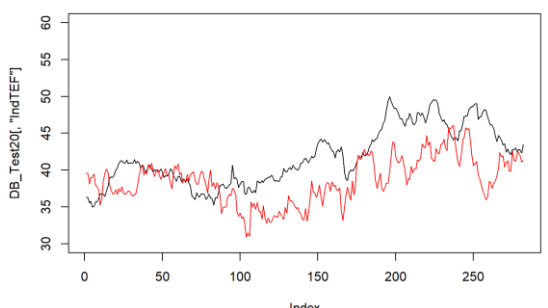
Modelo	El error estándar	R^2	Error medio cuadrático
MAX(5) 80% Training	7.507	0.8893	103.5175
MAX(5) 90% Training	7.879	0.8788	47.18739
MAX(10) 80% Training	3.687	0.9734	21.17588
MAX(10) 90% Training	3.607	0.9747	17.62934
MAX(15) 80% Training	2.918	0.9834	38.27777
MAX(15) 90% Training	2.912	0.9836	17.62934

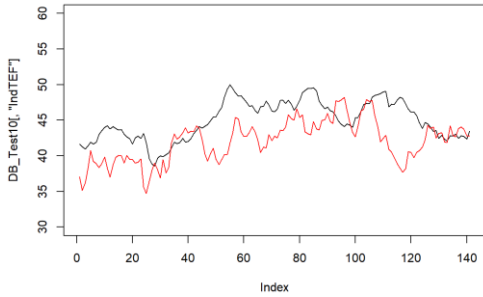
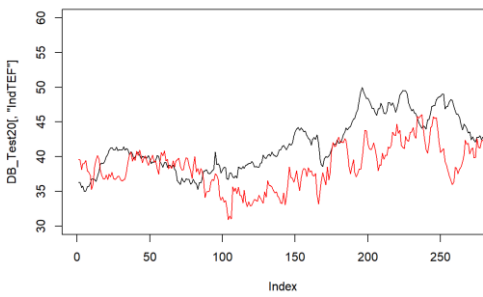
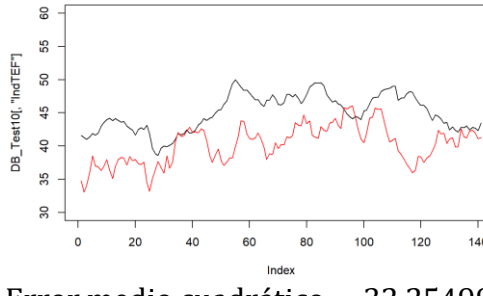
Tabla 7. Resumen el caso de los modelos de Regresión Lineal

Vemos que si aumentamos las variables explicativas, el error estándar va reduciendo pero R^2 y Error medio cuadrático no. Se ve que el modelo Max(10) 90% Training da el mejor resultado.

6.2.2. Redes Neuronales

Utilizando máximo 1000 iteraciones y el tamaño =10. Tenemos resultados:

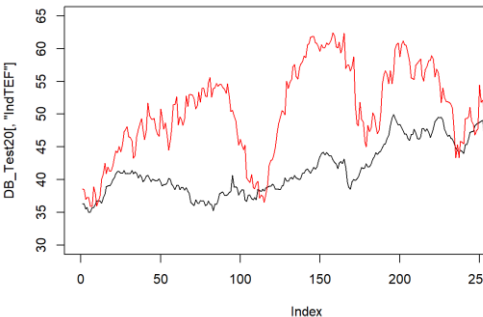
<p>Modelo estadístico Max(5) con 80% training</p>	<pre># weights: 71 initial value 5012616.034196 iter 10 value 530261.268523 iter 20 value 241719.434643 iter 30 value 154438.328404 iter 40 value 135236.826260 iter 50 value 95159.701149 iter 60 value 55107.806275 iter 70 value 52080.840022 final value 52080.833380 converged</pre>	 <p>Error medio cuadrático = 45.56839</p>
<p>Modelo estadístico Max(5) con 90% training</p>	<pre># weights: 71 initial value 5221102.137674 iter 10 value 264873.967944 iter 20 value 125866.866747 iter 30 value 99745.507295 iter 40 value 85228.609445 iter 50 value 68150.851530 iter 60 value 67984.068612 iter 70 value 63003.872807 iter 80 value 61829.319962 iter 90 value 61823.077941 iter 100 value 61546.153988 iter 110 value 61452.574339 iter 120 value 54417.319276 iter 130 value 49201.153281 iter 140 value 41296.094997 final value 41177.190565 converged</pre>	 <p>Error medio cuadrático = 108.6563</p>
<p>Modelo estadístico Max(10) con 80% training</p>	<pre># weights: 121 initial value 5036627.869158 iter 10 value 368334.450553 iter 20 value 171091.206581 iter 30 value 106451.904898 iter 40 value 65871.449188 iter 50 value 25149.215437 iter 60 value 17585.973986 iter 70 value 16049.031383 iter 80 value 15490.607678 iter 90 value 15311.520760 iter 100 value 15237.997320 iter 110 value 15190.912608 iter 120 value 15169.005687 iter 130 value 15156.499430 iter 140 value 15148.466396 iter 150 value 15141.747215 iter 160 value 15137.707568 iter 170 value 15135.322971 iter 180 value 15132.890545 iter 190 value 15131.186856 iter 200 value 15129.727439 iter 210 value 15128.567275 iter 220 value 15127.656938 iter 230 value 15127.157026 iter 240 value 15126.639248 iter 250 value 15126.254675 final value 15126.222322 converged</pre>	 <p>Error medio cuadrático = 22.11952</p>

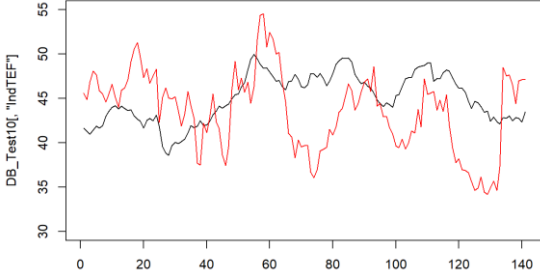
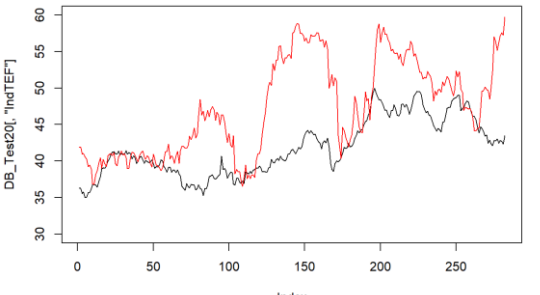
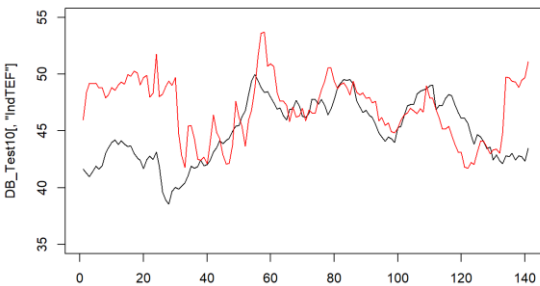
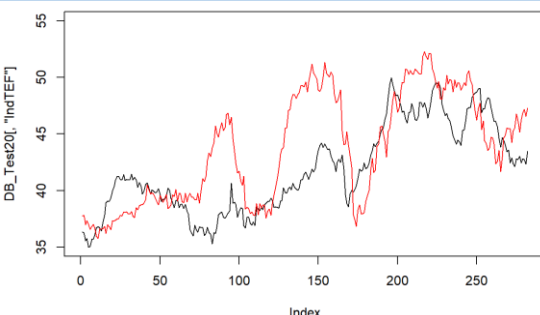
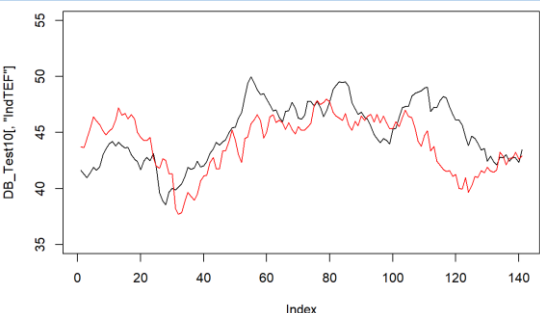
<p>Modelo estadístico Max(10) con 90% training</p>	<pre># weights: 121 initial value 5252400.743182 iter 10 value 368221.835793 iter 20 value 213971.705917 iter 30 value 210844.935907 iter 40 value 210554.075405 iter 50 value 210372.238210 iter 60 value 210072.168410 final value 209989.223683 converged</pre>	 <p>Error medio cuadrático = 55.01254</p>
<p>Modelo estadístico Max(15) con 80% training</p>	<pre># weights: 171 initial value 4987398.165026 iter 10 value 530062.069782 iter 20 value 309112.347856 iter 30 value 206228.617163 iter 40 value 194712.564787 iter 50 value 194547.179381 final value 194546.588882 converged</pre>	 <p>Error medio cuadrático = 22.11952</p>
<p>Modelo estadístico Max(15) con 90% training</p>	<pre># weights: 171 initial value 5191200.925984 final value 645927.197390 converged</pre>	 <p>Error medio cuadrático = 32.25499</p>

[13][16]

Se ve que el modelo Max(10) 80%-Training da el mejor resultado en este caso.

6.2.3. Máquinas de soporte vectorial (SVM)

<p>Modelo estadístico Max(5) con 80% training</p>	<pre>Parameters: SVM-Type: eps-regression SVM-Kernel: radial cost: 1 gamma: 0.2 epsilon: 0.1 Number of Support Vectors: 517</pre>	 <p>Error medio cuadrático = 116.0402</p>
---	--	---

<p>Modelo estadístico Max(5) con 90% training</p>	<p>Parameters: SVM-Type: eps-regression SVM-Kernel: radial cost: 1 gamma: 0.2 epsilon: 0.1</p> <p>Number of Support Vectors: 579</p>	 <p>Error medio cuadrático = 28.83045</p>
<p>Modelo estadístico Max(10) con 80% training</p>	<p>Parameters: SVM-Type: eps-regression SVM-Kernel: radial cost: 1 gamma: 0.1 epsilon: 0.1</p> <p>Number of Support Vectors: 327</p>	 <p>Error medio cuadrático = 56.40968</p>
<p>Modelo estadístico Max(10) con 90% training</p>	<p>Parameters: SVM-Type: eps-regression SVM-Kernel: radial cost: 1 gamma: 0.1 epsilon: 0.1</p> <p>Number of Support Vectors: 357</p>	 <p>Error medio cuadrático = 16.41906</p>
<p>Modelo estadístico Max(15) con 80% training</p>	<p>Parameters: SVM-Type: eps-regression SVM-Kernel: radial cost: 1 gamma: 0.0666667 epsilon: 0.1</p> <p>Number of Support Vectors: 224</p>	 <p>Error medio cuadrático = 16.31971</p>
<p>Modelo estadístico Max(15) con 90% training</p>	<p>Parameters: SVM-Type: eps-regression SVM-Kernel: radial cost: 1 gamma: 0.0666667 epsilon: 0.1</p> <p>Number of Support Vectors: 238</p>	 <p>Error medio cuadrático = 16.41906</p>

Se ve que el modelo Max(10) con 90%-Training y el modelo Max(15) con 90%-Training dan los mejores resultados en este caso. Estos dos modelos predicen muy bien la tendencia del precio de Telefónica.

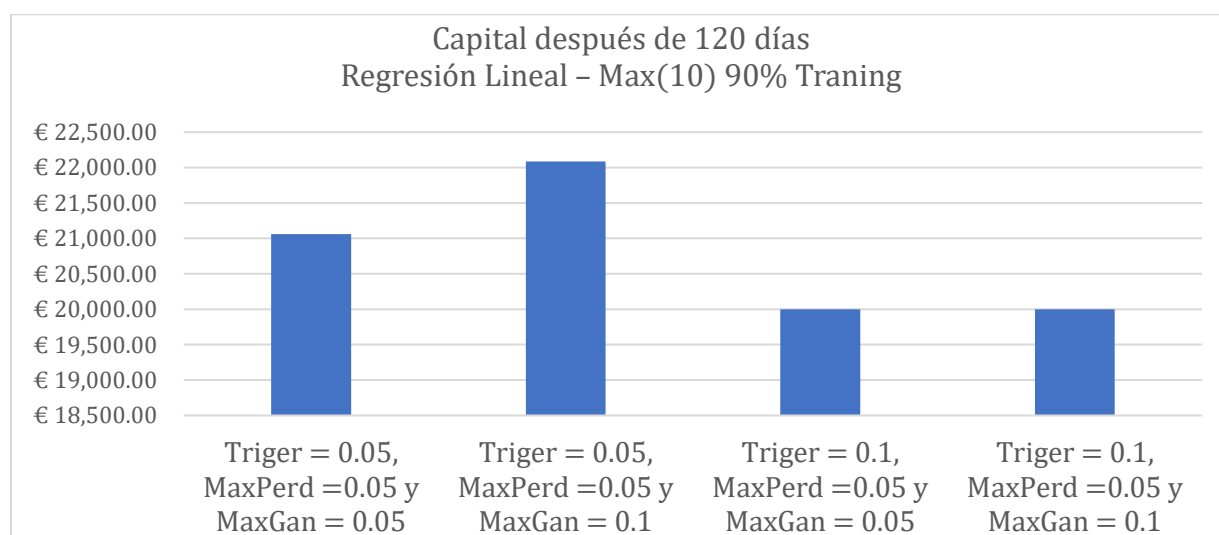
7. EXPERIMENTO COMPUTACIONAL DE SIMULACIÓN

Se realiza como parte del proyecto una simulación determinista consistente en variar una serie de parámetros relacionados tanto con las predicciones como con el sistema de trading para observar el desempeño económico de dichas predicciones. Estos parámetros son los siguientes:

- *Trigger*: Este parámetro hace referencia al límite superior e inferior del indicador a partir del cual se produce una señal de compra/venta de acciones. Puede tomar dos valores diferentes 0.05 y 0.1
- *Pérdida máxima en el sistema de trading (MaxPerd)*: Esta variable hace referencia a la disminución porcentual máxima permitida en el capital invertido que dispara una orden de venta sobre las acciones pertenecientes a esa inversión. Se permiten dos valores distintos, 0,05 o 5%.
- *Ganancia máxima en el sistema de trading (MaxGan)*: Esta variable hace referencia al aumento porcentual máximo permitida en el capital invertido que dispara una orden de venta sobre las acciones pertenecientes a esa inversión. Se permiten dos valores distintos, 0,05 o 5% y 0,1 o 10%.
- *Tiempo de inversión*: 120 días
- *El capital de inversión*: 20.000 €

7.1. Regresión Lineal – Max(10) 90% Training

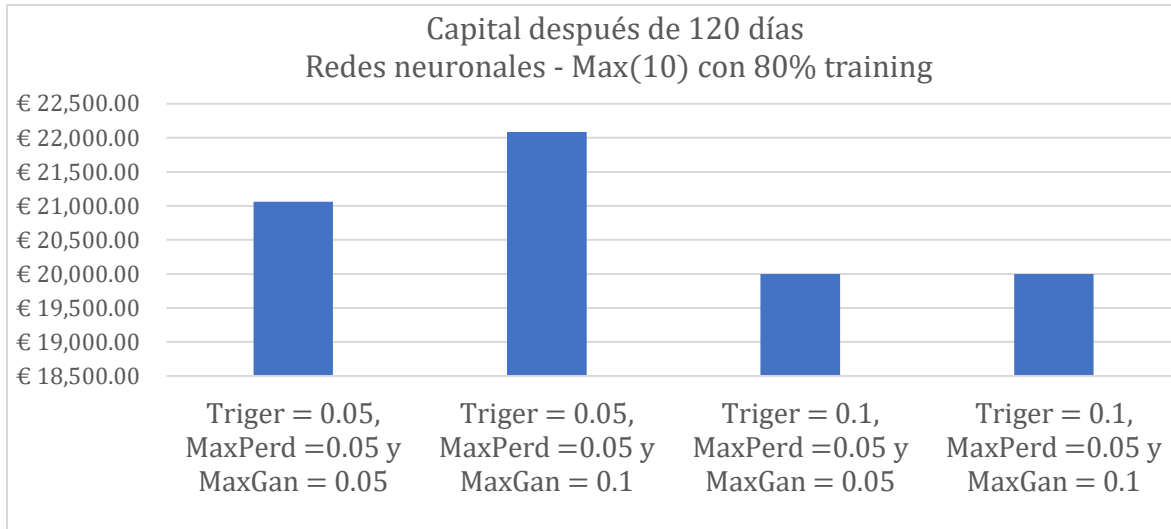
Para facilitar la simulación, se extraen los datos a Excel y se realiza en Excel (Anexo II), conseguimos los resultados después de 120 días de invertir:



Con el Trigger = 0.1, no nos da señal para comprar acciones en este caso.

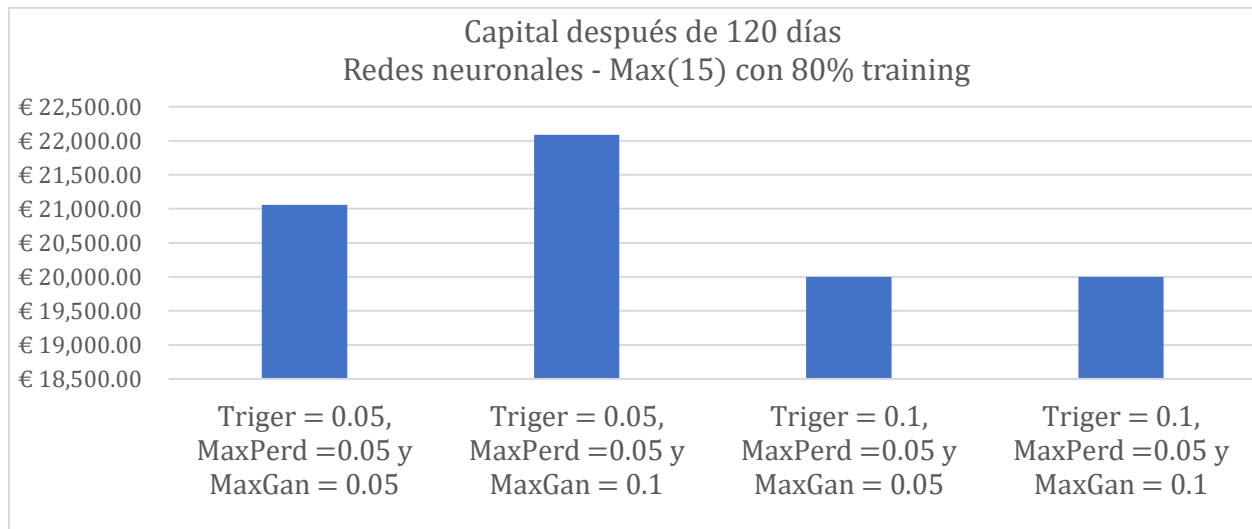
7.2. Redes neuronales - Modelo estadístico Max(10) con 80% training

Simulando el modelo en Excel (Anexo II), conseguimos los resultados después de 120 días de invertir:



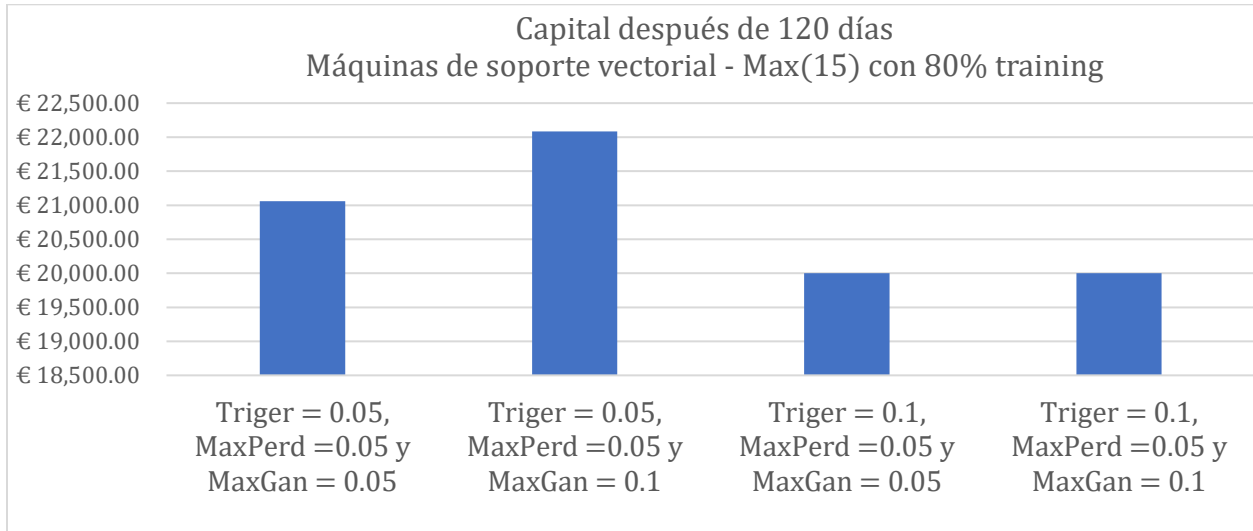
7.3. Redes neuronales - Modelo estadístico Max(15) con 80% training

Simulando el modelo en Excel (Anexo II), conseguimos los resultados después de 120 días de invertir:



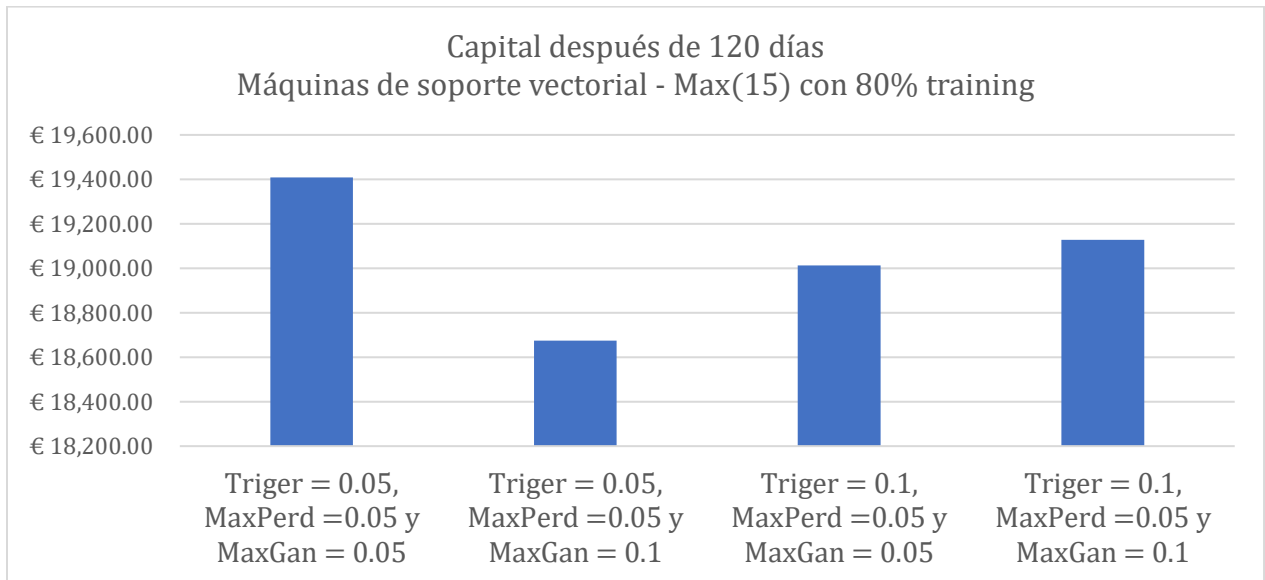
Hemos conseguido el mismo resultado con el caso de Max(10)

7.4. Máquinas de soporte vectorial (SVM) - Max(10) con 90% training



Con el Trigger = 0.1, no nos da señal para comprar acciones en este caso. (Anexo II)

7.5. Máquinas de soporte vectorial (SVM) - Max(15) con 90% training



En este caso, hemos pedido dinero. (Anexo II)

8. CONCLUSIONES

Como conclusión fundamental del proyecto hay que destacar el cumplimiento de los objetivos marcados.

En primer lugar, se han construido diversos modelos de predicción del precio de acciones a través de diversas técnicas de minería de datos y ello ha permitido apreciar un gran potencial en las mismas para su uso en el mundo de la inversión. Con respecto a lo anterior es destacable que, realizando una comparación entre técnicas, las que salen mejor resultados son redes neuronales. Por eso, es necesario un suavizado de esta técnica y probar un mayor número de capas en la red para ver la reacción en las variables repuesta. Por otro lado, la técnica conocida como Random Forest es la más estable de todas.

En segundo lugar, sabemos que es muy difícil predecir el mercado bursátil, estadísticamente hablando, siempre, debería ser respaldado por una estrategia de inversión, de trading, que se ajuste de forma adecuada a este modelo. Como ejemplo, del caso de SVM, que a pesar de presentar una precisión muy elevada no consiguió transformarla en rentabilidades.

En tercer lugar, se ha conseguido construir los modelos de predicción del precio de acciones a partir de herramientas gratuitas como R, también utilizamos Excel para la simulación. Además, el sistema construido muestra gran usabilidad y es capaz de adaptarse al análisis de muchos tipos de datos.

Por último, a partir de este trabajo se abren numerosas líneas de investigación al futuro con el objetivo de lograr máximas rentabilidades y más estables posibles.

9. LÍNERAS FUTURAS

Con el objetivo de mejorar el conocimiento y las posibilidades de aplicación de la técnica minería de datos a la predicción de los precios de acciones, se proponen a continuación una serie de investigación complementarias a este proyecto. Estas líneas están basadas en las limitaciones de fuente de materiales y las dificultades encontradas a la hora de llevar a cabo el trabajo.

En primer lugar, debería tener en cuenta un mayor número de variables, como podrían ser análisis fundamentales de la empresa, el precio de materias primas, datos macroeconómicos de los países, políticas, leyes, ratios bursátiles, noticias, opiniones en foros (text mining), etc. Todo esto cuenta con la dificultad en primer lugar de obtener estos datos y en segundo lugar con el inconveniente de que muchos de estos datos no tienen la misma periodicidad temporal.

Otra posibilidad podría ser incluir varias inversiones de diversas compañías para el sistema para contar con mayores alternativas de inversión.

Sería importante introducir nuevas estrategias de trading para ver cómo se comportan en función del modelo estadístico que haga las predicciones. Con la técnica de minería de datos, nos da los resultados muy bien pero si no es respaldado por una buena estrategia de inversión, no ganamos la rentabilidad a la hora de invertir.

Por último, en este proyecto se ha escogido una pequeña serie de técnicas de minería de datos del amplio conjunto existente. Otro avance relevante sería la inclusión de un mayor número de técnicas, los paquetes modernos para R para el mercado bursátil.

