

DECISIONES DE INVERSIÓN

AUTOMATIZADAS EN PYTHON

Joan Prieto Montes

El proyecto está orientado al campo de las finanzas. Las finanzas corresponden a un área de la economía que estudia la obtención y gestión del dinero, es decir, administra los recursos financieros. En este caso, nos centraremos en la inversión a través de algoritmos de trading (comprar y vender).

Este trabajo es para la empresa *iBroker*, que se dedica a actuar como un *broker*¹ online. Esta empresa tiene un producto denominado **Tradingmotion** (tradingmotion.com), que se trata de un *marketplace*² de sistemas automáticos. TradingMotion es una plataforma que contiene un conjunto de algoritmos de inversión automática que los clientes pueden contratar a cambio de una licencia mensual.

Las *Decisiones de Inversión Automatizadas en Python* son un seguimiento de procesos aplicados a la investigación y creación de un ranking de los mejores algoritmos de trading para invertir, los cuales pertenecen a la plataforma de TradingMotion.

¿Que es un algoritmo de trading?

Un **algoritmo** o **sistema de trading** es un programa informático suscrito al market data³ de un instrumento financiero, que opera de forma autónoma con el objetivo de obtener un rendimiento económico en el mercado financiero (CFDs de acciones, forex, criptomonedas, materias primas...).

¿Que hemos obtenido de TradingMotion?

Los datos que hemos obtenido son los rendimientos de los sistemas que pertenecen a la plataforma Tradingmotion. Los rendimientos representan los beneficios, pérdidas, porcentaje de sesiones ganadas o perdidas, ratio de acierto de los sistemas...

Con objetivo, de aprovechar los resultados de las simulaciones de los algoritmos que permanecen activos en la plataforma, para encontrar aquellos sistemas que obtienen grandes beneficios.

La finalidad del proyecto es puramente de investigación. Con el análisis como herramienta principal, y junto con unas hipótesis acertadas, se busca encontrar un conjunto de sistemas automáticos interesantes que cumplan nuestro objetivo.

¹El término broker hace referencia a la figura de un intermediario entre compradores y vendedores a cambio de una comisión cuando se ejecute la operación.

²Define el sitio en Internet donde se llevan a cabo interacciones comerciales entre diferentes empresas.

³En finanzas, los datos de mercado son datos del precio y están relacionados con la negociación para un instrumento financiero.

El objetivo ha sido escoger un conjunto de **sistemas automáticos** que pertenecen a la plataforma de *TradingMotion*, los cuales **no son irregulares** y **obtienen beneficios** en el mercado de futuros a lo largo del tiempo. Sabremos que algoritmos debemos seleccionar ya que a través de un amplio análisis, descubriremos que *algoritmos de predicción* son los mejores.

Procesos

Dividimos los procesos que hemos llevado a cabo en el proyecto en 5 pasos:

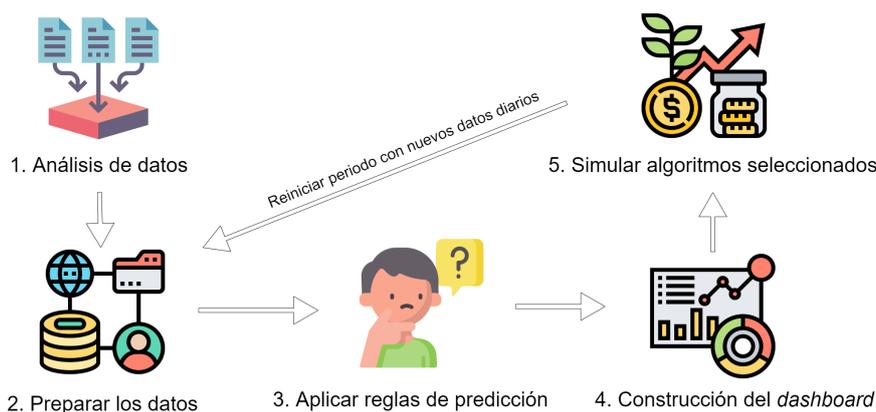


Figura 1: Procesos representativos del proyecto

1. **Analizar los datasets** (Exploratory Data Analysis): partimos de 3 datasets con información sobre las características, rendimientos y movimientos de los sistemas de trading. La tarea es mostrar gráficos relevantes que ayuden a entender cómo se comportan los sistemas a lo largo del tiempo.
2. **Preparación de los conjuntos de datos** para su posterior uso.
3. **Desarrollar hipótesis** para generar rankings de los sistemas ordenados de mayor a menos según su potencial.
4. **Construcción del portafolio** (dashboard). A través del *dashboard* podemos analizar como se han comportado los algoritmos de trading seleccionados por la hipótesis. El portafolio es un seguido de gráficos que nos describen los rendimientos de manera más visual.
5. **Simulación de los sistemas** seleccionados durante un periodo concreto.

