Programa ejecutivo de inteligencia artificial y Deep Learning

Adrián Grasa

Pedro Gómez

Raúl Estevas

TABLA DE CONTENIDO

INTRODUCCIÓN 4

SITUACIÓN ACTUAL (AS IS) DE LA OPORTUNIDAD DE NEGOCIO 6

2.1 Situación actual 6
2.2 Análisis de los competidores tradicionales 9
2.3 Análisis de empresas con uso de AI en trading algorítmico 13

DEFINICIÓN DE LA SOLUCIÓN (TO BE) 16

3.1. Identificación de la oportunidad 16
3.2. Definición conceptual del modelo y posicionamiento de la solución 17
3.3. Modelo de negocio 19
3.3.1 Misión, visión y valores 19
3.3.2 DAFO 21
3.3.3 Canvas 21
3.3.4 Objetivos estratégicos 23
3.3.5 Mapa estratégico 23
3.3.6 Cuadro de mando 25
3.4 Roadmap y plan de contingencias 29
3.5 Política de Inversión 30
3.6 Operativa en la ejecución de las ordenes 35
3.7 Business case 38
3.7.1 Ingresos e inversión inicial 40
3.7.2 Gastos 41
3.7.3 Riesgos analizados 41
3.7.4 Principales indicadores oportunidad de negocio 42
3.8 Procesos y planes de acción 42
3.8.1 Plan de marketing 42
3.8.2 Plan de RRHH 43
3.8.3 Plan de operación 44
3.9 Implementación tecnológica 45
3.9.1 Definición conceptual de la arquitectura 45
3.9.2 Tecnologías a utilizar 46
3.9.3 Arquitectura técnica de la solución 47
3.10 Fuentes de datos 49
3.10.1 Análisis fuentes de datos existentes 49
3.10.2 Fuentes y modelo de datos a utilizar 52
3.11 Simulación de métodos analíticos 56

3.11.1 Análisis y predicción de valores a partir de información financiera..... 56
3.11.2 Análisis y procesamiento de información no estructurada de noticias 63
3.11.3 Combinación de información estructurada y no estructurada para la predicción de cotizaciones 66
3.11.4 Simulación inversión basada en el uso de modelo de predicción de valores 69

CONCLUSIONES Y LÍNEAS DE MEJORA 73

BIBLIOGRAFÍA Y REFERENCIAS 76

INTRODUCCIÓN

Las finanzas y las actividades que lo rodean han sido a menudo objeto de fantasías y clichés a veces justificados. Sin embargo, actualmente estamos lejos de la imagen de las Bolsas de los años ochenta en la que los traders realizaban sus transacciones en voz alta. En los últimos casi 20 años hemos ingresado a un mundo hiper-competitivo basado en tecnología, en el que las computadoras ahora se utilizan para maximizar las ganancias. De hecho, en 1998 la U.S. Securities and Exchange Commission (SEC) autorizó el intercambio electrónico. De aquí en adelante se ha multiplicado el progreso en términos de velocidad y ejecución.

En España hemos asistido a un desarrollo en la misma línea, realizándose la última transacción tradicional en corros el 9 de julio del 2009, que tan solo funcionaba para un tipo muy reducido de valores, pasando desde este momento todas las transacciones a ser realizadas mediante sistema electrónico.

El proyecto ARP nace con el ánimo de aprovechar los últimos avances tecnológicos y especialmente aquellos relacionados con la Inteligencia Artificial y de manera específica el Deep Learning, creando un sistema de gestión de carteras basado en estas tecnologías y apoyado en el Big Data.

Por un lado, buscamos obtener rentabilidades superiores a los índices de referencia de manera constante permitiendo que nuestros clientes obtengan los mejores retornos de su inversión de manera sostenida en el tiempo, de acuerdo con su nivel de riesgo.

Por otro, buscamos dar solución a una problemática generalizada en los países más desarrollados y que tiene un profundo impacto en nuestro país y es aquella relacionada con las pensiones. Nuestra solución pretende gestionar, no solo fondos de inversión, si no también planes de pensiones de manera privada que sean complementarios a las pensiones públicas y puedan

proporcionar una renta adecuada a los trabajadores en el momento en que estos comiencen su jubilación.

SITUACIÓN ACTUAL (AS IS) DE LA OPORTUNIDAD DE NEGOCIO

2.1 SITUACIÓN ACTUAL

El ahorro colectivo en España continua inmerso en un proceso de cambio. Si consideramos que a nivel global se ha dejado la crisis financiera de 2008-2014, podemos empezar a vislumbrar como queda la configuración de la sociedad tras este periodo de crisis. Hemos asistido a un traspaso progresivo desde las instituciones públicas a los individuos de su responsabilidad sobre su bienestar presente y futuro.

Podemos destacar cuatro elementos cambiantes en el ahorro colectivo en España:

1- La situación de los tipos de interés en mínimos históricos ha llevado a los españoles a reducir drásticamente la inversión en vehículos monetarios y garantizados, adicionalmente hemos asistido a una cierta recuperación del hábito de los españoles a dedicar una parte significativa de su renta disponible en renta al ahorro inmobiliario en detrimento del ahorro financiero.

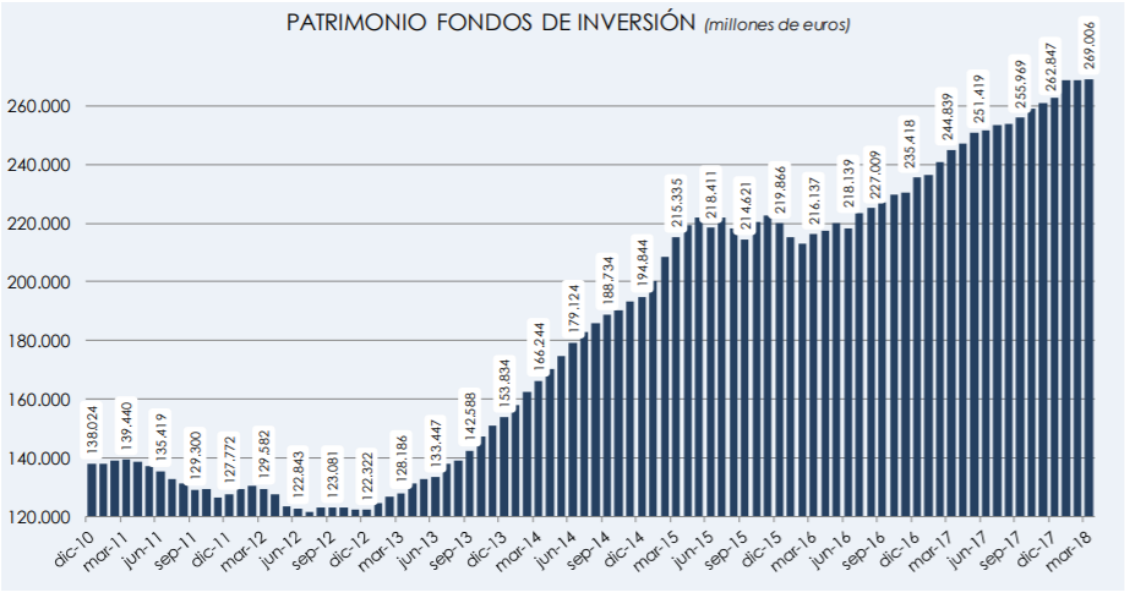
2- El segundo elemento que tiene un impacto en el ahorro colectivo es el debate en torno a las pensiones. La evolución prevista del Fondo de Reservas de las pensiones, podemos ver agotada en los próximos años, puede servir como detonante para reabrir el debate de las pensiones e incrementar la concienciación de la población respecto de las pensiones, su forma de financiación y la necesidad de ahorro finalista.

3- En tercer lugar, la creciente penetración de las gestoras internacionales que van ganado cuota de mercado en España y que están provocando cambios en el tipo de producto en el que inviertan los inversores españoles y en la forma de distribución. Por un lado, el tipo de producto que ofrecen este tipo de gestoras es significativamente diferente al que estaba habituado el inversor español. Por otro lado, la penetración de estas gestoras viene favorecida por los cambios en la distribución de sus

productos por otros canales diferentes a los canales tradicionales de la banca, estando enfocado a canales on-line o mediante Empresas de Asesoramiento Financiero Independiente EAFIs.

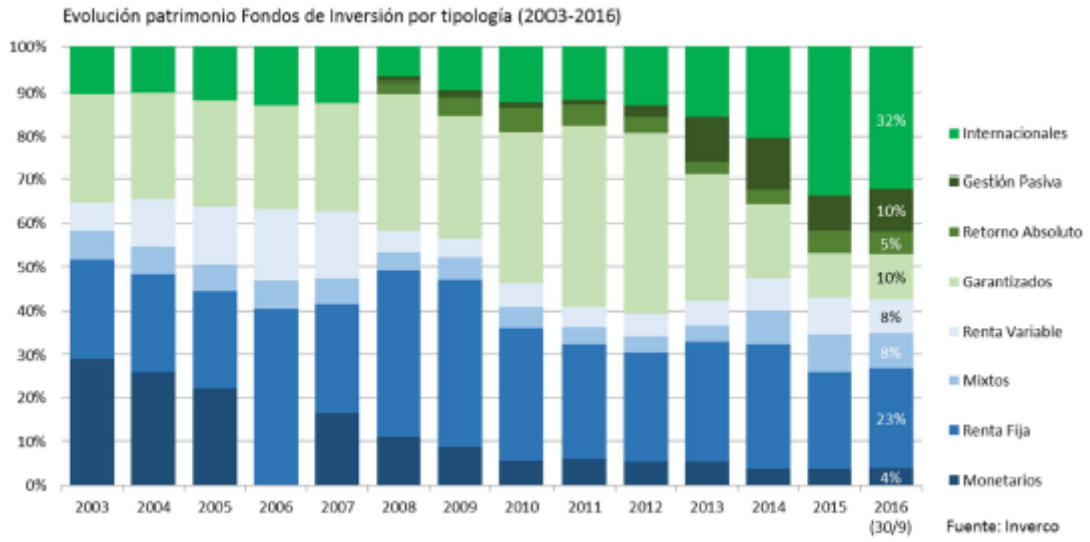
4- Por último y motivado en gran medida por los excesos cometidos durante la crisis financiera, se han incrementado las medidas regulatorias tanto para las entidades (DFA, EMIR, etc.), como destinadas a proporcionar una mayor protección al inversor (Mifid II), esto ha llevado a cambios en la estructura de las entidades, un mayor coste derivado del cumplimento de estas, así como en la gama de productos ofrecidos al inversor.

Como podemos observar en el siguiente gráfico el patrimonio gestionado por fondos de inversión en España ha sobrepasado el volumen de activos anteriores a la crisis, destacando además la irrupción de las gestoras de fondos internacionales, con un volumen cada vez mayor de los activos gestionados.



Por distribución de activos gestionados podemos observar en el siguiente grafico que cada vez hay una menor tendencia a invertir en activos con menor riesgo y por lo tanto menor rentabilidad como son los fondos monetarios, renta fija y especialmente garantizados, así mismo se sigue confirmando la situación indicada anteriormente y es que los fondos de inversión internacionales cada vez

tienen una mayor cuota de mercado en las preferencias de inversión de los españoles.

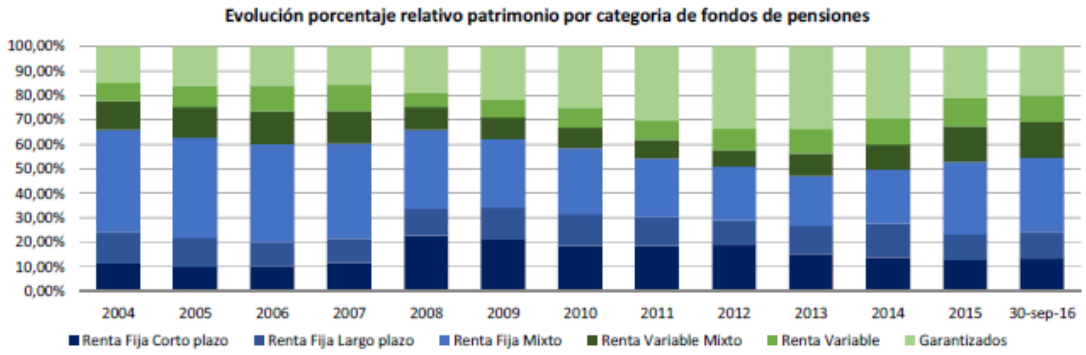


Un capítulo aparte por su relevancia, que entendemos va a incrementarse aún más en los próximos años, son los productos de inversión social, ante el previsible agotamiento del Fondo de Reserva de las pensiones podría ser un detonante para reabrir el debate de las pensiones e incrementar la concienciación de la población respecto a las mismas, su forma de financiación y la necesidad de ahorro finalista.

Si cada vez va a ser más importante el ahorro privado durante la vida del trabajador para poder mantener un nivel de vida similar al previo a la jubilación, los planes de pensiones deberían adoptar un papel cada vez más importante como complemento a las pensiones públicas, estos datos son avalados por las estadísticas que muestran un crecimiento de los patrimonios gestionados por este tipo de vehículos de ahorro cercano al 70% desde el 2003 hasta la actualidad.

En el siguiente grafico podemos observar la evolución de los diferentes tipos de activos gestionados por los fondos de pensiones, y en ella podemos comprobar los mismos patrones que habíamos observado hasta ahora y es que las categorías con menos riesgo han sufrido una disminución en comparación con otra que implican más riesgo, esto ha hecho que el inversor tenga que tomar más riesgo

para obtener mayor rentabilidad, lo que ha hecho por lo tanto que tengan que estar más informados sobre las características de los fondos y de sus inversiones.



2.2 ANÁLISIS DE LOS COMPETIDORES TRADICIONALES

En los ultimos años y dado el mayor apetito de los inversores por lo fondos como vehiculo de ahorro, hemos asistido a un incremento de la entidades que comercializan este tipo de vehiculos, hemos clasificado este tipo de entidades tradicionales de distribudoras en cuatro categorias:

1- *Aseguradoras*
2- *Entidades Bancarias Españoles*
3- *Boutiques Españolas*
4- *Gestoras Internacionales*

Los productos que distribuyen estas entidades podríamos dividirlos en:

- Fondos de inversión (FIM): son Instituciones de Inversión Colectiva (IIC) que tienen como objetivo proporcionar un rendimiento, mediante la gestión de las aportaciones de los partícipes y su inversión en bienes, valores u otros instrumentos financieros.

- Planes de pensiones: Son productos de ahorro a largo plazo, cuyo objetivo es acumular un capital determinado a fecha de jubilación que permita completar la pensión pública de jubilación. Dentro de este tipo de activo

podríamos diferenciar los Fondos de Pensiones de Empleo (FPE) establecido por la empresa del trabajador y que se integrará con los planes de pensiones del sistema de empleo y los Fondos de Pensiones Personales (FPI) cuyo ámbito de actuación se limitará al desarrollo de los Planes del sistema individual.

En el siguiente cuadro podemos ver las rentabilidades medias de los fondos de inversión ofrecidos y de los planes de pensiones individuales ofrecidos por las entidades tradicionales, comparando esto con las rentabilidades del Ibex35 vemos que las entidades tradicionales no son capaces de batir al índice de referencia.