10. BIBLIOGRAFÍA

- [1] Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani, *An Introduction to Statistical Learning – with Applications in R*. Second Edition. 2021.
- [2] Renato Di Lorenzo, *Basic Technical Analysis of Financial Markets*. 2013.
- [3] Pablo Gregori. La presentación de *Redes Neuronales*. Departamento de Matemática. Universidad Jaume de Castellón.
- [4] Trevor H., Robert T., Jerome F., *The Elements of Statistical Learning*. Second Edition. 2017.
- [5] https://es.wikipedia.org/wiki/IBEX_35
- [6] Italo Elola; Natalia Silva, *Análisis de la relación entre el Beta y el Tamaño de las empresas*.
- [7] <https://rdr.io/cran/>
- [8] Alfonso Herrero, Espinosa De Los Monteros, *Análisis de Las Teorías de Inversión en Bolsa*. Universidad Complutense de Madrid.
- [9] <https://es.investing.com/>
- [10] John J. Murphy, *Technical Analysis of The Financial Markets*. 1999.
- [11] Nicolas Heredia García, *Predicción del precio de acciones mediante técnicas de minería de datos*. Universidad Politécnica de Madrid. 2016.
- [12] Breiman, Friedman, Stone, & Olshen, *Classification and Regression Trees*. 1984.
- [13] Brian Ripley, William Venables, NNET package. 2022.
- [14] Andy Liaw, Matthew Wiener, Random Forest package. 2022.
- [15] <https://cran.r-project.org/doc/contrib/YanchangZhao-refcard-data-mining.pdf>
- [16] Marcus W. Beck, Package 'NeuralNetTools'. 2022

ANEXO I. R - Markdown File

Phuong Anh Pham

10/9/2022

1. Leer Base de Datos

```
DB = read.csv(file = "20170213-20220816-DATA_BASE_lag_R.csv")  
head(DB)
```


##	i..No.	Date	Day	Month	Year	IndTEF	TEF_delta_HL	IndTEF_lag5	IndTEF_lag8
##	1	13.02.2017	13	2	2017	91.29	1.19	89.00	88.93
##	2	14.02.2017	14	2	2017	91.96	0.96	88.77	89.91
##	3	15.02.2017	15	2	2017	93.06	1.34	89.10	90.35
##	4	16.02.2017	16	2	2017	93.25	0.99	91.41	89.00
##	5	17.02.2017	17	2	2017	93.24	0.97	90.30	88.77
##	6	20.02.2017	20	2	2017	94.01	0.73	91.29	89.10
##	IndTEF_lag13	IndTEF_lag20	IndTEF_lag30	IndBBVA_lag5	IndBBVA_lag8				
##	1	91.63	90.92	88.50	6.009	6.075			
##	2	91.20	90.86	89.36	5.997	6.016			
##	3	90.54	91.18	89.53	5.922	6.070			
##	4	89.54	90.61	91.43	5.943	6.009			
##	5	89.25	90.86	91.95	5.864	5.997			
##	6	88.93	89.64	91.76	5.950	5.922			
##	IndBBVA_lag13	IndBBVA_lag20	IndBBVA_lag30	IndBKT_lag5	IndBKT_lag8				
##	1	6.348	6.028	6.323	7.406	7.549			
##	2	6.296	6.018	6.446	7.244	7.546			
##	3	6.268	6.025	6.377	7.226	7.565			
##	4	6.198	6.006	6.399	7.284	7.406			
##	5	6.161	6.011	6.447	7.260	7.244			
##	6	6.075	5.975	6.418	7.310	7.226			
##	IndBKT_lag13	IndBKT_lag20	IndBKT_lag30	IndCABK_lag5	IndCABK_lag8				
##	1	7.671	7.499	7.401	3.46	3.40			
##	2	7.564	7.525	7.542	3.40	3.50			
##	3	7.531	7.619	7.539	3.36	3.54			
##	4	7.496	7.615	7.543	3.39	3.46			
##	5	7.435	7.631	7.571	3.33	3.40			
##	6	7.549	7.518	7.500	3.36	3.36			
##	IndCABK_lag13	IndCABK_lag20	IndCABK_lag30	IndSAN_lag5	IndSAN_lag8				
##	1	3.56	3.29	3.19	50.74	50.59			
##	2	3.50	3.25	3.29	49.73	51.39			
##	3	3.50	3.25	3.27	49.04	51.84			
##	4	3.40	3.32	3.31	49.82	50.74			
##	5	3.38	3.32	3.31	48.98	49.73			
##	6	3.40	3.29	3.29	49.86	49.04			
##	IndSAN_lag13	IndSAN_lag20	IndSAN_lag30	IndIDR_lag5	IndIDR_lag8	IndIDR_lag13			
##	1	52.19	49.22	48.49	10.440	10.180	10.430		
##	2	52.10	48.99	49.66	10.450	10.345	10.350		
##	3	52.05	49.16	49.89	10.545	10.580	10.385		
##	4	50.56	49.43	50.03	10.710	10.440	10.240		
##	5	50.03	49.88	50.31	10.725	10.450	10.185		
##	6	50.59	49.07	49.75	10.840	10.545	10.180		
##	IndIDR_lag20	IndIDR_lag30	IndAMA_lag5	IndAMA_lag8	IndAMA_lag13	IndAMA_lag20			
##	1	10.650	10.545	43.595	42.580	43.270	43.520		
##	2	10.695	10.610	43.210	43.395	43.870	43.250		
##	3	10.750	10.455	43.230	43.810	43.770	43.635		
##	4	10.650	10.570	43.100	43.595	42.805	43.565		
##	5	10.680	10.630	42.920	43.210	42.745	43.130		
##	6	10.450	10.630	43.070	43.230	42.580	43.175		
##	IndAMA_lag30	IndELE_lag5	IndELE_lag8	IndELE_lag13	IndELE_lag20	IndELE_lag30			
##	1	43.430	19.205	18.950	19.215	19.725	20.225		
##	2	43.505	19.190	19.185	19.345	19.650	20.200		

## 3	43.505	19.535	19.350	19.355	19.570	19.950
## 4	43.875	19.690	19.205	19.090	19.555	19.980
## 5	43.620	19.700	19.190	19.045	19.400	20.045
## 6	43.715	19.845	19.535	18.950	19.210	20.030
##	IndNTGY_lag5	IndNTGY_lag8	IndNTGY_lag13	IndNTGY_lag20	IndNTGY_lag30	
## 1	18.200	17.605	17.525	17.735	17.920	
## 2	18.345	18.065	17.670	17.790	17.855	
## 3	18.090	18.315	17.880	17.760	17.640	
## 4	18.240	18.200	17.830	17.680	17.605	
## 5	18.055	18.345	17.820	17.610	17.695	
## 6	18.195	18.090	17.605	17.440	17.550	
##	IndREP_lag5	IndREP_lag8	IndREP_lag13	IndREP_lag20	IndREP_lag30	IndITX_lag5
## 1	12.628	12.455	12.988	12.856	12.327	30.565
## 2	12.446	12.533	12.943	12.856	12.423	30.885
## 3	12.341	12.687	12.842	12.879	12.450	30.860
## 4	12.505	12.628	12.546	12.851	12.469	31.120
## 5	12.469	12.446	12.450	12.938	12.487	31.140
## 6	12.715	12.341	12.455	12.760	12.368	31.335
##	IndITX_lag8	IndITX_lag13	IndITX_lag20	IndITX_lag30	IndANA_lag5	IndANA_lag8
## 1	30.530	31.120	31.585	32.600	72.50	72.13
## 2	30.535	30.840	31.395	32.275	72.55	73.22
## 3	30.725	30.850	31.350	31.970	73.02	73.04
## 4	30.565	30.645	31.225	31.450	73.22	72.50
## 5	30.885	30.535	31.000	31.675	73.10	72.55
## 6	30.860	30.530	31.065	31.865	73.65	73.02
##	IndANA_lag13	IndANA_lag20	IndANA_lag30	IndACS_lag5	IndACS_lag8	IndACS_lag13
## 1	73.55	72.45	70.72	27.60	27.60	28.44
## 2	73.46	74.09	70.54	27.59	27.68	27.97
## 3	72.57	73.78	70.20	28.36	27.89	28.03
## 4	72.07	74.30	70.00	28.59	27.60	27.41
## 5	71.68	73.03	70.05	28.74	27.59	27.51
## 6	72.13	72.68	71.20	29.03	28.36	27.60
##	IndACS_lag20	IndACS_lag30	IndIBE_lag5	IndIBE_lag8	IndIBE_lag13	IndIBE_lag20
## 1	28.98	29.49	5.346	5.306	5.445	5.528
## 2	28.87	29.55	5.402	5.358	5.419	5.528
## 3	28.52	29.12	5.504	5.369	5.390	5.495
## 4	28.30	29.09	5.545	5.346	5.318	5.452
## 5	27.79	29.13	5.497	5.402	5.348	5.507
## 6	27.73	28.92	5.569	5.504	5.306	5.442
##	IndIBE_lag30	IndFCC_lag5	IndFCC_lag8	IndFCC_lag13	IndFCC_lag20	IndFCC_lag30
## 1	5.636	7.900	7.676	7.885	8.016	7.750
## 2	5.596	8.060	7.967	7.740	8.089	7.898
## 3	5.550	8.109	8.182	7.675	8.046	7.782
## 4	5.518	8.270	7.900	7.680	7.990	7.853
## 5	5.503	8.410	8.060	7.704	7.964	7.917
## 6	5.525	8.500	8.109	7.676	7.743	7.850
##	IndFER_lag5	IndFER_lag8	IndFER_lag13	IndFER_lag20	IndFER_lag30	IndTL5_lag5
## 1	16.04	15.86	16.28	16.85	16.29	11.261
## 2	15.97	15.96	16.09	16.78	16.74	11.291
## 3	16.20	16.10	16.09	16.59	16.65	11.281
## 4	16.44	16.04	15.97	16.35	16.85	11.425
## 5	16.39	15.97	15.85	16.23	16.95	11.365

## 6	16.48	16.20	15.86	16.19	17.11	11.375
##	IndTL5_lag8	IndTL5_lag13	IndTL5_lag20	IndTL5_lag30	IndREDE_lag5	IndREDE_lag8
## 1	11.316	11.207	11.460	11.152	16.660	16.400
## 2	11.346	11.291	11.395	11.117	16.865	16.595
## 3	11.385	11.291	11.256	11.102	17.175	16.635
## 4	11.261	11.306	11.097	11.167	17.270	16.660
## 5	11.291	11.236	11.102	11.187	17.170	16.865
## 6	11.281	11.316	11.087	11.072	17.315	17.175
##	IndREDE_lag13	IndREDE_lag20	IndREDE_lag30	IndSCYR_lag5	IndSCYR_lag8	
## 1	16.820	17.250	17.950	2.159	2.196	
## 2	16.700	17.340	17.570	2.136	2.199	
## 3	16.660	17.255	17.200	2.147	2.207	
## 4	16.535	17.100	17.345	2.181	2.159	
## 5	16.530	17.000	17.200	2.172	2.136	
## 6	16.400	16.850	17.275	2.227	2.147	
##	IndSCYR_lag13	IndSCYR_lag20	IndSCYR_lag30	IndMAP_lag5	IndMAP_lag8	
## 1	2.283	2.121	2.138	2.880	2.860	
## 2	2.272	2.125	2.145	2.886	2.858	
## 3	2.263	2.133	2.136	2.838	2.898	
## 4	2.208	2.157	2.136	2.877	2.880	
## 5	2.181	2.181	2.138	2.884	2.886	
## 6	2.196	2.205	2.145	2.910	2.838	
##	IndMAP_lag13	IndMAP_lag20	IndMAP_lag30	IndGRLS_lag5	IndGRLS_lag8	
## 1	2.872	2.853	2.900	20.320	19.875	
## 2	2.903	2.822	2.957	20.550	20.000	
## 3	2.867	2.804	2.943	20.375	20.185	
## 4	2.820	2.746	2.955	20.745	20.320	
## 5	2.800	2.737	2.960	20.900	20.550	
## 6	2.860	2.720	2.919	20.880	20.375	
##	IndGRLS_lag13	IndGRLS_lag20	IndGRLS_lag30	IndICAG_lag5	IndICAG_lag8	
## 1	19.570	19.210	19.045	5.504	5.724	
## 2	19.880	19.180	19.040	5.519	5.607	
## 3	19.960	19.070	19.600	5.629	5.616	
## 4	19.810	19.800	19.600	5.743	5.504	
## 5	19.820	19.415	19.565	5.708	5.519	
## 6	19.875	19.300	19.480	5.833	5.629	
##	IndICAG_lag13	IndICAG_lag20	IndICAG_lag30	IndVIS_lag5	IndVIS_lag8	
## 1	5.831	5.484	5.257	47.540	46.870	
## 2	5.793	5.691	5.252	48.195	47.285	
## 3	5.839	5.645	5.279	48.575	47.515	
## 4	5.693	5.684	5.484	48.580	47.540	
## 5	5.55	5.701	5.443	48.345	48.195	
## 6	5.724	5.714	5.389	48.600	48.575	
##	IndVIS_lag13	IndVIS_lag20	IndVIS_lag30	IndACX_lag5	IndACX_lag8	IndACX_lag13
## 1	47.625	47.950	47.535	12.83	12.34	12.45
## 2	47.020	47.885	49.870	12.91	12.89	12.34
## 3	47.410	47.945	49.420	12.93	12.89	12.37
## 4	47.070	47.885	48.855	12.92	12.83	12.18
## 5	47.015	47.695	49.490	13.09	12.91	12.22
## 6	46.870	47.400	48.960	13.25	12.93	12.34
##	IndACX_lag20	IndACX_lag30	IndAENA_lag5	IndAENA_lag8	IndAENA_lag13	
## 1	12.17	12.15	132.80	133.60	133.85	

```
## 2      11.97      12.29      133.00      133.75      135.35
## 3      12.11      12.24      133.50      133.35      137.70
## 4      12.13      12.23      134.05      132.80      135.75
## 5      12.16      12.21      134.25      133.00      134.35
## 6      12.14      12.12      134.85      133.50      133.60
##  IndAENA_lag20 IndAENA_lag30 IndMRL_lag5 IndMRL_lag8 IndMRL_lag13 IndMRL_lag20
## 1      136.50      130.85      10.420      10.335      10.345      10.295
## 2      135.50      131.20      10.430      10.235      10.375      10.195
## 3      135.00      129.70      10.485      10.415      10.350      10.165
## 4      134.45      131.60      10.650      10.420      10.380      10.010
## 5      134.10      131.00      10.625      10.430      10.355      10.005
## 6      134.30      132.50      10.630      10.485      10.335      10.115
##  IndMRL_lag30
## 1      10.475
## 2      10.310
## 3      10.375
## 4      10.360
## 5      10.510
## 6      10.460
```

```
# Coger solo los datos del precio de acciones
DB_ind = DB[,c(6,11:137)]
head(DB_ind,3)
```

##	IndTEF	IndTEF_lag20	IndTEF_lag30	IndBBVA_lag5	IndBBVA_lag8	IndBBVA_lag13
## 1	91.29	90.92	88.50	6.009	6.075	6.348
## 2	91.96	90.86	89.36	5.997	6.016	6.296
## 3	93.06	91.18	89.53	5.922	6.070	6.268
##	IndBBVA_lag20	IndBBVA_lag30	IndBKT_lag5	IndBKT_lag8	IndBKT_lag13	IndBKT_lag20
## 1	6.028	6.323	7.406	7.549	7.671	7.499
## 2	6.018	6.446	7.244	7.546	7.564	7.525
## 3	6.025	6.377	7.226	7.565	7.531	7.619
##	IndBKT_lag30	IndCABK_lag5	IndCABK_lag8	IndCABK_lag13	IndCABK_lag20	
## 1	7.401	3.46	3.40	3.56	3.29	
## 2	7.542	3.40	3.50	3.50	3.25	
## 3	7.539	3.36	3.54	3.50	3.25	
##	IndCABK_lag30	IndSAN_lag5	IndSAN_lag8	IndSAN_lag13	IndSAN_lag20	IndSAN_lag30
## 1	3.19	50.74	50.59	52.19	49.22	48.49
## 2	3.29	49.73	51.39	52.10	48.99	49.66
## 3	3.27	49.04	51.84	52.05	49.16	49.89
##	IndIDR_lag5	IndIDR_lag8	IndIDR_lag13	IndIDR_lag20	IndIDR_lag30	IndAMA_lag5
## 1	10.440	10.180	10.430	10.650	10.545	43.595
## 2	10.450	10.345	10.350	10.695	10.610	43.210
## 3	10.545	10.580	10.385	10.750	10.455	43.230
##	IndAMA_lag8	IndAMA_lag13	IndAMA_lag20	IndAMA_lag30	IndELE_lag5	IndELE_lag8
## 1	42.580	43.27	43.520	43.430	19.205	18.950
## 2	43.395	43.87	43.250	43.505	19.190	19.185
## 3	43.810	43.77	43.635	43.505	19.535	19.350
##	IndELE_lag13	IndELE_lag20	IndELE_lag30	IndNTGY_lag5	IndNTGY_lag8	
## 1	19.215	19.725	20.225	18.200	17.605	
## 2	19.345	19.650	20.200	18.345	18.065	
## 3	19.355	19.570	19.950	18.090	18.315	
##	IndNTGY_lag13	IndNTGY_lag20	IndNTGY_lag30	IndREP_lag5	IndREP_lag8	
## 1	17.525	17.735	17.920	12.628	12.455	
## 2	17.670	17.790	17.855	12.446	12.533	
## 3	17.880	17.760	17.640	12.341	12.687	
##	IndREP_lag13	IndREP_lag20	IndREP_lag30	IndITX_lag5	IndITX_lag8	IndITX_lag13
## 1	12.988	12.856	12.327	30.565	30.530	31.12
## 2	12.943	12.856	12.423	30.885	30.535	30.84
## 3	12.842	12.879	12.450	30.860	30.725	30.85
##	IndITX_lag20	IndITX_lag30	IndANA_lag5	IndANA_lag8	IndANA_lag13	IndANA_lag20
## 1	31.585	32.600	72.50	72.13	73.55	72.45
## 2	31.395	32.275	72.55	73.22	73.46	74.09
## 3	31.350	31.970	73.02	73.04	72.57	73.78
##	IndANA_lag30	IndACS_lag5	IndACS_lag8	IndACS_lag13	IndACS_lag20	IndACS_lag30
## 1	70.72	27.60	27.60	28.44	28.98	29.49
## 2	70.54	27.59	27.68	27.97	28.87	29.55
## 3	70.20	28.36	27.89	28.03	28.52	29.12
##	IndIBE_lag5	IndIBE_lag8	IndIBE_lag13	IndIBE_lag20	IndIBE_lag30	IndFCC_lag5
## 1	5.346	5.306	5.445	5.528	5.636	7.900
## 2	5.402	5.358	5.419	5.528	5.596	8.060
## 3	5.504	5.369	5.390	5.495	5.550	8.109
##	IndFCC_lag8	IndFCC_lag13	IndFCC_lag20	IndFCC_lag30	IndFER_lag5	IndFER_lag8
## 1	7.676	7.885	8.016	7.750	16.04	15.86
## 2	7.967	7.740	8.089	7.898	15.97	15.96
## 3	8.182	7.675	8.046	7.782	16.20	16.10

```

## IndFER_lag13 IndFER_lag20 IndFER_lag30 IndTL5_lag5 IndTL5_lag8 IndTL5_lag13
## 1 16.28 16.85 16.29 11.261 11.316 11.207
## 2 16.09 16.78 16.74 11.291 11.346 11.291
## 3 16.09 16.59 16.65 11.281 11.385 11.291
## IndTL5_lag20 IndTL5_lag30 IndREDE_lag5 IndREDE_lag8 IndREDE_lag13
## 1 11.460 11.152 16.660 16.400 16.82
## 2 11.395 11.117 16.865 16.595 16.70
## 3 11.256 11.102 17.175 16.635 16.66
## IndREDE_lag20 IndREDE_lag30 IndSCYR_lag5 IndSCYR_lag8 IndSCYR_lag13
## 1 17.250 17.95 2.159 2.196 2.283
## 2 17.340 17.57 2.136 2.199 2.272
## 3 17.255 17.20 2.147 2.207 2.263
## IndSCYR_lag20 IndSCYR_lag30 IndMAP_lag5 IndMAP_lag8 IndMAP_lag13 IndMAP_lag20
## 1 2.121 2.138 2.880 2.860 2.872 2.853
## 2 2.125 2.145 2.886 2.858 2.903 2.822
## 3 2.133 2.136 2.838 2.898 2.867 2.804
## IndMAP_lag30 IndGRLS_lag5 IndGRLS_lag8 IndGRLS_lag13 IndGRLS_lag20
## 1 2.900 20.320 19.875 19.57 19.21
## 2 2.957 20.550 20.000 19.88 19.18
## 3 2.943 20.375 20.185 19.96 19.07
## IndGRLS_lag30 IndICAG_lag5 IndICAG_lag8 IndICAG_lag13 IndICAG_lag20
## 1 19.045 5.504 5.724 5.831 5.484
## 2 19.040 5.519 5.607 5.793 5.691
## 3 19.600 5.629 5.616 5.839 5.645
## IndICAG_lag30 IndVIS_lag5 IndVIS_lag8 IndVIS_lag13 IndVIS_lag20 IndVIS_lag30
## 1 5.257 47.540 46.870 47.625 47.950 47.535
## 2 5.252 48.195 47.285 47.020 47.885 49.870
## 3 5.279 48.575 47.515 47.410 47.945 49.420
## IndACX_lag5 IndACX_lag8 IndACX_lag13 IndACX_lag20 IndACX_lag30 IndAENA_lag5
## 1 12.83 12.34 12.45 12.17 12.15 132.8
## 2 12.91 12.89 12.34 11.97 12.29 133.0
## 3 12.93 12.89 12.37 12.11 12.24 133.5
## IndAENA_lag8 IndAENA_lag13 IndAENA_lag20 IndAENA_lag30 IndMRL_lag5
## 1 133.60 133.85 136.5 130.85 10.420
## 2 133.75 135.35 135.5 131.20 10.430
## 3 133.35 137.70 135.0 129.70 10.485
## IndMRL_lag8 IndMRL_lag13 IndMRL_lag20 IndMRL_lag30
## 1 10.335 10.345 10.295 10.475
## 2 10.235 10.375 10.195 10.310
## 3 10.415 10.350 10.165 10.375

```

2. Analizar los datos

1. Random Forest

```
library (randomForest)
```

```
## randomForest 4.6-14
```

```
## Type rfNews() to see new features/changes/bug fixes.
```

```
set.seed (1)
rf.IndTEF <- randomForest(IndTEF ~ ., data = DB_ind,importance = TRUE)
```

```
rf.IndTEF
```

```
##
## Call:
## randomForest(formula = IndTEF ~ ., data = DB_ind, importance = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 42
##
##           Mean of squared residuals: 0.8877111
##           % Var explained: 99.82
```

```
importance(rf.IndTEF)
```

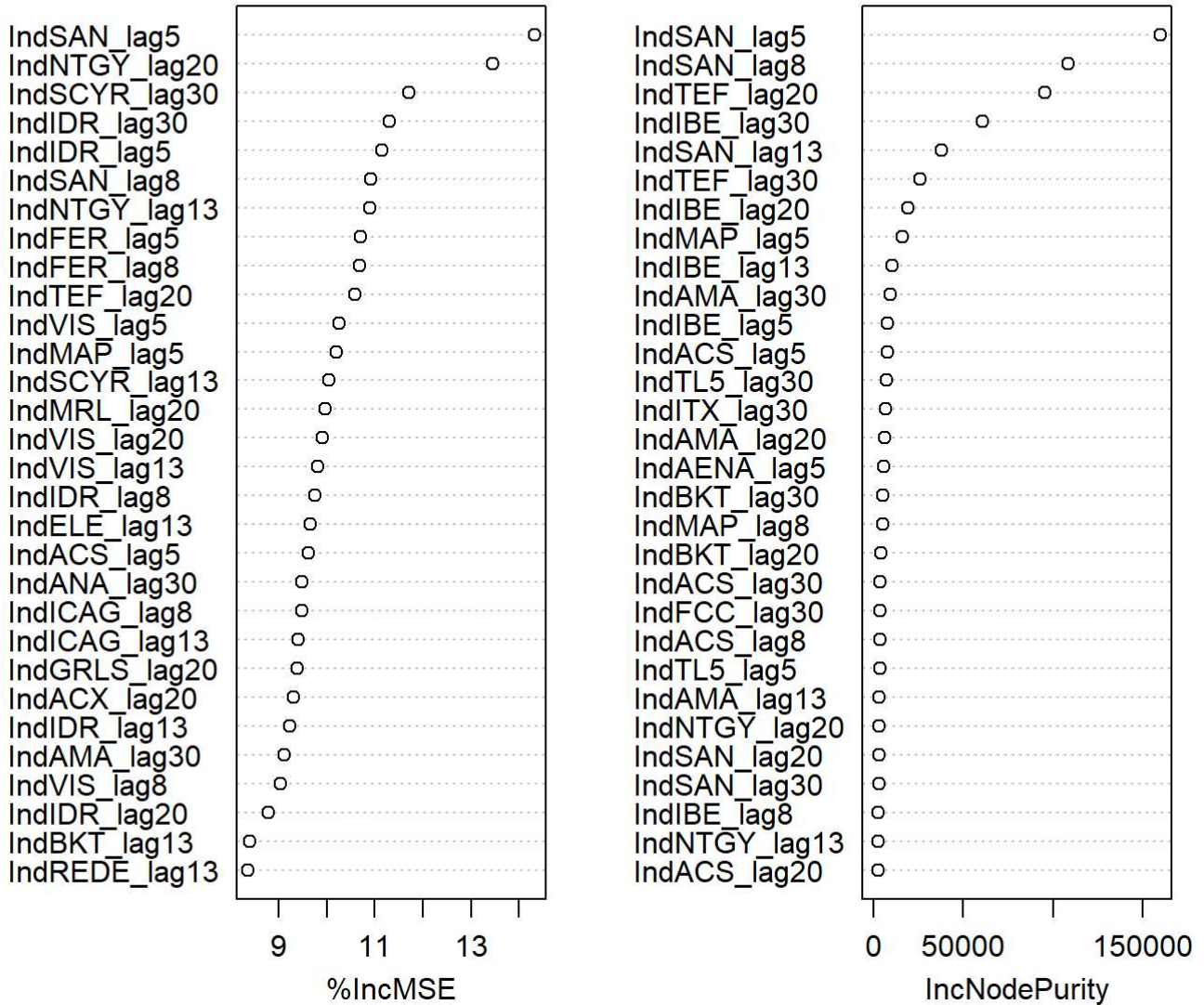
##	%IncMSE	IncNodePurity
## IndTEF_lag20	10.583222	95558.61148
## IndTEF_lag30	6.575947	25881.22851
## IndBBVA_lag5	5.716756	435.40838
## IndBBVA_lag8	6.179506	237.54609
## IndBBVA_lag13	6.788205	162.52673
## IndBBVA_lag20	2.730273	31.87549
## IndBBVA_lag30	6.467374	388.62610
## IndBKT_lag5	8.108913	127.94969
## IndBKT_lag8	1.237202	1132.34760
## IndBKT_lag13	8.389574	119.01633
## IndBKT_lag20	2.026364	4186.53259
## IndBKT_lag30	2.280537	5103.15218
## IndCABK_lag5	6.636856	694.50214
## IndCABK_lag8	6.020974	408.89684
## IndCABK_lag13	6.031852	326.51402
## IndCABK_lag20	5.630213	72.82247
## IndCABK_lag30	7.394441	192.49284
## IndSAN_lag5	14.323059	159721.29043
## IndSAN_lag8	10.905371	108220.37499
## IndSAN_lag13	5.949847	37700.61350
## IndSAN_lag20	1.705365	3126.03054
## IndSAN_lag30	1.780020	3093.98171
## IndIDR_lag5	11.151957	1408.58332
## IndIDR_lag8	9.746361	1304.63782
## IndIDR_lag13	9.236084	701.98405
## IndIDR_lag20	8.777470	437.11978
## IndIDR_lag30	11.301337	816.14409
## IndAMA_lag5	6.117479	2114.81718
## IndAMA_lag8	5.080460	2247.07776
## IndAMA_lag13	5.091088	3177.10069
## IndAMA_lag20	7.665101	5970.07400
## IndAMA_lag30	9.115643	9050.62071
## IndELE_lag5	3.755278	120.43067
## IndELE_lag8	7.360539	76.27809
## IndELE_lag13	9.653815	134.71936
## IndELE_lag20	4.505521	70.29828
## IndELE_lag30	3.929369	97.56549
## IndNTGY_lag5	4.540969	340.51562
## IndNTGY_lag8	8.083294	1345.20397
## IndNTGY_lag13	10.890464	2539.36795
## IndNTGY_lag20	13.458005	3142.71536
## IndNTGY_lag30	8.055158	879.80955
## IndREP_lag5	7.075206	1431.01543
## IndREP_lag8	4.888709	1319.86637
## IndREP_lag13	4.985348	1480.90750
## IndREP_lag20	4.959411	293.84125
## IndREP_lag30	6.384694	176.44652
## IndITX_lag5	2.834875	586.92925
## IndITX_lag8	3.103532	683.29895
## IndITX_lag13	4.126563	1380.41741
## IndITX_lag20	4.481803	2051.26465

## IndITX_lag30	7.407952	6653.46029
## IndANA_lag5	4.250965	264.03203
## IndANA_lag8	3.827282	169.61425
## IndANA_lag13	4.623324	176.09865
## IndANA_lag20	5.822266	123.73550
## IndANA_lag30	9.475327	562.14395
## IndACS_lag5	9.614963	7729.95101
## IndACS_lag8	6.222604	3432.17574
## IndACS_lag13	3.137202	877.91789
## IndACS_lag20	5.167839	2522.08642
## IndACS_lag30	6.294414	3695.91399
## IndIBE_lag5	2.963807	7741.43349
## IndIBE_lag8	2.105410	2589.75118
## IndIBE_lag13	3.232000	10363.10812
## IndIBE_lag20	4.303025	19077.24363
## IndIBE_lag30	8.179822	60871.93427
## IndFCC_lag5	4.172047	49.87432
## IndFCC_lag8	1.615469	170.50697
## IndFCC_lag13	5.827358	36.52982
## IndFCC_lag20	3.957879	1655.07289
## IndFCC_lag30	7.480171	3455.02058
## IndFER_lag5	10.697590	1074.77734
## IndFER_lag8	10.681425	1052.37777
## IndFER_lag13	5.213355	141.08862
## IndFER_lag20	3.246996	100.87231
## IndFER_lag30	3.474200	188.63100
## IndTL5_lag5	3.183913	3345.91014
## IndTL5_lag8	2.571409	2070.47903
## IndTL5_lag13	4.910614	501.74353
## IndTL5_lag20	3.716668	1137.87561
## IndTL5_lag30	7.366264	7008.59547
## IndREDE_lag5	7.374985	80.96442
## IndREDE_lag8	5.985506	48.18872
## IndREDE_lag13	8.358254	51.11936
## IndREDE_lag20	5.598089	44.44519
## IndREDE_lag30	3.990524	39.58106
## IndSCYR_lag5	7.723515	90.52793
## IndSCYR_lag8	7.992456	91.72698
## IndSCYR_lag13	10.050816	97.57312
## IndSCYR_lag20	7.617355	104.81122
## IndSCYR_lag30	11.708636	82.26061
## IndMAP_lag5	10.192410	15992.31528
## IndMAP_lag8	5.186557	4876.98383
## IndMAP_lag13	2.652120	641.81577
## IndMAP_lag20	8.143965	167.86313
## IndMAP_lag30	5.433348	49.62237
## IndGRLS_lag5	6.718208	153.03039
## IndGRLS_lag8	6.098670	180.92112
## IndGRLS_lag13	7.542055	390.97993
## IndGRLS_lag20	9.381846	289.82860
## IndGRLS_lag30	4.248085	214.89691
## IndICAG_lag5	3.513169	216.02082

```
## IndICAG_lag8 9.472539 113.42021
## IndICAG_lag13 9.405107 66.78923
## IndICAG_lag20 3.782830 89.87978
## IndICAG_lag30 6.517522 174.07857
## IndVIS_lag5 10.254376 168.88862
## IndVIS_lag8 9.039437 115.75041
## IndVIS_lag13 9.812942 223.34618
## IndVIS_lag20 9.905910 181.21148
## IndVIS_lag30 5.365046 172.70884
## IndACX_lag5 8.315901 39.31296
## IndACX_lag8 8.259555 32.95809
## IndACX_lag13 6.082058 40.05640
## IndACX_lag20 9.300100 109.44397
## IndACX_lag30 7.698368 415.27638
## IndAENA_lag5 8.025503 5648.05247
## IndAENA_lag8 4.579745 1888.25943
## IndAENA_lag13 1.419869 149.35879
## IndAENA_lag20 5.389820 58.61999
## IndAENA_lag30 4.142691 85.58450
## IndMRL_lag5 4.146565 1084.53861
## IndMRL_lag8 3.195438 550.95261
## IndMRL_lag13 5.237287 56.45324
## IndMRL_lag20 9.969069 202.62098
## IndMRL_lag30 4.848300 80.56913
```

```
varImpPlot(rf.IndTEF)
```

rf.IndTEF



MAX(5)

Se escogen las 5 variables más relevantes.

```
library (randomForest)
set.seed (1)
rfMax5.IndTEF <- randomForest(IndTEF ~ IndSAN_lag5+IndNTGY_lag20+IndSCYR_lag30+IndIDR_lag30+IndIDR_lag5, data = DB_ind,importance = TRUE)
rfMax5.IndTEF
```

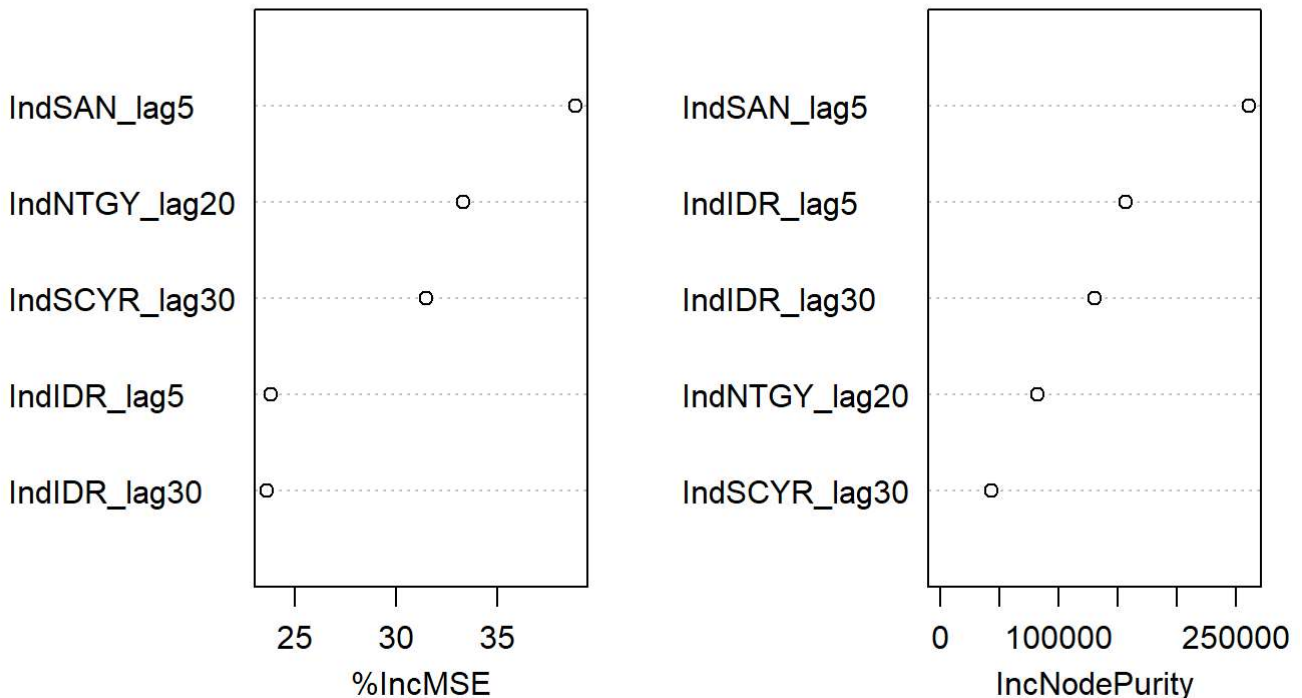
```
##
## Call:
## randomForest(formula = IndTEF ~ IndSAN_lag5 + IndNTGY_lag20 + IndSCYR_lag30 + IndIDR_la
g30 + IndIDR_lag5, data = DB_ind, importance = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 1
##
##           Mean of squared residuals: 5.865576
##           % Var explained: 98.78
```

```
importance(rfMax5.IndTEF)
```

```
##           %IncMSE  IncNodePurity
## IndSAN_lag5  38.86505    261239.28
## IndNTGY_lag20 33.33561     82565.10
## IndSCYR_lag30 31.48948     43123.62
## IndIDR_lag30  23.60938    130879.53
## IndIDR_lag5   23.82147    156830.87
```

```
varImpPlot(rfMax5.IndTEF)
```

rfMax5.IndTEF



MAX(10)