He decidido utilizar **Python** para ejecutar todos los procesos del proyecto, ya que es un lenguaje muy versátil, permite hacer muchas cosas con pocas líneas de código y la gran cantidad de

bibliotecas, facilita el manejo del lenguaje. Además, la gestión de datos a través de *Python* es tan habitual, que podemos encontrar todo tipo de guías, libros y artículos que sirven de soporte técnico. Todo el entorno está montado a través de VirtualEnv, este es un entorno virtual que permite instalar paquetes sin interferir en otras versiones ya instaladas en el python del sistema operativo.

Para desarrollar he utilizado Jupyter Notebooks, esta es una aplicación cliente-servidor que permite crear, ejecutar y compartir documentos en Python. La ventaja de estos notebooks, es que están dedicados al análisis de datos, ya que se puede combinar perfectamente el código con los textos explicativos. Así, se permite ejecutar a tiempo real y en memoria todo tipo de código en python, ecuaciones y una gran variedad de gráficos.

Generar hipótesis

La mejor manera de estudiar los rendimientos de los algoritmos, es a través de la aplicación de un método científico orientado a investigar como se comportan los sistemas de inversión y trading.

Esta idea es la que desarrolla *David Aronson* en su libro *Evidence Based Technical Analysis*. Él propone utilizar este método hipotético-deductivo para evaluar la eficacia de las señales en un sistema de trading.

La teoría de este método se divide en varias fases:

1. **Observación**
2. **Generar hipótesis**
3. **Predicción**
4. **Verificación**
5. **Conclusión**



Figura 2: Esquema fases predicción.

Tras un periodo de estudio, formulé 5 hipótesis:

- Clasificación de sistemas según la diferencia de la distribución *out-sample* comparada con la *in-sample* aplicando la prueba de T de Student, + filtro de análisis de riesgos (Sharpe ratio).

- Ranking de sistemas ordenados por profit-factor, Rating TM y Sharpe Ratio, con un mínimo de 1 año funcionando en activo.
- Sistemas con volatilidad baja, un buen profit factor y un filtro Sortino Ratio.
- Sistemas antiguos (+ de 2 años en activo) con grandes beneficios a lo largo del tiempo, filtrados por Sharpe Ratio, máximo drawdown y TMRating.
- Sistemas con mayor profit-factor, con mas de 3 meses en activo y un filtro de drawdown (Sistemas nuevos).

Simulación de hipótesis

Una vez implementadas todas las hipótesis, debemos compararlas de alguna forma. Para ello he desarrollado un *script* que compara los rendimientos de todas las hipótesis en un mismo periodo de tiempo y devuelve un listado con los *IDs* de los sistemas escogidos por cada hipótesis.

Hemos ejecutado el *script* con la fecha entre 15/12/19 y el 15/05/20, y este es el resultado:

- Sistemas más Nuevos: [22090, 22041, 21778, 22054, 22168, 22055, 17654, 21352, 21697, 17519]
- Comparación de distribuciones: [22150, 17919, 21361, 22293, 21031, 22183, 22266, 21443, 21704, 21341]
- Sistemas antiguos: [17654, 17519, 21147, 21218, 17653, 12373, 11088, 12296, 17521, 21212]
- Volatilidad baja: [21692, 22054, 21778, 21874, 22090]
- Profit Factor: [22103, 21828, 21766, 21635, 21733, 21702, 21251, 21608, 21734, 21978]



Figura 3: Ranking de *performance* de las hipótesis

	Hipotesis	Rendimiento(€)
Antiquity Systems	Sistemas antiguos	81809.5474
Newest Systems	Sistemas mas Nuevos	46927.5806
Low Volatility Systems	Volatilidad baja	43729.3078
T-test	Comparacion de distribuciones	19438.0361
Ranking Profit Factor	Profit Factor	-874.4735

Figura 4: Tabla de *performance* de las hipótesis

En este caso la hipótesis de los sistemas que llevan activos más de 2 años, es la que ha generado más dinero, y la hipótesis del ranking por *profit factor* la que menos. Esto depende mucho del tramo sobre el cual simulamos.

Conclusión

Tras simular las hipótesis planteadas, he concluido que el diseño y la implementación de los procesos del proyecto se ha llevado a cabo de forma acertada.

Es la simplicidad del trabajo la que permite que el resultado haya sido muy positivo, ya que al basarnos en el análisis de los sistemas, junto a un buen filtraje de estos mediante hipótesis y algoritmos que no son muy complejos, permitimos obtener un resultado claro.

En conclusión, tanto el entorno seleccionado como los resultados finales, han cumplido con las motivaciones, propósitos y objetivos planteados al inicio del proyecto. La gestión, analítica y predicción de datos orientados a decisiones de inversión, supone una gran ventaja de cara a seleccionar los sistemas automáticos sobre los que invertir.