

Proyecto 1: Estructura de Mercado

IA en Finanzas
M.U Inteligencia Artificial Aplicada
fcalle@grupobme.es
Fecha Entrega: 11 de Marzo de 2024.

1. Exploración de los datos (2pt)

En este ejercicio vamos a explorar los datos tick a tick del activo Santander del año 2018.

1. En el fichero `datosSAN2018.zip` encontrarás un fichero csv para cada día con cada una de las negociaciones que se realizaron, con la fecha, el precio y el volumen negociado. Carga todos los ficheros y júntalos en un único DataFrame de pandas con el que trabajaremos durante toda la práctica.
2. Usando los datos del apartado anterior calcula las velas (open, high, low, close, vol) diarias y mensuales.
3. Haz un gráfico de velas diarias y mensuales.
4. Para un día elegido por ti, calcula el precio ponderado por volumen (VWAP) desde el inicio del día, en cada instante de tiempo.
5. Obtén una gráfico parecido a la Figura 1: donde puedas ver las negociaciones, el VWAP y volumen acumulado a lo largo del día. ¿Qué conclusiones puedes extraer?

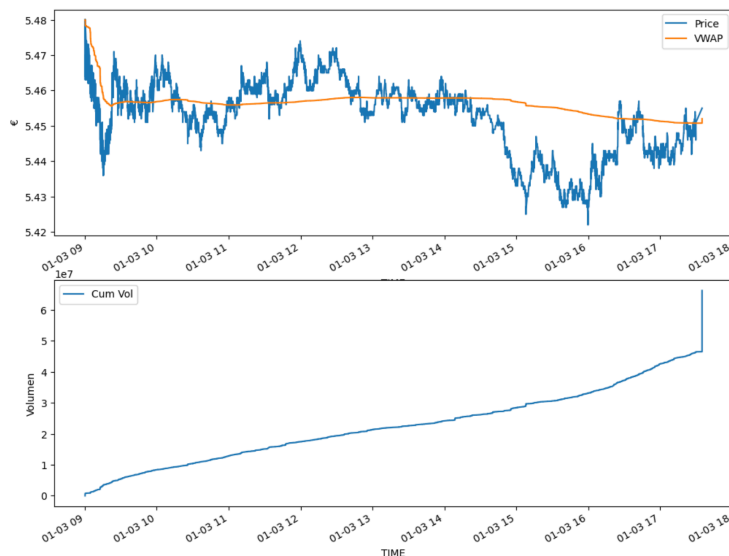


Figura 1: Precio, VWAP y volumen acumulado.

2. Cálculo y estimación del volumen diario (4pt)

En este ejercicio calcularemos el volumen diario y intentaremos realizar un modelo de aprendizaje automático para predecirla.

1. Calcula para cada día el volumen negociado en intervalos de 30 minutos. Normalizaremos esta curva con el volumen total diario para obtener un porcentaje por cada intervalo. Realiza un gráfico para un día en concreto, obtendrás algo parecido a la Figura 2.
2. Calcula el volumen medio diario por intervalos de 30 minutos de todo el año, ¿qué conclusiones puedes extraer?.
3. Realiza un modelo de regresión donde se estime para cada día el porcentaje de volumen de cada intervalo. Se tendrán en cuenta los siguientes puntos:
 - a) Como características de entrada al modelo puedes empezar por usar la curva de los cinco días anteriores. Posteriormente puedes diseñar características que tengan en cuenta otros factores.
 - b) Puedes usar el modelo que creas conveniente desde una regresión lineal, una red neuronal o modelos autorregresivos.
 - c) Realiza una correcta división en entrenamiento, validación y test de los datos, teniendo en cuenta que estamos tratando con un problema de series temporales.
 - d) Obtén las métricas adecuadas para evaluar la validez del modelo.
 - e) Se valorará el uso de diferentes alternativas tanto en el modelo elegido, como en las características de entrada al mismo.

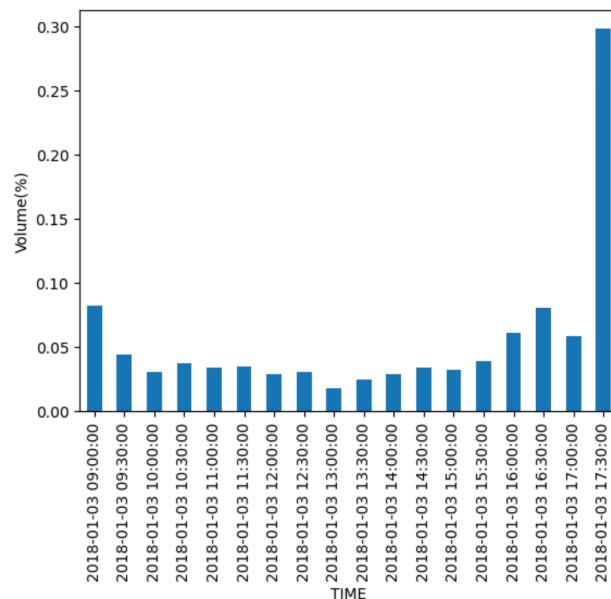


Figura 2: Volumen para el día 3 de enero de 2018 en intervalos de 30 minutos.

3. Optimización de Carteras (4pt)

Utilizando los datos del IBEX 35 disponibles en `IBEX.csv`, que contienen los precios de cierre de todos los activos que formaron parte del índice durante el periodo 2003-2022, realizaremos los siguientes algoritmos:

1. Ponderación Equitativa (Equally Weighted).
2. Ponderación por Volatilidad Inversa.

3. Optimización mediante la Teoría Moderna de Carteras (Modern Portfolio Theory).
4. Optimización mediante Hierarchical Risk Parity (HRP).

Los rebalances se llevarán a cabo cada 10 días, y compararemos los resultados con el índice IBEX con Dividendos (archivo `IBEX_div.csv`).