RENTABILIDADES MEDIAS ANUALES PONDERADAS DE LOS FONDOS DE INVERSIÓN (31/08/18)

Tipo de Fondo	Mes	2018	1 Año	3 Años	5 Años	10 Años	15 Años	20 Años	25 Años
MONETARIOS	-0,08	-0,41	-0,53	-0,25	0,04	0,71	1,10	1,45	2,39
RENTA FIJA EURO CORTO PLAZO	-0,25	-0,95	-0,92	-0,21	0,16	0,71	1,15	1,51	2,43
RENTA FIJA EURO LARGO PLAZO	-0,48	-1,21	-1,01	0,40	1,73	2,50	2,19	2,42	3,31
RENTA FIJA MIXTA EURO	-0,92	-1,76	-1,37	0,03	1,59	1,63	2,02	1,80	3,03
RENTA VARIABLE MIXTA EURO	-1,72	-2,01	-1,56	0,54	3,83	2,00	3,07	1,89	3,71
RENTA VARIABLE NACIONAL EURO	-2,41	0,56	1,60	3,67	7,59	2,83	5,71	3,58	6,47
RENTA FIJA INTERNACIONAL	-0,55	-1,34	-1,32	-0,27	2,00	2,39	1,87	2,18	2,95
RENTA FIJA MIXTA INTERNACIONAL	-0,43	-1,20	-0,46	-0,04	1,04	0,81	1,16	1,11	2,34
RENTA VARIABLE MIXTA INTERNACIONAL	-0,34	-0,40	1,47	1,42	2,56	2,42	2,48	1,38	3,14
RENTA VARIABLE EURO RESTO	-3,25	-1,54	1,12	3,30	6,25	4,64	5,28	2,71	5,62
RENTA VBLE. INTERNACIONAL EUROPA	-1,83	-0,92	2,55	2,33	6,10	3,11	4,04	1,46	4,25
RENTA VBLE. INTERNACIONAL EEUU	3,52	9,27	17,26	12,00	12,87	9,05	5,91	4,07	2,55
RENTA VBLE. INTERNACIONAL JAPÓN	-0,26	-1,06	11,09	6,24	9,66	4,69	2,93	0,48	-0,78
RENTA VBLE. INTERNACIONAL EMERGENTES	-3,19	-6,00	-1,62	7,36	4,53	1,78	6,76	5,55	4,19
RENTA VBLE. INTERNACIONAL RESTO	-0,63	1,16	6,13	6,21	8,00	7,29	5,52	2,39	4,06
GLOBALES	-0,44	-0,44	1,91	2,06	3,64	2,37	2,39	1,79	3,26
GARANTIZADOS DE RENDIMIENTO FIJO	-0,96	-1,25	-1,07	-0,14	0,73	1,96	2,00	2,27	
GARANTIZADOS DE RENDIMIENTO VARIABLE	-1,20	-1,73	-1,40	0,12	1,36	1,67	1,93	2,08	
DE GARANTÍA PARCIAL	-0,65	-0,50	-0,10	0,40	3,25				
DE GESTIÓN PASIVA	-1,16	-0,35	0,07	1,22	3,33				
RETORNO ABSOLUTO	-0,41	-1,66	-1,17	-0,03	0,77				
FONDOS DE INVERSIÓN LIBRE (FIL)	-1,09	1,01	3,79	6,06	7,41				
FONDOS DE FIL	-0,40	1,72	1,47	-0,66	1,44				
TOTALES FONDOS	**-0,69**	**-0,73**	**0,43**	**1,10**	**2,16**	**2,06**	**2,28**	**1,82**	**2,86**

Rentabilidades medias anuales ponderadas. NETAS de gastos y comisiones (%)

SISTEMA INDIVIDUAL	26 AÑOS	25 AÑOS	20 AÑOS	15 AÑOS	10 AÑOS	5 AÑOS	3 AÑOS	1 AÑO
Renta Fija Corto Plazo	2,30	1,74	1,64	0,86	0,89	0,22	-0,35	**-1,43**
Renta Fija Largo Plazo	3,70	3,27	1,95	1,92	2,40	1,80	0,26	**-1,72**
Renta Fija Mixta	3,88	3,45	1,91	1,86	1,57	1,68	0,41	**-0,81**
Renta Variable Mixta	5,21	4,68	3,38	3,83	3,41	4,08	2,13	**1,89**
Renta Variable			2,16	4,52	4,71	8,39	5,44	**5,54**
Garantizados				2,81	3,23	4,77	1,70	**-0,36**
TOTAL PLANES	**3,90**	**3,47**	**2,29**	**2,54**	**2,45**	**3,24**	**1,47**	**0,54**

RENTABILIDAD DEL IBEX 35	
	IBEX 35 con dividendos Rentabilidad*
2007	10,71
2008	-36,50
2009	38,27
2010	-12,93
2011	-7,75
2012	2,78
2013	27,75
2014	8,62
2015	-3,55
2016	2,60
2017	11,25
Rentabilidad 10 años	10,74

** Rentabilidad calculada sobre el cierre del año anterior (%)*

Este problema no solo es propio de las entidades nacionales si no que se trata de un fenómeno generalizado, con un porcentaje de fondos que son capaces de superar de manera consistente al índice S&P en sus diferentes categorías, en el mejor de los casos solo un 6'88% de dichos fondos con capaces de hacerlo.

Fund category	Benchmark	Percentage of funds that outperformed their benchmark for three-years	Percentage of funds that outperformed in the subsequent year	Percentage of funds that outperformed for two subsequent years	Percentage of funds that outperformed for three subsequent years
Large-cap	S&P 500	30.83%	33.93%	13.62%	5.17%
Mid-cap	S&P MidCap 400	25.65%	30.39%	10.35%	3.24%
Small-cap	S&P SmallCap 600	30.58%	35.25%	13.30%	4.60%
International	S&P 700	23.89%	36.68%	15.56%	6.88%
Emerging markets	S&P/IFCI Composite	25.24%	38.39%	15.48%	5.22%

Si a esto le sumamos las comisiones que se les aplican a los clientes, observamos que la rentabilidad de los fondos disminuye aún más, efecto que se agrava por

el efecto de la inflación, lo que nos llevaría en ocasiones a rentabilidades negativas.

Debemos destacar que el impacto de las comisiones en la rentabilidad del cliente tiene mayor impacto en los fondos a más largo plazo, en este caso los planes de pensiones, esto puede poner en duda la capacidad de estos como los vehículos apropiados para vencer a la inflación y lo que es más importante como elemento de ahorro para la jubilación.

Categoría Estudio	Comisiones medias totales
INV - Internacional Globales Universe	1,0
INV - Internacional Renta Fija Mixta	1,2
INV - Internacional Renta Variable EEUU	1,7
INV - Internacional Renta Variable Emerg	1,7
INV - Internacional Renta Variable Euro	1,8
INV - Internacional Renta Variable Europ	1,7
INV - Internacional Renta Variable Japon	1,7
INV - Internacional Renta Variable Mixta	1,4
INV - Internacional Renta Variable Mundo	1,7
INV - Monetarios Universe	0,4
INV - Renta Fija Euro Corto Plaza Univer	0,7
INV - Renta Fija Euro Large Plaza Univer	0,7
INV - Renta Fija Internacional Universe	0,9
INV - Renta Fija Mixta Euro Universe	1,3
INV - Renta Variable Mixta Euro Universe	1,7
INV - Renta Variable Nacional Euro Unive	1,8
INV - Retorno Absoluto Universe	1,0
FPE	0,4
IPF - Renta Fija Corto Plazo Universe	1,1
IPF - Renta Fija Largo Plazo Universe	1,2
IPF - Renta Fija Mixta Universe	1,3
IPF - Renta Variable Mixta Universe	1,3
IPF - Renta Variable Universe	1,4

Tabla 1: Comisión media de gestión y depósito por tipo de IIC y categoría de riesgo

Tabla – inflación histórica España (IPC) – por año

inflación anual (dic resp. dic)	inflación
IPC España 2017	1,11 %
IPC España 2016	1,57 %
IPC España 2015	0,02 %
IPC España 2014	-1,04 %
IPC España 2013	0,25 %
IPC España 2012	2,87 %
IPC España 2011	2,38 %
IPC España 2010	2,99 %
IPC España 2009	0,79 %
IPC España 2008	1,43 %

En conclusión, podemos decir que a pesar de la creciente competencia dentro del mundo de los fondos y el cada vez mayor número de activos gestionados por los mismos, es que estos no obtienen rentabilidades superiores a los índices de

referencia, situación que se agrava aún más por el efecto de las comisiones cargadas a los clientes.

2.3 ANÁLISIS DE EMPRESAS CON USO DE AI EN TRADING ALGORÍTMICO

Se ha realizado también el análisis de algunas startups las cuales utilizan la inteligencia artificial en su estrategia de trading algorítmico:

Empresa	Descripción
sentient	Se trata de una de las startups basadas en IA que más fondos ha recibido. Tiene diversas soluciones de finanzas (trading), retail y salud en las que se hacen uso de inteligencia artificial. En lo referente a trading su solución se basa de la creación de un IA trader distribuido en el que los diferentes nodos serían los usuarios de la solución. Han planteado crear una spin-off de la empresa que se enfocara exclusivamente en esta solución.
Alpaca	Esta empresa ha creado una plataforma de trading algorítmico basada en análisis técnico. La plataforma le permite identificar ciertos patrones que considere bajistas o alcistas y luego construir una estrategia comercial basada en la

	apariencia de esos patrones. El componente "AI" ayuda a reconocer los patrones a medida que se forman todo ello en tiempo real.
WALNUT ALGORITHMS	Walnut Algorithms, compañía de investigación y desarrollo, enfocada en aplicar los últimos avances en ciencia de datos e investigación de aprendizaje automático a los mercados financieros. Proporciona un análisis de las ganancias de rendimiento esperadas que se pueden lograr a través de las tecnologías de aprendizaje automático e implementa modelos predictivos y herramientas de apoyo a la toma de decisiones.
BINATIX Machine Intelligence	Startup la cual ha creado algoritmos de trading basados en Deep Learning capaces de asesorar sobre estrategias de inversión. Su modelo de negocio se basa en la venta de estos sistemas de asesoramiento a diferentes hedge funds.
Kavout	Se trata de una empresa que ha creado una plataforma de inversión la cual ofrece asesoramiento a diferentes niveles de inversores.

	Ofrece asesoramiento sobre inversiones utilizando algoritmos de AI los cuales analizan el escaneo de archivos históricos y cotizaciones de acciones de la SEC y examinando millones de puntos de datos cada segundo mientras analiza las acciones utilizando un enfoque objetivo.
eye capital	Se trata de una empresa argentina que utiliza técnicas IA y NLP (para interpretación de noticias financieras de forma automática en tiempo real) con el fin de optimización de porfolio y selección de valores. Su solución está especializada en técnicas de Low Latency Algorithms, High Frequency Trading y Risk Management.

Adicional se ha realizado un análisis sobre el tamaño de estas empresas en cuanto a número de empleados, así como información sobre rondas de financiación que han obtenido:

Compañía	Tamaño (empleados)	Rondas financiación	Localización	Año fundación
Sentient	95	6 rondas (174.4M $)	USA	2006
Alpaca	15	5 rondas (2.8M $)	USA	2015
Walnut algorithms	23	4 rondas (1.5M €)	UE (Francia)	2014
Binatix	5	N/A	USA	2012
Kavolut	11-50	N/A	USA	2015
Eye Capital	1-10	N/A	UK	N/A

DEFINICIÓN DE LA SOLUCIÓN (TO BE)

3.1. IDENTIFICACIÓN DE LA OPORTUNIDAD

Tal como hemos puesto de manifiesto anteriormente los fondos de inversión tradicionales no son capaces de ofrecer rentabilidades superiores a los índices de referencia de manera consistente y en algunos casos ni siquiera son capaces de batir incluso a la inflación.

Ante este pobre comportamiento de la gran mayoría de los fondos y planes tradicionales de inversión hemos asistido al interés de los inversores por nuevos productos de inversión como ETF´s (Fondos cotizados), Fondos de gestión pasiva, Hedge Funds, etc.

Dentro de esta categoría de nuevos vehículos de ahorro, hemos asistido desde hace varios años a la incorporación de algoritmos de trading automático en los mercado financieros (renta variable, renta fija y derivados), estos sistemas utiliza operaciones aritméticas matemáticas o lógicas con diferentes grados de sofisticación que pueden ir desde los promedios móviles simples y otros métodos de análisis técnicos, hasta modelos estocásticos que requieren capacidades computacionales altas para manejar el volumen de datos de entrada y la complejidad en las operaciones y reglas incluidas en el algoritmo.

Una de las principales ventajas del trading algorítmico es que, al estar guiado por un conjunto de reglas o procedimientos que no involucran la emocionalidad humana, permiten evitar los sesgos conductuales más comunes entre los inversionistas, como son el exceso de confianza, los sesgos heurísticos, la aversión a la ambigüedad y la aversión al riesgo miope, entre otros, identificados por el campo de las finanzas conductuales.

No obstante, la aplicación de la Inteligencia Artificial (IA) en este campo aún no está extendida, lo que implica que existe una enorme oportunidad de crecimiento dentro de esta área en base a la aplicación de esta disciplina y en su capacidad de aprendizaje.

Nuestro reto es por lo tanto crear un modelo que sea capaz de predecir el comportamiento de los activos, proporcionando recomendaciones de inversión en base a estas predicciones, estando dotado el sistema, así mismo, de flexibilidad para ir ajustando el riesgo de forma sistemática con el paso del tiempo, de tal manera que el modelo tenga capacidad para adaptarse de forma autónoma a la situación del mercado y todo ello sin la intervención del ser humano.

Nuestro objetivo, por lo tanto, es crear un sistema de trading utilizando IA que sea capaz de predecir el comportamiento de los valores de renta variable de manera que obtengamos rentabilidades superiores a nuestro benchmark y con un nivel de riesgo inferior. Adicionalmente nuestra filosofía de inversión persigue obtener estas rentabilidades en el largo plazo, de manera que podamos establecer una relación de confianza con los inversores y podamos satisfacer sus expectativas de rentabilidad.

3.2. DEFINICIÓN CONCEPTUAL DEL MODELO Y POSICIONAMIENTO DE LA SOLUCIÓN

En lo referente al posicionamiento se ha realizado un análisis sobre la viabilidad inicial de las diferentes posibilidades. Pese a plantear inicialmente una idea sobre un posicionamiento de nuestra solución orientada a la venta directa al cliente final (pequeño/mediano inversor). Esta opción pronto fue descartada como una de las opciones al menos en una primera fase.

Tras un estudio más en detalle de la oportunidad se observó que un modelo B2C en este sector hacía necesaria una inversión inicial muy elevada para poder llegar a tener un MVP y poder testarlo. Estamos dentro en un sector con actores consolidados lo cual dificulta su entrada. Se trata de un modelo complejo ya que de nosotros dependería tanto la adquisición de clientes como la estructuración de toda la operativa (licencia de bróker dealer, servicios de collateral agent, escrow, estructuración de los feeder funds, etc.) las cuales son laboriosas y caras debido a los aspectos legales y regulatorios.

Una vez realizado este análisis y teniendo claro que para una primera fase la opción más viable sería un modelo B2B junto con un socio ya presente en el sector se nos han presentado dos posibles opciones de desarrollo:

- **Modelo B2B (cliente):** En este modelo nuestro cliente realmente sería un fondo de inversión ya existente y nuestro producto sería la utilización de nuestro modelo en modo servicio. En este caso la captación de clientes giraría en torno a la búsqueda de fondos los cuales utilizarían nuestro servicio, el cual ofrecerían a sus clientes.

  Se trata de un modelo más sencillo y con una inversión inferior necesaria para su desarrollo y explotación. En este modelo nos enfocaríamos en lo que realmente consideramos nuestro valor añadido (algoritmia) y dejaríamos la responsabilidad de la parte regulatoria y de adquisición de clientes a nuestro socio, el cual ya sería experto.

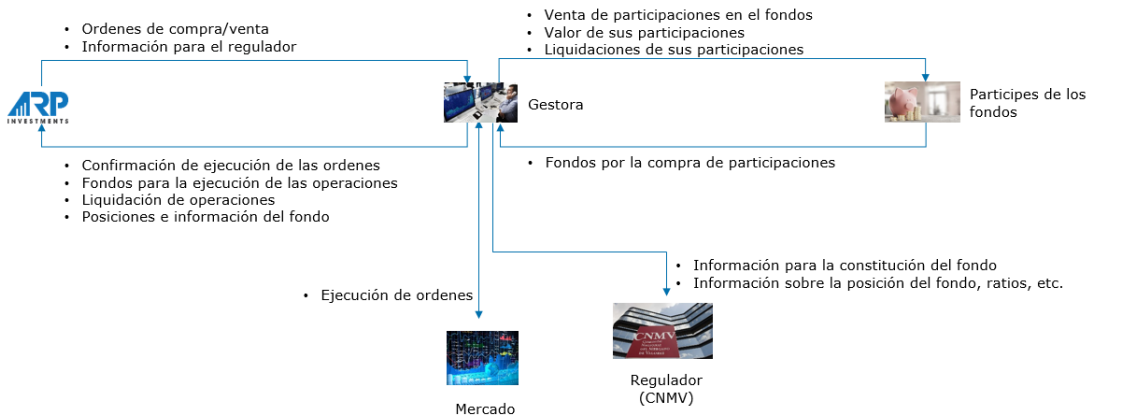
  La limitación que presenta esta opción se encuentra en que los fondos con los que se realizaría la asociación son los "dueños" reales del cliente final y se corre el riesgo (limitado) de que en algún momento puedan desarrollar una solución similar y llegar a prescindir de nuestros servicios.

- **Modelo B2B (partnership):** Consistiría en la búsqueda de un fondo "partner" junto con el cual venderíamos nuestro servicio a clientes finales (los beneficios se repartirían entre las partes). En esta opción nuestros clientes, como en el caso B2C, sería el cliente final con la diferencia que en este caso tanto la imagen de marca y los aspectos regulatorios ya estarían cubiertos por parte del partner.

Teniendo en cuenta el análisis planteado se ha decidido que el modelo a seguir será el modelo B2B en la opción partnership. Esta decisión se ha tomado debido a que se ha considerado la opción más simple de implementar en una primera fase y así poder centrarnos en lo que realmente sería nuestro valor añadido: el desarrollo de modelos de trading algorítmico utilizando tecnologías basadas en IA.

