



IA PARA INVERSIÓN

Memoria del proyecto final

28/06/2024

Anneka Halasz Aran
Big Data & Inteligencia Artificial
GeeksHub

Problema al que nos enfrentamos

Para este proyecto nos basamos en el problema de una empresa de inversión.

¿Debo invertir en esta empresa?

Con una respuesta de SI o No

Luego de una búsqueda sobre el tema me gusto definir el sí se puede invertir en una empresa o no a “¿Esta empresa va a subir o a bajar?” ya que invertir se puede hacer a los dos sentidos.

También me ha gustado el definir el tiempo de esta pregunta, ya que en inversión existen las inversiones a largo plazo y a corto plazo.

Cuando invertimos a corto plazo es más fácil eliminar las variables sociales (como tendencia o cambios en la empresa, noticias sobre ella que hacen que la gente compre o venda acciones, etc) y podemos centrarnos más en estadísticas, tendencias matemáticas y gráficos.

He elegido una IA como solución porque me parece que por la cantidad de datos que se utilizan y por el tiempo, ya que las personas que invierten pueden hacer hasta operaciones de pocos minutos o incluso minutos, por ello creo que les iría mejor la automatización que ofrece la IA frente al BI.

Planteamiento de soluciones

Aunque hay muchas formas de abordar el problema planteado, creo que el notebook escogido da una buena visión de cómo diferentes estrategias no se pueden aplicar de cualquier forma y cómo esto nos ayudaría en un futuro para mejorar y entrenar mejor nuestra IA a cometer menos errores y a entender mejor el mercado.

1. Un notebook entrenado con LSTM

Este notebook valora varios aspectos técnicos como el MA (moving average) y tiene en cuenta estos valores comparándolos con otras empresas, en el caso del sector tecnológico, para entender que el mercado se mueve a la par.

<https://www.kaggle.com/code/faressayah/stock-market-analysis-prediction-using-lstm/notebook>

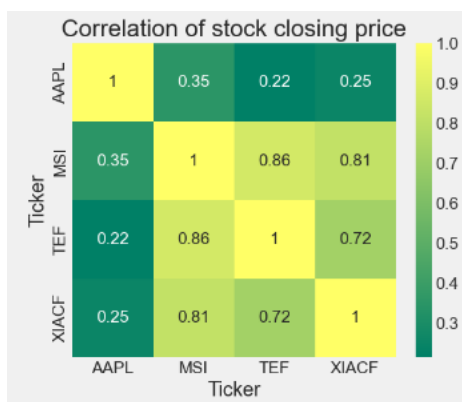
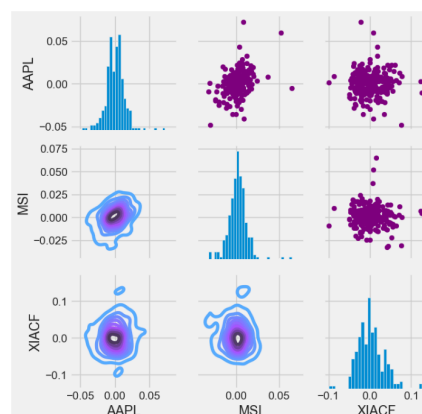
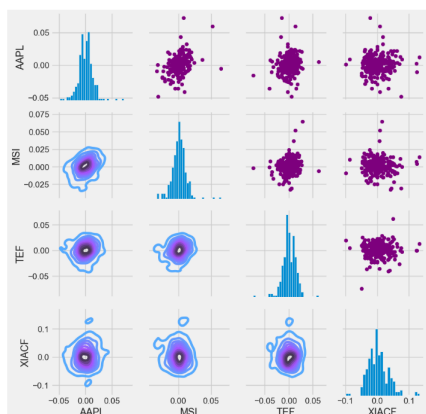
Al investigar el dataset me he dado cuenta de que las 4 empresas elegidas són de las que más cotizan, y en bolsa las empresas que más cotizan suelen tener movimiento entre ellas; no siempre es así, pero suelen estar influenciadas.

Para valorar si sería una solución viable, se ha probado con 3 empresas del mismo sector, el de venta de móviles; pero de diferente cotización (Xiaomi, Apple y Motorola)

Sería muy interesante ver el movimiento que tienen estas empresas con sectores directamente relacionados con ellas, como las empresas de comunicación (Telefónica, Pepephone, Orange...)

Se ha probado con Telefónica para ver si en las empresas se podía ver un movimiento positivo, y aunque los resultados han variado y hay algo menos de dispersión, no se puede apreciar un cambio lo suficientemente notable.

Comparativa con gráficos:



También me ha gustado este notebook por la elección de entrenamiento, no creo que sea la mejor pero si que soluciona un problema que creo que es importante

abordar y es la memoria de aprendizaje, ya que el mercado suele tener tendencias y repetir patrones, es relevante que nuestro modelo.