Se escogen las 10 variables más relevantes

```
set.seed (2)
rfMax10.IndTEF <- randomForest(IndTEF ~ IndSAN_lag5+IndNTGY_lag20+IndSCYR_lag30+IndIDR_lag30+Ind
IDR_lag5+IndSAN_lag8+IndNTGY_lag13+IndFER_lag5+IndFER_lag8+IndTEF_lag20,data = DB_ind, importanc
e = TRUE)
rfMax10.IndTEF
```

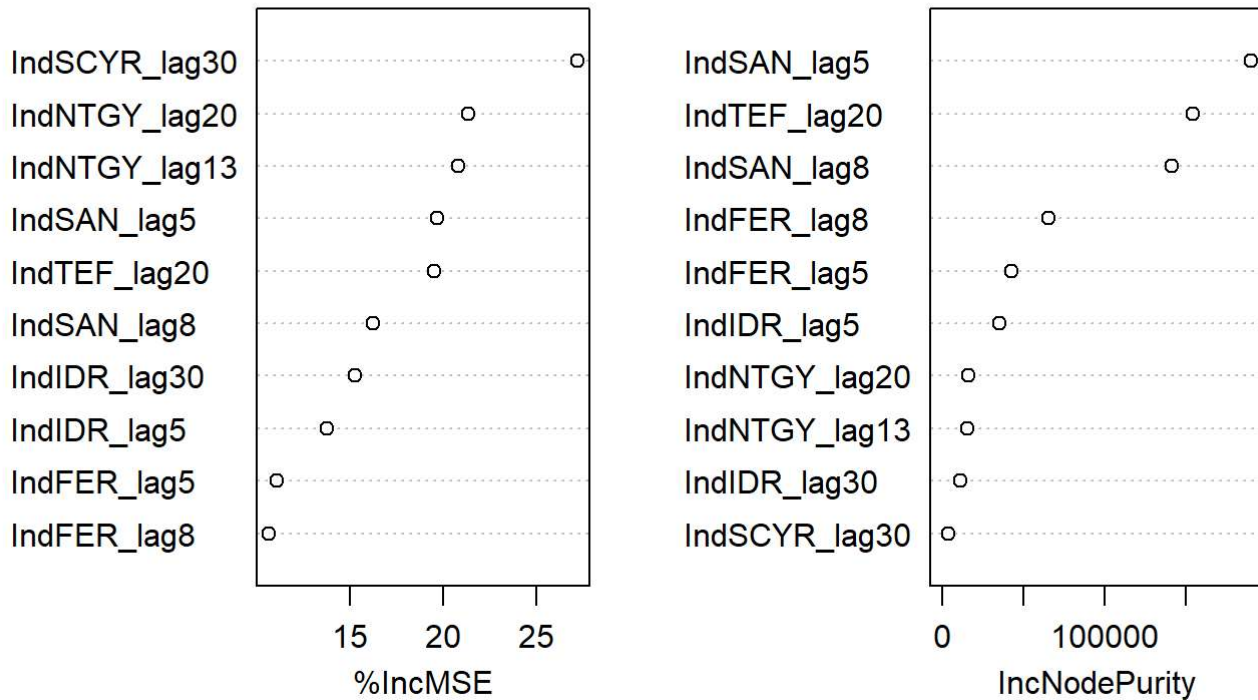
```
##
## Call:
## randomForest(formula = IndTEF ~ IndSAN_lag5 + IndNTGY_lag20 + IndSCYR_lag30 + IndIDR_la
g30 + IndIDR_lag5 + IndSAN_lag8 + IndNTGY_lag13 + IndFER_lag5 + IndFER_lag8 + IndTEF_lag20,
data = DB_ind, importance = TRUE)
## Type of random forest: regression
## Number of trees: 500
## No. of variables tried at each split: 3
##
## Mean of squared residuals: 1.690655
## % Var explained: 99.65
```

```
importance(rfMax10.IndTEF)
```

```
## %IncMSE IncNodePurity
## IndSAN_lag5 19.65548 190204.569
## IndNTGY_lag20 21.32367 15789.429
## IndSCYR_lag30 27.16029 3513.558
## IndIDR_lag30 15.27910 11191.455
## IndIDR_lag5 13.74847 35280.481
## IndSAN_lag8 16.25134 141697.067
## IndNTGY_lag13 20.80677 15580.591
## IndFER_lag5 11.07695 42767.291
## IndFER_lag8 10.67821 65654.407
## IndTEF_lag20 19.48398 154177.168
```

```
varImpPlot(rfMax10.IndTEF)
```

rfMax10.IndTEF



MAX(15)

Se escogen las 15 variables más relevantes

```
set.seed (3)
rfMax15.IndTEF <- randomForest(IndTEF ~ IndSAN_lag5+IndNTGY_lag20+IndSCYR_lag30+IndIDR_lag30+Ind
IDR_lag5+IndSAN_lag8+IndNTGY_lag13+IndFER_lag5+IndFER_lag8+IndTEF_lag20+IndVIS_lag5+IndMAP_lag5+
IndSCYR_lag13+IndMRL_lag20+IndVIS_lag20,data = DB_ind, importance = TRUE)
rfMax15.IndTEF
```

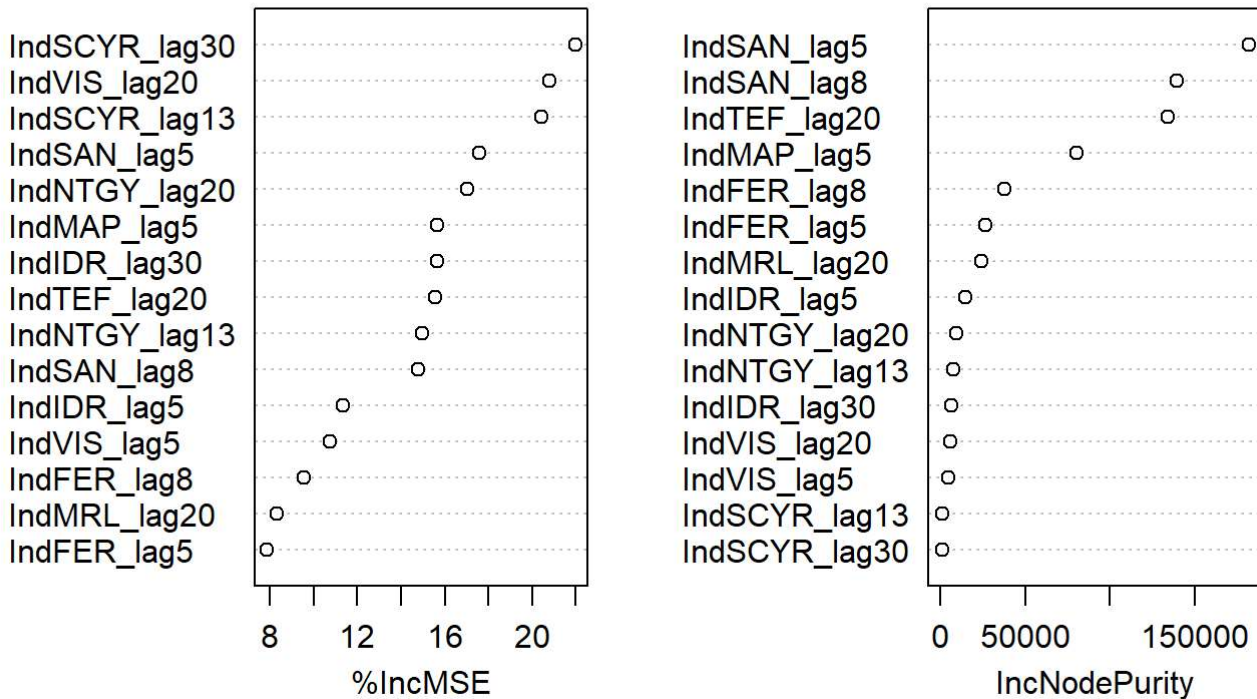
```
##
## Call:
## randomForest(formula = IndTEF ~ IndSAN_lag5 + IndNTGY_lag20 + IndSCYR_lag30 + IndIDR_la
g30 + IndIDR_lag5 + IndSAN_lag8 + IndNTGY_lag13 + IndFER_lag5 + IndFER_lag8 + IndTEF_lag20
+ IndVIS_lag5 + IndMAP_lag5 + IndSCYR_lag13 + IndMRL_lag20 + IndVIS_lag20, data = DB_i
nd, importance = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 5
##
##           Mean of squared residuals: 1.236503
##           % Var explained: 99.74
```

```
importance(rfMax15.IndTEF)
```

##	%IncMSE	IncNodePurity
## IndSAN_lag5	17.568669	182087.294
## IndNTGY_lag20	17.011889	9556.157
## IndSCYR_lag30	21.988401	1068.811
## IndIDR_lag30	15.648776	6425.238
## IndIDR_lag5	11.354452	14412.501
## IndSAN_lag8	14.758801	139623.451
## IndNTGY_lag13	14.948518	7847.059
## IndFER_lag5	7.850111	26537.165
## IndFER_lag8	9.531837	37936.243
## IndTEF_lag20	15.571676	134406.299
## IndVIS_lag5	10.747201	4839.476
## IndMAP_lag5	15.654279	80335.834
## IndSCYR_lag13	20.427642	1196.677
## IndMRL_lag20	8.305666	24426.896
## IndVIS_lag20	20.825664	5907.965

```
varImpPlot(rfMax15.IndTEF)
```

rfMax15.IndTEF



3. Aplicar las técnicas de minería de datos

Crear las bases de datos para Training y Test

```
# 80% Training, 20% Test
DB_Training80 = DB_ind[1:1126,]
DB_Test20 = DB_ind[1127:1408,]

# 90% Training, 10% Test
DB_Training90 = DB_ind[1:1267,]
DB_Test10 = DB_ind[1268:1408,]
```

3.1 Regresión Lineal

a. MAX(5)

Analizamos con la base de las 5 mejores variables

80% training

```
reg_lm_IndTEF_80Training = lm(formula = IndTEF ~ IndSAN_lag5+IndNTGY_lag20+IndSCYR_lag30+IndIDR_
lag30+IndIDR_lag5, data =DB_Training80)
summary(reg_lm_IndTEF_80Training)
```

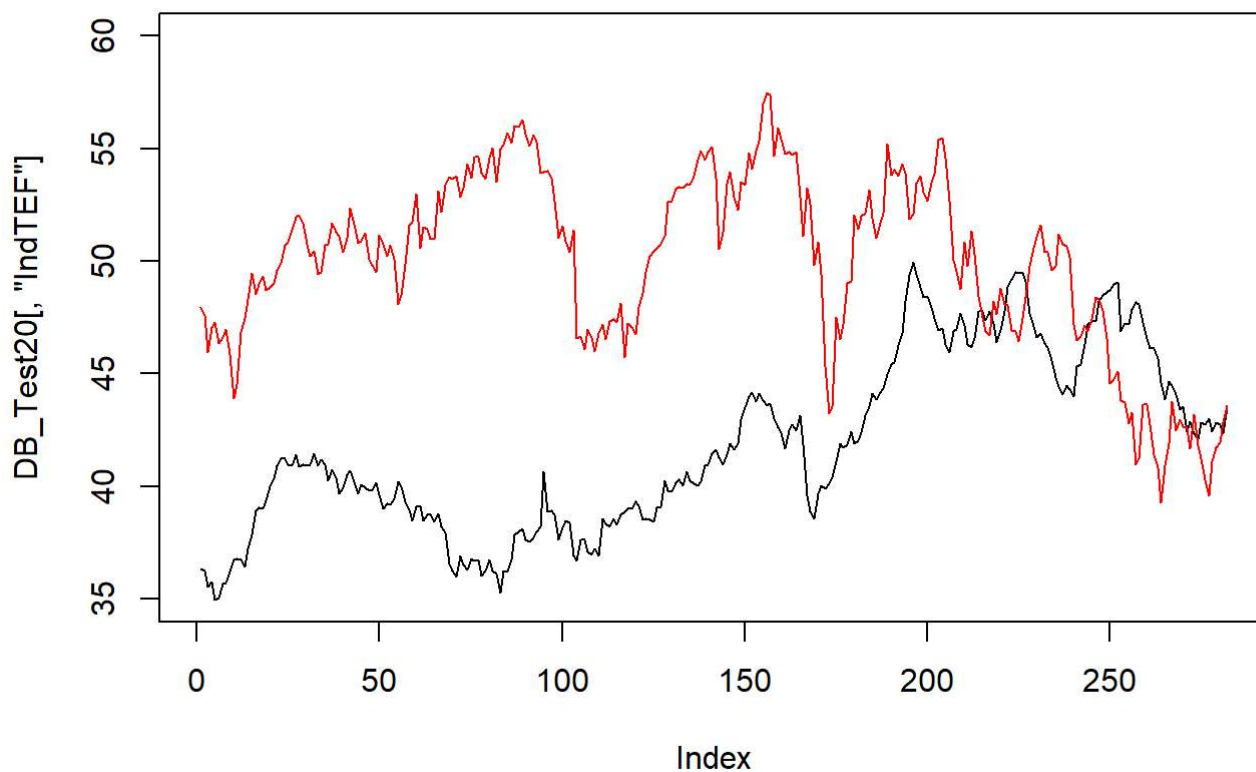
```
##
## Call:
## lm(formula = IndTEF ~ IndSAN_lag5 + IndNTGY_lag20 + IndSCYR_lag30 +
##     IndIDR_lag30 + IndIDR_lag5, data = DB_Training80)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -17.9692  -5.3187   0.1551   4.2916  22.6533
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -3.70117    2.28988  -1.616  0.106307
## IndSAN_lag5    1.35601    0.04730  28.665 < 2e-16 ***
## IndNTGY_lag20  0.35328    0.09726   3.632 0.000294 ***
## IndSCYR_lag30 -7.40386    1.10243  -6.716 2.96e-11 ***
## IndIDR_lag30   0.50564    0.29031   1.742 0.081829 .
## IndIDR_lag5    1.71720    0.34364   4.997 6.75e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.507 on 1120 degrees of freedom
## Multiple R-squared:  0.8893, Adjusted R-squared:  0.8888
## F-statistic: 1800 on 5 and 1120 DF, p-value: < 2.2e-16
```

```
predict_lm_IndTEF80 = predict.lm(reg_lm_IndTEF_80Training, newdata = DB_Test20[,c("IndSAN_lag5",
"IndNTGY_lag20","IndSCYR_lag30","IndIDR_lag30","IndIDR_lag5")], interval = "confidence")
head(predict_lm_IndTEF80)
```



```
##          fit      lwr      upr
## 1127 47.95926 47.02467 48.89386
## 1128 47.60859 46.66710 48.55008
## 1129 45.96831 45.03663 46.89999
## 1130 47.08928 46.12699 48.05157
## 1131 47.27230 46.32549 48.21910
## 1132 46.36408 45.43894 47.28922
```

```
plot(DB_Test20[,"IndTEF"], type = "l",ylim = c(35,60))
lines(predict_lm_IndTEF80[,"fit"], type = "l", col="red")
```



```
# Error media cuadrático
mean((DB_Test20[,"IndTEF"]-predict_lm_IndTEF80[,"fit"])^2)
```

```
## [1] 103.5175
```

90% training

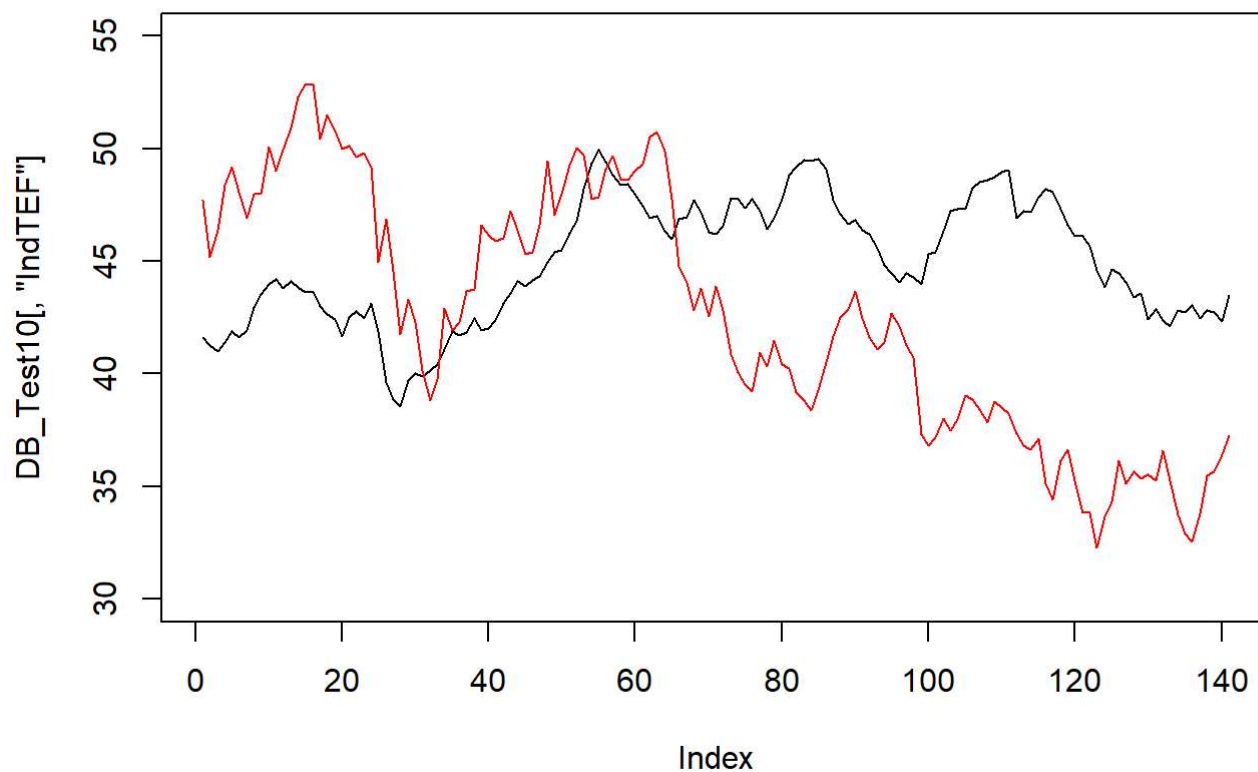
```
reg_lm_IndTEF_90Training = lm(formula = IndTEF ~ IndSAN_lag5+IndNTGY_lag20+IndSCYR_lag30+IndIDR_
lag30+IndIDR_lag5, data =DB_Training90)
summary(reg_lm_IndTEF_90Training)
```

```
##
## Call:
## lm(formula = IndTEF ~ IndSAN_lag5 + IndNTGY_lag20 + IndSCYR_lag30 +
##     IndIDR_lag30 + IndIDR_lag5, data = DB_Training90)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -18.9654  -6.6755   0.6273   5.1187  23.9940
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    6.75723    2.19438   3.079 0.002119 **
## IndSAN_lag5    1.65627    0.04141  39.998 < 2e-16 ***
## IndNTGY_lag20  0.04014    0.09495   0.423 0.672539
## IndSCYR_lag30 -8.75708    1.11456  -7.857 8.38e-15 ***
## IndIDR_lag30  0.95858    0.27872   3.439 0.000603 ***
## IndIDR_lag5   -0.21116    0.31804  -0.664 0.506860
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.879 on 1261 degrees of freedom
## Multiple R-squared:  0.8788, Adjusted R-squared:  0.8783
## F-statistic: 1829 on 5 and 1261 DF, p-value: < 2.2e-16
```

```
predict_lm_IndTEF90 = predict.lm(reg_lm_IndTEF_90Training, newdata = DB_Test10[,c("IndSAN_lag5",
"IndNTGY_lag20","IndSCYR_lag30","IndIDR_lag30","IndIDR_lag5")], interval = "confidence")
head(predict_lm_IndTEF90)
```

```
##           fit      lwr      upr
## 1268 47.69878 46.09562 49.30194
## 1269 45.18252 43.61972 46.74532
## 1270 46.37125 44.87489 47.86761
## 1271 48.38987 46.89005 49.88970
## 1272 49.17700 47.69694 50.65707
## 1273 48.07737 46.60527 49.54948
```

```
plot(DB_Test10["IndTEF"], type = "l",ylim = c(30,55))
lines(predict_lm_IndTEF90["fit"], type = "l", col="red")
```



```
# Error media cuadrático
mean((DB_Test10[, "IndTEF"]-predict_lm_IndTEF90[, "fit"])^2)
```

```
## [1] 47.18739
```

b. MAX(10)

Analizamos con la base de las 10 mejores variables

80%

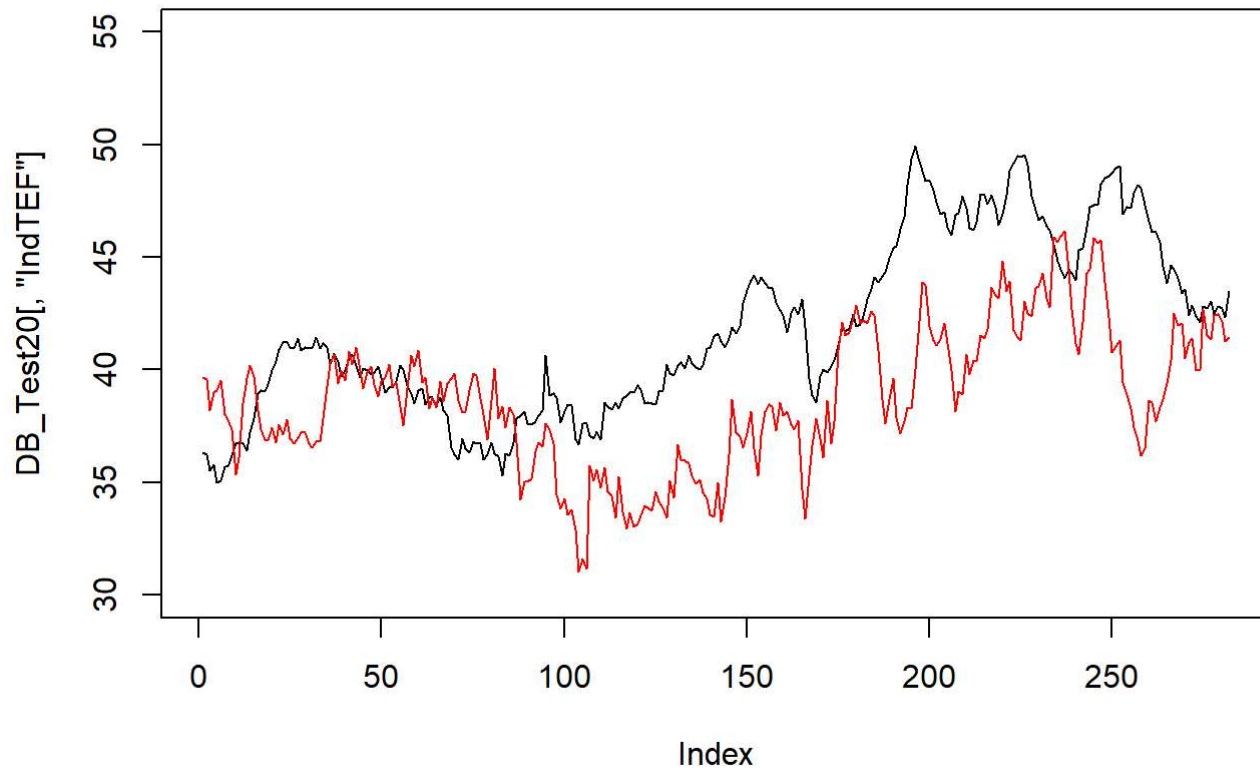
```
reg_lm_MAX10_IndTEF_80Training = lm(formula = IndTEF ~ IndSAN_lag5+IndNTGY_lag20+IndSCYR_lag30+IndIDR_lag30+IndIDR_lag5+IndSAN_lag8+IndNTGY_lag13+IndFER_lag5+IndFER_lag8+IndTEF_lag20, data =DB_Training80)
summary(reg_lm_MAX10_IndTEF_80Training)
```

```
##
## Call:
## lm(formula = IndTEF ~ IndSAN_lag5 + IndNTGY_lag20 + IndSCYR_lag30 +
##     IndIDR_lag30 + IndIDR_lag5 + IndSAN_lag8 + IndNTGY_lag13 +
##     IndFER_lag5 + IndFER_lag8 + IndTEF_lag20, data = DB_Training80)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -14.9784  -2.4954   0.0006   2.1522  11.4143
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   13.3963     1.4060   9.528 < 2e-16 ***
## IndSAN_lag5    0.8454     0.1018   8.302 2.93e-16 ***
## IndNTGY_lag20 -0.9542     0.1512  -6.313 3.95e-10 ***
## IndSCYR_lag30  0.3029     0.5593   0.542 0.58814
## IndIDR_lag30  -1.6196     0.1591 -10.179 < 2e-16 ***
## IndIDR_lag5    0.5512     0.1866   2.953 0.00321 **
## IndSAN_lag8   -0.5682     0.1027  -5.534 3.91e-08 ***
## IndNTGY_lag13  0.7588     0.1525   4.976 7.52e-07 ***
## IndFER_lag5    0.1683     0.2065   0.815 0.41515
## IndFER_lag8   -0.4219     0.2073  -2.035 0.04207 *
## IndTEF_lag20   0.9005     0.0159  56.623 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.687 on 1115 degrees of freedom
## Multiple R-squared:  0.9734, Adjusted R-squared:  0.9732
## F-statistic: 4084 on 10 and 1115 DF,  p-value: < 2.2e-16
```

```
predict_lm_MAX10_IndTEF80 = predict.lm(reg_lm_MAX10_IndTEF_80Training, newdata = DB_Test20[,c("IndSAN_lag5", "IndNTGY_lag20", "IndSCYR_lag30", "IndIDR_lag30", "IndIDR_lag5", "IndSAN_lag8", "IndNTGY_lag13", "IndFER_lag5", "IndFER_lag8", "IndTEF_lag20")], interval = "confidence")
head(predict_lm_MAX10_IndTEF80)
```

```
##           fit      lwr      upr
## 1127 39.65871 39.00689 40.31054
## 1128 39.55579 38.94472 40.16686
## 1129 38.19790 37.49200 38.90381
## 1130 38.99583 38.41007 39.58158
## 1131 39.08778 38.49143 39.68412
## 1132 39.52797 38.98329 40.07265
```

```
plot(DB_Test20[, "IndTEF"], type = "l", ylim = c(30, 55))
lines(predict_lm_MAX10_IndTEF80[, "fit"], type = "l", col = "red")
```



```
# Error media cuadrático
mean((DB_Test20[, "IndTEF"]-predict_lm_MAX10_IndTEF80[, "fit"])^2)
```

```
## [1] 21.17588
```

90% training

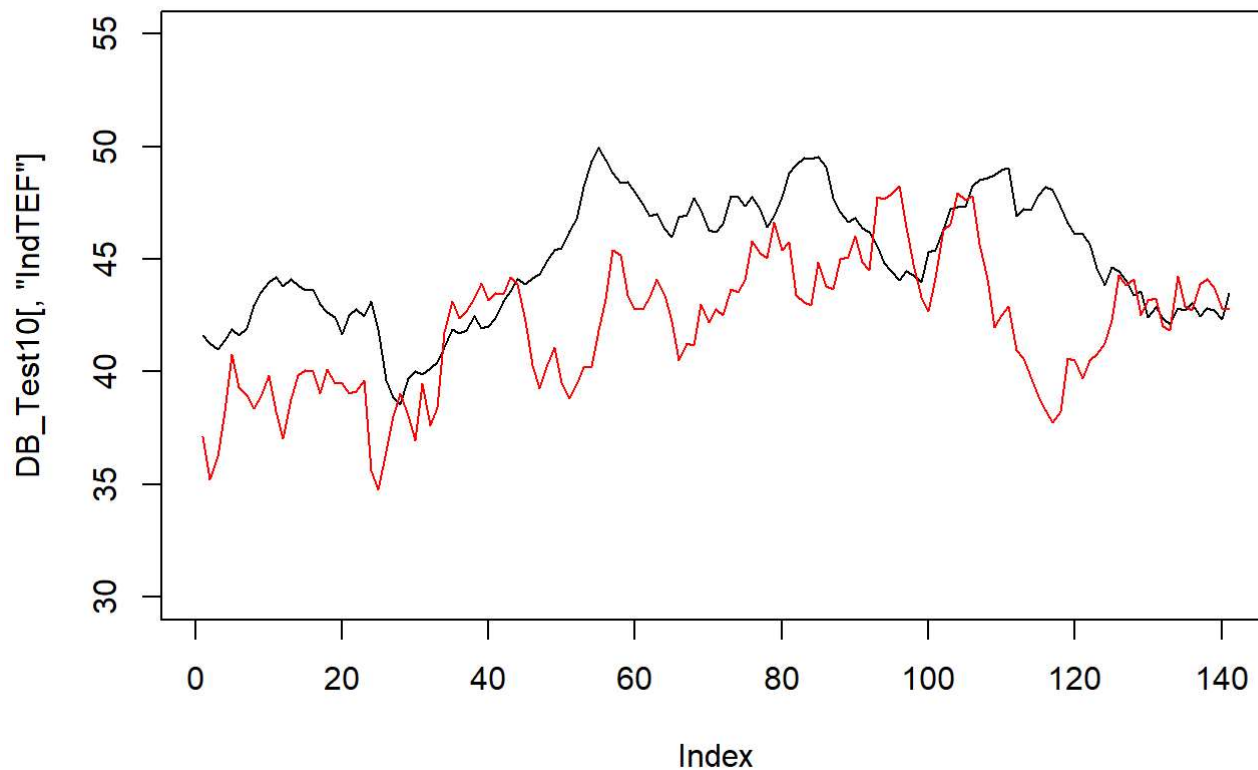
```
reg_lm_MAX10_IndTEF_90Training = lm(formula = IndTEF ~ IndSAN_lag5+IndNTGY_lag20+IndSCYR_lag30+IndIDR_lag30+IndIDR_lag5+IndSAN_lag8+IndNTGY_lag13+IndFER_lag5+IndFER_lag8+IndTEF_lag20, data = DB_Training90)
summary(reg_lm_MAX10_IndTEF_90Training)
```

```
##
## Call:
## lm(formula = IndTEF ~ IndSAN_lag5 + IndNTGY_lag20 + IndSCYR_lag30 +
##     IndIDR_lag30 + IndIDR_lag5 + IndSAN_lag8 + IndNTGY_lag13 +
##     IndFER_lag5 + IndFER_lag8 + IndTEF_lag20, data = DB_Training90)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -15.9009  -2.4399  -0.0755   2.2115  12.0862
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  10.62913    1.23185   8.629 < 2e-16 ***
## IndSAN_lag5    0.82019    0.09586   8.556 < 2e-16 ***
## IndNTGY_lag20 -0.99622    0.13797  -7.221 8.93e-13 ***
## IndSCYR_lag30 -0.16209    0.53007  -0.306  0.7598
## IndIDR_lag30  -1.36775    0.13712  -9.975 < 2e-16 ***
## IndIDR_lag5    0.83954    0.16339   5.138 3.21e-07 ***
## IndSAN_lag8   -0.58417    0.09650  -6.054 1.87e-09 ***
## IndNTGY_lag13  0.94318    0.13384   7.047 3.00e-12 ***
## IndFER_lag5    0.05124    0.19169   0.267  0.7893
## IndFER_lag8   -0.34109    0.19254  -1.772  0.0767 .
## IndTEF_lag20  0.86938    0.01362  63.838 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.607 on 1256 degrees of freedom
## Multiple R-squared:  0.9747, Adjusted R-squared:  0.9745
## F-statistic: 4839 on 10 and 1256 DF, p-value: < 2.2e-16
```

```
predict_lm_MAX10_IndTEF90 = predict.lm(reg_lm_MAX10_IndTEF_90Training, newdata = DB_Test10[,c("IndSAN_lag5", "IndNTGY_lag20", "IndSCYR_lag30", "IndIDR_lag30", "IndIDR_lag5", "IndSAN_lag8", "IndNTGY_lag13", "IndFER_lag5", "IndFER_lag8", "IndTEF_lag20")], interval = "confidence")
head(predict_lm_MAX10_IndTEF90)
```

```
##           fit      lwr      upr
## 1268 37.10968 36.25731 37.96204
## 1269 35.20090 34.29532 36.10649
## 1270 36.22453 35.30242 37.14663
## 1271 38.24571 37.30761 39.18380
## 1272 40.76191 39.74915 41.77467
## 1273 39.29296 38.40148 40.18444
```

```
plot(DB_Test10[, "IndTEF"], type = "l", ylim = c(30, 55))
lines(predict_lm_MAX10_IndTEF90[, "fit"], type = "l", col = "red")
```



```
# Error media cuadrático
mean((DB_Test10[, "IndTEF"]-predict_lm_MAX10_IndTEF90[, "fit"])^2)
```

```
## [1] 17.62934
```

c. MAX(15)