Una vez se consiguiera evolucionar nuestra solución y la compañía estuviera más consolidada, se podría pivotar hacia un modelo con una venta directa al cliente final. Pero como hemos indicado, estas opciones implican aspectos adicionales (captación de clientes, regulación...) los cuales requieren un know how y una inversión que se deberá ir adquiriendo en paralelo en una primera fase focalizada en la potenciación de nuestro principal valor diferencial.

La representación conceptual de los diferentes actores implicados en la solución sería la siguiente:



3.3. MODELO DE NEGOCIO

3.3.1 Misión, visión y valores

MISIÓN

Nuestra misión es la creación de fondos de gestión que superen de manera recurrente a los índices de referencia, todo esto haciendo uso de tecnologías innovadoras basadas en machine learning e inteligencia artificial y utilizando fuentes de datos financieras tradicionales y algunas fuentes alternativas como son noticias o RSS.

VISIÓN

La visión de nuestra empresa se centra en convertirse en una gestora de referencia a nivel nacional e internacional es los diferentes aspectos relacionados con la inversión en valores de diferentes mercados a partir del uso de las tecnologías descritas.

VALORES

- *Confianza*
- *Trato personalizado*
- *Sencillez*
- *Integridad profesional*
- *Excelencia: Know-how financiero y analítico*

3.3.2 DAFO

Fortalezas	Debilidades
• Algoritmo capaz de superar al índice de referencia. • Uso de IA como tecnología innovadora para la generación de las estrategias. • Aportaciones inicial reducida en comparación a otros fondos.	• Falta de imagen de marca. • Necesidad de asociación con un partner ya presente en el mercado
Oportunidades	**Amenazas**
• Capaces de ser independientes del ciclo económico. • Sin diferenciar por nivel de inversión de los clientes. • Solución tecnológica innovadora no utilizada en el sector de una forma generalizada	• Competencia en constante crecimiento. • Sector muy regulado. • Competencia consolidada. • Socio partner tome un control elevado sobre el servicio • Algoritmo no alcance los objetivos definidos.

3.3.3 Canvas

Aliados Clave	Actividades Clave	Propuesta de Valor	Relación con el Cliente	Segmentos de Clientes
- Socio financiero de referencia. - Compañías comercializadoras - Compañías obtención de datos financieros. - Compañías obtención da datos de noticias y RRSS. - Empresas de servicios cloud para despliegue de la solución.	- Gestión de fondos/ planes de pensiones ajustados a los diferentes niveles de riesgo del cliente. - Análisis de datos financieros y otras fuentes alternativas. - Creación de modelos basados en IA para la elaboración de inversiones. - Automatización de procesos	- Diseño y gestión de operativa de fondos de inversión que mejoran la rentabilidad respecto a los índices de referencia. - Modelo disruptivo	- Relación de confianza a largo plazo con un socio de referencia. - Acuerdos comerciales con los distribuidores de los fondos.	- Socio partner creador del fondo de inversión.
	Recursos Clave - Personal altamente cualificado con conocimientos financieros y en data science/IA - Fuentes de datos financieros - Fuentes de datos sobre noticias y RRSS. - Servicios cloud para el despliegue de la solución		**Canales** - **Venta:** Presencial - **Soporte**: Mail, teléfono.	

Estructura de Costes	Estructura de Ingresos
- RRHH (equipo de desarrollo de la solución, equipo de ventas...) - Costes de infraestructura (oficinas y equipos) - Costes de hosting de la solución y pago por uso servicios cloud. - Obtención de datos (financieros y noticias/RRSS). - Marketing	- Comisiones por gestión de fondos. - Comisión de performance.

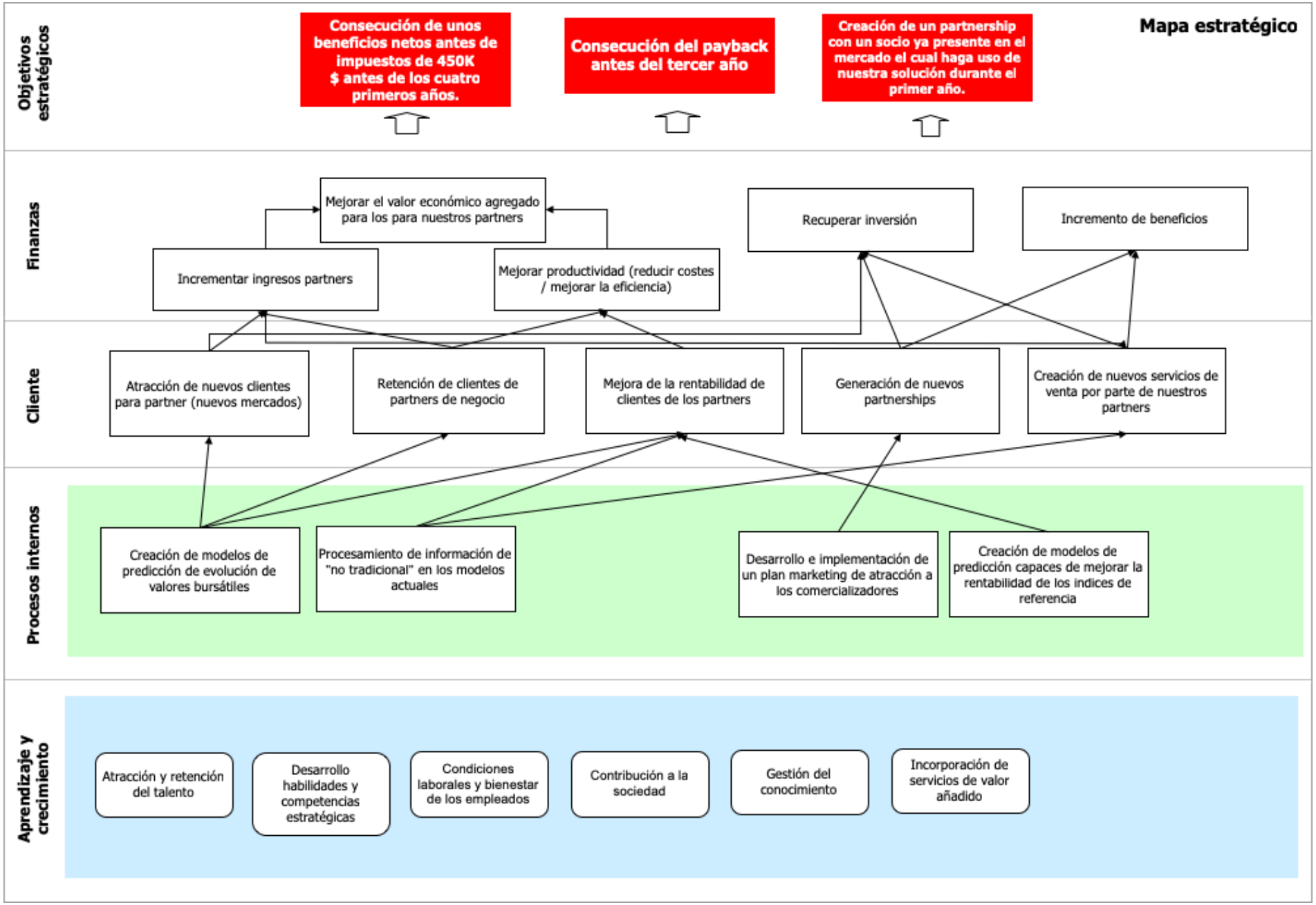
3.3.4 Objetivos estratégicos

Los objetivos estratégicos de nuestra compañía se centrarían en los siguientes puntos:

- *Consecución de unos beneficios antes de impuestos de 450K$ antes de los cuatro primeros años.*
- *Consecución del payback en el segundo año.*
- *Creación de un partnership con un socio ya presente en el mercado el cual haga uso de nuestra solución durante el primer año.*

Para lo cual se ha desarrollado el siguiente mapa estratégico con los diferentes indicadores.

3.3.5 Mapa estratégico


Mapa estratégico
Objetivos estratégicos
Consecución de unos beneficios netos antes de impuestos de 450K $ antes de los cuatro primeros años.
Consecución del payback antes del tercer año
Creación de un partnership con un socio ya presente en el mercado el cual haga uso de nuestra solución durante el primer año.
Finanzas
Mejorar el valor económico agregado para los para nuestros partners
Incrementar ingresos partners
Mejorar productividad (reducir costes / mejorar la eficiencia)
Recuperar inversión
Incremento de beneficios
Cliente
Atracción de nuevos clientes para partner (nuevos mercados)
Retención de clientes de partners de negocio
Mejora de la rentabilidad de clientes de los partners
Generación de nuevos partnerships
Creación de nuevos servicios de venta por parte de nuestros partners
Procesos internos
Creación de modelos de predicción de evolución de valores bursátiles
Procesamiento de información de "no tradicional" en los modelos actuales
Desarrollo e implementación de un plan marketing de atracción a los comercializadores
Creación de modelos de predicción capaces de mejorar la rentabilidad de los indices de referencia
Aprendizaje y crecimiento
Atracción y retención del talento
Desarrollo habilidades y competencias estratégicas
Condiciones laborales y bienestar de los empleados
Contribución a la sociedad
Gestión del conocimiento
Incorporación de servicios de valor añadido

3.3.6 Cuadro de mando

A partir del mapa estratégico se habría generado un cuadro de mando con los siguientes KPis:

PERSPECTIVA FINANCIERA

Respecto a la perspectiva financiera se pretende analizar cómo se obtendrían los objetivos desde un punto de vista financiero. Para la medición de esta perspectiva se plantea los siguientes objetivos que se van a medir con los siguientes indicadores:

Objetivo	Indicador	Target
Incrementar ingresos partners	Rentabilidad obtenida por parte de los partners	Mejora de un 10% respecto a los beneficios del partner
Mejorar productividad (reducir costes/ mejorar la eficiencia)	Reducción de costes en la obtención de las recomendaciones de inversión	Reducir los costes un 10% respecto a los modelos "tradicionales"
Recuperar inversión	Número y volumen de partners se ajusta a lo planificado	Conseguir al menos 1 acuerdo de colaboración con un volumen de negocio planteado en el Business Case
Incremento de beneficios	El volumen de cartera de los partners coincide con lo planificado	La cartera medida de los partners se ajusta al Business Case

PERSPECTIVA DE CLIENTES

En este apartado se trataría de medir nuestra propuesta de valor para nuestros clientes (fondos de inversión, planes de pensiones…). Para la medición de esta perspectiva se plantea los siguientes objetivos que se van a medir con los siguientes indicadores:

Objetivo	Indicador	Target
Atracción de nuevos clientes para partner (nuevos mercados)	Incremento de clientes nuevos finales de nuevos partners	Aumento en el número de suscriptores de los fondos de inversión en un 5%.
Retención de clientes de partners de negocio	Tiempo en el que los clientes finales mantienen su inversión en los servicios de nuestros socios	Incremento en un 20% del tiempo de permanencia de los clientes finales
Mejora de la rentabilidad de clientes de los partner	Retorno de la de inversión de nuestro modelo	Mejora de un 10% respecto a los beneficios obtenidos por parte de los servicios actuales de partner
Generación de nuevos partnerships	El número de acuerdos se ajusta a lo planificado	Conseguir 4 acuerdos de distribución dentro de los 3 primeros años.
Creación de nuevos servicios de venta por parte de nuestros partners	Generación de nuevo valor por parte de nuestros clientes	Encuestas de satisfacción

PERSPECTIVA DE PROCESOS INTERNOS

Respecto a los procesos internos se pretende medir los puntos de nuestra empresa en los cuales debemos buscar la excelencia con el fin de conseguir los objetivos de los niveles superiores. Para la medición de esta perspectiva se plantea los siguientes objetivos que se van a medir con los siguientes indicadores:

Objetivo	Indicador	Target
Creación de modelos de predicción de evolución de valores bursátiles	Generación de modelos de predicción que expliquen la evolución de los valores bursátiles	Los modelos generados sean capaces de explicar la realidad con porcentajes por encima del 95%
Procesamiento de información de "no tradicional" en los modelos actuales	Inclusión de información no tradicional en los modelos de predicción	Inclusión de información de noticias y RRSS en nuestros modelos de predicción.
Desarrollo e implementación de un plan marketing de atracción a los comercializadores	Incorporación de nuevos comercializares	Conseguir 4 acuerdos de distribución dentro de los 3 primeros años.
Creación de modelos de predicción capaces de mejorar la rentabilidad de los índices de referencia	Generación de modelos de predicción capaces de mejorar la rentabilidad respecto a los índices de referencia.	Los modelos generados sean capaces mejorar un 10% respecto a los beneficios obtenidos por parte de los índices de referencia

PERSPECTIVA DE DESARROLLO, FUTURO E INNOVACIÓN (INTANGIBLES)

En este apartado de medirían los aspectos necesarios para mantener la excelencia de cara a futuro. Para la medición de esta perspectiva se plantea los siguientes objetivos que se van a medir con los siguientes indicadores:

Objetivo	Indicador	Target
Atracción y retención del talento	Índice de rotación en la empresa	Rotación anual de la empresa por debajo del 0% en los 3 primeros años
Desarrollo habilidades y competencias estratégicas	Conseguir certificaciones técnicas y financieras en las áreas necesarias	Conseguir que todos los empleados/socios tengan al menos una certificación.
Condiciones laborales y bienestar de los empleados	Encuestas de clima laboral	Mejorar continua en las encuestas de clima
Contribución a la sociedad	Repercusión de una parte de los beneficios en la sociedad-	Dedicación de un 2% de los beneficios a acciones sociales
Gestión del conocimiento	Planes de formación adaptado a las necesidades de la empresa	Dedicación de un 5% del horario de los empleados a la formación de estos.
Incorporación de servicios de valor añadido	Desarrollo de nuevos servicios	Dedicación de un 20% del horario de los empleados a tareas de innovación

3.4 ROADMAP Y PLAN DE CONTINGENCIAS

El proceso de introducción de nuestra solución en el mercado se articularia en base a los siguientes puntos:

1) Identificación de posibles gestoras que pudieran actuar como nuestros socios de referencia, para ello deberían implementar una lista con todas aquellas a las que nuestra solución pudiera ser de interés, teniendo como objetivo prioritario a gestoras pequeñas y especialmente medianas por volumen de fondos gestionados, ya que el acceso a este tipo de entidades para la demostración de nuestra solución sería más fácil que aquellas más grandes. Una vez realizada esta primera selección realizaríamos un shortlist de 10 gestoras a las que contactaríamos de manera preferente en base a los fondos ofrecidos por la entidad y el nivel del desarrollo tecnológico de las mismas, siendo este último requisito una pauta fundamental en la identificación de las gestoras de referencia, ya que las mismas estarían más dispuestas a aceptar la implantación de este tipo de tecnología que aquellas basadas en modelos más tradicionales.

2) Contactar con las gestoras identificadas previamente y concertar citas con las mismas de manera que se pudiera realizar una prueba de concepto de nuestra solución. Durante 6 meses, y para demostrar el funcionamiento de nuestra solución, enviaríamos las recomendaciones de mercado a aquellas gestoras que hubieran demostrado interés en implementar nuestra herramienta como base para la gestión del fondo.

3) Tras este periodo de tiempo, que entendemos sería suficiente para demostrar las bondades de nuestro sistema, seleccionaríamos a aquella gestora que nos proporcionara las condiciones que se mejor alinearan a los objetivos de nuestro modelo de negocio.

4) Una vez realizada la selección de nuestro socio de referencia sería necesario la realización de un acuerdo entre ambas partes en las que se establecieran los puntos de dicha relación mediante un contrato que implicaría el uso de asesoría jurídica. La intención sería que esta relación

no fuera de exclusividad o si finalmente lo fuera tuviera un periodo de tiempo limitado.

5) La implantación del sistema implicaría realizar adaptaciones de nuestro modelo a la gestora, como por ejemplo como se conectaría nuestra herramienta con la mesa de operaciones mediante el desarrollo de una API o mediante un interfaz de usuario, etc.

6) Tras el establecimiento de la relación contractual sería necesario la realización de los trámites administrativos necesarios para la implantación de nuestro fondo con la CNMV para poder empezar a operar, estos trámites tal como hemos definido en nuestro modelo de negocio sería llevados a cabo por nuestro socio de referencia pero necesitaría colaboración de nuestra parte para la misma.

En caso de que no se pudiera establecer este roadmap tendríamos los siguientes planes de contingencia:

1) Ofrecer nuestra solución como una API a múltiples actores del mercado desde gestoras a Empresas de Asesoramiento Financiero (EAFI), o incluso a particulares.

2) Desarrollo de una herramienta de trading de alta frecuencia lo que nos llevaría a un modelo de negocio basado en la operativa diaria en lugar de la operativa a largo plazo basada en una cartera de valores.

3.5 POLÍTICA DE INVERSIÓN

Nuestra idea de negocio ha partido del desarrollo de un sistema de trading algorítmico en conjunto con el uso de Inteligencia Artificial, que nos proporcione una previsión sobre el comportamiento futuro de los valores, de manera que podamos construir una cartera en base a estas estimaciones.