	Close	Predictions
Date		
2022-07-13 00:00:00-04:00	145.490005	146.457565
2022-07-14 00:00:00-04:00	148.470001	146.872879
2022-07-15 00:00:00-04:00	150.169998	147.586197
2022-07-18 00:00:00-04:00	147.070007	148.572937
2022-07-19 00:00:00-04:00	151.000000	148.995255
...
2023-01-24 00:00:00-05:00	142.529999	138.565536
2023-01-25 00:00:00-05:00	141.860001	140.022110
2023-01-26 00:00:00-05:00	143.960007	141.225128
2023-01-27 00:00:00-05:00	145.929993	142.469315
2023-01-30 00:00:00-05:00	143.000000	143.833130

Foto 1: Notebook original

	Close	Predictions
Date		
2023-11-13	184.800003	191.415817
2023-11-14	187.440002	192.501358
2023-11-15	188.009995	193.693634
2023-11-16	189.710007	194.777161
2023-11-17	189.690002	195.951950
...
2024-06-24	208.139999	220.659164
2024-06-25	209.070007	219.239655
2024-06-26	213.250000	218.275436
2024-06-27	214.100006	218.588684
2024-06-28	210.619995	219.523224

157 rows x 2 columns

Foto2: Nuestro notebook

Al final, aunque la correlación no es el mejor método de entrenamiento o de visualización de patrones, si que es cierto que el modelo LSTM funciona.

Nuestras predicciones han sido mayores al precio de cierre mientras que en el notebook original, las predicciones son más bajas que el precio de cierre.

2. Un bot impulsado por IA para operar en el mercado (APP o web)

Como hemos visto que el notebook anterior no ha dado unos resultados muy positivos (aunque haya funcionado el LSTM) he intentado buscar otra solución que además creo que seria mas factible como solución real para una página o aplicación de inversión.

Aunque se han encontrado muchos tipos de recursos para crear un bot de IA para páginas de trading, he encontrado un notebook en Kaggle que me ha parecido bastante más útil.

Aunque el retorno que nos da es algo más bajo que otros bots, parece fallar menos y ser más seguro, con una inversión bastante pequeña (menos de 200 euros)

<https://www.kaggle.com/code/trooperog/ai-trading-bot/notebook>

No se ha podido terminar de desarrollar por falta de tiempo, pero creo que en un futuro se podrían ofrecer diferentes soluciones según las necesidades y presupuesto de la empresa y que un bot seria lo mas eficaz para operar en tiempo real y no tener que estar constantemente pendiente del mercado.

También se ha trabajado en un bot IA escrito con ayuda de Blackbox, una IA de código mencionada en los recursos utilizados

3. Indicadores que den una predicción, pero que no operen en el mercado

También se podrían implementar soluciones como el uso de índices calculados con IA, dan aproximaciones del precio, pero no hacen operaciones por nosotros.

Dan más autoridad al usuario a la hora de tomar decisiones, pero el usuario tiene que estar mucho más pendiente del mercado.

Recursos

I. Kaggle

Hemos utilizado varios notebooks de Kaggle en nuestro trabajo

El notebook de correlación:

<https://www.kaggle.com/code/faressayah/stock-market-analysis-prediction-using-lstm/notebook#Data-Project---Stock-Market-Analysis>

El notebook del bot: <https://www.kaggle.com/code/trooperog/ai-trading-bot/notebook>

Para entender la predicción de stocks mejor:

<https://www.kaggle.com/code/dbdmobile/predict-stocks-using-lstm/notebook>

II. Blackbox IA

Para ayudarme con temas de código se ha utilizado una IA llamada Blackbox, por una parte, porque me gusta que el código sea más limpio y por que no tengo tanta experiencia en programación.

III. Yahoo Finance (yfinance)

En un principio quería mostrar el proceso de limpieza y recolección de datos para mostrar los puntos aprendidos pero debido al tiempo del proyecto esto no ha sido posible, así que se ha decidido investigar y recolectar la data de yfinance (Yahoo finance), una API gratuita.

Explicación de conceptos: indicadores y fórmulas matemáticas

El hecho de intentar simplificar una situación real para automatizarla, en base a fórmulas específicas para esa situación se le llama crear modelos matemáticos. Se utiliza, por ejemplo, con la ley de Ohm. Entrarían aquí los modelos lineales o de regresión, para dar ejemplos más generalizados.

Los modelos matemáticos también se utilizan en inversión, se les conoce como indicadores. Suelen ayudar a dar información menos sesgada del mercado.

Dentro de lo que es el análisis técnico en inversión tenemos varios indicadores que nos pueden dar información de si esta empresa va a subir o bajar:

- **Moving Average (MA)**

- Nos da una tendencia del mercado y nos ayuda a determinar los niveles de apoyo y resistencia. Si se utiliza de forma simple nos ayuda a valorar los índices en inversiones a largo plazo y se utiliza de forma exponencial, nos ayuda en las inversiones a corto plazo.

$$SMA = \frac{A_1 + A_2 + \dots + A_n}{n}$$

$$SMA = \frac{P_1 + P_2 + \dots + P_n}{n}$$

- P es el precio de cierre.
- n es el número de períodos.