Analizamos con la base de las 15 mejores variables

80% training

```
reg_lm_MAX15_IndTEF_80Training = lm(formula = IndTEF ~ IndSAN_lag5+IndNTGY_lag20+IndSCYR_lag30+IndIDR_lag30+IndIDR_lag5+IndSAN_lag8+IndNTGY_lag13+IndFER_lag5+IndFER_lag8+IndTEF_lag20+IndVIS_lag5+IndMAP_lag5+IndSCYR_lag13+IndMRL_lag20+IndVIS_lag20, data =DB_Training80)
summary(reg_lm_MAX15_IndTEF_80Training)
```

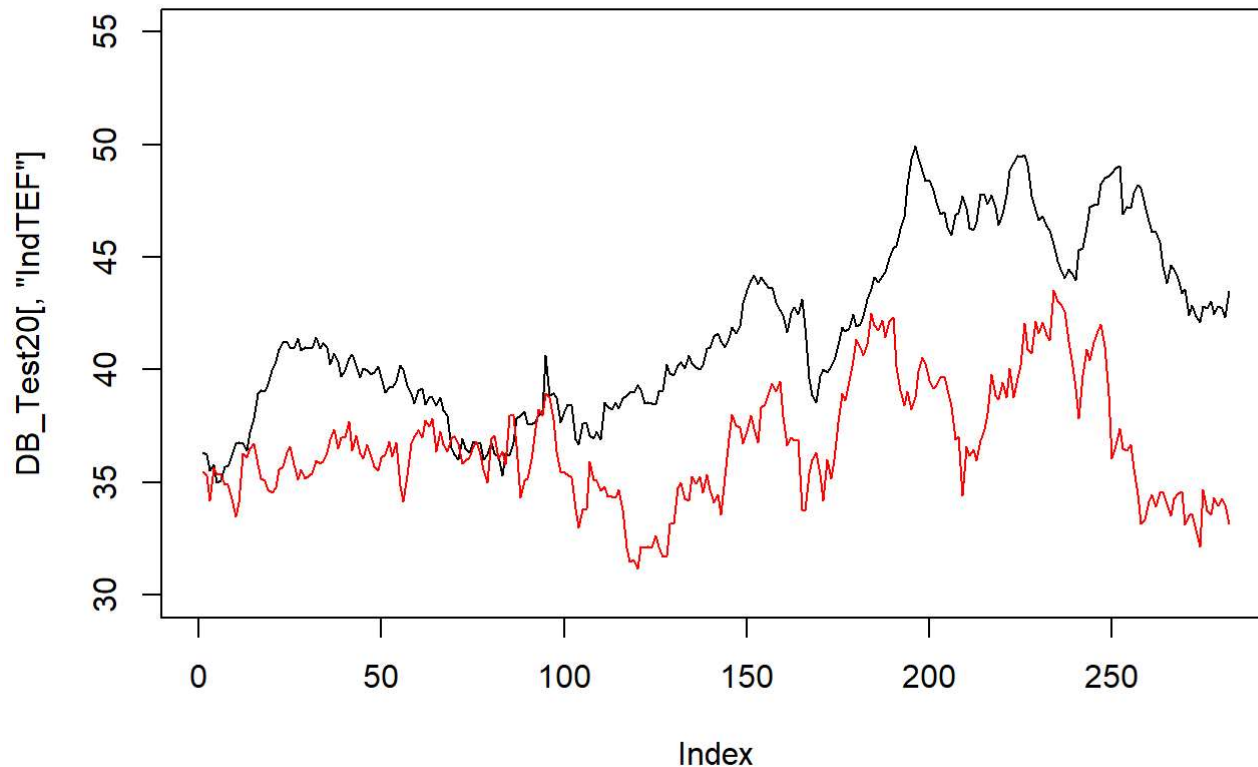
```
##
## Call:
## lm(formula = IndTEF ~ IndSAN_lag5 + IndNTGY_lag20 + IndSCYR_lag30 +
##     IndIDR_lag30 + IndIDR_lag5 + IndSAN_lag8 + IndNTGY_lag13 +
##     IndFER_lag5 + IndFER_lag8 + IndTEF_lag20 + IndVIS_lag5 +
##     IndMAP_lag5 + IndSCYR_lag13 + IndMRL_lag20 + IndVIS_lag20,
##     data = DB_Training80)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -9.8477 -1.8658  0.0663  1.8532 11.3189
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  44.78269    2.65119   16.892 < 2e-16 ***
## IndSAN_lag5    0.43567    0.08297    5.251 1.81e-07 ***
## IndNTGY_lag20 -0.81386    0.12674   -6.422 2.00e-10 ***
## IndSCYR_lag30  2.04027    0.56593    3.605 0.000326 ***
## IndIDR_lag30  -1.27764    0.14353   -8.902 < 2e-16 ***
## IndIDR_lag5    1.16803    0.15736    7.423 2.28e-13 ***
## IndSAN_lag8   -0.31690    0.08315   -3.811 0.000146 ***
## IndNTGY_lag13  0.55921    0.12514    4.469 8.67e-06 ***
## IndFER_lag5   -0.76660    0.16914   -4.532 6.47e-06 ***
## IndFER_lag8   -0.14147    0.17185   -0.823 0.410578
## IndTEF_lag20  0.52374    0.02041   25.662 < 2e-16 ***
## IndVIS_lag5   -0.21071    0.03919   -5.376 9.26e-08 ***
## IndMAP_lag5   15.26285    0.79430   19.215 < 2e-16 ***
## IndSCYR_lag13 -1.39252    0.69972   -1.990 0.046823 *
## IndMRL_lag20  -0.93039    0.12858   -7.236 8.63e-13 ***
## IndVIS_lag20  -0.21170    0.03971   -5.331 1.18e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.918 on 1110 degrees of freedom
## Multiple R-squared:  0.9834, Adjusted R-squared:  0.9832
## F-statistic: 4390 on 15 and 1110 DF, p-value: < 2.2e-16
```

```
predict_lm_MAX15_IndTEF80 = predict.lm(reg_lm_MAX15_IndTEF_80Training, newdata = DB_Test20[,c("IndSAN_lag5", "IndNTGY_lag20", "IndSCYR_lag30", "IndIDR_lag30", "IndIDR_lag5", "IndSAN_lag8", "IndNTGY_lag13", "IndFER_lag5", "IndFER_lag8", "IndTEF_lag20", "IndVIS_lag5", "IndMAP_lag5", "IndSCYR_lag13", "IndMRL_lag20", "IndVIS_lag20")], interval = "confidence")
head(predict_lm_MAX15_IndTEF80)
```

```
##           fit      lwr      upr
## 1127 35.46658 34.76423 36.16892
## 1128 35.28956 34.60223 35.97690
## 1129 34.18787 33.42392 34.95182
## 1130 35.64486 34.97329 36.31643
## 1131 35.36757 34.69104 36.04409
## 1132 35.36978 34.74341 35.99615
```



```
plot(DB_Test20["IndTEF"], type = "l",ylim = c(30,55))
lines(predict_lm_MAX15_IndTEF80["fit"], type = "l", col="red")
```



```
# Error media cuadrático
mean((DB_Test20["IndTEF"]-predict_lm_MAX15_IndTEF80["fit"])^2)
```

```
## [1] 38.27777
```

90% training

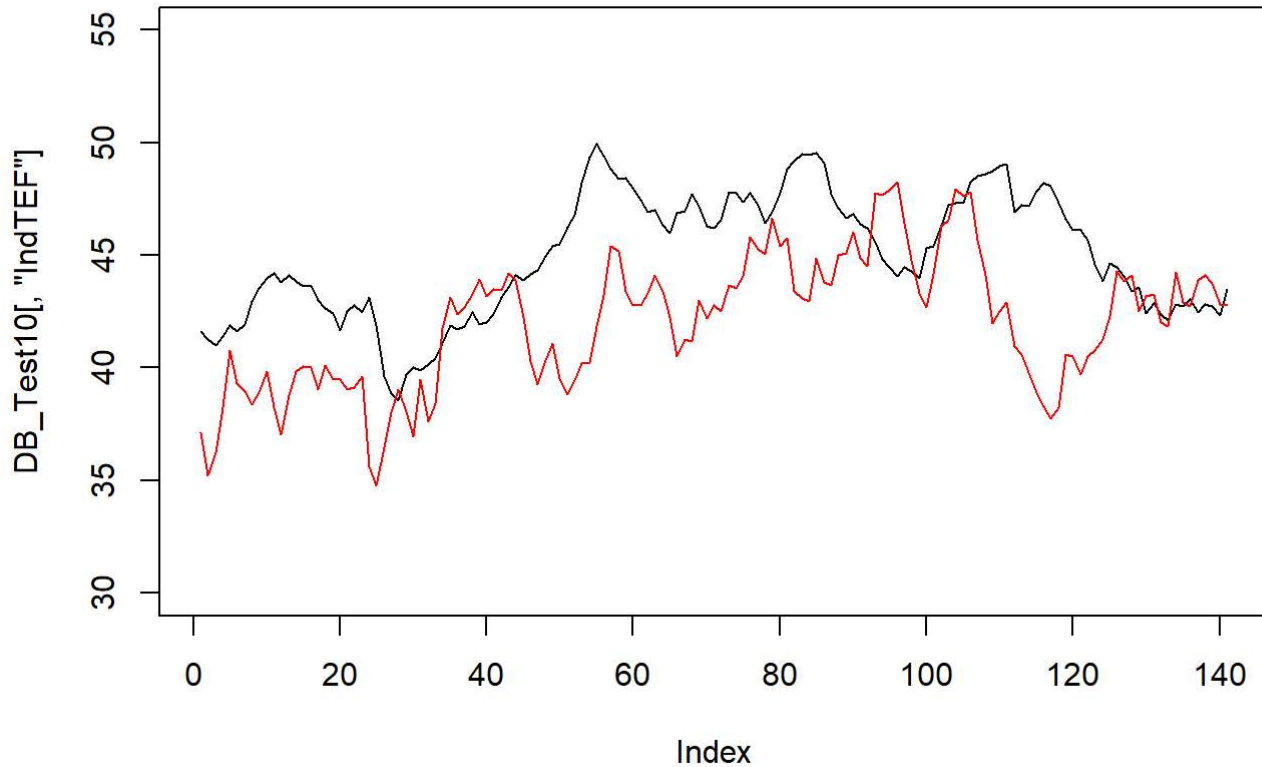
```
reg_lm_MAX15_IndTEF_90Training = lm(formula = IndTEF ~ IndSAN_lag5+IndNTGY_lag20+IndSCYR_lag30+IndIDR_lag30+IndIDR_lag5+IndSAN_lag8+IndNTGY_lag13+IndFER_lag5+IndFER_lag8+IndTEF_lag20+IndVIS_lag5+IndMAP_lag5+IndSCYR_lag13+IndMRL_lag20+IndVIS_lag20, data = DB_Training90)
summary(reg_lm_MAX15_IndTEF_90Training)
```

```
##
## Call:
## lm(formula = IndTEF ~ IndSAN_lag5 + IndNTGY_lag20 + IndSCYR_lag30 +
##     IndIDR_lag30 + IndIDR_lag5 + IndSAN_lag8 + IndNTGY_lag13 +
##     IndFER_lag5 + IndFER_lag8 + IndTEF_lag20 + IndVIS_lag5 +
##     IndMAP_lag5 + IndSCYR_lag13 + IndMRL_lag20 + IndVIS_lag20,
##     data = DB_Training90)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -10.1341  -1.7747   0.0177   1.8260  11.1900
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  37.55857    2.50127  15.016 < 2e-16 ***
## IndSAN_lag5    0.43834    0.07945   5.517 4.19e-08 ***
## IndNTGY_lag20 -0.76473    0.11912  -6.420 1.93e-10 ***
## IndSCYR_lag30  2.51123    0.53814   4.666 3.39e-06 ***
## IndIDR_lag30  -1.05370    0.12295  -8.571 < 2e-16 ***
## IndIDR_lag5    1.36167    0.14264   9.546 < 2e-16 ***
## IndSAN_lag8   -0.30572    0.07942  -3.849 0.000124 ***
## IndNTGY_lag13  0.77782    0.11175   6.960 5.46e-12 ***
## IndFER_lag5   -0.70337    0.15882  -4.429 1.03e-05 ***
## IndFER_lag8   -0.09320    0.16196  -0.575 0.565086
## IndTEF_lag20  0.52512    0.01980  26.519 < 2e-16 ***
## IndVIS_lag5   -0.18063    0.03765  -4.797 1.80e-06 ***
## IndMAP_lag5   13.98306    0.75413  18.542 < 2e-16 ***
## IndSCYR_lag13 -2.28052    0.65343  -3.490 0.000500 ***
## IndMRL_lag20  -1.21108    0.11436 -10.590 < 2e-16 ***
## IndVIS_lag20  -0.20914    0.03781  -5.531 3.87e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.912 on 1251 degrees of freedom
## Multiple R-squared:  0.9836, Adjusted R-squared:  0.9834
## F-statistic: 4993 on 15 and 1251 DF, p-value: < 2.2e-16
```

```
predict_lm_MAX15_IndTEF90 = predict.lm(reg_lm_MAX10_IndTEF_90Training, newdata = DB_Test10[,c("IndSAN_lag5", "IndNTGY_lag20", "IndSCYR_lag30", "IndIDR_lag30", "IndIDR_lag5", "IndSAN_lag8", "IndNTGY_lag13", "IndFER_lag5", "IndFER_lag8", "IndTEF_lag20", "IndVIS_lag5", "IndMAP_lag5", "IndSCYR_lag13", "IndMRL_lag20", "IndVIS_lag20")], interval = "confidence")
head(predict_lm_MAX15_IndTEF90)
```

```
##           fit      lwr      upr
## 1268 37.10968 36.25731 37.96204
## 1269 35.20090 34.29532 36.10649
## 1270 36.22453 35.30242 37.14663
## 1271 38.24571 37.30761 39.18380
## 1272 40.76191 39.74915 41.77467
## 1273 39.29296 38.40148 40.18444
```

```
plot(DB_Test10["IndTEF"], type = "l",ylim = c(30,55))
lines(predict_lm_MAX15_IndTEF90["fit"], type = "l", col="red")
```



```
# Error media cuadrático
mean((DB_Test10["IndTEF"]-predict_lm_MAX15_IndTEF90["fit"])^2)
```

```
## [1] 17.62934
```

3.2 Redes neuronales

```
require(ISLR2)
```

```
## Loading required package: ISLR2
```

```
require(nnet)
```

```
## Loading required package: nnet
```

a. MAX(5)

80% training

```
set.seed(12)
red10_80Train = nnet(formula = IndTEF ~ IndSAN_lag5+IndNTGY_lag20+IndSCYR_lag30+IndIDR_lag30+Ind
IDR_lag5, data=DB_Training80, size = 10, linout = TRUE, maxit = 1000)
```

```
## # weights: 71
## initial value 5012616.034196
## iter 10 value 530261.268523
## iter 20 value 241719.434643
## iter 30 value 154438.328404
## iter 40 value 135236.826260
## iter 50 value 95159.701149
## iter 60 value 55107.806275
## iter 70 value 52080.840022
## final value 52080.833380
## converged
```

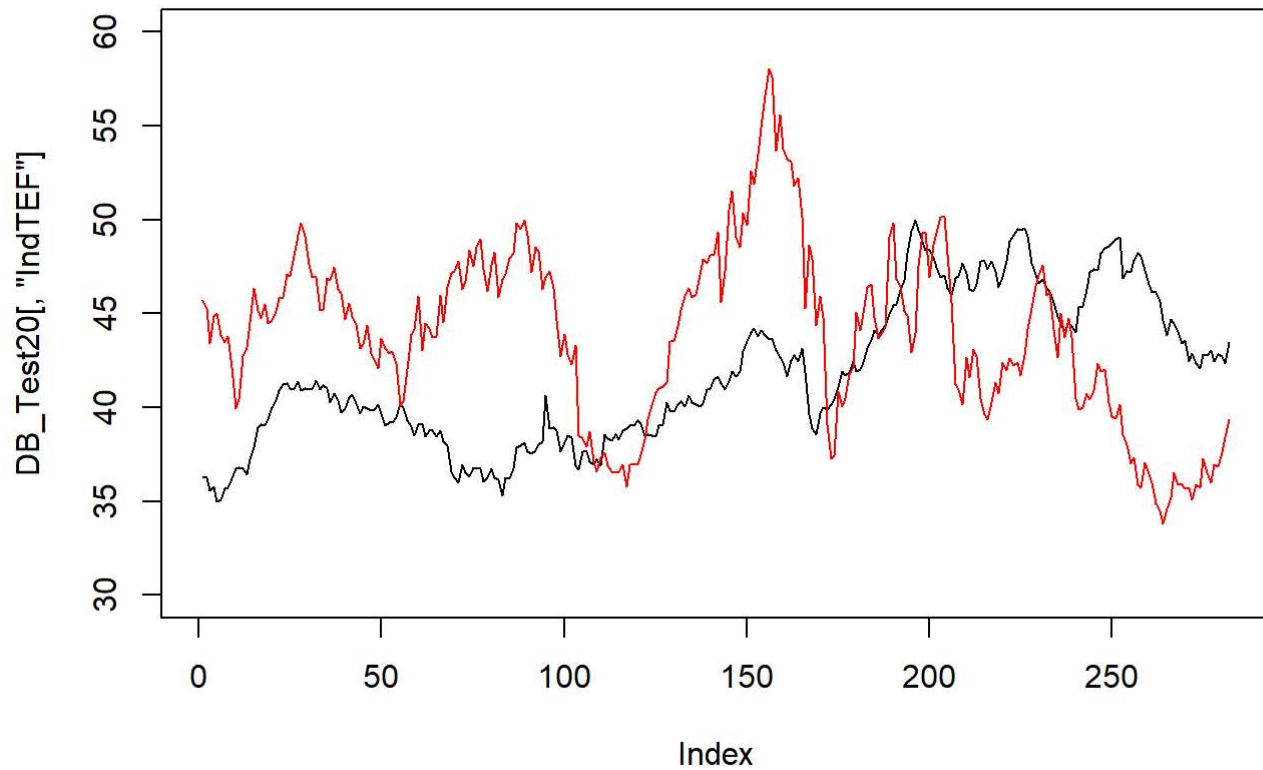
```
pred10_nnet_80Train = predict(object = red10_80Train, newdata = DB_Test20[,c("IndSAN_lag5", "IndN
TGY_lag20", "IndSCYR_lag30", "IndIDR_lag30", "IndIDR_lag5")])
head(pred10_nnet_80Train)
```

```
##          [,1]
## 1127 45.69727
## 1128 45.27750
## 1129 43.43720
## 1130 44.86410
## 1131 45.01145
## 1132 43.84727
```

```
ecm10_80Train = mean((DB_Test20[, "IndTEF"]-pred10_nnet_80Train)^2)
ecm10_80Train
```

```
## [1] 45.56839
```

```
plot(DB_Test20[, "IndTEF"], type = "l", ylim = c(30, 60) )
lines(pred10_nnet_80Train, type = "l", col="red")
```



90% training

```
set.seed(12)
red10_90Train = nnet(formula = IndTEF ~ IndSAN_lag5+IndNTGY_lag20+IndSCYR_lag30+IndIDR_lag30+Ind
IDR_lag5, data=DB_Training90, size = 10, linout = TRUE, maxit = 1000)
```

```
## # weights: 71
## initial value 5221102.137674
## iter 10 value 264873.967944
## iter 20 value 125866.866747
## iter 30 value 99745.507295
## iter 40 value 85228.609445
## iter 50 value 68150.851530
## iter 60 value 67984.068612
## iter 70 value 63003.872807
## iter 80 value 61829.319962
## iter 90 value 61823.077941
## iter 100 value 61546.153988
## iter 110 value 61452.574339
## iter 120 value 54417.319276
## iter 130 value 49201.153281
## iter 140 value 41296.094997
## final value 41177.190565
## converged
```

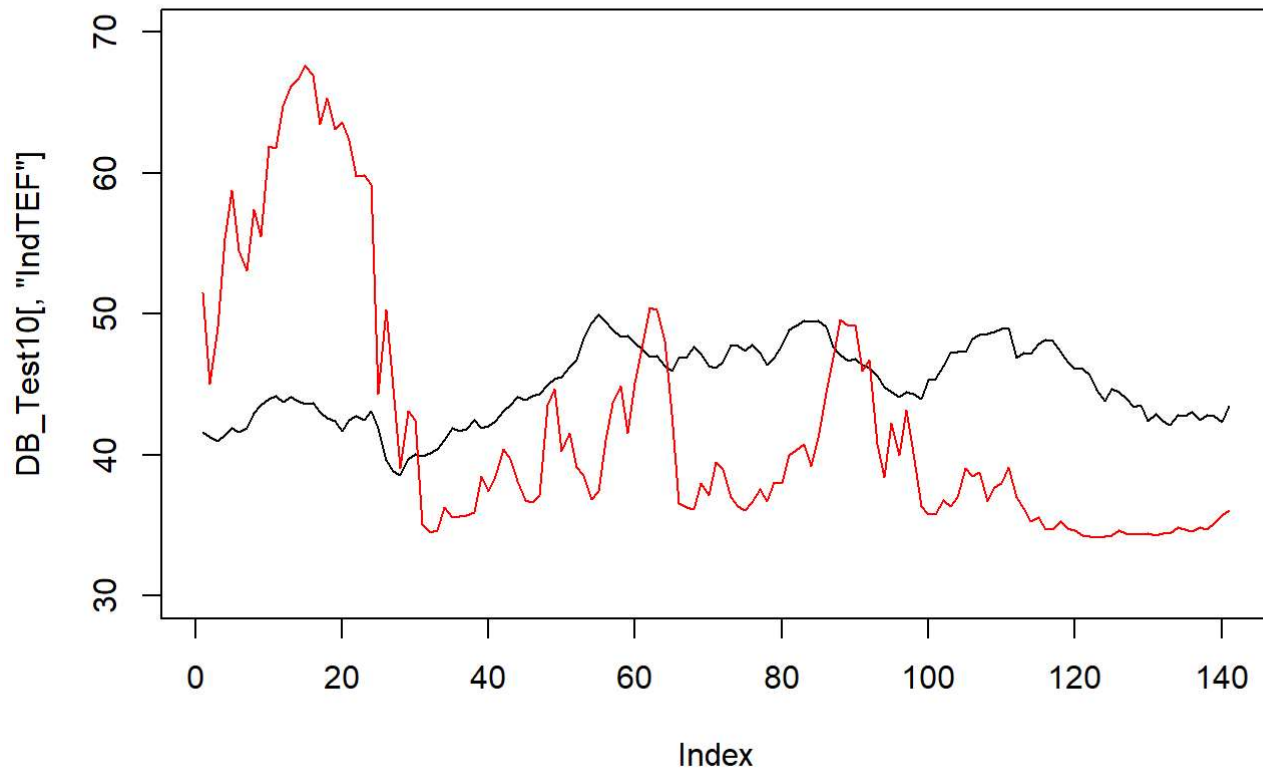
```
pred10_nnet_90Train = predict(object = red10_90Train, newdata = DB_Test10[,c("IndSAN_lag5", "IndN
TGY_lag20", "IndSCYR_lag30", "IndIDR_lag30", "IndIDR_lag5")])
head(pred10_nnet_90Train)
```

```
##          [,1]
## 1268 51.55034
## 1269 45.02100
## 1270 49.08836
## 1271 55.33520
## 1272 58.76458
## 1273 54.52057
```

```
ecm10_90Train = mean((DB_Test10[, "IndTEF"] - pred10_nnet_90Train)^2)
ecm10_90Train
```

```
## [1] 108.6563
```

```
plot(DB_Test10[, "IndTEF"], type = "l", ylim = c(30, 70) )
lines(pred10_nnet_90Train, type = "l", col = "red")
```



b. MAX(10)

80% training

```
set.seed(12)
red10_80Train_Max10 = nnet(formula = IndTEF ~ IndSAN_lag5+IndNTGY_lag20+IndSCYR_lag30+IndIDR_lag
30+IndIDR_lag5+IndSAN_lag8+IndNTGY_lag13+IndFER_lag5+IndFER_lag8+IndTEF_lag20, data=DB_Training8
0, size = 10, linout = TRUE, maxit = 1000)
```

```
## # weights: 121
## initial value 5036627.869158
## iter 10 value 368334.450553
## iter 20 value 171091.206581
## iter 30 value 106451.904898
## iter 40 value 65871.449188
## iter 50 value 25149.215437
## iter 60 value 17585.973986
## iter 70 value 16049.031383
## iter 80 value 15490.607678
## iter 90 value 15311.520760
## iter 100 value 15237.997320
## iter 110 value 15190.912608
## iter 120 value 15169.005687
## iter 130 value 15156.499430
## iter 140 value 15148.466396
## iter 150 value 15141.747215
## iter 160 value 15137.707568
## iter 170 value 15135.322971
## iter 180 value 15132.890545
## iter 190 value 15131.186856
## iter 200 value 15129.727439
## iter 210 value 15128.567275
## iter 220 value 15127.656938
## iter 230 value 15127.157026
## iter 240 value 15126.639248
## iter 250 value 15126.254675
## final value 15126.222322
## converged
```

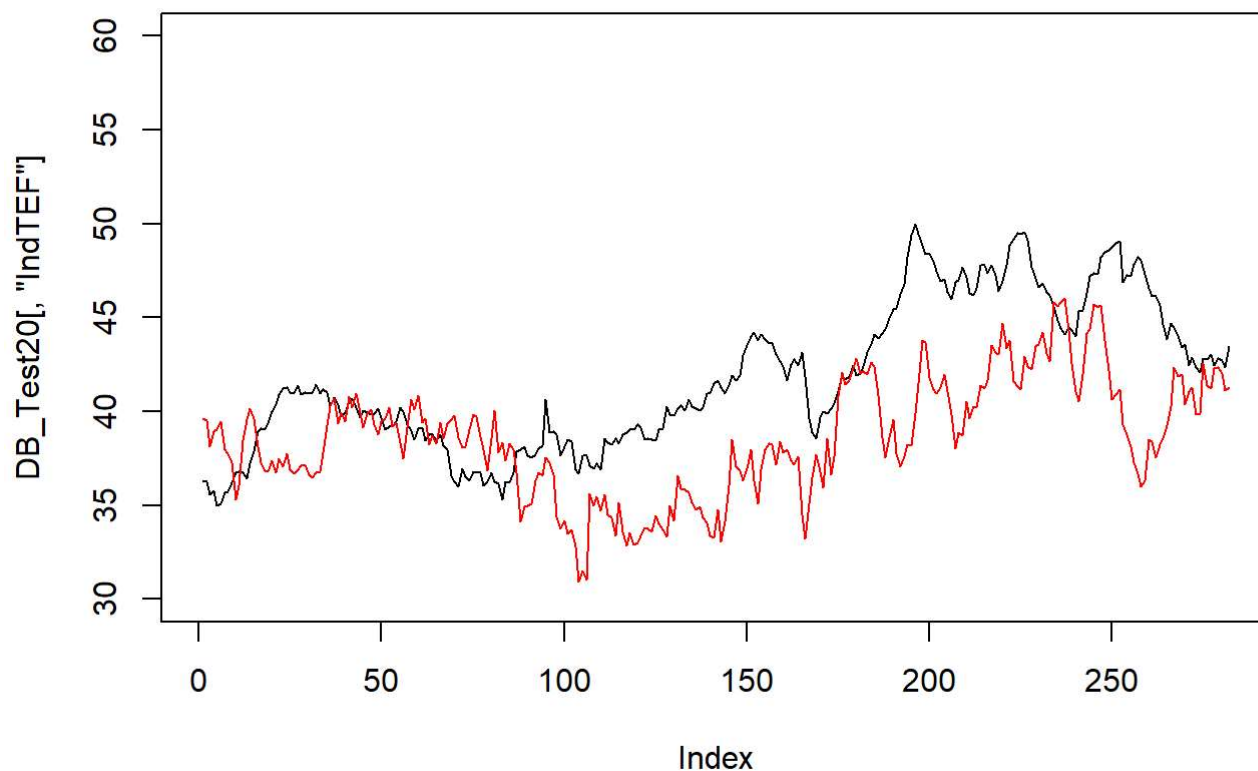
```
pred10_nnet_80Train_Max10 = predict(object = red10_80Train_Max10, newdata = DB_Test20[,c("IndSAN
_lag5", "IndNTGY_lag20", "IndSCYR_lag30", "IndIDR_lag30", "IndIDR_lag5", "IndSAN_lag8", "IndNTGY_lag1
3", "IndFER_lag5", "IndFER_lag8", "IndTEF_lag20")])
head(pred10_nnet_80Train_Max10)
```

```
##          [,1]
## 1127 39.60489
## 1128 39.50389
## 1129 38.13067
## 1130 38.94620
## 1131 39.04446
## 1132 39.48633
```

```
ecm10_80Train_Max10 = mean((DB_Test20[, "IndTEF"]-pred10_nnet_80Train_Max10)^2)
ecm10_80Train_Max10
```

```
## [1] 22.11952
```

```
plot(DB_Test20[, "IndTEF"], type = "l", ylim = c(30,60) )
lines(pred10_nnet_80Train_Max10, type = "l", col="red")
```