A lo hora de definir nuestra cartera de inversión debemos en primer lugar establecer el número de valores que van a componer nuestra cartera, definido entre 20 y 30, la razón de esta selección del número de valores se debe a nuestro deseo de reducir el riesgo no sistemático.

Recordemos que el riesgo de cualquier cartera de inversión está compuesto por el riesgo sistemático o riesgo del mercado, todos los activos que forman un mercado financiero sufren este tipo de riesgo, por lo que todas las cotizaciones se acaban moviendo de forma más o menos paralela, unas más y otras menos. Cuando a causa de cualquier circunstancia sube el mercado, la práctica totalidad de los precios de los activos suben, y cuando baja el mercado, lo mismo.

El otro tipo de riesgo al que está expuesto nuestra cartera de acciones es al riesgo no sistemático o específico de nuestras acciones, este es aquél que resulta de factores propios y específicos de cada instrumento. En el caso del mercado de acciones, por ejemplo, podemos decir que el riesgo no sistemático, son aquellas situaciones que afectan de manera particular a esa empresa y no al resto, como por ejemplo el descubrimiento de un nuevo producto que haga el producto que tiene una empresa obsoleta.

Benjamin Graham en su libro El inversor Inteligente, uno de los libros más influyentes en el mundo de la inversión, dice que una diversificación adecuada para una cartera de inversión se puede obtener teniendo de 10 a 30 acciones de empresas diferentes, no obstante el mínimo de 20 valores para invertir nos viene dado por la normativa vigente, que impone como inversión máxima en una determinada acción de un 5% del capital del fondo.

Esto tiene la ventaja tal como hemos indicado antes de reducir el riego no sistemático, no obstante, y como contrapartida debemos indicar que esto nos impide hacer una apuesta mayor en unos pocos valores que nos podrían proporcionar una mayor rentabilidad, aunque unida la misma a un mayor riesgo.

Otra de los requerimientos que impone el regulador es que al menos un 1% del capital del fondo debe estar en liquidez de manera que se puedan hacer frente a las retiradas de capital solicitadas por los partícipes en el fondo, por lo tanto, el máximo de inversión al que podemos aspirar será de un 99% del capital obtenido.

Con respecto a la composición de los activos de nuestra cartera, planteamos diferentes niveles de inversión en renta variable y renta fija en función del nivel de riesgo que quiera asumir el inversor:

Perfil conservador: Inversión de un 25% en renta variable y el 75% en renta fija soberana con grado investment grade.

Perfil moderado: Inversión de un 50% en renta variable y un 50% en renta fija soberana con grado investment grade.

Perfil arriesgado: Inversión de un 90% en renta variable y 10% en renta fija soberana con grado investment grade.

En el caso de fondos de pensiones y dado que este es un vehículo de ahorro para la jubilación el perfil de riesgo máximo que se podría asumir seria moderado.

Con respecto a los grados de inversión que incluimos en la siguiente tabla, recordemos que el bono español a 10 años tiene una clasificación de riesgo moderado (BBB+ Fitch, BBB+ S&P y Moody's Baa3).

Fitch	S&P	Moody's	Rating grade description (Moody's)	
AAA	AAA	Aaa	Investment grade	Minimal credit risk
AA+ AA AA-	AA+ AA AA-	Aa1 Aa2 Aa3		Very low credit risk
A+ A A-	A+ A A-	A1 A2 A3		Low credit risk
BBB+ BBB BBB-	BBB+ BBB BBB-	Baa1 Baa2 Baa3		Moderate credit risk
BB+ BB BB-	BB+ BB BB-	Ba1 Ba2 Ba3	Speculative grade	Substantial credit risk
B+ B B-	B+ B B-	B1 B2 B3		High credit risk
CCC+ CCC CCC-	CCC+ CCC CCC-	Caa1 Caa2 Caa3		Very high credit risk
CC C	CC C	Ca		In or near default, with possibility of recovery
DDD DD D	SD D	C		In default, with little chance of recovery

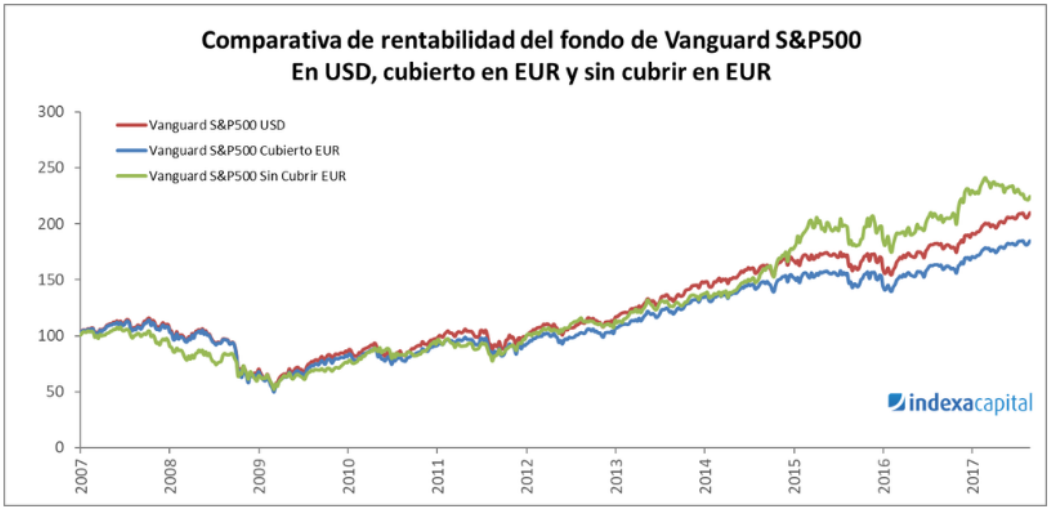
La inclusión de la renta fija dentro de nuestras carteras tiene como ventaja añadida que reduciría el riesgo sistemático de nuestro fondo debido a la correlación inversa de este tipo de activos con la renta variable.

Tal como podemos ver en el siguiente gráfico, en el que se muestra el retorno del SP500 y el Bloomberg Barclays US Aggregate Bond Index, este índice representa bonos invesment grade denominados en dólares, pudiendo el efecto de la diversificación ser aún mayor si incluyéramos bonos europeos denominados en euros como podría ser el caso del bono alemán o español a 10 años.

Con respecto a la propia generación de órdenes por nuestra solución, esta va proponer los valores a elegir en dos horizontes temporales diferentes, a un día y a siete días, el criterio para proponernos para la primer selección de valores que van a componer nuestra cartera, será aquellos que tengan una mayor posibilidad de apreciación a una semana, en base a esta propuesta seleccionaremos un mínimo de 20 valores, no invirtiendo en ningún de ellos más de un 5% del capital disponible en el fondo para la inversión en renta variable.

Con respecto a la selección de valores de renta fija proponemos un mix de al 50% entre renta fija a 10 años alemana de mayor calidad crediticia que la española y por lo tanto menor rentabilidad y renta fija a 10 años española con mayor riesgo y por lo tanto mayor rentabilidad.

Con respecto a la cobertura del riesgo divisa mediante derivados, no vamos a implementar este tipo de operaciones en el fondo, en base a que la rentabilidad de los fondos con este tipo de riesgo cubierto es inferior al de aquellos que si ponen en práctica este tipo de operativa debido a las comisiones cobradas y que las divisas reserva (Euro, dólar, libra, etc.) muestran una acusada reversión a la media.



Una vez que se hubiera lanzando esta primera selección de los valores que componen el fondo y basado en los dos horizontes temporales antes expuestos y teniendo en consideración que la ejecución de órdenes de compra-venta de acciones lleva aparejada una serie de costes (cánones de bolsa, comisiones, etc.) que deben ser pagados con cargo al capital comprometido por los partícipes en el fondo y que por lo tanto afectan a la rentabilidades obtenidas en la gestión de la cartera, adicionalmente nuestra filosofía de inversión siempre es el largo plazo y la ejecución de operaciones intradía.

3.6 OPERATIVA EN LA EJECUCIÓN DE LAS ORDENES

En base a lo dicho en el apartado anterior tendríamos las siguientes reglas para emitir órdenes de compraventa de los valores que componen nuestra cartera, para el rebalanceo de los activos que componen el fondo:

1. No hay límites en el número de valores que pueden entrar o salir del fondo, pudiendo rebalancearse el total de activos que la componen.

2. El universo de inversión es todo el S&P500 por lo que no se realizará discriminaciones en cuanto al tamaño/capitalización de los valores a incluir

en el fondo ni tampoco discriminaciones sectoriales (Bancos, Tecnológicas, etc.), es decir, no se tendrá en cuenta otro factor para incluir un valor en el fondo que aquellas recomendaciones realizadas por el motor de generación de órdenes.

3. Aquellos valores que tuvieran una perspectiva negativa, es decir, que el modelo indique que el valor de los mismo va a estar por debajo de su valor actual a una semana y que tengan una perspectiva negativa al día siguiente deberán ser descartados de la cartera o su peso en la misma reducido, siempre y cuando el motor nos indique a siete días que la bajada del valor sea mayor a un 1% y que el modelo nos indique que la probabilidad de bajada sea mayor o igual a un 97%, siendo excluidos primero aquellos cuya pérdida de valor sean más acusada, no obstante en caso de que la pérdida de valor de los activos en cartera, sea inferior a los valores propuestos por el modelo, no sea realizará acción alguna.

4. Aquellos valores que tuvieran una perspectiva positiva, es decir, que el modelo indique que el valor de los mismo va a estar por encima de su valor actual a una semana y que tengan una perspectiva positiva al día siguiente deberán ser incluidos en la cartera o si los mismos ya están incluidos incrementar su peso en la misma, siempre y cuando el motor nos indique a siete días que la subida del valor sea mayor a un 1% y que el modelo nos indique que la probabilidad de subida sea mayor o igual a un 97%, siendo incluidos primero aquellos cuyo incremento de valor sean más acusado, no obstante en caso de que el incremento de valor de los activos en cartera, sea superior a los valores propuestos por el modelo no sea realizará acción alguna.

5. Aquellos valores que tuvieran una perspectiva negativa a una semana y que tengan una perspectiva positiva al día siguiente, no serán candidatos para incluirse en la cartera.

6. Aquellos valores que tuvieran una perspectiva positiva a una semana y que tengan una perspectiva negativa al día siguiente, serán un posible candidato para incluirse en la cartera.

7. En caso de que el modelo indique una bajada generalizada del mercado, se podrá realizar una cobertura con derivados (Futuros, swaps, opciones, etc.) del SP500 por un importe no superior al total de los fondos gestionados, la razón para ello es que el uso de este tipo de activo financiero nos podría llevar a una sobre cobertura de nuestras carteras. Para implementar este tipo de coberturas la bajada en la cotización de los valores debe ser mayor a un 10% y debe afectar al menos a un 10% del capital invertido.

1 día	7 días	Porcentaje de subida	Porcentaje de bajada	Probabilidad de subida	Probabilidad de bajada	Compra / Venta
+	+	> 1%	N/A	> 97%	N/A	Comprar
−	−	N/A	> 1%	N/A	> 97%	Vender
+	−	N/A	> 1%	N/A	> 97%	Vender
−	+	> 1%	N/A	> 97%	N/A	Comprar

En la operativa antes expuesta, el papel de nuestro socio de referencia sería la ejecución de las órdenes de compraventa para la rotación de los valores de nuestra cartera, el envío de las órdenes a la mesa de ejecución se realizaría automáticamente mediante mail, con las ordenes seleccionadas en base a los precios de cierre del día anterior y teniendo en cuenta los parámetros antes definidos.

Con respecto a los parámetros antes indicados, estos no son estáticos, es decir, que pueden variar e irse adaptando a medida que el negocio evolucione, al igual que nuestro propio algoritmo, aunque por supuesto condicionados a cumplir la normativa vigente, especificada por el regulador español (CNMV).

Con esta política de envío de las órdenes a nuestro socio de referencias, conseguiríamos precios de ejecución por debajo del mercado, lo que redundaría en un incremento de nuestra rentabilidad, un incremento de los ingresos de nuestro socio y adicionalmente al ser nuestra ordenes ejecutadas en la mesas de renta variable de nuestro socio esto implicaría la confidencialidad de nuestras operaciones, esto es de suma importancia ya que nuestro valor se basa

precisamente en la selección de acciones y que esta operativa no se puede replicar.

3.7 BUSINESS CASE

A continuación, se detalla el Business case en el que se aborda la viabilidad del proyecto.

En primer lugar, se presenta una tabla detallada con las principales partidas de gastos e ingresos que se han considerado en el proyecto:

	Año 1	Año 2	Año 3	Año 4	Total
Gastos					
Tecnología	7.000 €	9.500 €	12.000 €	12.000 €	**40.500 €**
Servicios cloud/Infraestuctura	5.000 €	7.500 €	10.000 €	10.000 €	**32.500 €**
Fuentes de datos	2.000 €	2.000 €	2.000 €	2.000 €	**8.000 €**
Gastos estructurales	305.000 €	307.000 €	346.000 €	356.000 €	**1.314.000 €**
RRHH	215.000 €	220.000 €	260.000 €	270.000 €	**965.000 €**
Marketing	50.000 €	45.000 €	40.000 €	40.000 €	**175.000 €**
Oficina	36.000 €	36.000 €	36.000 €	36.000 €	**144.000 €**
Renting equipo informáticos	4.000 €	6.000 €	10.000 €	10.000 €	**30.000 €**
Servicios externos	6.000 €	83.224 €	83.180 €	83.159 €	**255.563 €**
Administración y asesoría legal	6.000 €	6.000 €	6.000 €	6.000 €	**24.000 €**
Devolución crédito	- €	77.224 €	77.180 €	77.159 €	**231.563 €**
Total Gastos	318.000 €	399.724 €	441.180 €	451.159 €	**1.610.063 €**
Inversión inicial	200.000 €	- €	- €	- €	**240.000 €**
Ingresos					
Comisiones de gestión	105.000 €	420.000 €	630.000 €	1.050.000 €	**2.205.000 €**
Comisiones de éxito	54.000 €	162.000 €	162.000 €	378.000 €	**756.000 €**
Total ingresos	159.000 €	582.000 €	792.000 €	1.428.000 €	**2.961.000 €**

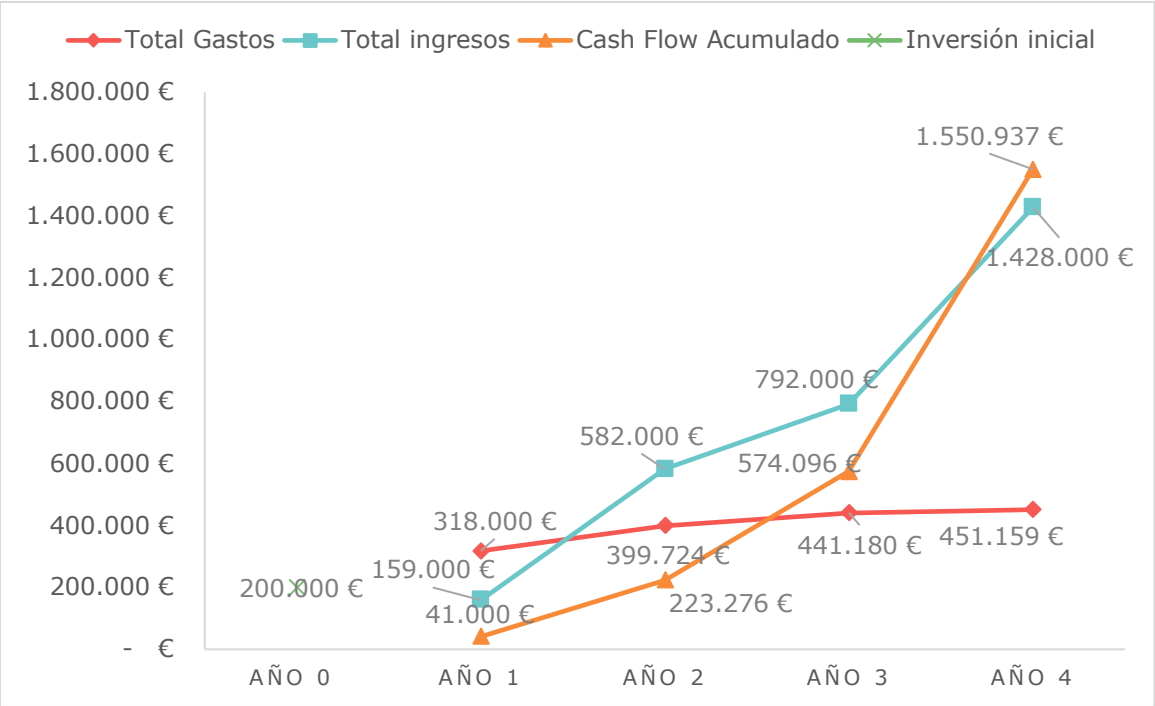
Esto da lugar a unos los siguientes beneficios anuales y acumulados:

	Año 1	Año 2	Año 3	Año 4
Beneficios	- 159.000 €	182.276 €	350.820 €	976.841 €
Beneficios acumulados	- 159.000 €	23.276 €	374.096 €	1.350.937 €

Y a un cash-flow:

	Año 1	Año 2	Año 3	Año 4
Cash Flow	41.000 €	182.276 €	350.820 €	976.841 €
Cash Flow Acumulado	41.000 €	223.276 €	574.096 €	1.550.937 €

Gráficamente quedaría del siguiente modo:



A continuación, detallaríamos cada una de las partidas.