- **Daily return**

- Se refiere al rendimiento diario de un activo financiero, que es la ganancia o pérdida experimentada en un solo día de operación

$$\text{Daily Return} = \frac{P_t - P_{t-1}}{P_{t-1}} \times 100$$

- P_t es el precio de cierre del activo al final del día actual.
- P_{t-1} es el precio de cierre del activo al final del día anterior.

- **Media Móvil Exponencial (EMA)**

- La Media Móvil Exponencial (EMA, por sus siglas en inglés) otorga más peso a los precios recientes.

$$EMA = P_t \times \frac{2}{n+1} + EMA_{\text{anterior}} \times \left(1 - \frac{2}{n+1}\right)$$

- P_t es el precio de cierre actual.
- n es el número de períodos.
- EMA_{anterior} es la EMA calculada en el período anterior.

- **Índice de Fuerza Relativa (RSI)**

- El RSI es un oscilador que mide la velocidad y el cambio de los movimientos de precio.

$$RSI = 100 - \frac{100}{1 + RS}$$

Donde RS es la media de las ganancias cerradas dividida por la media de las pérdidas cerradas durante un período n .

Cálculo de RS :

$$RS = \frac{\text{Media de las ganancias durante } n \text{ períodos}}{\text{Media de las pérdidas durante } n \text{ períodos}}$$

- **Bandas de Bollinger**

Las Bandas de Bollinger consisten en una media móvil y dos desviaciones estándar (una superior y otra inferior) alrededor de la media móvil.

- Banda superior: $B_{\text{sup}} = \text{SMA}(P, n) + k \times \sigma$
- Banda inferior: $B_{\text{inf}} = \text{SMA}(P, n) - k \times \sigma$

- $\text{SMA}(P, n)$ es la media móvil simple del precio P durante n períodos.
- σ es la desviación estándar del precio durante n períodos.
- k es un factor multiplicador (generalmente 2).

- **ADX (Average Directional Index)**

- El ADX mide la fuerza de la tendencia del mercado.

$$ADX = 100 \times \text{SMA}\left(\frac{|+DI - -DI|}{+DI + -DI}, n\right)$$

- $+DI$ es el indicador direccional positivo.
- $-DI$ es el indicador direccional negativo.
- n es el número de períodos.

- **MACD (Moving Average Convergence Divergence)**

- El MACD mide la diferencia entre dos medias móviles exponenciales (EMA).

$$\text{MACD} = \text{EMA}_{\text{rápida}} - \text{EMA}_{\text{lenta}}$$

- $\text{EMA}_{\text{rápida}}$ es típicamente la EMA de 12 períodos.
- $\text{EMA}_{\text{lenta}}$ es típicamente la EMA de 26 períodos.

- **Oscilador Estocástico**

- El oscilador estocástico compara el precio de cierre de un activo con su rango de precios durante un período específico.

$$\%K = \frac{(C - L_{14})}{(H_{14} - L_{14})} \times 100$$

- C es el precio de cierre más reciente.
- L_{14} es el precio más bajo en los últimos 14 períodos.
- H_{14} es el precio más alto en los últimos 14 períodos.

$$\%D = \text{SMA}(\%K, 3)$$

Donde:

- $\%D$ es la media móvil simple de $\%K$ durante 3 períodos.

- **Volumen en Balance (OBV - On-Balance Volume)**

- Se utiliza para confirmar tendencias mediante el análisis del volumen de operaciones.

- **Acumulación/Distribución (A/D)**

- Mide la relación entre el precio de cierre y el rango verdadero multiplicado por el volumen.

Todos estos indicadores se pueden clasificar:

Indicadores de Tendencia(SMA, EMA, MACD)

Identificar la dirección general en la que se mueve el precio de un activo en el mercado financiero

Indicadores de Momento(Oscilador Estocástico, ADX,CCI)

Miden la velocidad y el cambio de los movimientos de precios de un activo financiero

Indicadores de Volatilidad(Bandas de Bollinger)

Miden la variabilidad y el rango de los movimientos de precios de un activo financiero

Indicadores de Volumen(OBV, A/D)

Analizan la cantidad de activos que se compran y venden en un período de tiempo determinado

Aunque hay muchos tipo de indicadores, y por lo general no hay ninguno mejor que otro, si es cierto que algunos són más indicados para cierto tipo de estrategias o empresas; ya sea por que son más volátiles, por el precio de la acción, entre otras variables.

Esto se explica por que muchos de estos indicadores se tiene que tener en cuenta si quieres tener un análisis de mercado rico y útil, más si quieres ofrecer estrategias fuertes a clientes variados como usuarios de una APP o web.

¡Gracias por leerme!