90% training

```
set.seed(13)
red10_90Train_Max10 = nnet(formula = IndTEF ~ IndSAN_lag5+IndNTGY_lag20+IndSCYR_lag30+IndIDR_lag
30+IndIDR_lag5+IndSAN_lag8+IndNTGY_lag13+IndFER_lag5+IndFER_lag8+IndTEF_lag20, data=DB_Training9
0, size = 10, linout = TRUE, maxit = 1000)
```



```
## # weights: 121
## initial value 5329235.790071
## iter 10 value 495811.915459
## iter 20 value 486976.858041
## iter 30 value 65755.409154
## iter 40 value 35178.719049
## iter 50 value 19524.163381
## iter 60 value 17521.825737
## iter 70 value 16786.876595
## iter 80 value 16601.309930
## iter 90 value 16482.096232
## iter 100 value 16436.110070
## iter 110 value 16405.477163
## iter 120 value 16389.069165
## iter 130 value 16378.415982
## iter 140 value 16367.424899
## iter 150 value 16361.644611
## iter 160 value 16357.159518
## iter 170 value 16353.967074
## iter 180 value 16351.183580
## iter 190 value 16349.508740
## iter 200 value 16347.784000
## iter 210 value 16346.554582
## iter 220 value 16345.660642
## iter 230 value 16345.076058
## iter 240 value 16344.435102
## iter 250 value 16344.073806
## final value 16344.059184
## converged
```

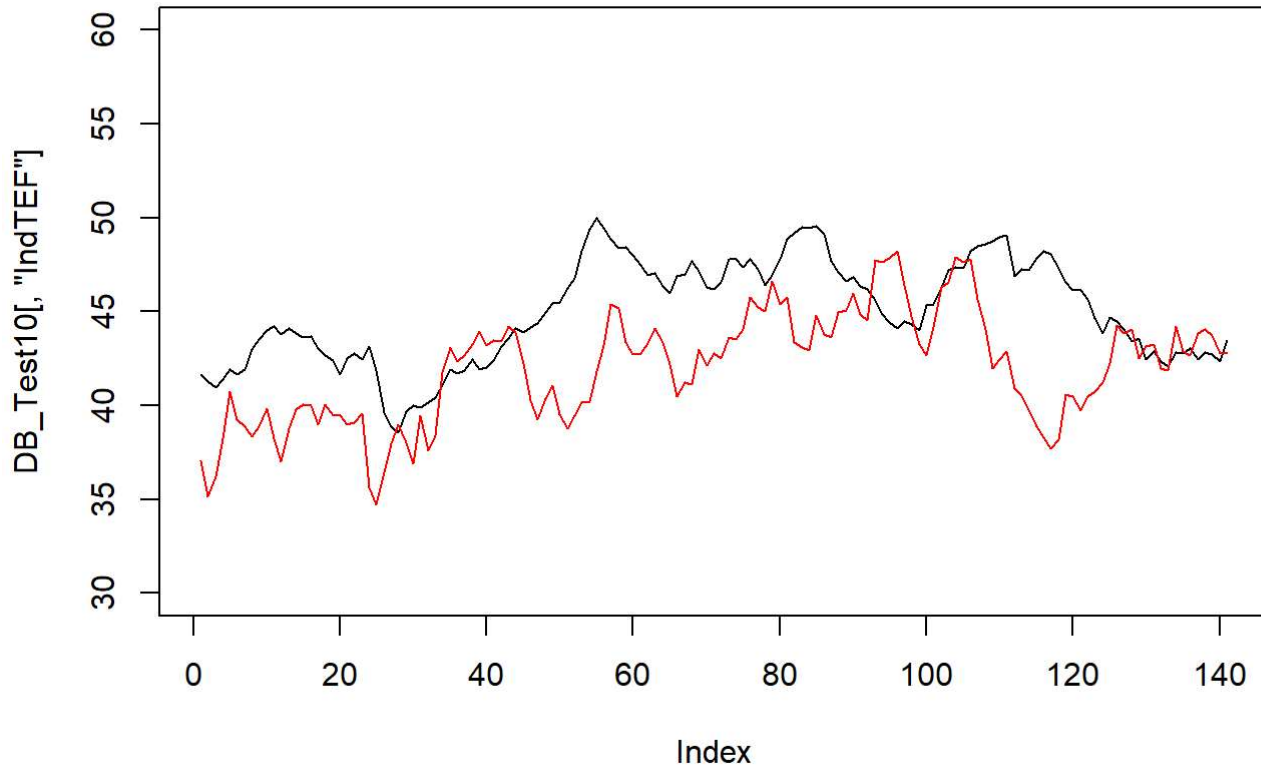
```
pred10_nnet_90Train_Max10 = predict(object = red10_90Train_Max10, newdata = DB_Test10[,c("IndSAN_lag5", "IndNTGY_lag20", "IndSCYR_lag30", "IndIDR_lag30", "IndIDR_lag5", "IndSAN_lag8", "IndNTGY_lag13", "IndFER_lag5", "IndFER_lag8", "IndTEF_lag20")])
head(pred10_nnet_90Train_Max10)
```

```
##           [,1]
## 1268 37.07663
## 1269 35.16322
## 1270 36.18983
## 1271 38.21446
## 1272 40.73227
## 1273 39.25819
```

```
ecm10_90Train_Max10 = mean((DB_Test10[, "IndTEF"] - pred10_nnet_80Train_Max10)^2)
ecm10_90Train_Max10
```

```
## [1] 55.01254
```

```
plot(DB_Test10[,"IndTEF"], type = "l",ylim = c(30,60) )
lines(pred10_nnet_90Train_Max10, type = "l", col="red")
```



b. MAX(15)

80% training

```
set.seed(12)
red10_80Train_Max15 = nnet(formula = IndTEF ~ IndSAN_lag5+IndNTGY_lag20+IndSCYR_lag30+IndIDR_lag
30+IndIDR_lag5+IndSAN_lag8+IndNTGY_lag13+IndFER_lag5+IndFER_lag8+IndTEF_lag20+IndVIS_lag5+IndMAP
_lag5+IndSCYR_lag13+IndMRL_lag20+IndVIS_lag20, data=DB_Training80, size = 10, linout = TRUE, max
it = 1000)
```

```
## # weights: 171
## initial value 4987398.165026
## iter 10 value 530062.069782
## iter 20 value 309112.347856
## iter 30 value 206228.617163
## iter 40 value 194712.564787
## iter 50 value 194547.179381
## final value 194546.588882
## converged
```

```

pred10_nnet_80Train_Max15 = predict(object = red10_80Train_Max10, newdata = DB_Test20[,c("IndSAN_lag5", "IndNTGY_lag20", "IndSCYR_lag30", "IndIDR_lag30", "IndIDR_lag5", "IndSAN_lag8", "IndNTGY_lag13", "IndFER_lag5", "IndFER_lag8", "IndTEF_lag20", "IndVIS_lag5", "IndMAP_lag5", "IndSCYR_lag13", "IndMRL_lag20", "IndVIS_lag20")])
head(pred10_nnet_80Train_Max15)

```

```

##           [,1]
## 1127 39.60489
## 1128 39.50389
## 1129 38.13067
## 1130 38.94620
## 1131 39.04446
## 1132 39.48633

```

```

ecm10_80Train_Max15 = mean((DB_Test20[,"IndTEF"]-pred10_nnet_80Train_Max15)^2)
ecm10_80Train_Max15

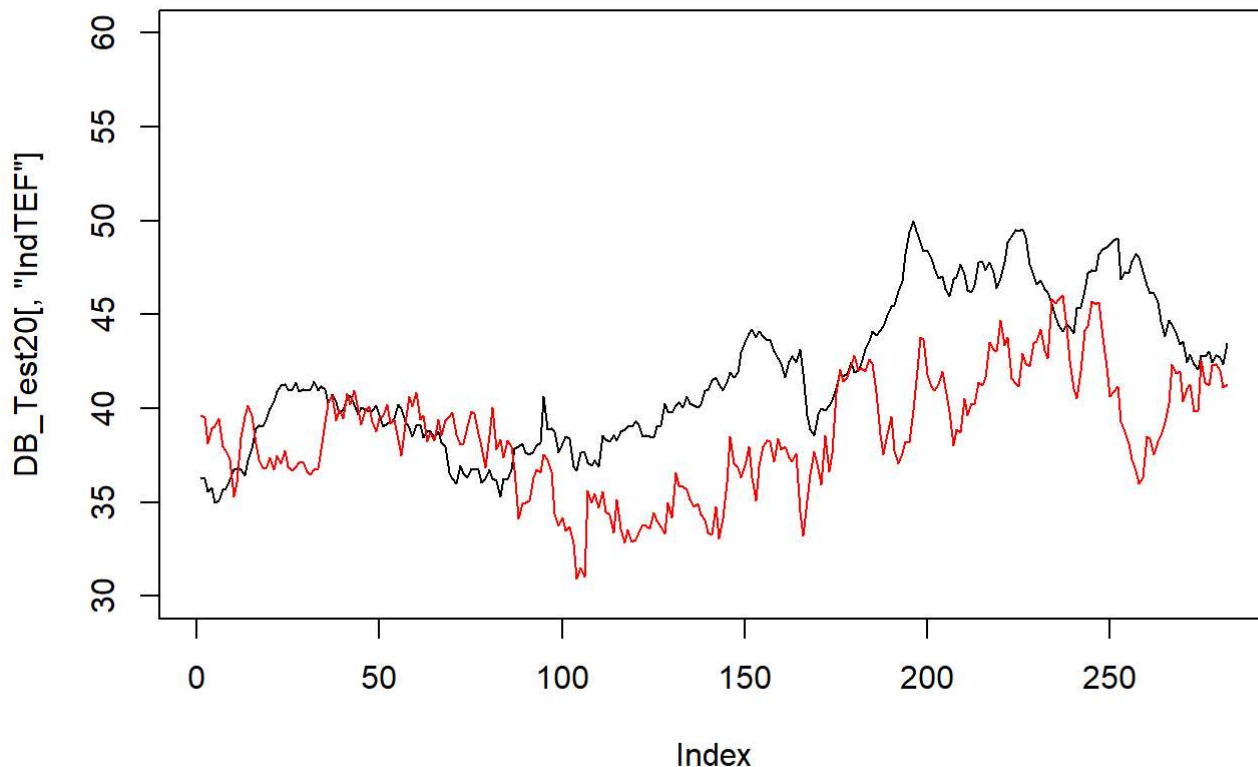
```

```
## [1] 22.11952
```

```

plot(DB_Test20[, "IndTEF"], type = "l", ylim = c(30, 60) )
lines(pred10_nnet_80Train_Max15, type = "l", col = "red")

```



```
set.seed(12)
red10_90Train_Max15 = nnet(formula = IndTEF ~ IndSAN_lag5+IndNTGY_lag20+IndSCYR_lag30+IndIDR_lag
30+IndIDR_lag5+IndSAN_lag8+IndNTGY_lag13+IndFER_lag5+IndFER_lag8+IndTEF_lag20+IndVIS_lag5+IndMAP
_lag5+IndSCYR_lag13+IndMRL_lag20+IndVIS_lag20, data=DB_Training90, size = 10, linout = TRUE, max
it = 1000)
```

```
## # weights: 171
## initial value 5191200.925984
## final value 645927.197390
## converged
```

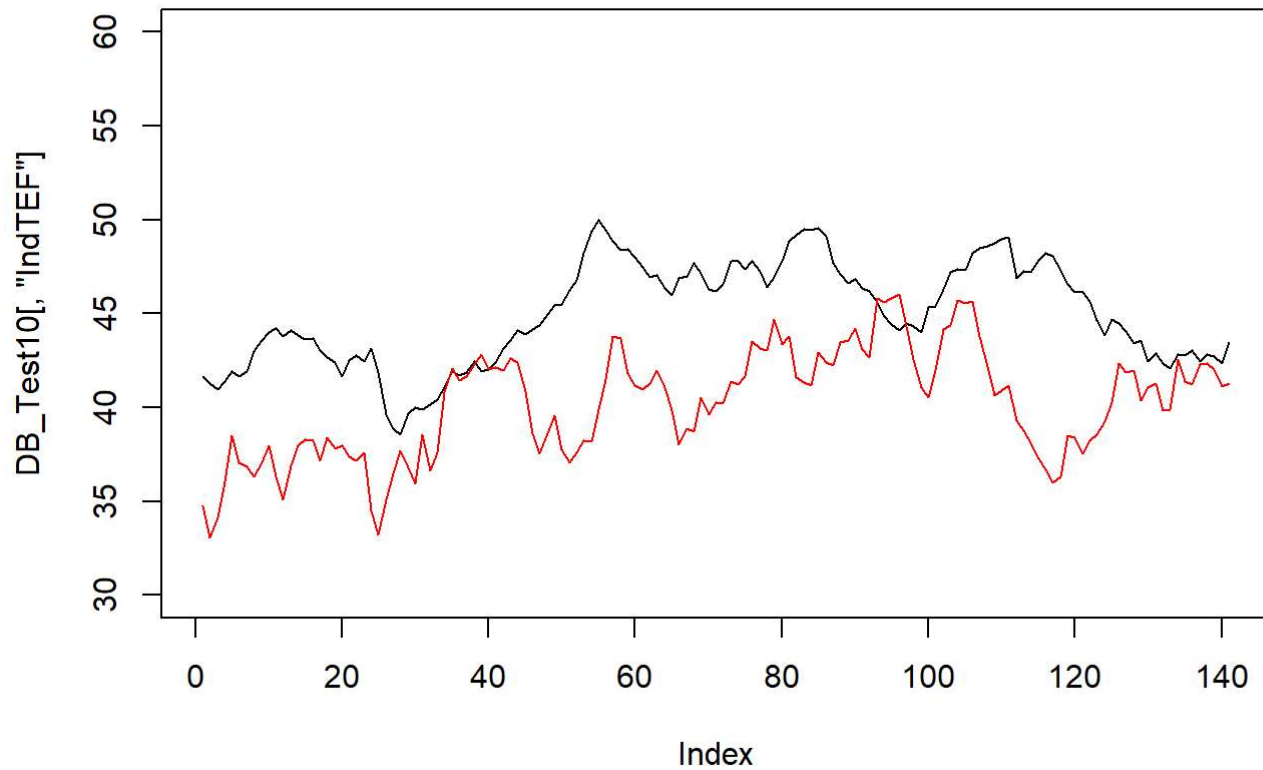
```
pred10_nnet_90Train_Max15 = predict(object = red10_80Train_Max10, newdata = DB_Test10[,c("IndSAN
_lag5", "IndNTGY_lag20", "IndSCYR_lag30", "IndIDR_lag30", "IndIDR_lag5", "IndSAN_lag8", "IndNTGY_lag1
3", "IndFER_lag5", "IndFER_lag8", "IndTEF_lag20", "IndVIS_lag5", "IndMAP_lag5", "IndSCYR_lag13", "IndMR
L_lag20", "IndVIS_lag20")])
head(pred10_nnet_90Train_Max15)
```

```
##          [,1]
## 1268 34.77577
## 1269 33.06486
## 1270 34.12064
## 1271 36.03930
## 1272 38.51917
## 1273 37.05167
```

```
ecm10_90Train_Max15 = mean((DB_Test10[,"IndTEF"]-pred10_nnet_90Train_Max15)^2)
ecm10_90Train_Max15
```

```
## [1] 32.25499
```

```
plot(DB_Test10["IndTEF"], type = "l", ylim = c(30,60) )
lines(pred10_nnet_90Train_Max15, type = "l", col="red")
```



3.3 Máquinas de soporte vectorial (SVM)

```
library(e1071)
```

a. MAX(5)

80% training

```
svm_Max5_Train80 = svm(IndTEF ~ IndSAN_lag5+IndNTGY_lag20+IndSCYR_lag30+IndIDR_lag30+IndIDR_lag  
5, data=DB_Training80)  
summary(svm_Max5_Train80)
```

```
##  
## Call:  
## svm(formula = IndTEF ~ IndSAN_lag5 + IndNTGY_lag20 + IndSCYR_lag30 +  
##   IndIDR_lag30 + IndIDR_lag5, data = DB_Training80)  
##  
##  
## Parameters:  
##   SVM-Type:  eps-regression  
##   SVM-Kernel: radial  
##     cost:  1  
##   gamma:  0.2  
##   epsilon: 0.1  
##  
##  
## Number of Support Vectors: 517
```

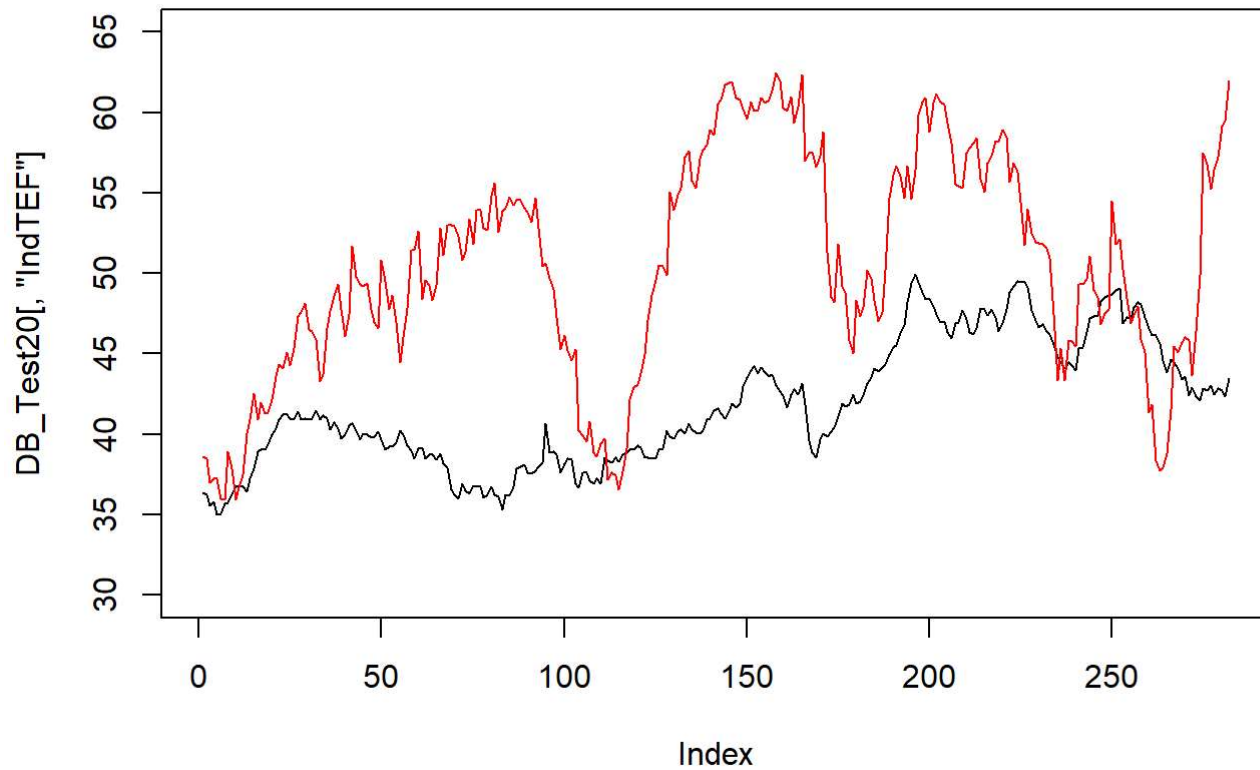
```
pred_svm_Max5_Train80 = predict(svm_Max5_Train80,newdata = DB_Test20[,c("IndSAN_lag5","IndNTGY_lag20","IndSCYR_lag30","IndIDR_lag30","IndIDR_lag5")])  
head(pred_svm_Max5_Train80)
```

```
##      1127      1128      1129      1130      1131      1132  
## 38.60077 38.45636 37.00787 37.24802 37.23370 35.98454
```

```
svm_mean_error2_80Train = mean((DB_Test20["IndTEF"]-pred_svm_Max5_Train80)^2)  
svm_mean_error2_80Train
```

```
## [1] 116.0402
```

```
plot(DB_Test20["IndTEF"], type = "l",ylim = c(30,65) )  
lines(pred_svm_Max5_Train80, type = "l", col="red")
```



90% training

```
svm_Max5_Train90 = svm(IndTEF ~ IndSAN_lag5+IndNTGY_lag20+IndSCYR_lag30+IndIDR_lag30+IndIDR_lag
5, data=DB_Training90)
summary(svm_Max5_Train90)
```

```
##
## Call:
## svm(formula = IndTEF ~ IndSAN_lag5 + IndNTGY_lag20 + IndSCYR_lag30 +
##     IndIDR_lag30 + IndIDR_lag5, data = DB_Training90)
##
##
## Parameters:
##   SVM-Type:  eps-regression
##   SVM-Kernel: radial
##     cost:    1
##     gamma:   0.2
##   epsilon:   0.1
##
##
## Number of Support Vectors:  579
```

```

pred_svm_Max5_Train90 = predict(svm_Max5_Train90,newdata = DB_Test10[,c("IndSAN_lag5","IndNTGY_lag20","IndSCYR_lag30","IndIDR_lag30","IndIDR_lag5")])
head(pred_svm_Max5_Train90)

```

```

##      1268      1269      1270      1271      1272      1273
## 45.59086 44.85200 46.81214 48.06211 47.65202 45.87055

```

```

svm_mean_error2_90Train = mean((DB_Test10[, "IndTEF"]-pred_svm_Max5_Train90)^2)
svm_mean_error2_90Train

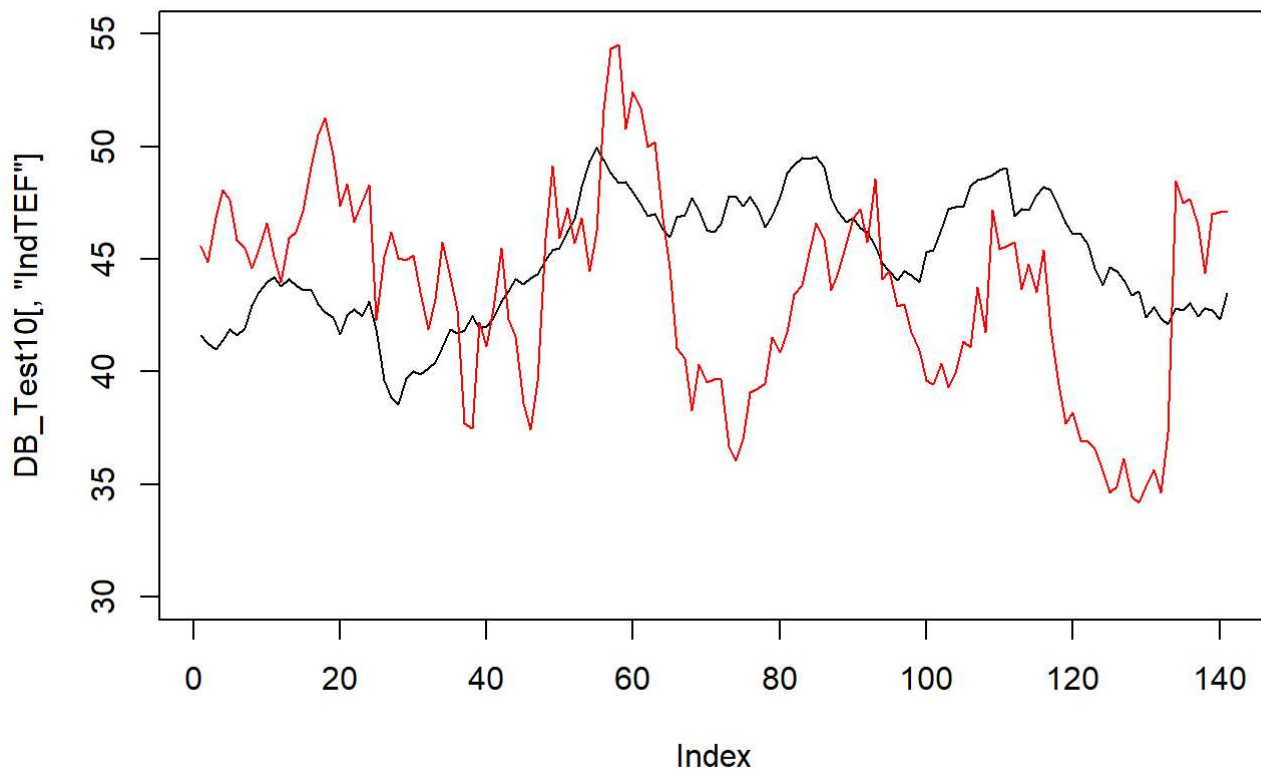
```

```
## [1] 28.83045
```

```

plot(DB_Test10[, "IndTEF"], type = "l",ylim = c(30,55) )
lines(pred_svm_Max5_Train90, type = "l", col="red")

```



a. MAX(10)

80% training

```

svm_Max10_Train80 = svm(IndTEF ~ IndSAN_lag5+IndNTGY_lag20+IndSCYR_lag30+IndIDR_lag30+IndIDR_lag5+IndSAN_lag8+IndNTGY_lag13+IndFER_lag5+IndFER_lag8+IndTEF_lag20, data=DB_Training80)
summary(svm_Max10_Train80)

```



```
##
## Call:
## svm(formula = IndTEF ~ IndSAN_lag5 + IndNTGY_lag20 + IndSCYR_lag30 +
##     IndIDR_lag30 + IndIDR_lag5 + IndSAN_lag8 + IndNTGY_lag13 + IndFER_lag5 +
##     IndFER_lag8 + IndTEF_lag20, data = DB_Training80)
##
##
## Parameters:
##   SVM-Type:  eps-regression
##   SVM-Kernel: radial
##     cost:    1
##     gamma:   0.1
##     epsilon: 0.1
##
##
## Number of Support Vectors: 327
```

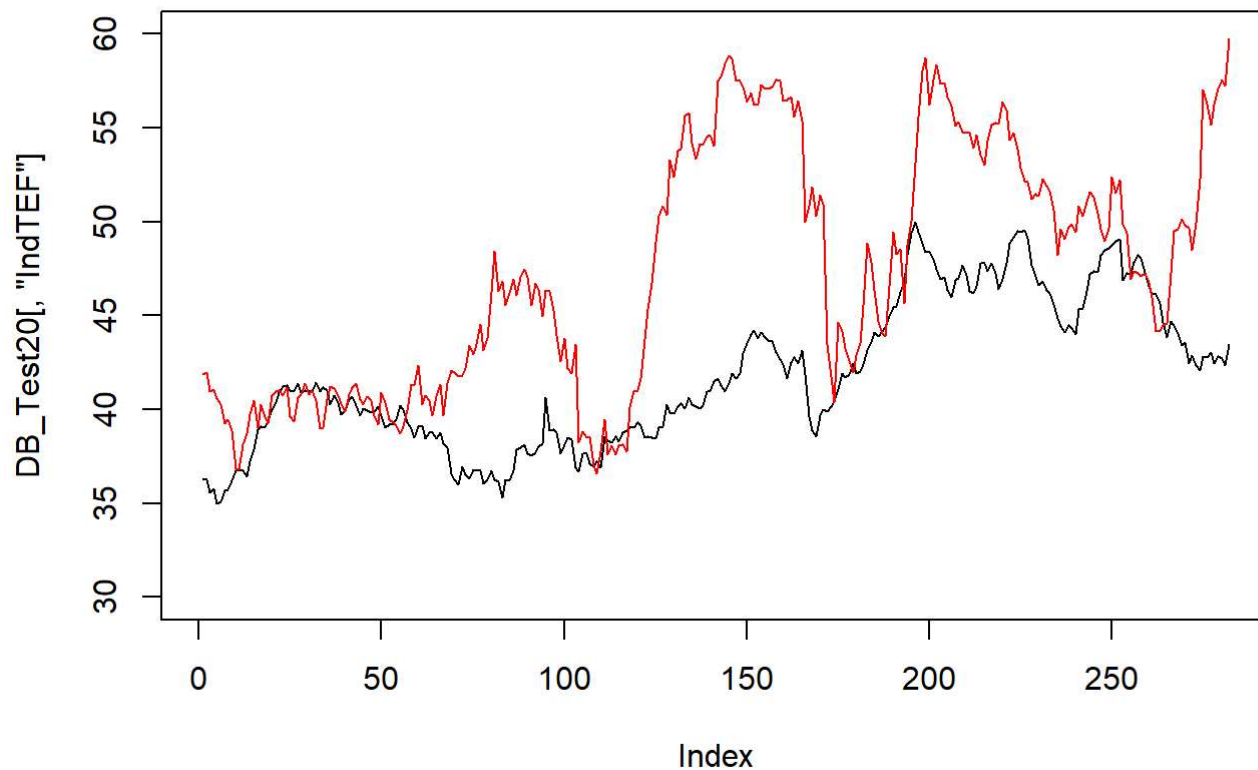
```
pred_svm_Max10_Train80 = predict(svm_Max10_Train80,newdata = DB_Test20[,c("IndSAN_lag5","IndNTGY_lag20","IndSCYR_lag30","IndIDR_lag30","IndIDR_lag5","IndSAN_lag8","IndNTGY_lag13","IndFER_lag5","IndFER_lag8","IndTEF_lag20")])
head(pred_svm_Max10_Train80)
```

```
##      1127      1128      1129      1130      1131      1132
## 41.86726 41.94837 40.94623 41.06955 40.57479 40.26728
```

```
svm_mean_error2_80Train_Max10 = mean((DB_Test20[,"IndTEF"]-pred_svm_Max10_Train80)^2)
svm_mean_error2_80Train_Max10
```

```
## [1] 56.40968
```

```
plot(DB_Test20[,"IndTEF"], type = "l",ylim = c(30,60) )
lines(pred_svm_Max10_Train80, type = "l", col="red")
```



90% training

```
svm_Max10_Train90 = svm(IndTEF ~ IndSAN_lag5+IndNTGY_lag20+IndSCYR_lag30+IndIDR_lag30+IndIDR_lag
5+IndSAN_lag8+IndNTGY_lag13+IndFER_lag5+IndFER_lag8+IndTEF_lag20, data=DB_Training90)
summary(svm_Max10_Train90)
```

```
##
## Call:
## svm(formula = IndTEF ~ IndSAN_lag5 + IndNTGY_lag20 + IndSCYR_lag30 +
##     IndIDR_lag30 + IndIDR_lag5 + IndSAN_lag8 + IndNTGY_lag13 + IndFER_lag5 +
##     IndFER_lag8 + IndTEF_lag20, data = DB_Training90)
##
##
## Parameters:
##   SVM-Type:  eps-regression
##   SVM-Kernel: radial
##     cost:    1
##    gamma:   0.1
##   epsilon:  0.1
##
##
## Number of Support Vectors:  357
```