3.7.1 Ingresos e inversión inicial

INVERSIÓN INICIAL

Tras un análisis de la oportunidad en detalle se validó que para la puesta en marcha del proyecto se hacía necesario de una inversión inicial de 200.000 €. Los cuales inicialmente se han planteado obtener a partir de un préstamo bancario con las siguientes condiciones:
Se trataría de un préstamo a 48 meses con unos 12 primeros meses de cadencia y un TAE 6%.

Todo ello con el fin de poder hacer frente a los primeros meses de funcionamiento de la empresa en la que no se percibirían ingresos (se ha considerado un periodo de 6 meses necesarios para finalizar la implementación de la solución y la negociación con un socio de referencia) y adicional con el fin de mantener un flujo de caja acumulado positivo durante los primeros años del proyecto.

INGRESOS

Los ingresos en este proyecto, al igual que en el resto de los gestores de fondos, proceden principalmente de las diferentes comisiones percibidas por la gestión de estos.

En este caso se realizó un análisis de los diferentes tipos de comisiones existentes, así como los límites máximos legales aplicables (se tomó España como país de referencia para este análisis) y se consideró que el reparto de estas se debería realizar al 50 % entre nuestra compañía y nuestro socio junto con la comercializadora.

Así pues, las comisiones que se consideraron fueron (en nuestro caso nos correspondería el 50 %):

- Comisión de gestión: 1,35%
- Comisión de éxito: 9%

Para el cálculo de la comisión de éxito se consideró que en media nuestros modelos tendrían una ganancia media anual del 8% sobre el capital invertido, (para esto se ha tenido en cuenta la ganancia histórica en bolsa y hemos aplicado nuestro KPI de mejora sobre el índice de referencia).
Adicional para el cálculo de los ingresos se analizó la capacidad de obtención de capital que sería gestionado a través de nuestra compañía y realizamos la siguiente estimación (esta estimación está basada en base al análisis del volumen de fondos gestionados actualmente en España http://www.inverco.es/38/39/101/2018/6):

- Año 1: 30 M € (Durante 6 meses)
- Año 2: 60 M €
- Año 3: 90 M €
- Año 4: 150 M €

Con estos volúmenes en el cuarto año representaríamos el 0,056% del mercado en España.

3.7.2 Gastos

El apartado se detalla y justifica de forma explícita en los siguientes apartados sobre los planes de acción.

3.7.3 Riesgos analizados

Para la obtención de los principales indicadores también se analizaron los principales riesgos en nuestro proyecto. Identificándose como principales tres de ellos:

- Algoritmo se comporta por debajo de las expectativas
- No conseguir socio financiero en el primer año
- La captación de comercializadoras se encuentra por debajo las expectativas.

Sobre estos tres riesgos se trató de cuantificarlos con el fin de poder ponderarlos dentro del resultado de nuestro de caso de negocio y se aplicaron las siguientes métricas:

Tipo de riesgo	Probabilidad	Impacto			Coeficiente de corrección		
		Inversión	Δ Ingresos	∇ Costes	Inversión	Δ Ingresos	∇ Costes
Algoritmo se coporta por debajo de las espectativas	40%	Muy Alto	Medio	Muy Bajo	0,10	0,02	0,01
No conseguir socio financiero en el primer año	20%	Muy Alto	Muy Alto	Muy Bajo	0,05	0,02	0,00
Captación de comercializadoras por debajo de nuestras espec	40%	Muy Bajo	Muy Alto	Muy Bajo	0,02	0,04	0,01

3.7.4 Principales indicadores oportunidad de negocio.

Los principales indicadores del proyecto, teniendo también en cuenta el impacto de los principales riesgos identificados, quedarían del siguiente modo:

ROI: 424 %

Payback: Año 2

VAN: 976.270 €

TIR: 124 %

3.8 PROCESOS Y PLANES DE ACCIÓN

3.8.1 Plan de marketing

Nuestras acciones dentro de este apartado estarían orientadas en varias direcciones: aquellas enfocadas a la consecución de una alianza con nuestro socio de estratégico, aquellas que tendríamos que lanzar para obtener un acuerdo con distribuidoras de fondos y aquellas enfocadas a plataformas de investigación tales como Morningstar.

Dentro de estas dos líneas de acción, el desembolso no sería excesivo ya que no tendríamos mejor carta de presentación que la propia rentabilidad obtenida por nuestro algoritmo, que podría ser fácilmente contrastado por nuestro posible socio, por lo tanto, las acciones de marketing irían enfocadas a la presentación de la solución a los actores identificados anteriormente.

En relación con nuestras acciones de marketing dirigidas a distribuidoras de fondos y plataformas de investigación de mercados, tendríamos que tener en cuenta la realización de acciones de comarketing junto con nuestro socio de referencia, tales como eventos de presentación de nuestros fondos y sobre todo de sus resultados, de manera que pudiéramos dar a conocer los mismos de una manera más amplia.

3.8.2 Plan de RRHH

Los valores de nuestra empresa se basan en las personas. Somos una empresa con vocación de crecimiento, aunque los primeros años deberán ser cautos en este capítulo.

Iniciaremos esta andadura apenas con un equipo de dirección que estará formado por los tres socios fundadores asumiendo los siguientes roles:

Pedro – CEO y CFO

Adrián – CDO y CIO

Raúl – CTO y COO

Los tres socios fundadores tendrán un salario, fijándose este en 60.000€ el primer año. Dicho salario se incrementaría en función de los resultados obtenidos.

En los primeros años no planteamos recibir bonus ya que los beneficios los reinvertiríamos en la empresa y en el incremento de la participación en el propio fondo. La participación con nuestro capital en el fondo tiene como beneficio dar una señal a inversores de nuestra confianza en nuestros fondos, ya que estamos compartiendo con ellos beneficios y riesgos.

Adicionalmente será necesaria la contratación inicial de una responsable de administración que percibirá una retribución bruta anual de 20.000 €.

También será necesaria la contratación de un desarrollador que haga las veces de técnico de sistemas, su retribución anual está en torno a los 30.000 € anuales.

El crecimiento previsto hará necesaria la contratación de un científico de datos al inicio del tercer año, esta persona será responsable por la evolución y mejora del modelo de recomendación.

3.8.3 Plan de operación

Para el inicio de la operación nuestro primer paso consiste en encontrar una fuente de financiación para arrancar la empresa. Para la cual buscaremos el apoyo de la banca o algún fondo de capital riesgo para conseguir el dinero que con permita iniciar nuestra andadura,

Nuestro siguiente paso será cerrar el acuerdo con nuestro socio tecnológico, para esto hemos elegido a IBM que con **Global Entrepreneur** su plan de apoyo al emprendedor y a los startups, nos aportará tanto el apoyo tecnológico como humano para iniciar los primeros pasos en IBM Cloud.

A través de la iniciativa Global Entrepreneur, IBM facilita a las empresas los siguientes recursos:

- Acceso sin coste al catálogo de software de IBM para prueba y desarrollo de soluciones on-premise o "en la nube".

- Soporte de un experto que les guiará en las diferentes fases de su proyecto de desarrollo.

- Participación en las competiciones SmartCamp, que tienen lugar alrededor del mundo y suponen una oportunidad única para recibir apoyo y asesoramiento de la mano de empresas de capital riesgo (venture capitals), inversores locales e internacionales, ejecutivos de la industria y emprendedores de referencia.

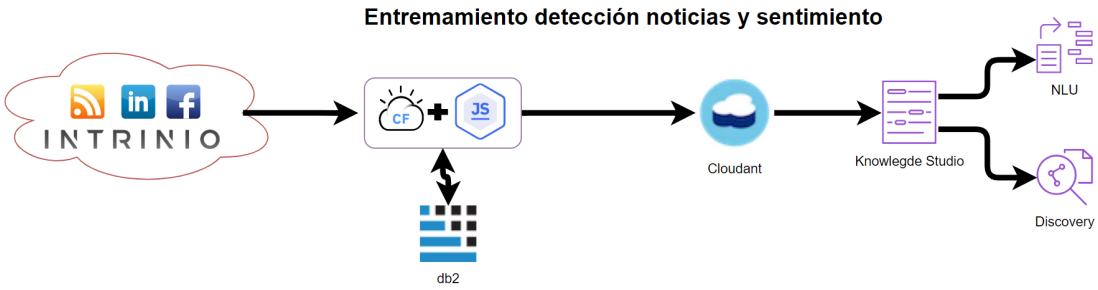
- Acceso a una nueva comunidad social, integrada en Developerworks Community, para colaborar y entrar en contacto con otros emprendedores y más de ocho millones de profesionales del sector de las tecnologías de la información en todo el mundo.

Una vez resulto el tema tecnológico, nos centraremos en la contratación del responsable de administración y del técnico desarrollador.

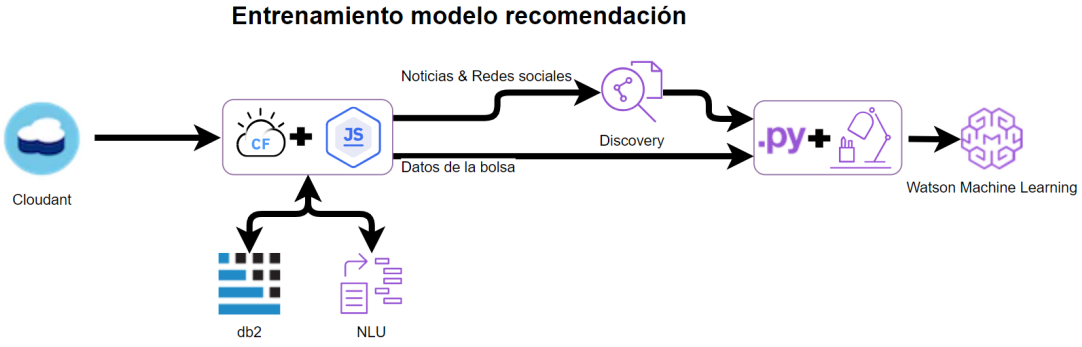
Posteriormente iniciaremos el paso más delicado, cerrar el acuerdo de colaboración con nuestro partner financiero, que es el que nos dará la capacidad de operar los fondos que componga nuestro modelo.

3.9 IMPLEMENTACIÓN TECNOLÓGICA

3.9.1 Definición conceptual de la arquitectura



- A través de una aplicación desarrollada en nodeJS se obtienen las noticias e información de valores de Intrinio y otras redes sociales.
- Toda la información obtenida es almacenada en un servidor Cloudant.
- Utilizamos Knowledge Studio para entrenar un nuevo dominio de reconocimiento de entidades, patrones, clasificación y detección de sentimiento.
- Desplegamos el nuevo modelo obtenido en el NLU y Watson Discovery.



- A través de una aplicación desarrollada en nodeJS se obtienen los datos almacenado en Cloudant.
- Utilizamos NLU para detectar las noticias que hablan de temas relevantes y son enviadas a Watson Discovery, Discovery analizará las noticias detectando relaciones, sentimiento, etc, dejando toda la información preparada para entrenar el modelo.
- Entrenamos el modelo desde Watson Studio con toda la información obtenida.
- Desplegamos el modelo en Watson Machine Learning.

3.9.2 Tecnologías a utilizar

	Tecnología	Funcionalidad
	Discovery	
	Watson Machine Learning	Plataforma de ejecución del modelo estadístico de predicción
	Natural Language Understanding	Analiza y valora las noticias de la bolsa
	Watson Studio	Plataforma para el entrenamiento del modelo de aprendizaje profundo
	Knowlegde Studio	Plataforma de entrenamiento de los dominios necesarios para NLU
	Db2	Base de datos Relacional

	Cloudant	Base de datos Big Data para el almacenamiento de las noticias y la información recibida de las diferentes bolsas
	Node JS	Integrador de los diferentes servicios que expondrá los resultados del modelo mediante API Rest
	Cloud Foundry Enterprise	Entorno de ejecución de la solución
	Keras	Librería de Deep Learning para el diseño de modelos
	Tensorflow	Librería de Deep Learning para el diseño de modelos

3.9.3 Arquitectura técnica de la solución

Nuestra plataforma se basará en el análisis diario y predicción de la tendencia de todas las acciones del S&P500, estimando las probabilidades de subir o bajar respecto a la posición actual, para realizar dichas estimaciones utilizaremos la información histórica del S&P500, así como todas las noticias financieras que hablen de los valores del índice, esto nos ayudará a predecir posibles comportamientos de los valores basados en noticias tanto positivas como negativas. Nuestro algoritmo tendrá en cuenta como otros tipos de noticias que no hablen directamente de los valores del S&P500 pueden afectar positiva o negativa a dichos valores, como pueden ser noticias políticas, climatológicas, sociales, deportivas, etc.

Para poder llevar a cabo estas tareas nuestro sistema se encargará de recopilar diariamente todo tipo de noticias de diferentes medios públicos, así como de recopilar toda la información financiera de los valores a través de nuestro proveedor Intrinio (con el que firmaremos un acuerdo de colaboración).

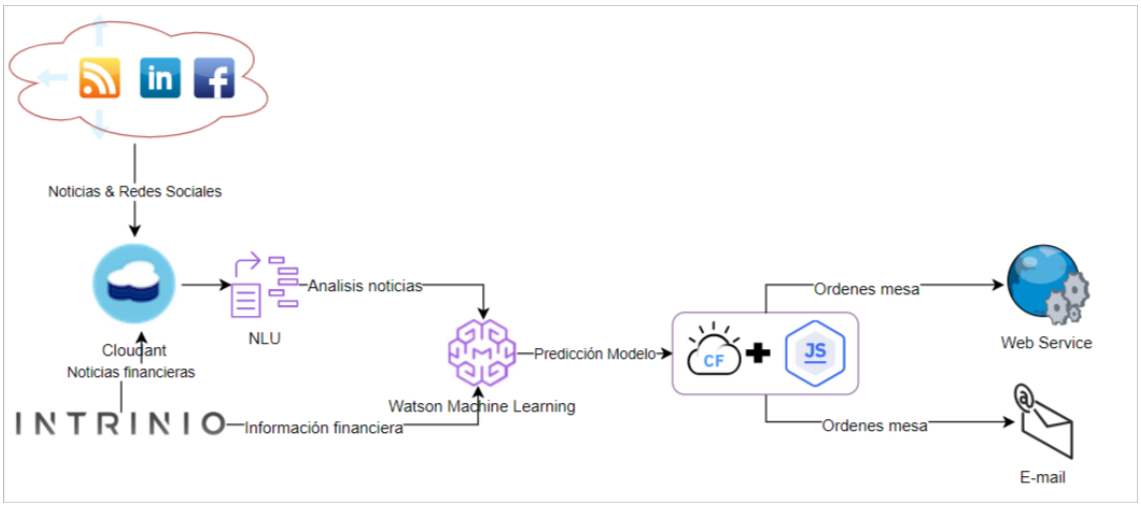
Toda la información y noticias recogidas serán almacenadas en nuestro almacén de datos basado en Cloudant, para posteriormente aplicar nuestro modelo de "Reconocimiento natural del lenguaje" (NLU) que nos devolverá si dichas noticias son relevantes para los valores del S&P500 y en qué medida (detalle descrito en el apartado 3.8), ayudando a predecir de forma más exacta el comportamiento de los valores.

Sobre los datos actuales de los valores en bolsa, conjugados con toda la información obtenida de las noticias, hemos desarrollado un modelo de Deep Learning (detalle descrito en el apartado 3.8) capaz de obtener 2 predicciones compuestas por el porcentaje de variación de los valores para el día siguiente y el porcentaje de variación de los valores para los próximos 7 días.

Esta información será utilizada para definir la cartera tipo recomendada, así como las actuaciones que haya que aplicar sobre las carteras tipo existentes actualmente. Recomendado cambiar los pesos entre los valores de la cartera, así como la venta de algunos de ellos o la compra de otros.

Toda esta información se convertirá en órdenes de compraventa que serán puestas a disposición de la mesa de nuestro partner, que se encargará de ejecutarlas y dejar configurada la cartera tal y como nuestro recomendador indique.

Para hacer llegar toda esta información a la mesa dispondremos de dos modos, uno mediante envío de correo electrónico a la mesa y el otro será un servicio web que se podrá integrar con los servicios de compra venta de la mesa para hacerle llegar la ordenes de manera automática.