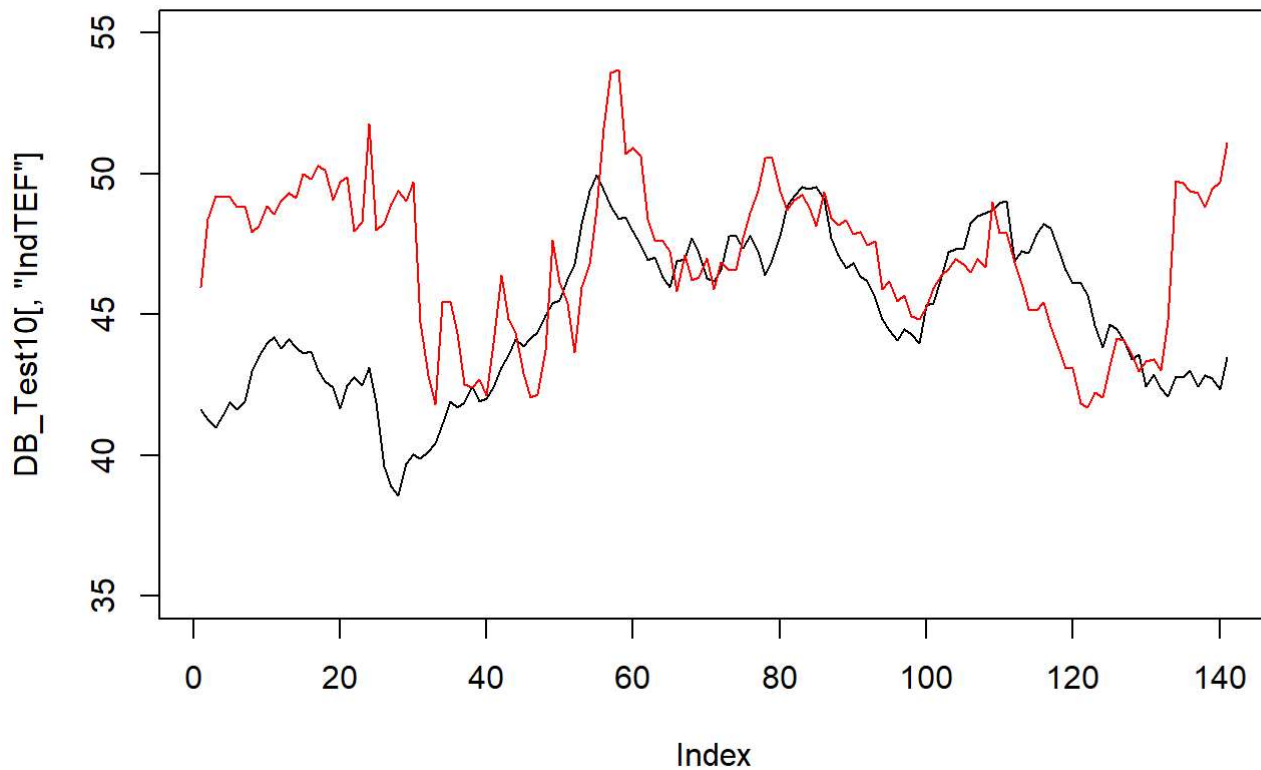
```
pred_svm_Max10_Train90 = predict(svm_Max10_Train90,newdata = DB_Test10[,c("IndSAN_lag5","IndNTGY_lag20","IndSCYR_lag30","IndIDR_lag30","IndIDR_lag5","IndSAN_lag8","IndNTGY_lag13","IndFER_lag5","IndFER_lag8","IndTEF_lag20")])
head(pred_svm_Max10_Train90)
```

```
##      1268      1269      1270      1271      1272      1273
## 45.97612 48.39181 49.20322 49.16302 49.19071 48.80953
```

```
svm_mean_error2_90Train_Max10 = mean((DB_Test10[,"IndTEF"]-pred_svm_Max10_Train90)^2)
svm_mean_error2_90Train_Max10
```

```
## [1] 16.41906
```

```
plot(DB_Test10[,"IndTEF"], type = "l",ylim = c(35,55) )
lines(pred_svm_Max10_Train90, type = "l", col="red")
```



a. MAX(15)

80% training

```
svm_Max15_Train80 = svm(IndTEF ~ IndSAN_lag5+IndNTGY_lag20+IndSCYR_lag30+IndIDR_lag30+IndIDR_lag
5+IndSAN_lag8+IndNTGY_lag13+IndFER_lag5+IndFER_lag8+IndTEF_lag20+IndVIS_lag5+IndMAP_lag5+IndSCYR
_lag13+IndMRL_lag20+IndVIS_lag20, data=DB_Training80)
summary(svm_Max15_Train80)
```

```
##
## Call:
## svm(formula = IndTEF ~ IndSAN_lag5 + IndNTGY_lag20 + IndSCYR_lag30 +
##     IndIDR_lag30 + IndIDR_lag5 + IndSAN_lag8 + IndNTGY_lag13 + IndFER_lag5 +
##     IndFER_lag8 + IndTEF_lag20 + IndVIS_lag5 + IndMAP_lag5 + IndSCYR_lag13 +
##     IndMRL_lag20 + IndVIS_lag20, data = DB_Training80)
##
##
## Parameters:
##   SVM-Type:  eps-regression
##   SVM-Kernel: radial
##     cost:    1
##     gamma:   0.06666667
##     epsilon: 0.1
##
##
## Number of Support Vectors: 224
```

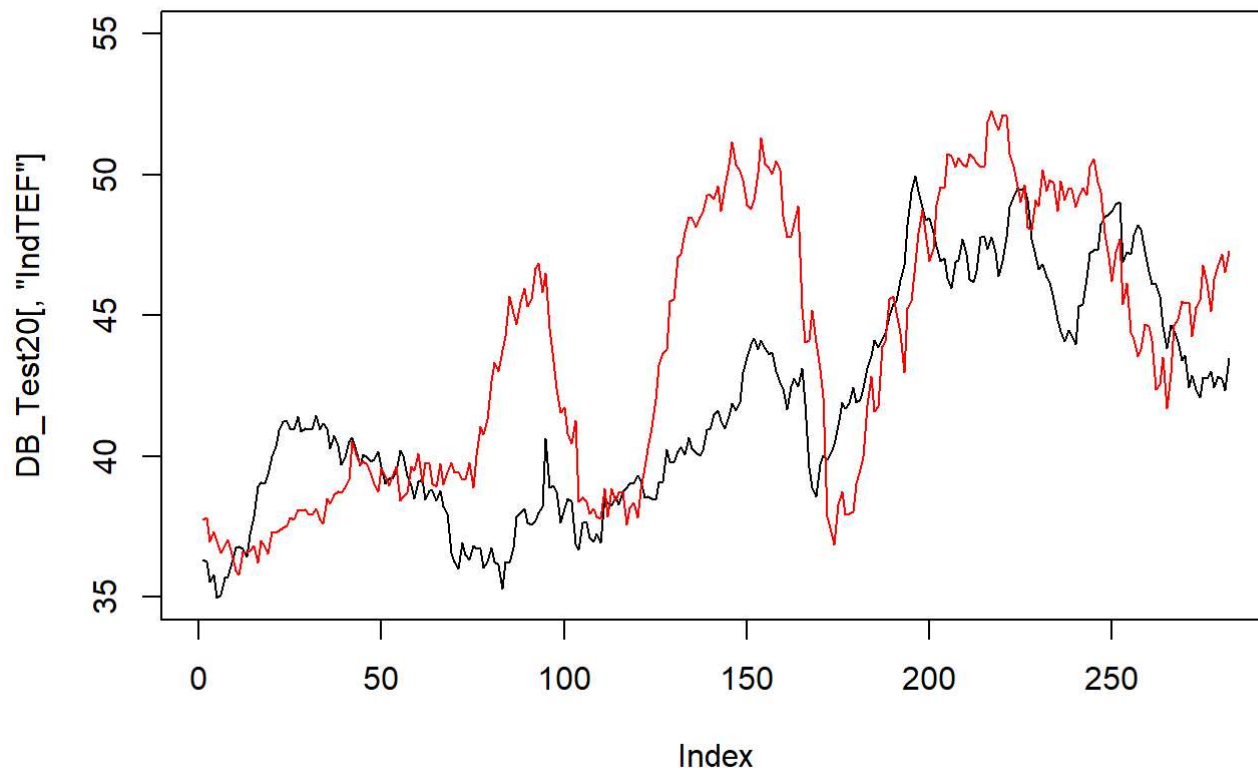
```
pred_svm_Max15_Train80 = predict(svm_Max15_Train80,newdata = DB_Test20[,c("IndSAN_lag5","IndNTGY
_lag20","IndSCYR_lag30","IndIDR_lag30","IndIDR_lag5","IndSAN_lag8","IndNTGY_lag13","IndFER_lag5"
,"IndFER_lag8","IndTEF_lag20","IndVIS_lag5","IndMAP_lag5","IndSCYR_lag13","IndMRL_lag20","IndVIS
_lag20")])
head(pred_svm_Max15_Train80)
```

```
##      1127      1128      1129      1130      1131      1132
## 37.74213 37.82783 36.96407 37.33632 37.00745 36.56902
```

```
svm_mean_error2_80Train_Max15 = mean((DB_Test20[,"IndTEF"]-pred_svm_Max15_Train80)^2)
svm_mean_error2_80Train_Max15
```

```
## [1] 16.31971
```

```
plot(DB_Test20[,"IndTEF"], type = "l",ylim = c(35,55) )
lines(pred_svm_Max15_Train80, type = "l", col="red")
```



90% training

```
svm_Max15_Train90 = svm(IndTEF ~ IndSAN_lag5+IndNTGY_lag20+IndSCYR_lag30+IndIDR_lag30+IndIDR_lag
5+IndSAN_lag8+IndNTGY_lag13+IndFER_lag5+IndFER_lag8+IndTEF_lag20+IndVIS_lag5+IndMAP_lag5+IndSCYR
_lag13+IndMRL_lag20+IndVIS_lag20, data=DB_Training90)
summary(svm_Max15_Train90)
```

```
##
## Call:
## svm(formula = IndTEF ~ IndSAN_lag5 + IndNTGY_lag20 + IndSCYR_lag30 +
##     IndIDR_lag30 + IndIDR_lag5 + IndSAN_lag8 + IndNTGY_lag13 + IndFER_lag5 +
##     IndFER_lag8 + IndTEF_lag20 + IndVIS_lag5 + IndMAP_lag5 + IndSCYR_lag13 +
##     IndMRL_lag20 + IndVIS_lag20, data = DB_Training90)
##
##
## Parameters:
##   SVM-Type:  eps-regression
##   SVM-Kernel: radial
##     cost: 1
##     gamma: 0.06666667
##     epsilon: 0.1
##
##
## Number of Support Vectors: 238
```

```

pred_svm_Max15_Train90 = predict(svm_Max15_Train90,newdata = DB_Test10[,c("IndSAN_lag5","IndNTGY_lag20","IndSCYR_lag30","IndIDR_lag30","IndIDR_lag5","IndSAN_lag8","IndNTGY_lag13","IndFER_lag5","IndFER_lag8","IndTEF_lag20","IndVIS_lag5","IndMAP_lag5","IndSCYR_lag13","IndMRL_lag20","IndVIS_lag20")])
head(pred_svm_Max15_Train90)

```

```

##      1268      1269      1270      1271      1272      1273
## 43.74436 43.66411 44.44773 45.36919 46.42319 46.02662

```

```

svm_mean_error2_90Train_Max15 = mean((DB_Test10["IndTEF"]-pred_svm_Max10_Train90)^2)
svm_mean_error2_90Train_Max15

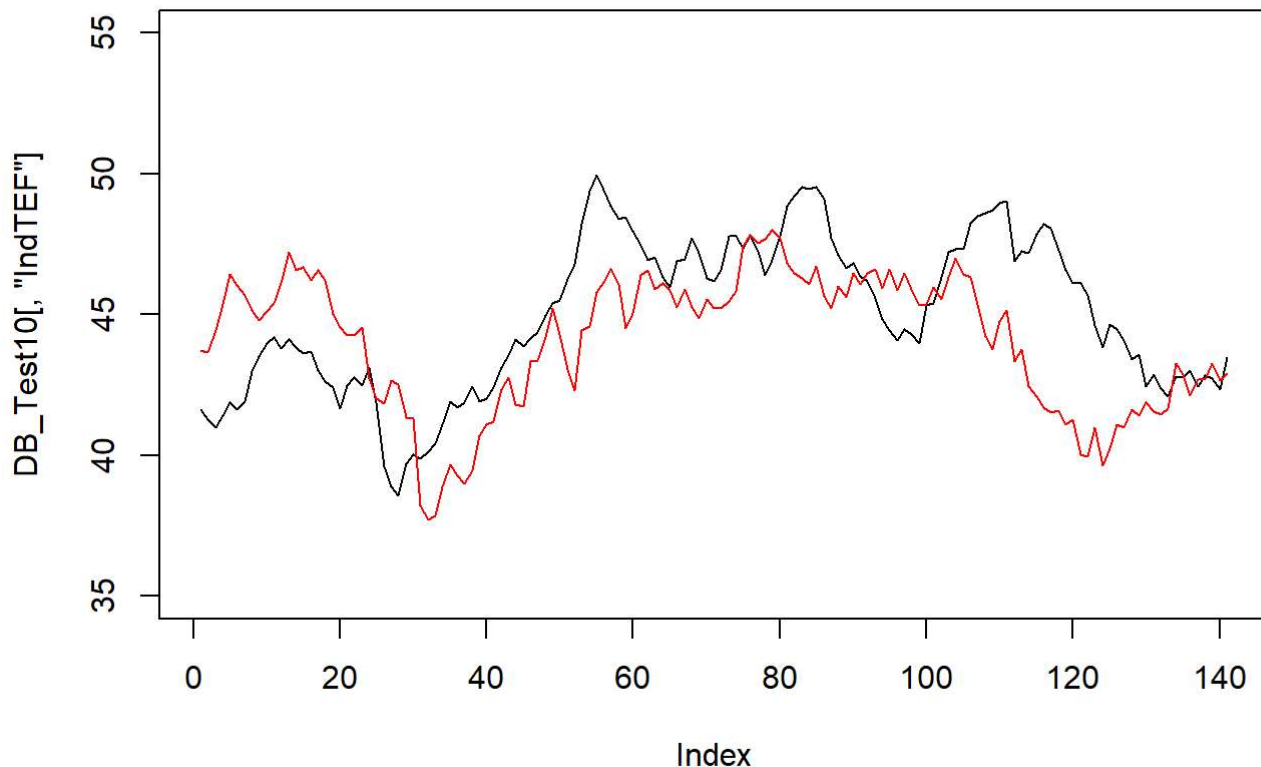
```

```
## [1] 16.41906
```

```

plot(DB_Test10["IndTEF"], type = "l",ylim = c(35,55) )
lines(pred_svm_Max15_Train90, type = "l", col="red")

```



4. Elegir los mejores modelos para la simulación y Extraer los datos para hacer simulación en Excel

4.1 Regresión Lineal – Max(10) 90% Training

```
# plot(predict_lm_MAX10_IndTEF90["fit"], type = "l")  
write.csv(predict_lm_MAX10_IndTEF90["fit"], file="Predict_lm_MAX10_IndTEF90.csv", row.names = FALSE)
```

4.2 Redes neuronales - Max(10) con 80% training

```
write.csv(pred10_nnet_80Train_Max10, file="Pred10_nnet_80Train_Max10.csv", row.names = FALSE)
```

4.3 Redes neuronales - Max(15) con 80% training

```
write.csv(pred10_nnet_80Train_Max15, file="pred10_nnet_80Train_Max15.csv", row.names = FALSE)
```

4.4 Máquinas de soporte vectorial (SVM) - Max(10) con 90% training

```
write.csv(pred_svm_Max10_Train90, file="Pred_svm_Max10_Train90.csv", row.names = FALSE)
```

4.5 Máquinas de soporte vectorial (SVM) - Max(15) con 90% training

```
write.csv(pred_svm_Max15_Train90, file="Pred_svm_Max15_Train90.csv", row.names = FALSE)
```


1334	46.97	41.2	46.9	-0.12	Vender	-	21,009	0 €	21,009	5% Vender	22,120	0 €	22,120	10% Esperar	Vender	20,000	0 €	20,000	0 Esperar
1335	47.7	41.2	47.0	-0.12	Vender	-	21,009	0 €	21,009	5% Vender	22,120	0 €	22,120	10% Esperar	Vender	20,000	0 €	20,000	0 Esperar
1336	47.17	43.0	46.3	-0.07	Vender	-	21,009	0 €	21,009	5% Vender	22,120	0 €	22,120	10% Esperar	Esperar	20,000	0 €	20,000	0 Esperar
1337	46.31	42.2	46.0	-0.08	Vender	-	21,009	0 €	21,009	5% Vender	22,120	0 €	22,120	10% Esperar	Esperar	20,000	0 €	20,000	0 Esperar
1338	46.19	42.8	46.9	-0.09	Vender	-	21,009	0 €	21,009	5% Vender	22,120	0 €	22,120	10% Esperar	Esperar	20,000	0 €	20,000	0 Esperar
1339	46.57	42.5	47.0	-0.09	Vender	-	21,009	0 €	21,009	5% Vender	22,120	0 €	22,120	10% Esperar	Esperar	20,000	0 €	20,000	0 Esperar
1340	47.8	43.7	47.7	-0.08	Vender	-	21,009	0 €	21,009	5% Vender	22,120	0 €	22,120	10% Esperar	Esperar	20,000	0 €	20,000	0 Esperar
1341	47.77	43.6	47.2	-0.08	Vender	-	21,009	0 €	21,009	5% Vender	22,120	0 €	22,120	10% Esperar	Esperar	20,000	0 €	20,000	0 Esperar
1342	47.37	44.1	46.3	-0.05	Esperar	-	21,009	0 €	21,009	5% Vender	22,120	0 €	22,120	10% Esperar	Esperar	20,000	0 €	20,000	0 Esperar
1343	47.78	45.8	46.2	-0.01	Esperar	-	21,009	0 €	21,009	5% Vender	22,120	0 €	22,120	10% Esperar	Esperar	20,000	0 €	20,000	0 Esperar
1344	47.25	45.3	46.6	-0.03	Esperar	-	21,009	0 €	21,009	5% Vender	22,120	0 €	22,120	10% Esperar	Esperar	20,000	0 €	20,000	0 Esperar
1345	46.41	45.1	47.8	-0.06	Vender	-	21,009	0 €	21,009	5% Vender	22,120	0 €	22,120	10% Esperar	Esperar	20,000	0 €	20,000	0 Esperar
1346	46.92	46.7	47.8	-0.02	Esperar	-	21,009	0 €	21,009	5% Vender	22,120	0 €	22,120	10% Esperar	Esperar	20,000	0 €	20,000	0 Esperar
1347	47.78	45.4	47.4	-0.04	Esperar	-	21,009	0 €	21,009	5% Vender	22,120	0 €	22,120	10% Esperar	Esperar	20,000	0 €	20,000	0 Esperar
1348	48.85	45.8	47.8	-0.04	Esperar	-	21,009	0 €	21,009	5% Vender	22,120	0 €	22,120	10% Esperar	Esperar	20,000	0 €	20,000	0 Esperar
1349	49.2	43.4	47.3	-0.08	Vender	-	21,009	0 €	21,009	5% Vender	22,120	0 €	22,120	10% Esperar	Esperar	20,000	0 €	20,000	0 Esperar
1350	49.52	43.1	46.4	-0.07	Vender	-	21,009	0 €	21,009	5% Vender	22,120	0 €	22,120	10% Esperar	Esperar	20,000	0 €	20,000	0 Esperar
1351	49.47	43.0	46.9	-0.08	Vender	-	21,009	0 €	21,009	5% Vender	22,120	0 €	22,120	10% Esperar	Esperar	20,000	0 €	20,000	0 Esperar
1352	49.53	44.9	47.8	-0.06	Vender	-	21,009	0 €	21,009	5% Vender	22,120	0 €	22,120	10% Esperar	Esperar	20,000	0 €	20,000	0 Esperar
1353	49.12	43.8	48.9	-0.10	Vender	-	21,009	0 €	21,009	5% Vender	22,120	0 €	22,120	10% Esperar	Vender	20,000	0 €	20,000	0 Esperar
1354	47.7	43.7	49.2	-0.11	Vender	-	21,009	0 €	21,009	5% Vender	22,120	0 €	22,120	10% Esperar	Vender	20,000	0 €	20,000	0 Esperar
1355	47.11	45.0	49.5	-0.09	Vender	-	21,009	0 €	21,009	5% Vender	22,120	0 €	22,120	10% Esperar	Esperar	20,000	0 €	20,000	0 Esperar
1356	46.64	45.1	49.5	-0.09	Vender	-	21,009	0 €	21,009	5% Vender	22,120	0 €	22,120	10% Esperar	Esperar	20,000	0 €	20,000	0 Esperar
1357	46.82	46.0	49.5	-0.07	Vender	-	21,009	0 €	21,009	5% Vender	22,120	0 €	22,120	10% Esperar	Esperar	20,000	0 €	20,000	0 Esperar
1358	46.37	44.9	49.1	-0.09	Vender	-	21,009	0 €	21,009	5% Vender	22,120	0 €	22,120	10% Esperar	Esperar	20,000	0 €	20,000	0 Esperar
1359	46.19	44.5	47.7	-0.07	Vender	-	21,009	0 €	21,009	5% Vender	22,120	0 €	22,120	10% Esperar	Esperar	20,000	0 €	20,000	0 Esperar
1360	45.58	47.8	47.1	0.01	Esperar	-	21,009	0 €	21,009	5% Vender	22,120	0 €	22,120	10% Esperar	Esperar	20,000	0 €	20,000	0 Esperar
1361	44.83	47.7	46.6	0.02	Esperar	-	21,009	0 €	21,009	5% Vender	22,120	0 €	22,120	10% Esperar	Esperar	20,000	0 €	20,000	0 Esperar
1362	44.45	47.9	46.8	0.02	Esperar	-	21,009	0 €	21,009	5% Vender	22,120	0 €	22,120	10% Esperar	Esperar	20,000	0 €	20,000	0 Esperar
1363	44.09	48.2	46.4	0.04	Esperar	-	21,009	0 €	21,009	5% Vender	22,120	0 €	22,120	10% Esperar	Esperar	20,000	0 €	20,000	0 Esperar
1364	44.48	46.4	46.2	0.01	Esperar	-	21,009	0 €	21,009	5% Vender	22,120	0 €	22,120	10% Esperar	Esperar	20,000	0 €	20,000	0 Esperar
1365	44.3	44.8	45.6	-0.02	Esperar	-	21,009	0 €	21,009	5% Vender	22,120	0 €	22,120	10% Esperar	Esperar	20,000	0 €	20,000	0 Esperar
1366	43.98	43.3	44.8	-0.03	Esperar	-	21,009	0 €	21,009	5% Vender	22,120	0 €	22,120	10% Esperar	Esperar	20,000	0 €	20,000	0 Esperar
1367	45.32	42.7	44.5	-0.04	Esperar	-	21,009	0 €	21,009	5% Vender	22,120	0 €	22,120	10% Esperar	Esperar	20,000	0 €	20,000	0 Esperar
1368	45.4	44.2	44.1	0.00	Esperar	-	21,009	0 €	21,009	5% Vender	22,120	0 €	22,120	10% Esperar	Esperar	20,000	0 €	20,000	0 Esperar
1369	46.32	46.3	44.5	0.04	Esperar	-	21,009	0 €	21,009	5% Vender	22,120	0 €	22,120	10% Esperar	Esperar	20,000	0 €	20,000	0 Esperar
1370	47.21	46.6	44.3	0.05	Comprar	Comprar	€ -	445 €	21,009	0% Esperar	€ -	469	22,120	0% Esperar	Esperar	20,000	0 €	20,000	0 Esperar
1371	47.34	47.9	44.0	0.09	Comprar	-	€ -	445 €	21,067	0% Esperar	€ -	469	22,180	0% Esperar	Esperar	20,000	0 €	20,000	0 Esperar
1372	47.33	47.6	45.3	0.05	Comprar	-	€ -	445 €	21,063	0% Esperar	€ -	469	22,176	0% Esperar	Esperar	20,000	0 €	20,000	0 Esperar
1373	48.24	47.8	45.4	0.05	Comprar	-	€ -	445 €	21,468	2% Esperar	€ -	469	22,602	2% Esperar	Esperar	20,000	0 €	20,000	0 Esperar

1374	48.51	45.7	46.3	-0.01	Esperar	-	€ -	445	€ 21,588	21,588	3% Esperar	€ -	469	22,729	22,729	3% Esperar	Esperar	20,000	0	€ -	20,000	0	Esperar
1375	48.62	44.1	47.2	-0.07	Vender	-	€ -	445	€ 21,637	21,637	3% Esperar	€ -	469	22,780	22,780	3% Esperar	Esperar	20,000	0	€ -	20,000	0	Esperar
1376	48.73	42.0	47.3	-0.11	Vender	-	€ -	445	€ 21,686	21,686	3% Esperar	€ -	469	22,832	22,832	3% Esperar	Vender	20,000	0	€ -	20,000	0	Esperar
1377	48.97	42.5	47.3	-0.10	Vender	-	€ -	445	€ 21,792	21,792	4% Esperar	€ -	469	22,944	22,944	4% Esperar	Vender	20,000	0	€ -	20,000	0	Esperar
1378	49.05	42.9	48.2	-0.11	Vender	-	€ -	445	€ 21,828	21,828	4% Esperar	€ -	469	22,982	22,982	4% Esperar	Vender	20,000	0	€ -	20,000	0	Esperar
1379	46.9	40.9	48.5	-0.16	Vender	-	€ -	445	€ 20,871	20,871	-1% Esperar	€ -	469	21,974	21,974	-1% Esperar	Vender	20,000	0	€ -	20,000	0	Esperar
1380	47.24	40.6	48.6	-0.17	Vender	-	€ -	445	€ 21,023	21,023	0% Esperar	€ -	469	22,134	22,134	0% Esperar	Vender	20,000	0	€ -	20,000	0	Esperar
1381	47.2	39.8	48.7	-0.18	Vender	-	€ -	445	€ 21,005	21,005	0% Esperar	€ -	469	22,115	22,115	0% Esperar	Vender	20,000	0	€ -	20,000	0	Esperar
1382	47.87	38.9	49.0	-0.21	Vender	-	€ -	445	€ 21,303	21,303	1% Esperar	€ -	469	22,429	22,429	1% Esperar	Vender	20,000	0	€ -	20,000	0	Esperar
1383	48.2	38.3	49.1	-0.22	Vender	-	€ -	445	€ 21,450	21,450	2% Esperar	€ -	469	22,583	22,583	2% Esperar	Vender	20,000	0	€ -	20,000	0	Esperar
1384	48.08	37.8	46.9	-0.19	Vender	-	€ -	445	€ 21,396	21,396	2% Esperar	€ -	469	22,527	22,527	2% Esperar	Vender	20,000	0	€ -	20,000	0	Esperar
1385	47.3	38.2	47.2	-0.19	Vender	-	€ -	445	€ 21,049	21,049	0% Esperar	€ -	469	22,162	22,162	0% Esperar	Vender	20,000	0	€ -	20,000	0	Esperar
1386	46.59	40.6	47.2	-0.14	Vender	-	€ -	445	€ 20,733	20,733	-1% Esperar	€ -	469	21,829	21,829	-1% Esperar	Vender	20,000	0	€ -	20,000	0	Esperar
1387	46.13	40.5	47.9	-0.15	Vender	-	€ -	445	€ 20,529	20,529	-2% Esperar	€ -	469	21,614	21,614	-2% Esperar	Vender	20,000	0	€ -	20,000	0	Esperar
1388	46.12	39.7	48.2	-0.18	Vender	-	€ -	445	€ 20,524	20,524	-2% Esperar	€ -	469	21,609	21,609	-2% Esperar	Vender	20,000	0	€ -	20,000	0	Esperar

Triger = 0.05, MaxPerd =0.05 y MaxGan = 0.05
Triger = 0.05, MaxPerd =0.05 y MaxGan = 0.1

€ 20,524.11
€ 21,608.84

Triger = 0.1, MaxPerd =0.05 y MaxGan = 0.05
Triger = 0.1, MaxPerd =0.05 y MaxGan = 0.1

€ 20,000.00
€ 20,000.00

1198	36.9	38.1	38.2	0.00	Esperar	-	21,431	0 €	21,431	0.00	Esperar	-	21,377	0 €	21,377	0.00	Esperar	-	22,157	0 €	22,157	0.00	Esperar						
1199	36.5	38.1	38.0	0.00	Esperar	-	21,431	0 €	21,431	0.00	Esperar	-	21,377	0 €	21,377	0.00	Esperar	-	22,157	0 €	22,157	0.00	Esperar						
1200	36.3	39.1	36.6	0.07	Comprar	Comprar	€	590	21,431	21,431	0.00	Esperar	Compra	€	21,377	0 €	21,377	0.00	Esperar	-	22,157	0 €	22,157	0.00	Esperar				
1201	36.8	39.8	36.3	0.10	Comprar	-	€	590	21,717	21,717	0.01	Esperar	-	€	21,377	0 €	21,377	0.00	Esperar	-	22,157	0 €	22,157	0.00	Esperar				
1202	36.7	39.7	36.0	0.10	Comprar	-	€	590	21,664	21,664	0.01	Esperar	-	€	21,377	0 €	21,377	0.00	Esperar	-	22,157	0 €	22,157	0.00	Esperar				
1203	36.7	38.7	36.9	0.05	Esperar	-	€	590	21,667	21,667	0.01	Esperar	-	€	582	21,377	21,377	0.00	Esperar	Comprar	€	603	22,157	22,157	0.00	Esperar			
1204	36.0	38.0	36.5	0.04	Esperar	-	€	590	21,254	21,254	-0.01	Esperar	-	€	582	21,380	21,380	0.00	Esperar	-	€	603	22,160	22,160	0.00	Esperar			
1205	36.3	36.9	36.3	0.02	Esperar	-	€	590	21,388	21,388	0.00	Esperar	-	€	582	20,973	20,973	-0.02	Esperar	-	€	603	21,738	21,738	-0.02	Esperar			
1206	36.8	38.6	36.8	0.05	Esperar	-	€	590	21,683	21,683	0.01	Esperar	-	€	582	21,105	21,105	-0.01	Esperar	-	€	603	21,875	21,875	-0.01	Esperar			
1207	36.3	40.0	36.7	0.09	Comprar	-	€	590	21,396	21,396	0.00	Esperar	-	€	582	21,396	21,396	0.00	Esperar	-	€	603	22,177	22,177	0.00	Esperar			
1208	36.1	37.8	36.7	0.03	Esperar	-	€	590	21,317	21,317	-0.01	Esperar	-	€	582	21,113	21,113	-0.01	Esperar	-	€	603	21,884	21,884	-0.01	Esperar			
1209	35.3	38.3	36.0	0.06	Comprar	-	€	590	20,817	20,817	-0.03	Esperar	-	€	582	21,035	21,035	-0.02	Esperar	-	€	603	21,802	21,802	-0.02	Esperar			
1210	36.3	37.4	36.3	0.03	Esperar	-	€	590	21,396	21,396	0.00	Esperar	-	€	582	20,541	20,541	-0.04	Esperar	-	€	603	21,291	21,291	-0.04	Esperar			
1211	36.2	38.3	36.8	0.04	Esperar	-	€	590	21,360	21,360	0.00	Esperar	-	€	582	21,113	21,113	-0.01	Esperar	-	€	603	21,884	21,884	-0.01	Esperar			
1212	36.8	38.0	36.3	0.05	Esperar	-	€	590	21,689	21,689	0.01	Esperar	-	€	582	21,077	21,077	-0.01	Esperar	-	€	603	21,846	21,846	-0.01	Esperar			
1213	37.8	36.7	36.1	0.03	Esperar	-	€	590	22,328	22,328	0.04	Esperar	-	€	582	21,402	21,402	0.00	Esperar	-	€	603	22,183	22,183	0.00	Esperar			
1214	38.0	34.1	35.3	-0.03	Esperar	-	€	590	22,414	22,414	0.05	Esperar	-	€	582	22,032	22,032	0.03	Esperar	-	€	603	22,837	22,837	0.03	Esperar			
1215	38.1	34.9	36.3	-0.04	Esperar	-	€	590	22,498	22,498	0.05	Esperar	-	€	582	22,201	22,201	0.04	Esperar	-	€	603	22,924	22,924	0.03	Esperar			
1216	37.6	35.0	36.2	-0.03	Esperar	-	€	590	22,206	22,206	0.04	Esperar	-	€	582	21,912	21,912	0.03	Esperar	-	€	603	23,011	23,011	0.04	Esperar			
1217	37.6	35.1	36.8	-0.05	Esperar	-	€	590	22,158	22,158	0.03	Esperar	-	€	582	21,865	21,865	0.02	Esperar	-	€	603	22,712	22,712	0.03	Esperar			
1218	37.7	36.3	37.8	-0.04	Esperar	-	€	590	22,249	22,249	0.04	Esperar	-	€	582	21,954	21,954	0.03	Esperar	-	€	603	22,663	22,663	0.02	Esperar			
1219	38.0	36.7	38.0	-0.03	Esperar	-	€	590	22,416	22,416	0.05	Esperar	-	€	582	22,120	22,120	0.03	Esperar	-	€	603	22,927	22,927	0.03	Esperar			
1220	38.2	36.6	38.1	-0.04	Esperar	Vender	22,542	0 €	22,542	0.05	Vender	-	€	582	22,243	22,243	0.04	Esperar	-	€	603	23,055	23,055	0.04	Esperar				
1221	40.6	37.6	37.6	0.00	Esperar	-	22,542	0 €	22,542	0.00	Esperar	Vender	24,639	0 €	24,639	0.12	Vender	Esperar	-	23,664	0 €	23,664	0.11	Vender	24,528	0 €	24,528	0.11	Vender
1222	38.9	37.3	37.6	-0.01	Esperar	-	22,542	0 €	22,542	0.00	Esperar	-	24,639	0 €	24,639	0.00	Esperar	Esperar	-	23,664	0 €	23,664	0.00	Esperar	24,528	0 €	24,528	0.00	Esperar
1223	38.9	36.7	37.7	-0.03	Esperar	-	22,542	0 €	22,542	0.00	Esperar	-	24,639	0 €	24,639	0.00	Esperar	Esperar	-	23,664	0 €	23,664	0.00	Esperar	24,528	0 €	24,528	0.00	Esperar
1224	38.8	34.4	38.0	-0.10	Vender	-	22,542	0 €	22,542	0.00	Esperar	-	24,639	0 €	24,639	0.00	Esperar	Esperar	-	23,664	0 €	23,664	0.00	Esperar	24,528	0 €	24,528	0.00	Esperar
1225	37.6	33.7	38.2	-0.12	Vender	-	22,542	0 €	22,542	0.00	Esperar	-	24,639	0 €	24,639	0.00	Esperar	Vender	-	23,664	0 €	23,664	0.00	Esperar	24,528	0 €	24,528	0.00	Esperar
1226	38.1	34.2	40.6	-0.16	Vender	-	22,542	0 €	22,542	0.00	Esperar	-	24,639	0 €	24,639	0.00	Esperar	Vender	-	23,664	0 €	23,664	0.00	Esperar	24,528	0 €	24,528	0.00	Esperar
1227	38.5	33.5	38.9	-0.14	Vender	-	22,542	0 €	22,542	0.00	Esperar	-	24,639	0 €	24,639	0.00	Esperar	Vender	-	23,664	0 €	23,664	0.00	Esperar	24,528	0 €	24,528	0.00	Esperar
1228	38.4	33.7	38.9	-0.13	Vender	-	22,542	0 €	22,542	0.00	Esperar	-	24,639	0 €	24,639	0.00	Esperar	Vender	-	23,664	0 €	23,664	0.00	Esperar	24,528	0 €	24,528	0.00	Esperar
1229	36.9	32.8	38.8	-0.15	Vender	-	22,542	0 €	22,542	0.00	Esperar	-	24,639	0 €	24,639	0.00	Esperar	Vender	-	23,664	0 €	23,664	0.00	Esperar	24,528	0 €	24,528	0.00	Esperar
1230	36.7	30.9	37.6	-0.18	Vender	-	22,542	0 €	22,542	0.00	Esperar	-	24,639	0 €	24,639	0.00	Esperar	Vender	-	23,664	0 €	23,664	0.00	Esperar	24,528	0 €	24,528	0.00	Esperar
1231	37.6	31.5	38.1	-0.17	Vender	-	22,542	0 €	22,542	0.00	Esperar	-	24,639	0 €	24,639	0.00	Esperar	Vender	-	23,664	0 €	23,664	0.00	Esperar	24,528	0 €	24,528	0.00	Esperar
1232	37.6	31.0	38.5	-0.19	Vender	-	22,542	0 €	22,542	0.00	Esperar	-	24,639	0 €	24,639	0.00	Esperar	Vender	-	23,664	0 €	23,664	0.00	Esperar	24,528	0 €	24,528	0.00	Esperar
1233	37.1	35.6	38.4	-0.07	Vender	-	22,542	0 €	22,542	0.00	Esperar	-	24,639	0 €	24,639	0.00	Esperar	Vender	-	23,664	0 €	23,664	0.00	Esperar	24,528	0 €	24,528	0.00	Esperar
1234	37.0	35.0	36.9	-0.05	Vender	-	22,542	0 €	22,542	0.00	Esperar	-	24,639	0 €	24,639	0.00	Esperar	Esperar	-	23,664	0 €	23,664	0.00	Esperar	24,528	0 €	24,528	0.00	Esperar
1235	37.2	35.5	36.7	-0.03	Esperar	-	22,542	0 €	22,542	0.00	Esperar	-	24,639	0 €	24,639	0.00	Esperar	Esperar	-	23,664	0 €	23,664	0.00	Esperar	24,528	0 €	24,528	0.00	Esperar
1236	36.9	34.7	37.6	-0.08	Vender	-	22,542	0 €	22,542	0.00	Esperar	-	24,639	0 €	24,639	0.00	Esperar	Esperar	-	23,664	0 €	23,664	0.00	Esperar	24,528	0 €	24,528	0.00	Esperar
1237	38.6	35.6	37.6	-0.05	Vender	-	22,542	0 €	22,542	0.00	Esperar	-	24,639	0 €	24,639	0.00	Esperar	Esperar	-	23,664	0 €	23,664	0.00	Esperar	24,528	0 €	24,528	0.00	Esperar
1238	38.4	34.5	37.1	-0.07	Vender	-	22,542	0 €	22,542	0.00	Esperar	-	24,639	0 €	24,639	0.00	Esperar	Esperar	-	23,664	0 €	23,664	0.00	Esperar	24,528	0 €	24,528	0.00	Esperar

1239	38.2	34.3	37.0	-0.07	Vender	-	22,542	0 € -	22,542	0.00	Esperar	24,639	0 € -	24,639	0.00	Esperar	Esperar	-	23,664	0 € -	23,664	0.00	Esperar	-	24,528	0 € -	24,528	0.00	Esperar
1240	38.6	33.3	37.2	-0.10	Vender	-	22,542	0 € -	22,542	0.00	Esperar	24,639	0 € -	24,639	0.00	Esperar	Vender	-	23,664	0 € -	23,664	0.00	Esperar	-	24,528	0 € -	24,528	0.00	Esperar
1241	38.3	35.1	36.9	-0.05	Esperar	-	22,542	0 € -	22,542	0.00	Esperar	24,639	0 € -	24,639	0.00	Esperar	Esperar	-	23,664	0 € -	23,664	0.00	Esperar	-	24,528	0 € -	24,528	0.00	Esperar
1242	38.8	33.6	38.6	-0.13	Vender	-	22,542	0 € -	22,542	0.00	Esperar	24,639	0 € -	24,639	0.00	Esperar	Vender	-	23,664	0 € -	23,664	0.00	Esperar	-	24,528	0 € -	24,528	0.00	Esperar
1243	38.9	32.8	38.4	-0.14	Vender	-	22,542	0 € -	22,542	0.00	Esperar	24,639	0 € -	24,639	0.00	Esperar	Vender	-	23,664	0 € -	23,664	0.00	Esperar	-	24,528	0 € -	24,528	0.00	Esperar
1244	39.0	33.5	38.2	-0.12	Vender	-	22,542	0 € -	22,542	0.00	Esperar	24,639	0 € -	24,639	0.00	Esperar	Vender	-	23,664	0 € -	23,664	0.00	Esperar	-	24,528	0 € -	24,528	0.00	Esperar
1245	39.0	32.9	38.6	-0.15	Vender	-	22,542	0 € -	22,542	0.00	Esperar	24,639	0 € -	24,639	0.00	Esperar	Vender	-	23,664	0 € -	23,664	0.00	Esperar	-	24,528	0 € -	24,528	0.00	Esperar
1246	39.3	33.0	38.3	-0.14	Vender	-	22,542	0 € -	22,542	0.00	Esperar	24,639	0 € -	24,639	0.00	Esperar	Vender	-	23,664	0 € -	23,664	0.00	Esperar	-	24,528	0 € -	24,528	0.00	Esperar
1247	39.2	33.4	38.8	-0.14	Vender	-	22,542	0 € -	22,542	0.00	Esperar	24,639	0 € -	24,639	0.00	Esperar	Vender	-	23,664	0 € -	23,664	0.00	Esperar	-	24,528	0 € -	24,528	0.00	Esperar

Triger = 0.05, MaxPerd =0.05 y MaxGan = 0.05 € 22,541.54

Triger = 0.05, MaxPerd =0.05 y MaxGan = 0.05 € 24,638.54

Triger = 0.1, MaxPerd =0.05 y MaxGan = 0.05 € 23,664.36

Triger = 0.1, MaxPerd =0.05 y MaxGan = 0.1 € 24,527.98

III. Redes neuronales - Modelo estadístico Max(15) con 80% training

Index	Y real	Y pred	TEF la cada 5 días	Trigger	Señal compraventa con trigger = 0.05										Señal compraventa con trigger = 0.1																						
					MaxPerd y MaxGan = 0.05					MaxPerd = 0.05 y MaxGan = 0.1					MaxPerd y MaxGan = 0.05					MaxPerd = 0.05 y MaxGan = 0.1																	
					Señal compra venta con trigger = 0.05	Capital libre	Volu m de accio nes	€ en accio nes	Total de dinero	Ganar / Perder	Señal compra venta con MaxPerd y MaxGan = 0.05	Señal compra venta con MaxPerd y MaxGan = 0.1	Señal compra venta con MaxPerd y MaxGan = 0.05	Capital libre	Volum en de accio nes	€ en accio nes	Total de €	Ganar / Perder	Señal compra venta con MaxPerd y MaxGan = 0.05	Señal compra venta con MaxPerd y MaxGan = 0.1	Señal compra venta con MaxPerd y MaxGan = 0.05	Capital libre	Volum en de accio nes	€ en accio nes	Total de dinero	Ganar / Perder	Señal compra venta con MaxPerd y MaxGan = 0.05	Señal compra venta con MaxPerd y MaxGan = 0.1	Señal compra venta con MaxPerd y MaxGan = 0.05								
1127	36.3	39.6	36.7	0.08	Comprar	Comprar	€-	550	20,000	20,000	0.00	Esperar	Esperar	20,000	20,000	0.00	Esperar	Esperar	20,000	0 €-	20,000	0 €-	20,000	0.00	Esperar	20,000	0 €-	20,000	0.00	Esperar	20,000	0 €-	20,000	0.00	Esperar		
1128	36.3	39.5	36.3	0.09	Comprar	-	€-	550	19,955	19,955	0.00	Esperar	Esperar	19,955	19,955	0.00	Esperar	Esperar	20,000	0 €-	20,000	0 €-	20,000	0.00	Esperar	20,000	0 €-	20,000	0.00	Esperar	20,000	0 €-	20,000	0.00	Esperar		
1129	35.5	38.1	35.9	0.06	Comprar	-	€-	550	19,560	19,560	-0.02	Esperar	Esperar	19,560	19,560	-0.02	Esperar	Esperar	20,000	0 €-	20,000	0 €-	20,000	0.00	Esperar	20,000	0 €-	20,000	0.00	Esperar	20,000	0 €-	20,000	0.00	Esperar		
1130	35.8	38.9	36.2	0.07	Comprar	-	€-	550	19,693	19,693	-0.02	Esperar	Esperar	19,693	19,693	-0.02	Esperar	Esperar	20,000	0 €-	20,000	0 €-	20,000	0.00	Esperar	20,000	0 €-	20,000	0.00	Esperar	20,000	0 €-	20,000	0.00	Esperar		
1131	35.0	39.0	36.5	0.07	Comprar	-	€-	550	19,253	19,253	-0.04	Esperar	Esperar	19,253	19,253	-0.04	Esperar	Esperar	20,000	0 €-	20,000	0 €-	20,000	0.00	Esperar	20,000	0 €-	20,000	0.00	Esperar	20,000	0 €-	20,000	0.00	Esperar		
1132	35.1	39.5	36.3	0.09	Comprar	-	€-	550	19,303	19,303	-0.03	Esperar	Esperar	19,303	19,303	-0.03	Esperar	Esperar	20,000	0 €-	20,000	0 €-	20,000	0.00	Esperar	20,000	0 €-	20,000	0.00	Esperar	20,000	0 €-	20,000	0.00	Esperar		
1133	35.7	38.0	36.3	0.05	Esperar	-	€-	550	19,640	19,640	-0.02	Esperar	Esperar	19,640	19,640	-0.02	Esperar	Esperar	20,000	0 €-	20,000	0 €-	20,000	0.00	Esperar	20,000	0 €-	20,000	0.00	Esperar	20,000	0 €-	20,000	0.00	Esperar		
1134	35.7	37.7	35.5	0.06	Comprar	-	€-	550	19,664	19,664	-0.02	Esperar	Esperar	19,664	19,664	-0.02	Esperar	Esperar	20,000	0 €-	20,000	0 €-	20,000	0.00	Esperar	20,000	0 €-	20,000	0.00	Esperar	20,000	0 €-	20,000	0.00	Esperar		
1135	36.2	37.3	35.8	0.04	Esperar	-	€-	550	19,931	19,931	0.00	Esperar	Esperar	19,931	19,931	0.00	Esperar	Esperar	20,000	0 €-	20,000	0 €-	20,000	0.00	Esperar	20,000	0 €-	20,000	0.00	Esperar	20,000	0 €-	20,000	0.00	Esperar		
1136	36.7	35.3	35.0	0.01	Esperar	-	€-	550	20,220	20,220	0.01	Esperar	Esperar	20,220	20,220	0.01	Esperar	Esperar	20,000	0 €-	20,000	0 €-	20,000	0.00	Esperar	20,000	0 €-	20,000	0.00	Esperar	20,000	0 €-	20,000	0.00	Esperar		
1137	36.8	36.0	35.1	0.03	Esperar	-	€-	550	20,242	20,242	0.01	Esperar	Esperar	20,242	20,242	0.01	Esperar	Esperar	20,000	0 €-	20,000	0 €-	20,000	0.00	Esperar	20,000	0 €-	20,000	0.00	Esperar	20,000	0 €-	20,000	0.00	Esperar		
1138	36.7	38.5	35.7	0.08	Comprar	-	€-	550	20,217	20,217	0.01	Esperar	Esperar	20,217	20,217	0.01	Esperar	Esperar	20,000	0 €-	20,000	0 €-	20,000	0.00	Esperar	20,000	0 €-	20,000	0.00	Esperar	20,000	0 €-	20,000	0.00	Esperar		
1139	36.4	39.6	35.7	0.11	Comprar	-	€-	550	20,051	20,051	0.00	Esperar	Comprar	20,051	20,051	0.00	Esperar	Comprar	€-	549	20,000	20,000	0.00	Esperar	Comprar	€-	549	20,000	20,000	0.00	Esperar	€-	549	20,000	20,000	0.00	Esperar
1140	37.2	40.2	36.2	0.11	Comprar	-	€-	550	20,490	20,490	0.02	Esperar	Comprar	20,490	20,490	0.02	Esperar	Comprar	€-	549	20,439	20,439	0.02	Esperar	€-	549	20,439	20,439	0.02	Esperar	€-	549	20,439	20,439	0.02	Esperar	
1141	37.9	39.6	36.7	0.08	Comprar	-	€-	550	20,838	20,838	0.04	Esperar	Esperar	20,838	20,838	0.04	Esperar	Esperar	€-	549	20,785	20,785	0.04	Esperar	€-	549	20,785	20,785	0.04	Esperar	€-	549	20,785	20,785	0.04	Esperar	
1142	38.9	38.1	36.8	0.04	Esperar	Vender	21,431	0 €-	21,431	21,431	0.07	Vender	Esperar	21,431	21,431	0.07	Esperar	Esperar	21,377	0 €-	21,377	0 €-	21,377	0.07	Vender	€-	549	21,377	21,377	0.07	Esperar	€-	549	21,377	21,377	0.07	Esperar
1143	39.1	37.3	36.7	0.02	Esperar	-	21,431	0 €-	21,431	21,431	0.00	Esperar	Esperar	21,431	21,431	0.00	Esperar	Esperar	21,377	0 €-	21,377	0 €-	21,377	0.00	Esperar	€-	549	21,451	21,451	0.07	Esperar	€-	549	21,451	21,451	0.07	Esperar
1144	39.0	36.8	36.4	0.01	Esperar	-	21,431	0 €-	21,431	21,431	0.00	Esperar	Esperar	21,431	21,431	0.00	Esperar	Esperar	21,377	0 €-	21,377	0 €-	21,377	0.00	Esperar	€-	549	21,427	21,427	0.07	Esperar	€-	549	21,427	21,427	0.07	Esperar
1145	39.4	36.8	37.2	-0.01	Esperar	-	21,431	0 €-	21,431	21,431	0.00	Esperar	Esperar	21,431	21,431	0.00	Esperar	Esperar	21,377	0 €-	21,377	0 €-	21,377	0.00	Esperar	€-	549	21,610	21,610	0.08	Esperar	€-	549	21,610	21,610	0.08	Esperar
1146	40.0	37.4	37.9	-0.01	Esperar	-	21,431	0 €-	21,431	21,431	0.00	Esperar	Esperar	21,431	21,431	0.00	Esperar	Esperar	21,377	0 €-	21,377	0 €-	21,377	0.00	Esperar	€-	549	21,959	21,959	0.10	Esperar	€-	549	21,959	21,959	0.10	Esperar
1147	40.4	36.7	38.9	-0.06	Vender	-	21,431	0 €-	21,431	21,431	0.00	Esperar	Esperar	21,431	21,431	0.00	Esperar	Esperar	21,377	0 €-	21,377	0 €-	21,377	0.00	Esperar	Vender	22,157	0 €-	22,157	0.11	Vender	22,157	0 €-	22,157	0.11	Vender	
1148	40.9	37.5	39.1	-0.04	Vender	-	21,431	0 €-	21,431	21,431	0.00	Esperar	Esperar	21,431	21,431	0.00	Esperar	Esperar	21,377	0 €-	21,377	0 €-	21,377	0.00	Esperar	-	22,157	0 €-	22,157	0.00	Esperar	-	22,157	0.00	Esperar		
1149	41.2	37.1	39.0	-0.05	Vender	-	21,431	0 €-	21,431	21,431	0.00	Esperar	Esperar	21,431	21,431	0.00	Esperar	Esperar	21,377	0 €-	21,377	0 €-	21,377	0.00	Esperar	-	22,157	0 €-	22,157	0.00	Esperar	-	22,157	0.00	Esperar		
1150	41.3	37.7	39.4	-0.04	Esperar	-	21,431	0 €-	21,431	21,431	0.00	Esperar	Esperar	21,431	21,431	0.00	Esperar	Esperar	21,377	0 €-	21,377	0 €-	21,377	0.00	Esperar	-	22,157	0 €-	22,157	0.00	Esperar	-	22,157	0.00	Esperar		
1151	41.0	36.9	40.0	-0.08	Vender	-	21,431	0 €-	21,431	21,431	0.00	Esperar	Esperar	21,431	21,431	0.00	Esperar	Esperar	21,377	0 €-	21,377	0 €-	21,377	0.00	Esperar	-	22,157	0 €-	22,157	0.00	Esperar	-	22,157	0.00	Esperar		
1152	40.9	36.7	40.4	-0.09	Vender	-	21,431	0 €-	21,431	21,431	0.00	Esperar	Esperar	21,431	21,431	0.00	Esperar	Esperar	21,377	0 €-	21,377	0 €-	21,377	0.00	Esperar	-	22,157	0 €-	22,157	0.00	Esperar	-	22,157	0.00	Esperar		
1153	41.4	36.9	40.9	-0.10	Vender	-	21,431	0 €-	21,431	21,431	0.00	Esperar	Esperar	21,431	21,431	0.00	Esperar	Esperar	21,377	0 €-	21,377	0 €-	21,377	0.00	Esperar	-	22,157	0 €-	22,157	0.00	Esperar	-	22,157	0.00	Esperar		
1154	40.9	37.1	41.2	-0.10	Vender	-	21,431	0 €-	21,431	21,431	0.00	Esperar	Esperar	21,431	21,431	0.00	Esperar	Esperar	21,377	0 €-	21,377	0 €-	21,377	0.00	Esperar	-	22,157	0 €-	22,157	0.00	Esperar	-	22,157	0.00	Esperar		

1197	36.0	38.6	38.8	0.00	Esperar	-	21,431	0 €	21,431	0.00	Esperar	-	22,014	0 €	22,014	0.00	Esperar	-	21,377	0 €	21,377	0.00	Esperar	-	22,157	0 €	22,157	0.00	Esperar		
1198	36.9	38.1	38.2	0.00	Esperar	-	21,431	0 €	21,431	0.00	Esperar	-	22,014	0 €	22,014	0.00	Esperar	-	21,377	0 €	21,377	0.00	Esperar	-	22,157	0 €	22,157	0.00	Esperar		
1199	36.5	38.1	38.0	0.00	Esperar	-	21,431	0 €	21,431	0.00	Esperar	-	22,014	0 €	22,014	0.00	Esperar	-	21,377	0 €	21,377	0.00	Esperar	-	22,157	0 €	22,157	0.00	Esperar		
1200	36.3	39.1	36.6	0.07	Comprar	Comprar	€	590	21,431	21,431	0.00	Esperar	22,014	606	22,014	22,014	0.00	Esperar	-	21,377	0 €	21,377	0.00	Esperar	-	22,157	0 €	22,157	0.00	Esperar	
1201	36.8	39.8	36.3	0.10	Comprar	-	€	590	21,721	21,721	0.01	Esperar	€	606	22,312	22,312	0.01	Esperar	-	21,377	0 €	21,377	0.00	Esperar	-	22,157	0 €	22,157	0.00	Esperar	
1202	36.7	39.7	36.0	0.10	Comprar	-	€	590	21,667	21,667	0.01	Esperar	€	606	22,257	22,257	0.01	Esperar	Comprar	Comprar	€	582	21,377	21,377	0.00	Esperar	603	22,157	22,157	0.00	Esperar
1203	36.7	38.7	36.9	0.05	Esperar	-	€	590	21,670	21,670	0.01	Esperar	€	606	22,260	22,260	0.01	Esperar	-	€	582	21,380	21,380	0.00	Esperar	603	22,160	22,160	0.00	Esperar	
1204	36.0	38.0	36.5	0.04	Esperar	-	€	590	21,258	21,258	-0.01	Esperar	€	606	21,836	21,836	-0.01	Esperar	-	€	582	20,973	20,973	-0.02	Esperar	603	21,738	21,738	-0.02	Esperar	
1205	36.3	36.9	36.3	0.02	Esperar	-	€	590	21,392	21,392	0.00	Esperar	€	606	21,974	21,974	0.00	Esperar	-	€	582	21,105	21,105	-0.01	Esperar	603	21,875	21,875	-0.01	Esperar	
1206	36.8	38.6	36.8	0.05	Esperar	-	€	590	21,687	21,687	0.01	Esperar	€	606	22,277	22,277	0.01	Esperar	-	€	582	21,396	21,396	0.00	Esperar	603	22,177	22,177	0.00	Esperar	
1207	36.3	40.0	36.7	0.09	Comprar	-	€	590	21,400	21,400	0.00	Esperar	€	606	21,982	21,982	0.00	Esperar	-	€	582	21,113	21,113	-0.01	Esperar	603	21,884	21,884	-0.01	Esperar	
1208	36.1	37.8	36.7	0.03	Esperar	-	€	590	21,320	21,320	-0.01	Esperar	€	606	21,901	21,901	-0.01	Esperar	-	€	582	21,035	21,035	-0.02	Esperar	603	21,802	21,802	-0.02	Esperar	
1209	35.3	38.3	36.0	0.06	Comprar	-	€	590	20,820	20,820	-0.03	Esperar	€	606	21,387	21,387	-0.03	Esperar	-	€	582	20,541	20,541	-0.04	Esperar	603	21,291	21,291	-0.04	Esperar	
1210	36.3	37.4	36.3	0.03	Esperar	-	€	590	21,400	21,400	0.00	Esperar	€	606	21,982	21,982	0.00	Esperar	-	€	582	21,113	21,113	-0.01	Esperar	603	21,884	21,884	-0.01	Esperar	
1211	36.2	38.3	36.8	0.04	Esperar	-	€	590	21,363	21,363	0.00	Esperar	€	606	21,945	21,945	0.00	Esperar	-	€	582	21,077	21,077	-0.01	Esperar	603	21,846	21,846	-0.01	Esperar	
1212	36.8	38.0	36.3	0.05	Esperar	-	€	590	21,692	21,692	0.01	Esperar	€	606	22,283	22,283	0.01	Esperar	-	€	582	21,402	21,402	0.00	Esperar	603	22,183	22,183	0.00	Esperar	
1213	37.8	36.7	36.1	0.02	Esperar	-	€	590	22,332	22,332	0.04	Esperar	€	606	22,939	22,939	0.04	Esperar	-	€	582	22,032	22,032	0.03	Esperar	603	22,837	22,837	0.03	Esperar	
1214	38.0	34.1	35.3	-0.03	Esperar	-	€	590	22,417	22,417	0.05	Esperar	€	606	23,027	23,027	0.05	Esperar	-	€	582	22,117	22,117	0.03	Esperar	603	22,924	22,924	0.03	Esperar	
1215	38.1	34.9	36.3	-0.04	Esperar	-	€	590	22,502	22,502	0.05	Esperar	€	606	23,115	23,115	0.05	Esperar	-	€	582	22,201	22,201	0.04	Esperar	603	23,011	23,011	0.04	Esperar	
1216	37.6	35.0	36.2	-0.03	Esperar	-	€	590	22,209	22,209	0.04	Esperar	€	606	22,814	22,814	0.04	Esperar	-	€	582	21,912	21,912	0.03	Esperar	603	22,712	22,712	0.03	Esperar	
1217	37.6	35.1	36.8	-0.05	Esperar	-	€	590	22,162	22,162	0.03	Esperar	€	606	22,765	22,765	0.03	Esperar	-	€	582	21,865	21,865	0.02	Esperar	603	22,663	22,663	0.02	Esperar	
1218	37.7	36.3	37.8	-0.04	Esperar	-	€	590	22,252	22,252	0.04	Esperar	€	606	22,858	22,858	0.04	Esperar	-	€	582	21,954	21,954	0.03	Esperar	603	22,756	22,756	0.03	Esperar	
1219	38.0	36.7	38.0	-0.03	Esperar	-	€	590	22,420	22,420	0.05	Esperar	€	606	23,030	23,030	0.05	Esperar	-	€	582	22,120	22,120	0.03	Esperar	603	22,927	22,927	0.03	Esperar	
1220	38.2	36.6	38.1	-0.04	Esperar	Vender	22,545	0 €	22,545	0.05	Vender	-	€	606	23,159	23,159	0.05	Esperar	-	€	582	22,243	22,243	0.04	Esperar	603	23,055	23,055	0.04	Esperar	
1221	40.6	37.6	37.6	0.00	Esperar	-	22,545	0 €	22,545	0.00	Esperar	Vender	24,639	0 €	24,639	0.12	Vender	Esperar	Vender	23,664	0 €	23,664	0.11	Vender	24,528	0 €	24,528	0.11	Vender		
1222	38.9	37.3	37.6	-0.01	Esperar	-	22,545	0 €	22,545	0.00	Esperar	-	24,639	0 €	24,639	0.00	Esperar	-	€	23,664	0 €	23,664	0.00	Esperar	-	24,528	0 €	24,528	0.00	Esperar	
1223	38.9	36.7	37.7	-0.03	Esperar	-	22,545	0 €	22,545	0.00	Esperar	-	24,639	0 €	24,639	0.00	Esperar	-	€	23,664	0 €	23,664	0.00	Esperar	-	24,528	0 €	24,528	0.00	Esperar	
1224	38.8	34.4	38.0	-0.10	Vender	-	22,545	0 €	22,545	0.00	Esperar	-	24,639	0 €	24,639	0.00	Esperar	-	€	23,664	0 €	23,664	0.00	Esperar	-	24,528	0 €	24,528	0.00	Esperar	
1225	37.6	33.7	38.2	-0.12	Vender	-	22,545	0 €	22,545	0.00	Esperar	-	24,639	0 €	24,639	0.00	Esperar	-	€	23,664	0 €	23,664	0.00	Esperar	-	24,528	0 €	24,528	0.00	Esperar	
1226	38.1	34.2	40.6	-0.16	Vender	-	22,545	0 €	22,545	0.00	Esperar	-	24,639	0 €	24,639	0.00	Esperar	-	€	23,664	0 €	23,664	0.00	Esperar	-	24,528	0 €	24,528	0.00	Esperar	
1227	38.5	33.5	38.9	-0.14	Vender	-	22,545	0 €	22,545	0.00	Esperar	-	24,639	0 €	24,639	0.00	Esperar	-	€	23,664	0 €	23,664	0.00	Esperar	-	24,528	0 €	24,528	0.00	Esperar	
1228	38.4	33.7	38.9	-0.13	Vender	-	22,545	0 €	22,545	0.00	Esperar	-	24,639	0 €	24,639	0.00	Esperar	-	€	23,664	0 €	23,664	0.00	Esperar	-	24,528	0 €	24,528	0.00	Esperar	
1229	36.9	32.8	38.8	-0.15	Vender	-	22,545	0 €	22,545	0.00	Esperar	-	24,639	0 €	24,639	0.00	Esperar	-	€	23,664	0 €	23,664	0.00	Esperar	-	24,528	0 €	24,528	0.00	Esperar	
1230	36.7	30.9	37.6	-0.18	Vender	-	22,545	0 €	22,545	0.00	Esperar	-	24,639	0 €	24,639	0.00	Esperar	-	€	23,664	0 €	23,664	0.00	Esperar	-	24,528	0 €	24,528	0.00	Esperar	
1231	37.6	31.5	38.1	-0.17	Vender	-	22,545	0 €	22,545	0.00	Esperar	-	24,639	0 €	24,639	0.00	Esperar	-	€	23,664	0 €	23,664	0.00	Esperar	-	24,528	0 €	24,528	0.00	Esperar	
1232	37.6	31.0	38.5	-0.19	Vender	-	22,545	0 €	22,545	0.00	Esperar	-	24,639	0 €	24,639	0.00	Esperar	-	€	23,664	0 €	23,664	0.00	Esperar	-	24,528	0 €	24,528	0.00	Esperar	
1233	37.1	35.6	38.4	-0.07	Vender	-	22,545	0 €	22,545	0.00	Esperar	-	24,639	0 €	24,639	0.00	Esperar	-	€	23,664	0 €	23,664	0.00	Esperar	-	24,528	0 €	24,528	0.00	Esperar	
1234	37.0	35.0	36.9	-0.05	Vender	-	22,545	0 €	22,545	0.00	Esperar	-	24,639	0 €	24,639	0.00	Esperar	-	€	23,664	0 €	23,664	0.00	Esperar	-	24,528	0 €	24,528	0.00	Esperar	
1235	37.2	35.5	36.7	-0.03	Esperar	-	22,545	0 €	22,545	0.00	Esperar	-	24,639	0 €	24,639	0.00	Esperar	-	€	23,664	0 €	23,664	0.00	Esperar	-	24,528	0 €	24,528	0.00	Esperar	
1236	36.9	34.7	37.6	-0.08	Vender	-	22,545	0 €	22,545	0.00	Esperar	-	24,639	0 €	24,639	0.00	Esperar	-	€	23,664	0 €	23,664	0.00	Esperar	-	24,528	0 €	24,528	0.00	Esperar	
1237	38.6	35.6	37.6	-0.05	Vender	-	22,545	0 €	22,545	0.00	Esperar	-	24,639	0 €	24,639	0.00	Esperar	-	€	23,664	0 €	23,664	0.00	Esperar	-	24,528	0 €	24,528	0.00	Esperar	