3.10 FUENTES DE DATOS

3.10.1 Análisis fuentes de datos existentes

Se han analizado posibles fuentes de datos necesarias para el buen funcionamiento de los modelos a desarrollar:

INFORMACIÓN DE VALORES:

API de FRED: La fuente de datos sería la Reserva Federal de Datos Económicos de Estados Unidos. Es un servicio web que contiene una serie cronológica de macroeconomía y economía regional de los Estados Unidos actualizada con regularidad y frecuencia anual, trimestral, mensual, semanal y diaria. FRED también agrega datos económicos de distintas fuentes, la mayoría de las cuales son organismos gubernamentales estadounidenses.

FRED proporciona esos datos a través de una API. Se pueden utilizar gratuitamente simplemente dándose de alta para obtener una clave del API. Más información en la página de documentos de la API

Información sobre la API: https://research.stlouisfed.org/docs/api/fred/

Quandl: Se trata una plataforma de datos que ofrece acceso gratuito a fuentes de datos financieros públicos a través de API. La cantidad de conjuntos de datos disponibles de distintas fuentes es realmente enorme y la información facilitada también es muy variada: datos sobre acciones, datos sobre economía general, beneficios empresariales, estadísticas por país, materias primas, etc.

Información sobre la API: https://www.quandl.com/

EODHistoricalData.com: Se trata de una API la cual dispone información histórica sobre valores de diferentes mercados.

Información sobre la API: https://eodhistoricaldata.com/knowledgebase/api-for-historical-data-and-volumes/

Alpha Vantage: Se trata de una API gratuita con información financiera en tiempo real e histórica sobre diferentes valores:

Información sobre la API: h https://www.alphavantage.co/documentation/

Barchart: Similar a la anterior siendo una API de pago

Información sobre la API: https://www.barchart.com/

Intrinio: Se trata de un servicio que proporciona información financiera variada sobre valores de bolsa. La información facilitada por este servicio estaría relacionada con información estructurada (valores de cierre de un día, máximos, mínimos…) como información no estructurada (links a noticias financieras de los valores a analizar).

Información sobre el servicio: https://intrinio.com/

INFORMACIÓN NO ESTRUCTURADA:

Newsapi.org: Se trata de un API la cual permite obtener noticias de los principales medios de información en múltiples países y sobre una temática definida. La API facilita el siguiente tipo de información de cada noticia (Titulo, autor, fuente, headlines, url y fecha de publicación).

Información sobre la API: https://newsapi.org/

Webhose.io: Se trata de una API la cual facilita información similar a la anterior con el añadido de que facilita información adicional sobre RRSS (como ha sido compartida en RRSS), el tipo de fuente…

Información sobre la API: https://webhose.io/products/news-feeds/

Intrinio: Ya descrito en el apartado anterior.

Información sobre el servicio: https://intrinio.com/

Sobre este listado se ha realizado un análisis de las posibles fuentes de datos para poder analizar cuáles serían sus ventajas e inconvenientes:

INFORMACIÓN ESTRUCTURADA FINANCIERA:

Servicio	Obtención de datos	Formatos	Coste	Información no estructurada
FRED	- API	- JSON - XML	Gratuita	No
Quandl	- Descarga directa - API	- JSON - XML - CSV - Excel	Pago por consulta / dataset	No
EODHistoricalData	- API	- Json - CSV	Cuota mensual	No
Alpha Vantage	- API	- Json - CSV	Gratuita	No
Barchart	- API	- Json	Pago por consulta	No

Intrinio	- Descarga directa - API	- Json - Csv - Excel - XML	Cuota mensual y pago por descarga dataset	Si, incluye links a noticias financieras sobre una compañía consultada

INFORMACIÓN NO ESTRUCTURADA (NOTICIAS/RRSS):

Servicio	Obtención de datos	Formatos	Coste	Tipo de información
Newsapi.org	- API	- JSON	Cuota mensual	Noticias (temáticas)
Webhose.io	- API	- JSON	Cuota mensual	Noticias con información de su difusión en Redes Sociales
Intrinio	- API - Descarga directa	- Json - CSV - Excel - XML	Cuota mensual y pago por descarga dataset	Noticias financieras vinculadas a un valor bursátil

3.10.2 Fuentes y modelo de datos a utilizar

Con el análisis desarrollado en el apartado anterior y habiendo validado la calidad de los datos y el histórico de información que seríamos capaces de obtener, el servicio que nos parecería más completo sería el que ofrece Intrinio.

Otro de los motivos para utilizar este servicio es que realmente dispone de una amplia red de partners los cuales van ampliando constantemente la cantidad de información disponible gracias a la creación de nuevos partnerships que permiten hacer uso de nuevos tipos de informaciones disponibles (información bursátil, información financiera de una compañía, ETFs, estimaciones financieras, información bursátil en tiempo real…). A continuación, se muestra un listado de los principales partners de fuentes de datos disponibles:

No obstante, no se descarta, como una posible mejora, una combinación con otros de los servicios. Esto aportaría un gran valor ya que algunos de estos servicios ofrecen información diferencial como puede ser el caso de webhose.io que permiten conocer como la información sobre noticias de un valor bursátil es difundida a través de las redes sociales.

A continuación, se muestra el listado de información y los datos a utilizar en nuestro modelo. Como se ha comentado anteriormente y se detallará más en los siguientes apartados, nuestro sistema tiene dos partes claramente diferenciadas: Una primera parte de análisis de información no estructurada de noticias y una segunda parte en la que se combinaría el resultado del análisis de información procedente del análisis de noticias (resultado obtenido ya sería información estructurada) y la información financiera "tradicional" sobre cotizaciones (información estructurada). Así pues, se van a detallar los datos de entrada y de salida de cada una de las partes del sistema descrito.

ANÁLISIS DE NOTICIAS

En esta parte la información de entrada sería la procedente del API de Intrinio sobre noticias financieras en la que a partir del nombre de una cotización y de un rango de fechas de búsqueda, devuelve el listado de noticias sobre dicha cotización en el periodo indicado. Por tanto, de este proveedor obtendríamos la siguiente información de cada noticia:

- *Titulo*

- *Fecha de publicación*
- *URL*
- *Resumen de la noticia.*

De la información disponible el servicio de análisis de noticias (detallado en el próximo apartado) podríamos introducir como entrada el resumen de la noticia. Sin embargo, dado que el servicio también acepta el envío de urls para su análisis, utilizaremos la url facilitada por Intrinio para analizar la noticia completa.

El resultado del análisis de las noticias devuelve numerosa información. Sin embargo, tras la realización de un análisis de esta, la información que se utilizará como entrada del modelo de predicción de valores es la relativa al análisis de sentimiento y emociones la cual se describe a continuación:

- **Análisis de sentimiento:** Valor entre -1 y 1 representando el valor -1 un sentimiento negativo y 1 un sentimiento positivo.
- **Análisis de emociones:** Valor entre 0 y 1 representando la probabilidad de que una emoción concreta se encuentre reflejada en una noticia. Las emociones disponibles son las siguientes:
    - Alegría
    - Enfado
    - Disgusto
    - Tristeza
    - Temor

MODELO DE PREDICCIÓN DE VALORES

En esta parte se describen los datos que formarían parte de la entrada de nuestro modelo, así como el resultado el resultado de este. Como parámetros de entrada tendríamos dos grupos de datos:

- Información bursátil de una cotización (procedente del API de Intrinio):
    - Máximo en un día.
    - Mínimo en un día.
    - Precios de apertura en un día.
    - Precios de cierre en un día.
    - Volumen de acciones

- Resultado análisis noticias del día:
    - Sentimiento: Mediana, máximo y mínimo del día
    - Emoción (alegría): Mediana, máximo y mínimo del día
    - Emoción (enfado): Mediana, máximo y mínimo del día
    - Emoción (disgusto): Mediana, máximo y mínimo del día
    - Emoción (tristeza): Mediana, máximo y mínimo del día
    - Emoción (temor): Mediana, máximo y mínimo del día

Dado que en un día concreto pueden existir diversas noticias sobre una cotización, sobre el resultado del análisis de noticias descrito en el apartado anterior se realizaría un preprocesado del resultado para utilizarlo como valores de entrada del modelo de predicción, con el fin de poder obtener más información. El detalle de este procesamiento previo se describe en el siguiente apartado de la memoria.

Como resultado de este modelo, lo que tendríamos es el valor de la cotización en un día siguiente o en varios días posteriores para ser capaces de identificar tendencias de dicha cotización.

Otra información útil que nos facilita el modelo es la precisión de este y de sus predicciones. Por tanto, podríamos utilizar dicha precisión para poder calcular el rango de variación posible de la predicción.

Como comentario general hay que indicar que el listado de datos descritos es un subconjunto de toda la información financiera estructurada disponible. Estos datos han sido utilizados para desarrollar la prueba de concepto descrita a continuación. No obstante, el modelo descrito no presentaría ningún problema, si dicha información estructurada fuera ampliada con nuevos valores estructurados.

3.11 SIMULACIÓN DE MÉTODOS ANALÍTICOS

De cara a la elaboración de los métodos analíticos que van a ser capaces de predecir los valores y tendencias de las acciones a partir las cuales nuestro sistema operará dentro de los fondos que gestionemos se han utilizado diferentes tecnologías.

En primer lugar, para tratar de abordar el problema principal "predicción de los valores de una acción en un periodo de tiempo futuro" se trató de dividir el caso en problemas más pequeños. Así pues, identificamos tres partes claramente diferenciadas:

- *Análisis y predicción de valores a partir de información estructurada financiera.*
- *Análisis y procesamiento de información no estructurada procedente de noticias para la mejora de las predicciones sobre los valores.*
- *Análisis y predicción de valores a partir de la combinación de la información estructurada financiera y no estructurada procedente de noticias.*

Por tanto, vamos a realizar una descripción de cada una de las fases con el fin de mostrar el proceso completo de construcción de un modelo el cual es capaz de solventar la problemática planteada.

3.11.1 Análisis y predicción de valores a partir de información financiera

En un primer momento se trató de analizar información estructurada financiera "tradicional" a partir de la cual durante mucho tiempo se ha venido implementado diferentes modelos de predicción de valores. Algunos ejemplos de esta información podrían ser:

- *Máximos y mínimos de una cotización.*

- *Precios de cierre*
- *Precios de apertura*
- *Volumen de acciones*
- *Beneficios*
- *Ingresos*

En nuestro caso, a partir de esta información hemos tratado de crear un modelo de predicción de valores, pero en este caso utilizando el Deep Learning como herramienta.

Para ello, tras un análisis de las diferentes posibilidades que ofrecía este tipo de modelos, nos decantamos por la utilización de redes neuronales recurrentes LSTM (Long Short-Terma Memory). Las cuales están especializadas en el análisis de series temporales. Pese a que los valores de una cotización no pueden tratarse exactamente de una serie temporal, se tomó como punto de partida una vez se analizaron los resultados de estas.

Las redes neuronales recurrentes LSTM se caracterizan porque sus valores de entrada pueden ser diferentes tipos de variables, pero entre ellas destaca el factor temporal. Es decir, podemos tomar como variables de entrada los valores de una misma variable, tomados en diferentes espacios de tiempos pasados respecto a la predicción que se desea obtener. Unas de estas variables de entrada pude ser incluso la propia variable objetivo en momentos temporales pasados.

Esta característica de este tipo de redes es muy valiosa de cara al problema que queremos abordar ya que las predicciones en bolsa también se ven influenciadas por la información de días anteriores. Son redes neuronales "con memoria"

Sobre esta base se empezaron a construir diferentes modelos con diferentes informaciones disponibles y diferentes arquitecturas de la red neuronal:

Predicción del valor de cierre de una cotización

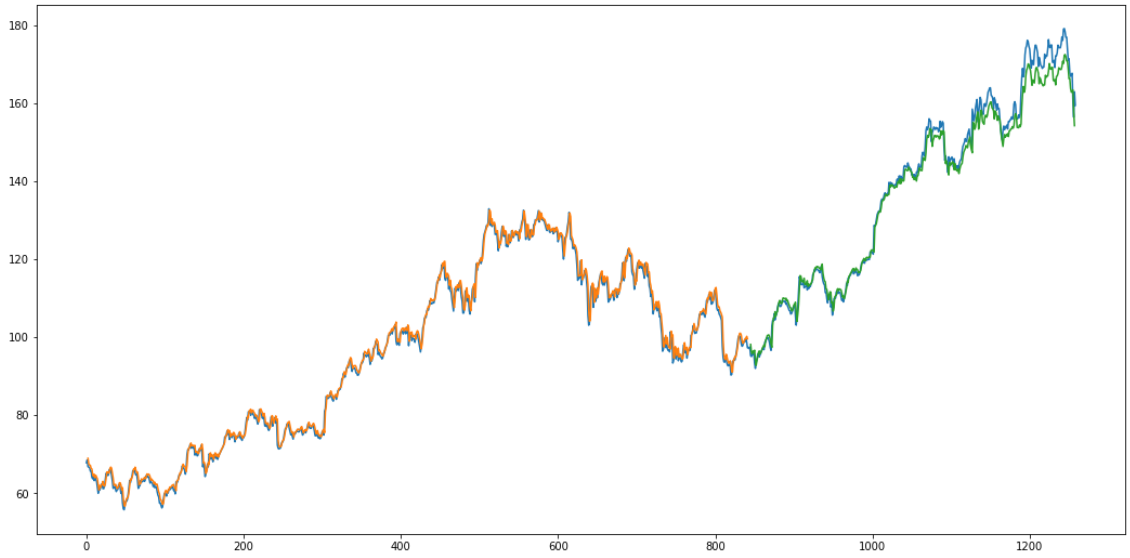
Nuestro primer objetivo planteado fue la creación de modelos capaces de predecir el valor de cierre de una cotización para día siguiente.

Con esta premisa el primer modelo que construimos y por tanto el más simple, fue un modelo capaz de predecir el valor de cierre de una cotización a partir de los valores de cierre de los dos días anteriores. Así pues, la arquitectura de nuestra primera red neuronal fue la siguiente:

- Capa de entrada: 2 neuronas (los valores de cierre de los dos días anteriores).
- Capa oculta: 5 neuronas LSTM.
- Capa de salida: 1 neurona con el resultado.

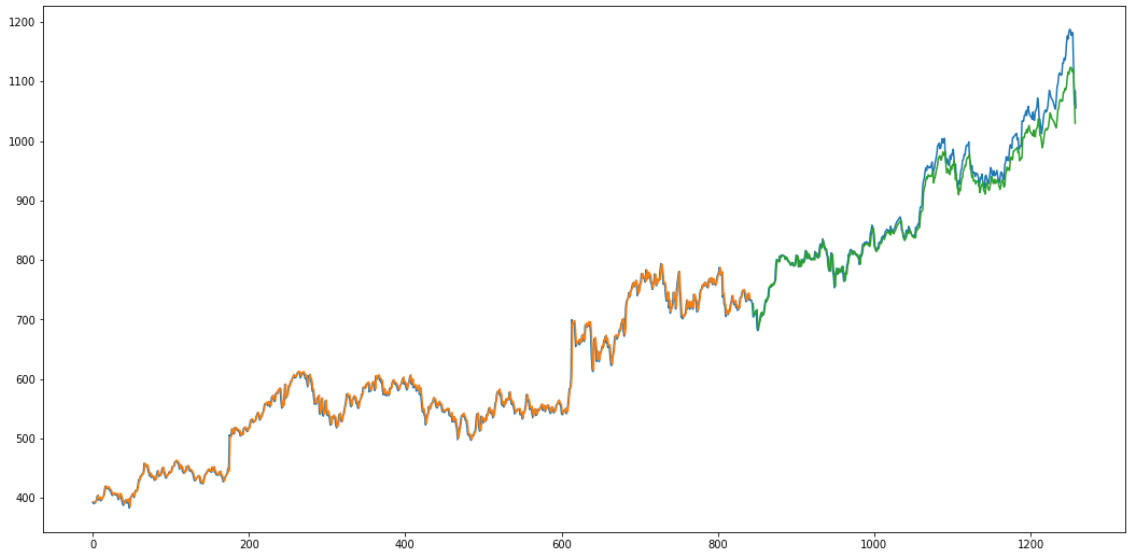
Sobre esta estructura y dividiendo nuestro dataset en un 67% de datos de entrenamiento y un 33% de test, se entrenaron diferentes modelos para diferentes cotizaciones. Adicionalmente hemos de indicar que se han utilizado técnicas de dropout para tratar de evitar el sobre-entrenamiento del modelo. A continuación, se muestran los resultados:

Cotización de Apple:




```
Test R2: 0.985
```

Cotización de Google:




```
Test R2: 0.965
```

La traducción de las gráficas sería la siguiente:

- La gráfica azul representa el valor real a lo largo del tiempo (nuestro dataset).
- La gráfica naranja representa el resultado de nuestro modelo sobre los valores de entrenamiento utilizados para ajustar el modelo.
- La gráfica verde representa el resultado del nuestro modelo sobre los valores de test utilizados para medir la validez del modelo.

Como se puede ver los resultados de estas presentan un valor de su R^2 entorno al 0,96- 0,98. Eso se podría traducir aproximadamente como que nuestros modelos serían capaces de explicar la realidad en el 96% - 98%.