1238	38.4	34.5	37.1	-0.07	Vender	-	22,545	0 € -	22,545	0.00	Esperar	24,639	0 € -	24,639	0.00	Esperar	Esperar	-	23,664	0 € -	23,664	0.00	Esperar	-	24,528	0 € -	24,528	0.00	Esperar
1239	38.2	34.3	37.0	-0.07	Vender	-	22,545	0 € -	22,545	0.00	Esperar	24,639	0 € -	24,639	0.00	Esperar	Esperar	-	23,664	0 € -	23,664	0.00	Esperar	-	24,528	0 € -	24,528	0.00	Esperar
1240	38.6	33.3	37.2	-0.10	Vender	-	22,545	0 € -	22,545	0.00	Esperar	24,639	0 € -	24,639	0.00	Esperar	Vender	-	23,664	0 € -	23,664	0.00	Esperar	-	24,528	0 € -	24,528	0.00	Esperar
1241	38.3	35.1	36.9	-0.05	Esperar	-	22,545	0 € -	22,545	0.00	Esperar	24,639	0 € -	24,639	0.00	Esperar	Esperar	-	23,664	0 € -	23,664	0.00	Esperar	-	24,528	0 € -	24,528	0.00	Esperar
1242	38.8	33.6	38.6	-0.13	Vender	-	22,545	0 € -	22,545	0.00	Esperar	24,639	0 € -	24,639	0.00	Esperar	Vender	-	23,664	0 € -	23,664	0.00	Esperar	-	24,528	0 € -	24,528	0.00	Esperar
1243	38.9	32.8	38.4	-0.14	Vender	-	22,545	0 € -	22,545	0.00	Esperar	24,639	0 € -	24,639	0.00	Esperar	Vender	-	23,664	0 € -	23,664	0.00	Esperar	-	24,528	0 € -	24,528	0.00	Esperar
1244	39.0	33.5	38.2	-0.12	Vender	-	22,545	0 € -	22,545	0.00	Esperar	24,639	0 € -	24,639	0.00	Esperar	Vender	-	23,664	0 € -	23,664	0.00	Esperar	-	24,528	0 € -	24,528	0.00	Esperar
1245	39.0	32.9	38.6	-0.15	Vender	-	22,545	0 € -	22,545	0.00	Esperar	24,639	0 € -	24,639	0.00	Esperar	Vender	-	23,664	0 € -	23,664	0.00	Esperar	-	24,528	0 € -	24,528	0.00	Esperar
1246	39.3	33.0	38.3	-0.14	Vender	-	22,545	0 € -	22,545	0.00	Esperar	24,639	0 € -	24,639	0.00	Esperar	Vender	-	23,664	0 € -	23,664	0.00	Esperar	-	24,528	0 € -	24,528	0.00	Esperar
1247	39.2	33.4	38.8	-0.14	Vender	-	22,545	0 € -	22,545	0.00	Esperar	24,639	0 € -	24,639	0.00	Esperar	Vender	-	23,664	0 € -	23,664	0.00	Esperar	-	24,528	0 € -	24,528	0.00	Esperar

Triger = 0.05, MaxPerd = 0.05 y MaxGan = 0.05 € 22,545.15
 Triger = 0.05, MaxPerd = 0.05 y MaxGan = 0.05 € 24,638.54
 Triger = 0.1, MaxPerd = 0.05 y MaxGan = 0.05 € 23,664.36
 Triger = 0.1, MaxPerd = 0.05 y MaxGan = 0.1 € 24,527.98

1296	39.7	36.8	43.1	-0.15	Vender	-	20,000	0 €	20,000	0.00	Esperar	-	20,000	0 €	20,000	0 €	20,000	0.00	Esperar	-	20,000	0 €	20,000	0.00	Esperar	
1297	40.0	35.9	41.9	-0.14	Vender	-	20,000	0 €	20,000	0.00	Esperar	-	20,000	0 €	20,000	0 €	20,000	0.00	Esperar	-	20,000	0 €	20,000	0.00	Esperar	
1298	39.9	38.5	39.6	-0.03	Esperar	-	20,000	0 €	20,000	0.00	Esperar	-	20,000	0 €	20,000	0 €	20,000	0.00	Esperar	-	20,000	0 €	20,000	0.00	Esperar	
1299	40.1	36.6	38.9	-0.06	Vender	-	20,000	0 €	20,000	0.00	Esperar	-	20,000	0 €	20,000	0 €	20,000	0.00	Esperar	-	20,000	0 €	20,000	0.00	Esperar	
1300	40.4	37.6	38.6	-0.02	Esperar	-	20,000	0 €	20,000	0.00	Esperar	-	20,000	0 €	20,000	0 €	20,000	0.00	Esperar	-	20,000	0 €	20,000	0.00	Esperar	
1301	41.1	40.8	39.7	0.03	Esperar	-	20,000	0 €	20,000	0.00	Esperar	-	20,000	0 €	20,000	0 €	20,000	0.00	Esperar	-	20,000	0 €	20,000	0.00	Esperar	
1302	41.9	42.1	40.0	0.05	Comprar	Comprar	€	477	20,000	20,000	0.00	Esperar	20,000	€	477	20,000	20,000	0.00	Esperar	Esperar	-	20,000	0 €	20,000	0.00	Esperar
1303	41.7	41.5	39.9	0.04	Esperar	-	€	477	19,909	19,909	0.00	Esperar	-	€	477	19,909	19,909	0.00	Esperar	Esperar	-	20,000	0 €	20,000	0.00	Esperar
1304	41.9	41.6	40.1	0.04	Esperar	-	€	477	19,981	19,981	0.00	Esperar	-	€	477	19,981	19,981	0.00	Esperar	Esperar	-	20,000	0 €	20,000	0.00	Esperar
1305	42.5	42.3	40.4	0.05	Esperar	-	€	477	20,263	20,263	0.01	Esperar	-	€	477	20,263	20,263	0.01	Esperar	Esperar	-	20,000	0 €	20,000	0.00	Esperar
1306	41.9	42.8	41.1	0.04	Esperar	-	€	477	20,007	20,007	0.00	Esperar	-	€	477	20,007	20,007	0.00	Esperar	Esperar	-	20,000	0 €	20,000	0.00	Esperar
1307	42.0	42.0	41.9	0.00	Esperar	-	€	477	20,062	20,062	0.00	Esperar	-	€	477	20,062	20,062	0.00	Esperar	Esperar	-	20,000	0 €	20,000	0.00	Esperar
1308	42.4	42.1	41.7	0.01	Esperar	-	€	477	20,236	20,236	0.01	Esperar	-	€	477	20,236	20,236	0.01	Esperar	Esperar	-	20,000	0 €	20,000	0.00	Esperar
1309	43.1	42.0	41.9	0.00	Esperar	-	€	477	20,592	20,592	0.03	Esperar	-	€	477	20,592	20,592	0.03	Esperar	Esperar	-	20,000	0 €	20,000	0.00	Esperar
1310	43.6	42.6	42.5	0.00	Esperar	-	€	477	20,792	20,792	0.04	Esperar	-	€	477	20,792	20,792	0.04	Esperar	Esperar	-	20,000	0 €	20,000	0.00	Esperar
1311	44.1	42.4	41.9	0.01	Esperar	Vender	21,060	0 €	21,060	0.05	Vender	-	€	477	21,060	21,060	0.05	Esperar	Esperar	-	20,000	0 €	20,000	0.00	Esperar	
1312	43.9	40.8	42.0	-0.03	Esperar	-	21,060	0 €	21,060	0.00	Esperar	-	€	477	20,947	20,947	0.05	Esperar	Esperar	-	20,000	0 €	20,000	0.00	Esperar	
1313	44.2	38.6	42.4	-0.09	Vender	-	21,060	0 €	21,060	0.00	Esperar	-	€	477	21,084	21,084	0.05	Esperar	Esperar	-	20,000	0 €	20,000	0.00	Esperar	
1314	44.4	37.5	43.1	-0.13	Vender	-	21,060	0 €	21,060	0.00	Esperar	-	€	477	21,174	21,174	0.06	Esperar	Vender	-	20,000	0 €	20,000	0.00	Esperar	
1315	45.0	38.6	43.6	-0.11	Vender	-	21,060	0 €	21,060	0.00	Esperar	-	€	477	21,465	21,465	0.07	Esperar	Vender	-	20,000	0 €	20,000	0.00	Esperar	
1316	45.4	39.5	44.1	-0.10	Vender	-	21,060	0 €	21,060	0.00	Esperar	-	€	477	21,680	21,680	0.08	Esperar	Vender	-	20,000	0 €	20,000	0.00	Esperar	
1317	45.5	37.8	43.9	-0.14	Vender	-	21,060	0 €	21,060	0.00	Esperar	-	€	477	21,718	21,718	0.09	Esperar	Vender	-	20,000	0 €	20,000	0.00	Esperar	
1318	46.3	37.1	44.2	-0.16	Vender	-	21,060	0 €	21,060	0.00	Esperar	Vender	22,086	€	477	21,718	21,718	0.10	Vender	Vender	-	20,000	0 €	20,000	0.00	Esperar
1319	46.8	37.6	44.4	-0.15	Vender	-	21,060	0 €	21,060	0.00	Esperar	-	22,086	€	477	21,718	21,718	0.10	Vender	Vender	-	20,000	0 €	20,000	0.00	Esperar
1320	48.2	38.2	45.0	-0.15	Vender	-	21,060	0 €	21,060	0.00	Esperar	-	22,086	€	477	21,718	21,718	0.10	Vender	Vender	-	20,000	0 €	20,000	0.00	Esperar
1321	49.4	38.2	45.4	-0.16	Vender	-	21,060	0 €	21,060	0.00	Esperar	-	22,086	€	477	21,718	21,718	0.10	Vender	Vender	-	20,000	0 €	20,000	0.00	Esperar
1322	50.0	39.9	45.5	-0.12	Vender	-	21,060	0 €	21,060	0.00	Esperar	-	22,086	€	477	21,718	21,718	0.10	Vender	Vender	-	20,000	0 €	20,000	0.00	Esperar
1323	49.4	41.5	46.3	-0.10	Vender	-	21,060	0 €	21,060	0.00	Esperar	-	22,086	€	477	21,718	21,718	0.10	Vender	Vender	-	20,000	0 €	20,000	0.00	Esperar
1324	48.9	43.8	46.8	-0.06	Vender	-	21,060	0 €	21,060	0.00	Esperar	-	22,086	€	477	21,718	21,718	0.10	Vender	Vender	-	20,000	0 €	20,000	0.00	Esperar
1325	48.4	43.7	48.2	-0.09	Vender	-	21,060	0 €	21,060	0.00	Esperar	-	22,086	€	477	21,718	21,718	0.10	Vender	Vender	-	20,000	0 €	20,000	0.00	Esperar
1326	48.5	41.8	49.4	-0.15	Vender	-	21,060	0 €	21,060	0.00	Esperar	-	22,086	€	477	21,718	21,718	0.10	Vender	Vender	-	20,000	0 €	20,000	0.00	Esperar
1327	48.0	41.2	50.0	-0.18	Vender	-	21,060	0 €	21,060	0.00	Esperar	-	22,086	€	477	21,718	21,718	0.10	Vender	Vender	-	20,000	0 €	20,000	0.00	Esperar
1328	47.5	41.0	49.4	-0.17	Vender	-	21,060	0 €	21,060	0.00	Esperar	-	22,086	€	477	21,718	21,718	0.10	Vender	Vender	-	20,000	0 €	20,000	0.00	Esperar
1329	46.9	41.3	48.9	-0.16	Vender	-	21,060	0 €	21,060	0.00	Esperar	-	22,086	€	477	21,718	21,718	0.10	Vender	Vender	-	20,000	0 €	20,000	0.00	Esperar
1330	47.0	42.0	48.4	-0.13	Vender	-	21,060	0 €	21,060	0.00	Esperar	-	22,086	€	477	21,718	21,718	0.10	Vender	Vender	-	20,000	0 €	20,000	0.00	Esperar
1331	46.3	41.1	48.5	-0.15	Vender	-	21,060	0 €	21,060	0.00	Esperar	-	22,086	€	477	21,718	21,718	0.10	Vender	Vender	-	20,000	0 €	20,000	0.00	Esperar
1332	46.0	39.8	48.0	-0.17	Vender	-	21,060	0 €	21,060	0.00	Esperar	-	22,086	€	477	21,718	21,718	0.10	Vender	Vender	-	20,000	0 €	20,000	0.00	Esperar
1333	46.9	38.0	47.5	-0.20	Vender	-	21,060	0 €	21,060	0.00	Esperar	-	22,086	€	477	21,718	21,718	0.10	Vender	Vender	-	20,000	0 €	20,000	0.00	Esperar
1334	47.0	38.9	46.9	-0.17	Vender	-	21,060	0 €	21,060	0.00	Esperar	-	22,086	€	477	21,718	21,718	0.10	Vender	Vender	-	20,000	0 €	20,000	0.00	Esperar
1335	47.7	38.7	47.0	-0.18	Vender	-	21,060	0 €	21,060	0.00	Esperar	-	22,086	€	477	21,718	21,718	0.10	Vender	Vender	-	20,000	0 €	20,000	0.00	Esperar
1336	47.2	40.5	46.3	-0.13	Vender	-	21,060	0 €	21,060	0.00	Esperar	-	22,086	€	477	21,718	21,718	0.10	Vender	Vender	-	20,000	0 €	20,000	0.00	Esperar

1379	46.9	39.3	48.5	-0.19	Vender	-	21,060	0 €-	21,060	0.00	Esperar	22,086	0 €-	22,086	0.00	Esperar	Vender	-	20,000	0 €-	20,000	0.00	Esperar	-	20,000	0 €-	20,000	0.00	Esperar
1380	47.2	38.8	48.6	-0.20	Vender	-	21,060	0 €-	21,060	0.00	Esperar	22,086	0 €-	22,086	0.00	Esperar	Vender	-	20,000	0 €-	20,000	0.00	Esperar	-	20,000	0 €-	20,000	0.00	Esperar
1381	47.2	38.1	48.7	-0.22	Vender	-	21,060	0 €-	21,060	0.00	Esperar	22,086	0 €-	22,086	0.00	Esperar	Vender	-	20,000	0 €-	20,000	0.00	Esperar	-	20,000	0 €-	20,000	0.00	Esperar
1382	47.9	37.3	49.0	-0.24	Vender	-	21,060	0 €-	21,060	0.00	Esperar	22,086	0 €-	22,086	0.00	Esperar	Vender	-	20,000	0 €-	20,000	0.00	Esperar	-	20,000	0 €-	20,000	0.00	Esperar
1383	48.2	36.7	49.1	-0.25	Vender	-	21,060	0 €-	21,060	0.00	Esperar	22,086	0 €-	22,086	0.00	Esperar	Vender	-	20,000	0 €-	20,000	0.00	Esperar	-	20,000	0 €-	20,000	0.00	Esperar
1384	48.1	36.0	46.9	-0.23	Vender	-	21,060	0 €-	21,060	0.00	Esperar	22,086	0 €-	22,086	0.00	Esperar	Vender	-	20,000	0 €-	20,000	0.00	Esperar	-	20,000	0 €-	20,000	0.00	Esperar
1385	47.3	36.3	47.2	-0.23	Vender	-	21,060	0 €-	21,060	0.00	Esperar	22,086	0 €-	22,086	0.00	Esperar	Vender	-	20,000	0 €-	20,000	0.00	Esperar	-	20,000	0 €-	20,000	0.00	Esperar
1386	46.6	38.5	47.2	-0.18	Vender	-	21,060	0 €-	21,060	0.00	Esperar	22,086	0 €-	22,086	0.00	Esperar	Vender	-	20,000	0 €-	20,000	0.00	Esperar	-	20,000	0 €-	20,000	0.00	Esperar
1387	46.1	38.4	47.9	-0.20	Vender	-	21,060	0 €-	21,060	0.00	Esperar	22,086	0 €-	22,086	0.00	Esperar	Vender	-	20,000	0 €-	20,000	0.00	Esperar	-	20,000	0 €-	20,000	0.00	Esperar
1388	46.1	37.5	48.2	-0.22	Vender	-	21,060	0 €-	21,060	0.00	Esperar	22,086	0 €-	22,086	0.00	Esperar	Vender	-	20,000	0 €-	20,000	0.00	Esperar	-	20,000	0 €-	20,000	0.00	Esperar

Trigger = 0.05, MaxPerd = 0.05 y MaxGan = 0.05 € 21,059.67
 Trigger = 0.05, MaxPerd = 0.05 y MaxGan = 0.05 € 22,085.92
 Trigger = 0.1, MaxPerd = 0.05 y MaxGan = 0.05 € 20,000.00
 Trigger = 0.1, MaxPerd = 0.05 y MaxGan = 0.1 € 20,000.00

V. Máquinas de soporte vectorial (SVM) - Max(15) con 90% training

Index	Y real	Y pred	TEF la cada 5 días	Trigger	Señal compraventa con trigger = 0.05										Señal compraventa con trigger = 0.1																			
					MaxPerd = 0.05 y MaxGan = 0.05					MaxPerd = 0.05 y MaxGan = 0.1					MaxPerd = 0.05 y MaxGan = 0.05					MaxPerd = 0.05 y MaxGan = 0.1														
					Señal compra venta con trigger = 0.05	Capital libre	Volum de acciones	€ en acciones	Total de dinero	Ganar / Perder	Señal compra venta con MaxPerd y MaxGan n = 0.05	Capital libre	Volum de acciones	€ en acciones	Total de dinero	Ganar / Perder	Señal compra venta con MaxPerd y MaxGan n = 0.05	Capital libre	Volum de acciones	€ en acciones	Total de dinero	Ganar / Perder	Señal compra venta con MaxPerd y MaxGan n = 0.05	Capital libre	Volum de acciones	€ en acciones	Total de dinero	Ganar / Perder	Señal compra venta con MaxPerd y MaxGan n = 0.05					
1268	41.6	43.7	40.0	0.09	Comprar	0	480.42	20.000	20.000	20.000	0.00	Esperar	0	480.42	20.000	20.000	20.000	0.00	Esperar	-	20.000	0	0	20.000	0.00	Esperar	-	20.000	0	0	20.000	0.00	Esperar	
1269	41.3	43.7	40.2	0.09	Comprar	-	480.42	19.834	19.834	19.834	-0.01	Esperar	-	480.42	19.834	19.834	19.834	-0.01	Esperar	-	20.000	0	0	20.000	0.00	Esperar	-	20.000	0	0	20.000	0.00	Esperar	
1270	41.0	44.4	40.9	0.09	Comprar	-	480.42	19.688	19.688	19.688	-0.02	Esperar	-	480.42	19.688	19.688	19.688	-0.02	Esperar	-	20.000	0	0	20.000	0.00	Esperar	-	20.000	0	0	20.000	0.00	Esperar	
1271	41.4	45.4	41.0	0.11	Comprar	-	480.42	19.894	19.894	19.894	-0.01	Esperar	-	480.42	19.894	19.894	19.894	-0.01	Esperar	Comprar	0	482.975	20.000	20.000	0.00	Esperar	Comprar	0	483	20.000	20.000	0.00	Esperar	
1272	41.9	46.4	41.5	0.12	Comprar	-	480.42	20.127	20.127	20.127	0.00	Esperar	-	480.42	20.127	20.127	20.127	0.00	Esperar	Comprar	0	482.975	20.234	20.234	0.01	Esperar	-	0	483	20.234	20.234	0.01	Esperar	
1273	41.6	46.0	41.6	0.11	Comprar	-	480.42	20.005	20.005	20.005	0.00	Esperar	-	480.42	20.005	20.005	20.005	0.00	Esperar	Comprar	0	482.975	20.111	20.111	0.01	Esperar	-	0	483	20.111	20.111	0.01	Esperar	
1274	41.9	45.7	41.3	0.11	Comprar	-	480.42	20.139	20.139	20.139	0.01	Esperar	-	480.42	20.139	20.139	20.139	0.01	Esperar	Comprar	0	482.975	20.246	20.246	0.01	Esperar	-	0	483	20.246	20.246	0.01	Esperar	
1275	43.0	45.2	41.0	0.10	Comprar	-	480.42	20.649	20.649	20.649	0.03	Esperar	-	480.42	20.649	20.649	20.649	0.03	Esperar	Comprar	0	482.975	20.758	20.758	0.04	Esperar	-	0	483	20.758	20.758	0.04	Esperar	
1276	43.5	44.8	41.4	0.08	Comprar	-	480.42	20.910	20.910	20.910	0.05	Esperar	-	480.42	20.910	20.910	20.910	0.05	Esperar	Vender	21.021	0	0	21.021	0.05	Vender	-	0	483	21.021	21.021	0.05	Esperar	
1277	44.0	45.1	41.9	0.08	Comprar	Vender	21.129	0	21.129	21.129	0.06	Esperar	-	480.42	21.129	21.129	21.129	0.06	Esperar	Esperar	21.021	0	0	21.021	0.00	Esperar	-	0	483	21.241	21.241	0.06	Esperar	
1278	44.2	45.4	41.6	0.09	Comprar	Comprar	0	478.14	21.129	21.129	0.00	Esperar	-	480.42	21.230	21.230	21.230	0.06	Esperar	Esperar	21.021	0	0	21.021	0.00	Esperar	-	0	483	21.343	21.343	0.07	Esperar	
1279	43.8	46.1	41.9	0.10	Comprar	-	478.14	20.938	20.938	20.938	-0.01	Esperar	-	480.42	21.038	21.038	21.038	0.05	Esperar	Comprar	0	480.052	21.021	21.021	0.00	Esperar	-	0	483	21.149	21.149	0.06	Esperar	
1280	44.1	47.2	43.0	0.10	Comprar	-	478.14	21.100	21.100	21.100	0.00	Esperar	-	480.42	21.201	21.201	21.201	0.06	Esperar	Esperar	0	480.052	21.185	21.185	0.01	Esperar	-	0	483	21.314	21.314	0.07	Esperar	
1281	43.8	46.6	43.5	0.07	Comprar	-	478.14	20.962	20.962	20.962	-0.01	Esperar	-	480.42	21.062	21.062	21.062	0.05	Esperar	Esperar	0	480.052	21.045	21.045	0.00	Esperar	-	0	483	21.174	21.174	0.06	Esperar	
1282	43.6	46.7	44.0	0.06	Comprar	-	478.14	20.854	20.854	20.854	-0.01	Esperar	-	480.42	20.954	20.954	20.954	0.05	Esperar	Esperar	0	480.052	20.937	20.937	0.00	Esperar	-	0	483	21.065	21.065	0.05	Esperar	
1283	43.7	46.2	44.2	0.05	Esperar	-	478.14	20.890	20.890	20.890	-0.01	Esperar	-	480.42	20.990	20.990	20.990	0.05	Esperar	Esperar	0	480.052	20.973	20.973	0.00	Esperar	-	0	483	21.101	21.101	0.06	Esperar	
1284	43.0	46.6	43.8	0.06	Comprar	-	478.14	20.572	20.572	20.572	-0.03	Esperar	-	480.42	20.670	20.670	20.670	0.03	Esperar	Esperar	0	480.052	20.654	20.654	-0.02	Esperar	-	0	483	20.780	20.780	0.04	Esperar	
1285	42.6	46.2	44.1	0.05	Esperar	-	478.14	20.388	20.388	20.388	-0.04	Esperar	-	480.42	20.485	20.485	20.485	0.02	Esperar	Esperar	0	480.052	20.469	20.469	-0.03	Esperar	-	0	483	20.594	20.594	0.03	Esperar	
1286	42.4	45.0	43.8	0.03	Esperar	-	478.14	20.280	20.280	20.280	-0.04	Esperar	-	480.42	20.377	20.377	20.377	0.02	Esperar	Esperar	0	480.052	20.361	20.361	-0.03	Esperar	-	0	483	20.485	20.485	0.02	Esperar	
1287	41.7	44.6	43.6	0.02	Esperar	Vender	19.922	0	19.922	19.922	-0.06	Vender	-	480.42	20.017	20.017	20.017	0.00	Esperar	Esperar	0	480.052	20.001	20.001	-0.05	Esperar	-	0	483	20.123	20.123	0.01	Esperar	
1288	42.5	44.3	43.7	0.01	Esperar	-	19.922	0	19.922	19.922	0.00	Esperar	-	480.42	20.418	20.418	20.418	0.02	Esperar	Esperar	0	480.052	20.402	20.402	-0.03	Esperar	-	0	483	20.526	20.526	0.03	Esperar	
1289	42.8	44.3	43.0	0.03	Esperar	-	19.922	0	19.922	19.922	0.00	Esperar	-	480.42	20.543	20.543	20.543	0.03	Esperar	Esperar	0	480.052	20.527	20.527	-0.02	Esperar	-	0	483	20.652	20.652	0.03	Esperar	
1290	42.5	44.6	42.6	0.05	Esperar	-	19.922	0	19.922	19.922	0.00	Esperar	-	480.42	20.406	20.406	20.406	0.02	Esperar	Esperar	0	480.052	20.390	20.390	-0.03	Esperar	-	0	483	20.514	20.514	0.03	Esperar	
1291	43.1	42.7	42.4	0.01	Esperar	-	19.922	0	19.922	19.922	0.00	Esperar	-	480.42	20.723	20.723	20.723	0.04	Esperar	Esperar	0	480.052	20.707	20.707	-0.01	Esperar	-	0	483	20.833	20.833	0.04	Esperar	
1292	41.9	42.0	41.7	0.01	Esperar	-	19.922	0	19.922	19.922	0.00	Esperar	-	480.42	20.120	20.120	20.120	0.01	Esperar	Esperar	0	480.052	20.105	20.105	-0.04	Esperar	-	0	483	20.227	20.227	0.01	Esperar	
1293	39.6	41.8	42.5	-0.02	Esperar	-	19.922	0	19.922	19.922	0.00	Esperar	-	480.42	19.027	19.027	19.027	-0.05	Esperar	Esperar	19.012	0	0	19.012	-0.10	Vender	Vender	19.128	0	0	19.128	-0.04	Esperar	
1294	38.9	42.7	42.8	0.00	Esperar	-	19.922	0	19.922	19.922	0.00	Esperar	-	18.674	0	0	18.674	-0.07	Vender	Esperar	19.012	0	0	19.012	0.00	Esperar	-	0	19.128	0	0	19.128	0.00	Esperar
1295	38.6	42.5	42.5	0.00	Esperar	-	19.922	0	19.922	19.922	0.00	Esperar	-	18.674	0	0	18.674	0.00	Esperar	Esperar	19.012	0	0	19.012	0.00	Esperar	-	0	19.128	0	0	19.128	0.00	Esperar

1380	47.2	43.8	48.6	-0.10	Vender	-	0	420.82	19,880	19,880	19,880	0.00	Esperar	18,674	0	0	18,674	0.00	Esperar	0.00	Esperar	Vender	-	19,012	0	0	19,012	0.00	Esperar	-	19,128	0	0	19,128	0.00	Esperar
1381	47.2	42.5	48.7	-0.13	Vender	-	0	420.82	19,863	19,863	19,863	0.00	Esperar	18,674	0	0	18,674	0.00	Esperar	0.00	Esperar	Vender	-	19,012	0	0	19,012	0.00	Esperar	-	19,128	0	0	19,128	0.00	Esperar
1382	47.9	42.1	49.0	-0.14	Vender	-	0	420.82	20,145	20,145	20,145	0.01	Esperar	18,674	0	0	18,674	0.00	Esperar	0.00	Esperar	Vender	-	19,012	0	0	19,012	0.00	Esperar	-	19,128	0	0	19,128	0.00	Esperar
1383	48.2	41.7	49.1	-0.15	Vender	-	0	420.82	20,284	20,284	20,284	0.02	Esperar	18,674	0	0	18,674	0.00	Esperar	0.00	Esperar	Vender	-	19,012	0	0	19,012	0.00	Esperar	-	19,128	0	0	19,128	0.00	Esperar
1384	48.1	41.5	46.9	-0.11	Vender	-	0	420.82	20,233	20,233	20,233	0.00	Esperar	18,674	0	0	18,674	0.00	Esperar	0.00	Esperar	Vender	-	19,012	0	0	19,012	0.00	Esperar	-	19,128	0	0	19,128	0.00	Esperar
1385	47.3	41.6	47.2	-0.12	Vender	-	0	420.82	19,905	19,905	19,905	0.00	Esperar	18,674	0	0	18,674	0.00	Esperar	0.00	Esperar	Vender	-	19,012	0	0	19,012	0.00	Esperar	-	19,128	0	0	19,128	0.00	Esperar
1386	46.6	41.1	47.2	-0.13	Vender	-	0	420.82	19,606	19,606	19,606	-0.02	Esperar	18,674	0	0	18,674	0.00	Esperar	-0.02	Esperar	Vender	-	19,012	0	0	19,012	0.00	Esperar	-	19,128	0	0	19,128	0.00	Esperar
1387	46.1	41.3	47.9	-0.14	Vender	-	0	420.82	19,412	19,412	19,412	-0.03	Esperar	18,674	0	0	18,674	0.00	Esperar	-0.03	Esperar	Vender	-	19,012	0	0	19,012	0.00	Esperar	-	19,128	0	0	19,128	0.00	Esperar
1388	46.1	40.0	48.2	-0.17	Vender	-	0	420.82	19,408	19,408	19,408	-0.03	Esperar	18,674	0	0	18,674	0.00	Esperar	-0.03	Esperar	Vender	-	19,012	0	0	19,012	0.00	Esperar	-	19,128	0	0	19,128	0.00	Esperar

Triger = 0.05, MaxPerd =0.05 y MaxGan = 0.05 € 19,408.29
 Triger = 0.05, MaxPerd =0.05 y MaxGan = 0.05 € 18,674.03
 Triger = 0.1, MaxPerd =0.05 y MaxGan = 0.05 € 19,012.47
 Triger = 0.1, MaxPerd =0.05 y MaxGan = 0.1 € 19,128.23