A partir de estos valores los cuales era esperanzadores, debían se mejorados con el fin de buscar una mayor precisión en nuestras predicciones (un error de un 4% en bolsa es demasiado elevado).

Predicción del valor de cierre de una cotización del siguiente día a partir de información financiera adicional

Con el fin de mejorar los resultados anteriores lo que hicimos fue incluir valores adicionales (en este caso también estructurados). Para ello se tomó:

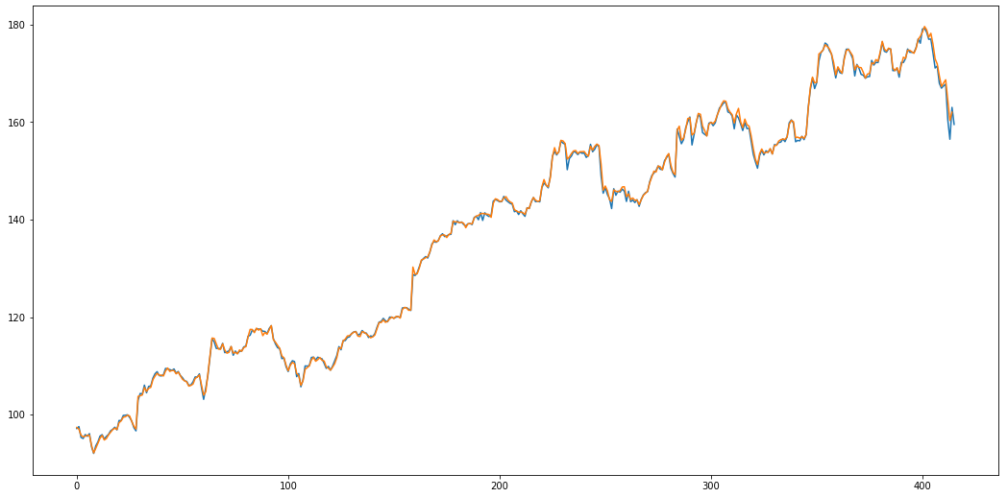
- *Máximo del día anterior*
- *Mínimo del día anterior*
- *Valor de apertura del día anterior*
- *Valor de cierre del día anterior*
- *Volumen de acciones del día anterior*

Sobre esta información se construyó la siguiente arquitectura de red neuronal:

- *Capa de entrada: 5 neuronas (variables indicadas).*
- *Capa oculta: 50 neuronas LSTM.*
- *Capa de salida: 1 neurona con el resultado.*

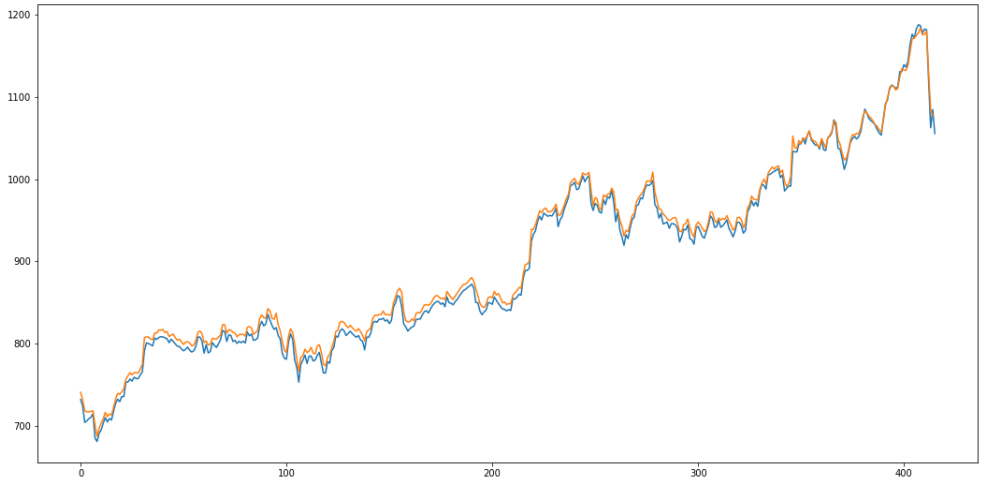
A partir de esta estructura y dividiendo el dataset del mismo modo que en el caso anterior y utilizando técnicas de dropout para tratar de evitar el sobre-entrenamiento del modelo, se obtuvieron los siguientes resultados:

Cotización de Apple:



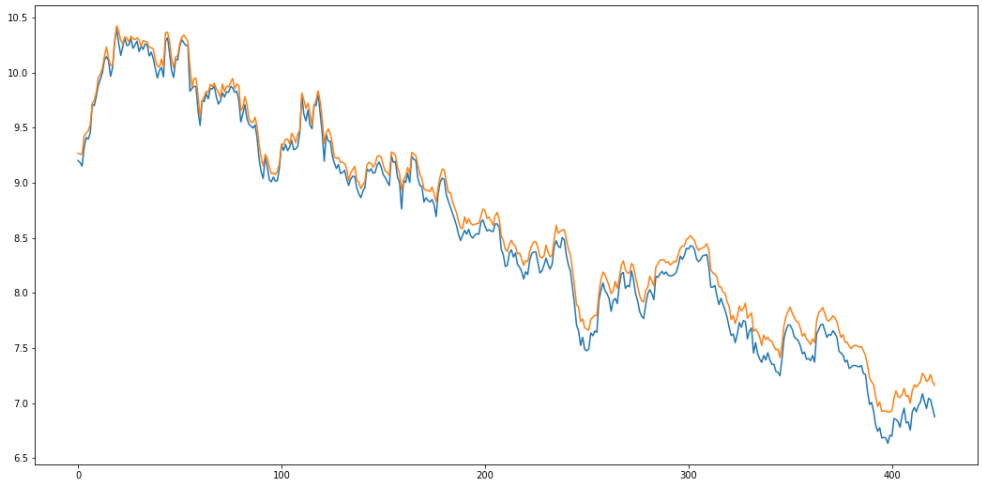
```
Test R2: 0.996
```

Cotización de Google:




```
Test R2: 0.994
```

Cotización de Telefónica:




```
Test R2: 0.989
```

Como se puede ver los valores del R^2 con este nuevo modelo ha mejorado respecto al modelo anterior con valores entorno al 0,99 y superándolo en varios casos. Con esta situación tenemos un modelo capaz de predecir el valor de cierre del día siguiente con una precisión entorno al 99%.

Sin embargo, nuestro objetivo final de cara a la gestión de fondos no es tanto tener con una precisión elevada el valor de la cotización del día siguiente, más bien sería poder tener predicciones de la tendencia que va a llevar una cotización en un medio plazo (por ejemplo, en la próxima semana). Esta información podría ser usada para la implementación de un sistema de alertas en caso de predecir una variación elevada en el día siguiente.

Predicción de la tendencia del valor de cierre de una cotización en varios días siguientes al día actual

Como se comentaba en el apartado anterior el siguiente objetivo fue construir un modelo capaz de predecir el valor de una acción en varios días siguientes al valor actual visualizando así su tendencia.

Para la construcción de este modelo aprovechamos la característica de las redes neuronales recurrentes LTSM por la cual son capaces de predecir los valores de una serie temporal no solo para la siguiente unidad de tiempo si no para sucesivos valores futuros.

En este caso se construyó una arquitectura de red neuronal con una capa con 5 neuronas LTSM en la capa oculta para una predicción de siete días en el futuro y como en los casos anteriores se incluyeron técnicas de dropout para evitar el sobre-entrenamiento del modelo. Con esto se obtuvieron los siguientes resultados:

Cotización	R^2 día t+1	R^2 día t+2	R^2 día t+3	R^2 día t+4	R^2 día t+5	R^2 día t+6	R^2 día t+7
Apple	0,996	0,991	0,987	0,982	0,977	0,971	0,965
Google	0,994	0,987	0,979	0,972	0,962	0,954	0,944

Como podemos ver la precisión de nuestro modelo va empeorando (como era esperado) conforme nuestras predicciones se van alejando de los valores reales de los cuales disponemos. La precisión en el séptimo día estaba entorno al 94%-

96%. Con esta situación la precisión había descendido, sin embargo, éramos capaces de predecir tendencias de una cotización con una cierta precisión.

En este punto lo que planteamos era como podíamos mejorar las predicciones de las tendencias de las cotizaciones. Para ello planteamos la inclusión de nuevas variables "no tradicionales" dentro de nuestro modelo como veremos en los siguientes apartados.

3.11.2 Análisis y procesamiento de información no estructurada de noticias

Este apartado se focaliza en el análisis de sentimiento y emociones de noticias financieras sobre las cotizaciones que estamos analizando. Para ello hemos utilizado la herramienta de Watson de IBM, Natural Language Understanding. Por la cual, a partir de una noticia, el sistema es capaz de devolvernos los resultados sobre los sentimientos y emociones presentes en una noticia y las probabilidades de estas. Esto ya se podría tratar como información estructurada la cual podría incluirse junto con el resto de información estructurada financiera la cual ya disponíamos en nuestro modelo.

De cara a poder utilizar este servicio se ha realizado una serie de simulaciones con el servicio propio de IBM. La compañía dispone de una versión de prueba la cual nos ha permitido realizar ensayos, validar su funcionamiento y confirmar si dicha funcionalidad de adaptaba a las necesidades de nuestro proyecto.

El servicio puede consumirse en modo servicio web y su uso es bastante simple. A partir del envío de un texto o de la url de un texto el servicio es capaz diversos elementos de este:

- Sentimiento: Positivo o negativo.
- Emociones: Alegría, enfado, disgusto, tristeza o temor.

- Palabras clave.
- Entidades.
- Temática.
- Conceptos.
- Análisis semántico.

En nuestro caso, lo que inicialmente utilizaríamos para enriquecer nuestro modelo sería tanto el análisis de sentimiento, como el análisis de emociones. En el caso de las emociones el resultado devuelto sería un valor entre 0 y 1 en el que representa la probabilidad del resultado presentado siendo 1 el valor máximo que indica la presencia de ese sentimiento/ emoción dentro del texto. En el caso del sentimiento el resultado varía en un valor entre -1 y 1, siendo -1 un sentimiento negativo y 1 positivo.

Por ello, el modo en el que utilizaríamos el servicio sería, a partir de las noticias del día de una compañía, enviaríamos el texto o la url de la noticia (Intrinio devuelve ambos valores) y obtendríamos los resultados de sentimientos y emociones de esta para combinarlo con el resultado del resto de noticias analizadas.

Para demostrar el uso de este servicio se ha utilizado una noticia de ejemplos:

URL noticia: https://finance.yahoo.com/news/qualcomm-says-apple-7-billion-012941302.html?.tsrc=rss&guccounter=1

Texto:

(Reuters) - Qualcomm Inc said in a court hearing on Friday that Apple Inc is $7 billion (£5.5 billion) behind in patent royalty payments to the mobile chip firm, which for years supplied parts for the iPhone.

Apple and Qualcomm are involved in a series legal actions. Apple has accused Qualcomm of unfair patent licensing practices. Qualcomm, the world's largest mobile phone chip supplier, has in turn accused Apple of patent infringement.

Qualcomm made its comments about the size of Apple's unpaid royalties in a hearing in one of the cases in federal court in San Diego. Apple has disputed the amount of royalties that it owes to Qualcomm.

Apple has argued that Qualcomm is forcing it to pay for the same patents twice, once when Apple uses Qualcomm's chip in iPhones and then again through patent royalties. Qualcomm has argued that its practices are legal and that Apple is seeking to destroy Qualcomm's business model after agreeing to it for years.

Tras analizar dicha noticia con el servicio NLU de IBM Watson el resultado fue el siguiente:

Análisis	Resultado
Sentimiento	Negativo (-0.29)
Emociones (Alegría)	0.15
Emociones (Enfado)	0.57
Emociones (Disgusto)	0.16
Emociones (Tristeza)	0.20
Emociones (Temor)	0.05

A partir de esta información el resultado podría combinarse con el resultado del resto de noticias de esa compañía. Con estos resultados habríamos transformado la información no estructurada de una noticia, en información estructurada la cual también puede combinarse con la información financiera la cual se había descrito en el apartado anterior.

Una posible línea de mejora de este módulo podría ser la realización de un nuevo reentrenamiento de este sistema adaptándolo a un contexto más ajustado sobre el análisis de noticias financieras.

3.11.3 Combinación de información estructurada y no estructurada para la predicción de cotizaciones

En esta parte del análisis se trató de combinar la información financiera que veíamos en el primer apartado con la cual ya teníamos un modelo capaz de predecir tendencias de los valores de cierre de una cotización y tratar de mejorar su precisión con la inclusión de nuevas variables (en este momento información ya estructurada, como hemos visto en el segundo apartado) sobre el sentimiento de noticias financieras de un día concreto sobre la cotización que se está analizando.

Para la ejecución de esta combinación se ha partido del mismo modelo presentado en el primer apartado. Es decir, una red neuronal recurrente LSTM (Long Short-Term Memory), para conseguir realizar predicciones de los valores de una cotización durante los días siguientes y así poder identificar tendencias de estas y ejecutar las acciones de compra o venta según corresponda. La diferencia en este caso radica en la inclusión de nuevos valores procedentes del análisis de noticias con el fin de mejorar las predicciones obtenidas con los datos financieros mostrados en el primer apartado.

Los parámetros de entrada del modelo (para su entrenamiento y ejecución) serían:

- Datos financieros de la cotización analizada:
    - Máximo del día
    - Mínimo del día
    - Valor de apertura del día
    - Valor de cierre del día
    - Volumen de acciones del día

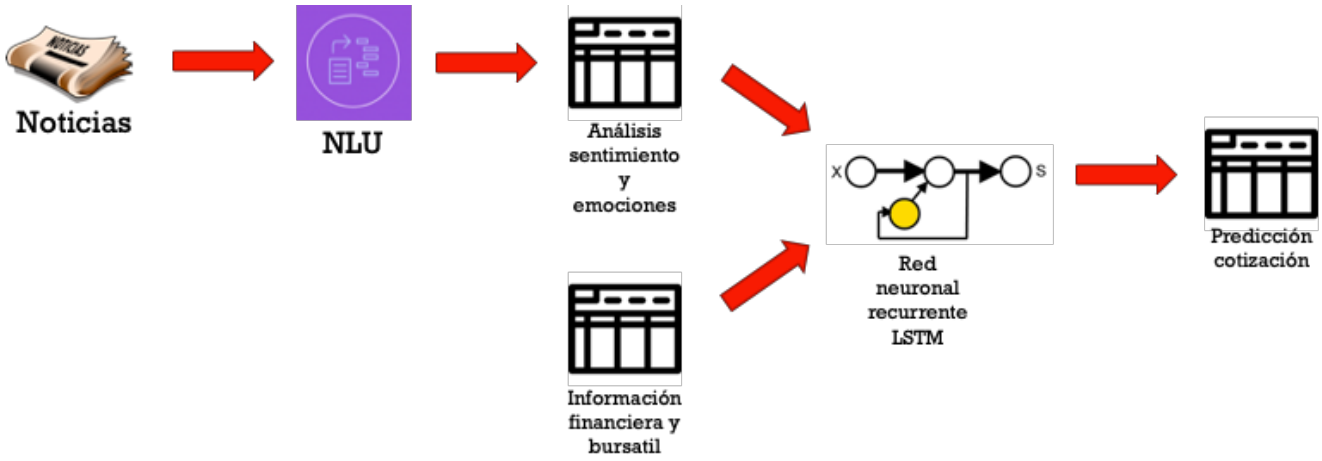
- Datos del análisis de sentimiento y emociones de noticias de la cotización analizada:
    - Sentimiento: Mediana, máximo y mínimo del día
    - Emoción (alegría): Mediana, máximo y mínimo del día
    - Emoción (enfado): Mediana, máximo y mínimo del día
    - Emoción (disgusto): Mediana, máximo y mínimo del día
    - Emoción (tristeza): Mediana, máximo y mínimo del día
    - Emoción (temor): Mediana, máximo y mínimo del día

En el caso del análisis de noticias utilizaríamos la mediana, el máximo y el mínimo de cada valor ya que para cada cotización podemos tener varias noticias dentro de un mismo día. Utilizaríamos la mediana y adicionalmente los máximos y mínimos ya que pueden ser noticias de importancia que se escaparían de la mediana del día. Representarían los "outliers" que se hubieran podido dar en el día y que probablemente fueran los que más información proporcionan de cara a las predicciones de los días siguientes.

A partir de esta información, podríamos obtener las predicciones de los siguientes días al igual que se realizaba en el primer apartado. En este caso las predicciones deberían mejorar gracias al enriquecimiento de la nueva información disponible de las noticias.

Como comentario del modelo planteado hay que indicar que, al igual que en el primer apartado, la información a tomar en cuenta podría ser del día anterior a predecir, pero también se podrían incluir en el modelo varios días previos para la realización de la predicción. Del mismo modo, el resultado inicialmente planteado sería tener una predicción de una semana para cada cotización. Sin embargo, en cuanto consigamos mejorar la calidad de las predicciones, el rango temporal de las predicciones también podría ampliarse.

De este modo, el flujo de procesamiento de la información dentro de nuestro modelo se puede ver reflejado en alto nivel en el siguiente gráfico:



Respecto al prototipo de la solución, como se ha descrito en los apartados y se ha podido implementar, ha sido un modelo de red recurrente LSTM con información financiera y bursátil y la validación del funcionamiento del servicio de NLU de Watson. Esto se debe en primer lugar a la falta actual de la información referente a las noticias de las cotizaciones (el servicio de Intrinio que facilita dicha información tiene un coste para su uso "masivo"). Otro motivo también está relacionado con el coste del uso "intensivo" de NLU. La versión gratuita disponible se trata de una demo del servicio. El elevado número de invocaciones necesarias para utilizarlo dentro de nuestro modelo también tendría un coste económico.

Así pues, los próximos pasos serían confirmar la mejora de las predicciones a partir de la información procedente del análisis de las noticias. Otro aspecto para tener en cuenta de cara a mejorar la precisión del modelo sería mejorar el tiempo y la capacidad de computo disponible. Con una mayor capacidad se podría ampliar el número de días previos a tomar en cuenta ya que este es uno de los factores que más incrementa el tiempo de computo. Todo esto podríamos conseguirlo a través del uso de algunas las funcionalidades cloud disponibles en los principales proveedores.

3.11.4 Simulación inversión basada en el uso de modelo de predicción de valores

Una vez elaborado y entrenado el modelo de predicción de valores descrito en los apartados anteriores se trató de aplicar nuestro modelo en proceso de decisión de inversión dentro de una selección de valores. Este sería el uso principal para el que se utilizaría nuestro modelo.

Para tratar de simplificar el problema de cara a la simulación lo que se hizo fue tomar como referencia tres cotizaciones del S&P500 durante un periodo de tiempo pasado (la rentabilidad real combinada de estos tres valores durante este periodo representaría al índice de referencia al cual debemos intentar batir).

Por otro lado, nuestra política de inversión se fundamentó en la inversión diaria en los dos mejores valores de acuerdo con la predicción del precio de cierre del día siguiente de nuestro modelo. Es cierto que esta política es una simplificación del modelo de inversión que tendría el fondo, tal y como se describió en la política del mismo. Sin embargo, esta simulación nos permite conocer y validar si nuestro modelo funciona correctamente. A continuación, se describe el detalle de la simulación realizada:

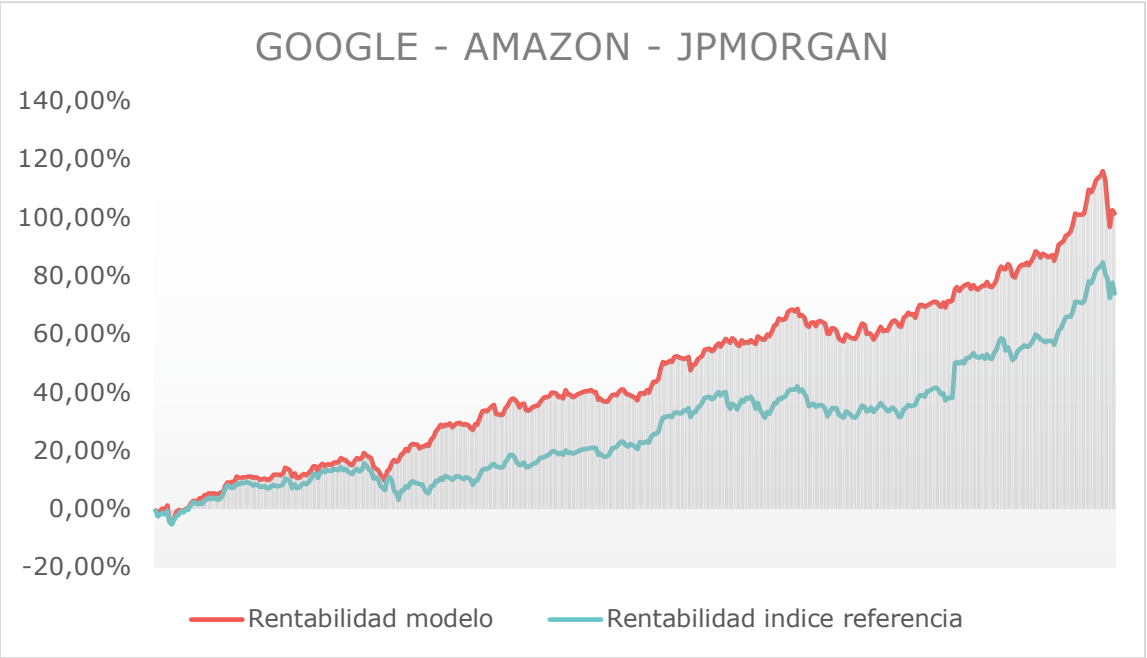
- ***Cotizaciones analizadas:*** *Google (GOOGL), Apple (AAPL) y JP Morgan (JPM).*
- ***Periodo de análisis:*** *Desde 16/6/2016 al 07/02/2018 (conjunto de test de nuestro modelo)*
- ***Rentabilidad de nuestro índice de referencia en el periodo descrito:*** *74,04 %*

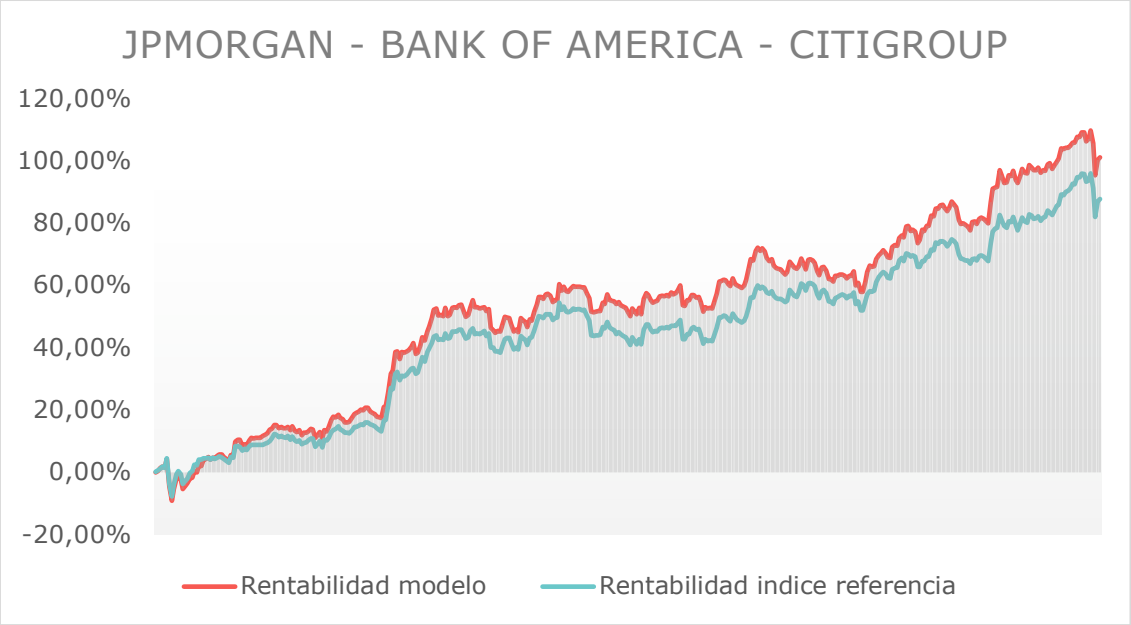
A partir de esta información real durante este periodo se realizó una simulación de la inversión siguiendo las reglas anteriormente descritas en función de las predicciones de nuestro modelo. **Como resultado de esta simulación de inversión obtuvimos una rentabilidad del 101,42% lo cual supone una mejora respecto a nuestro índice tomado como referencia del 36,98%.**

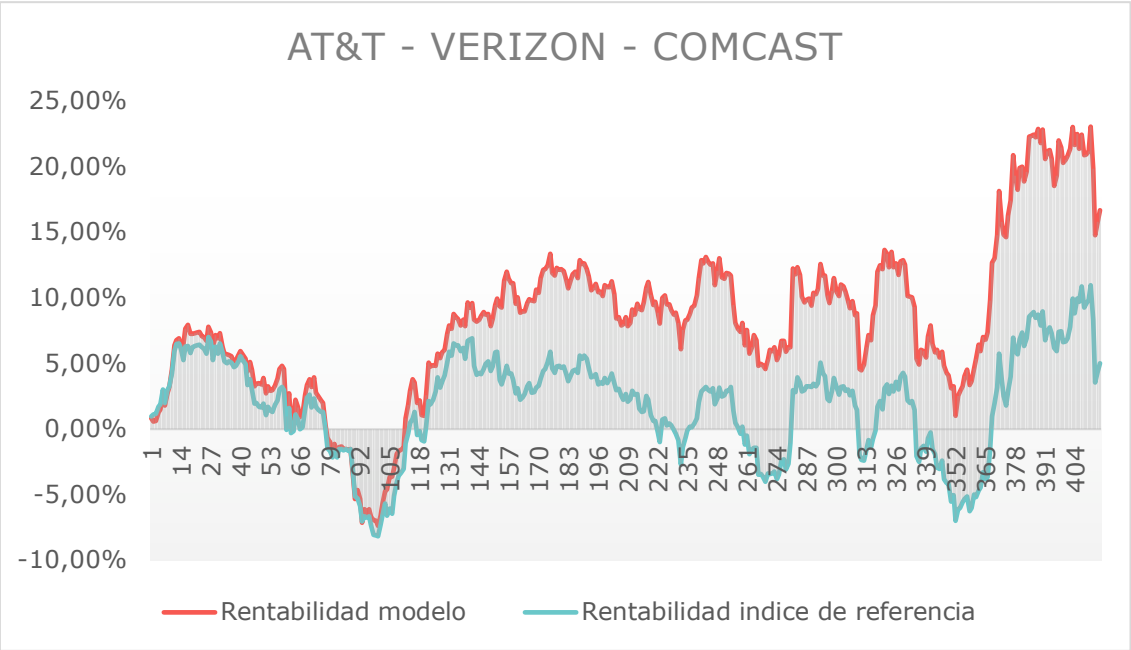
Esta misma simulación ha sido repetida utilizando diferentes valores en el mismo periodo de tiempo (desde 16/6/2016 al 07/02/2018) con el fin de confirmar el funcionamiento. A continuación, se resume el resultado de las simulaciones realizadas:

Valores utilizados	Rentabilidad índice de referencia	Rentabilidad simulación	Mejora respecto al índice de referencia
Google – Amazon - JP Morgan	74,04 %	101,42 %	36,98 %
JP Morgan – Bank of America – Citigroup	87,77 %	101,20 %	15,30 %
AT&T - Verizon - Comcast	5,02%	16,69 %	232,62 %

Adicional podemos ver la rentabilidad acumulada diaria de cada una de estas simulaciones frente a la rentabilidad del índice de referencia en cada uno de los casos:







Estas simulaciones se fundamentan en una simplificación de lo que sería la política de inversión con la que se gestionará nuestro fondo de inversión. No obstante, en base a los resultados obtenidos, podríamos afirmar que nuestro modelo planteado funciona y podría ser utilizado tomando un mayor rango de valores a analizar y aplicando las políticas de inversión necesarias, fundamentadas en el resultado de nuestras predicciones. Cabe destacar que

estas simulaciones habrían superado nuestro KPI de superar un 10% al índice de referencia.

CONCLUSIONES Y LÍNEAS DE MEJORA

ARP se presenta como un proyecto innovador dentro del sector en el que se enmarca. Si bien es cierto que la industria del trading y la inversión es un ámbito en el que la tecnología hace tiempo que emergió con el fin de agilizar y maximizar los beneficios (trading algorítmico, HFT…). La tecnología del Deep Learning aplicada con este fin no ha sido utilizada hasta el momento de manera amplia y generalizada. Tal y como se ha descrito anteriormente en la memoria, tan solo algunas startups americanas han sido identificadas desarrollando esta tecnología en el campo de la inversión. Sin embargo, todavía se trata de un nicho a desarrollar en Europa y en especial, España.

Otro elemento diferenciador es el uso de información no tradicional en la obtención de predicciones de valores. La tecnología utilizada permite incorporar información procedente de noticias financieras, RRSS… e incorporarlas como una variable más dentro de nuestro análisis. Todo esto gracias al análisis de sentimientos y emociones que nos ofrece NLU. Históricamente esta información no era utilizada en los análisis algorítmicos tradicionales y se fundamentaba más en el análisis e intuición del inversor con las limitaciones que esto conlleva (volumen de información, sesgos personales…).

Como posibles líneas futuras de mejora y evolución, adicional a lo descrito en el roadmap, se han planteado diversas alternativas las cuales se describen a continuación:

- **Evolución de los modelos**: Evolución del modelo planteado mediante el uso de nuevas fuentes de información y la utilización de nuevas tecnologías con el fin mejorar la precisión y la varianza del resultado de nuestras predicciones.

- **Desarrollo de modelos para su uso en HFT (High-Frequency Trading)**: En este caso se debería utilizar otro tipo de información en tiempo real (hemos identificado servicios que ofrecen este tipo de información) y habría que optimizar los tiempos de procesamiento de dicho modelo.

- **Expansión geográfica**: Inicialmente el proyecto se ha focalizado en su implantación en España junto con un socio ya presente en el sector. No obstante, un modelo similar se podría implantar en otras regiones de Europa o Estados Unidos.

- **Evolución hacia un modelo B2C**: Como se describió en el posicionamiento de la solución, el modelo inicialmente planteado (B2B) tenía como fin podernos focalizar en la parte en la cual teníamos ya conocimientos (desarrollo de modelos) y delegar el resto de los factores (regulación, captación de clientes…) en nuestro socio, ya presente en el mercado y con un mayor conocimiento del mismo. Lo que se plantearía en esta opción sería, una vez ya desarrollado, implantado y afianzado este modelo, evolucionar hacia un segundo de venta directa al cliente final, abordando desde nuestro rango de actuación todos los aspectos regulatorios y de captación sin necesidad de un socio. Este planteamiento solo sería viable habiendo demostrado al cliente final el buen funcionamiento de nuestra solución y habiendo creado una imagen de marca alrededor de ella.

- **Ampliación del modelo a derivados**: Una de las salidas naturales sería la aplicación del sistema de recomendación a la operativa con derivados. Aunque hay cambios conceptuales como la depositar garantías para relazar las operaciones de compra venta en el caso de los futuros no tendríamos problema para incluir este tipo de producto al ser los mismo también tendenciales y tener subyacentes basados entre otros en acciones y renta fija. En el caso de opciones si tendríamos que incluir modificaciones de más calado al tener las griegas de las opciones un parte importante en su valoración, no obstante las mismas seguirían teniendo como subyacente a activos que ya estamos considerando como antes hemos dicho. No obstante la inclusión de este tipo de activos nos llevaría a tener

que realizar cambios en nuestros perfiles de riesgos de los clientes y nuestra manera de operar.

BIBLIOGRAFÍA Y REFERENCIAS

Análisis mercado:

- *Análisis de fondos y oportunidad de negocio:*
    - *http://www.inverco.es*
    - *"Diez años de ahorro colectivo" ESADE*
    - *https://www.bolsasymercados.es/esp/Estudios-Publicaciones/Estadisticas*
    - *https://us.spindices.com/indices/equity/sp-500*
    - *https://www.morningstar.com/*
    - *http://www.expansion.com/*
    - *"El inversor inteligente" Benjamín Graham*
-
- *Análisis compañías (startups):*

    - *https://www.crunchbase.com/*
    - *https://www.sentient.ai/*
    - *https://www.alpaca.ai/*
    - *https://walnut.ai/en/*
    - *http://www.binatix.com/*
    - *https://www.kavout.com/*
    - *https://www.eye-capital.com/*

Tecnología:

- *Desarrollo modelos:*
    - *https://keras.io/*
    - *https://www.tensorflow.org/?hl=es*

- *Análisis solución:*
    - *IBM Watson: https://www.ibm.com/watson/*
    - *IA Platform Azure: https://azure.microsoft.com/id-id/overview/ai-platform/*
    - *Google AI: https://ai.google/tools/*
    - *Machine Learning en Amazon: https://aws.amazon.com/es/machine-learning/*

Análisis redes neuronales recurrentes LSTM:

- *Análisis redes LSTM:*

    - *http://colah.github.io/posts/2015-08-Understanding-LSTMs/*
    - *https://hackernoon.com/understanding-architecture-of-lstm-cell-from-scratch-with-code-8da40f0b71f4*
    - *https://keras.io/layers/recurrent/*

- *Redes LSTM aplicado en análisis de valores:*

    - *https://machinelearningmastery.com/time-series-prediction-lstm-recurrent-neural-networks-python-keras/*
    - *https://www.datacamp.com/community/tutorials/lstm-python-stock-market*