



**DATA SCIENCE
MACHINE LEARNING**

METADATA

BITCOIN



CODERHOUSE



COMISION 31490



LUCIANO BIZIN



21 OCTUBRE 2022





ÍNDICE

1-14

Bases de datos

Creación de bases de datos y modelar

15-17

Análisis univariado, bivariado y multivariado

Análisis exploratorio de datos (EDA)

17-22

Modelado

Modelos de Machine Learning

22-24

Librerías utilizadas

Manipulación de datos



Metadata

Aplicando modelos de ML a BTC

1. Bases de datos

1.1. *BTC_final*

Fecha de creación: 23 de julio de 2022

Diccionario y tipos de datos:

- Date: fecha en YYYY-MM-DD (datetime64 [ns])
- Price: precio de cierre de BTC en el día de la fecha (USD) (float) (Variable numérica continua)
- Open: precio de apertura de BTC en el día de la fecha (USD) (float) (Variable numérica continua)
- High: precio más alto de BTC en el día de la fecha (USD) (float) (Variable numérica continua)
- Low: precio más bajo de BTC en el día de la fecha (USD) (float) (Variable numérica continua)
- Vol.: volumen de BTC (Nº de intercambios) en el día de la fecha (float) (Variable numérica continua)
- Percentage_diff: diferencia % del precio de BTC con día anterior (float) (Variable numérica continua)
- Target: 1 -> precio del día de BTC subió, y 0 -> precio bajó (float) (Tratada como variable categórica).

Ubicación: https://drive.google.com/file/d/1M38n5So0-6r_Q0vhs1i5nIW-cpJgDlww/view?usp=sharing

Notebook: BTC_final.ipynb

Ubicación notebook: <https://drive.google.com/file/d/1k8hseoTHvi2q644RE-geU4UI0GYQnPNL/view?usp=sharing>

Base de datos (sin tratamiento): Bitcoin Historical Data (investing).csv

https://drive.google.com/file/d/1pnStUmNaW2CK2jKc_yCLaj2nsL7nLLOC/view?usp=sharing

Fuente de datos (sin tratamiento): <https://www.investing.com/crypto/bitcoin/historical-data>

1.2. *BTC_metals*

Fecha de creación: 23 de julio de 2022

Diccionario y tipos de datos:

- Date: fecha en YYYY-MM-DD (datetime64 [ns])
- Price: precio de cierre de BTC en el día de la fecha (USD) (float) (Variable numérica continua)
- Open: precio de apertura de BTC en el día de la fecha (USD) (float) (Variable numérica continua)
- High: precio más alto de BTC en el día de la fecha (USD) (float) (Variable numérica continua)
- Low: precio más bajo de BTC en el día de la fecha (USD) (float) (Variable numérica continua)
- Vol.: volumen de BTC (Nº de intercambios) en el día de la fecha (float) (Variable numérica continua)
- Percentage_diff: diferencia % del precio de BTC con día anterior (float) (Variable numérica continua)
- Target: 1 -> precio del día de BTC subió, y 0 -> precio bajó (float) (Tratada como variable categórica).
- Price_gold: indica el precio de cierre del oro en el día de la fecha (float) (USD) (variable numérica continua)
- Perc_diff_gold: diferencial porcentual del precio del oro de la fecha [x+1] con respecto a la fecha [x] (float) (variable numérica continua)
- Price_silver: indica el precio de cierre de la plata en el día de la fecha (float) (USD) (variable numérica continua)
- Perc_diff_silver: diferencial porcentual del precio de la plata en la fecha [x+1] con respecto a la fecha [x] (float) (variable numérica continua)
- Price_copper: indica el precio de cierre del cobre en el día de la fecha (float) (USD) (variable numérica continua)

- Perc_diff_copper: diferencial porcentual del precio del cobre en la fecha [x+1] con respecto a la fecha [x] (float) (variable numérica continua)
- Price_aluminium: indica el precio de cierre del aluminio en el día de la fecha (float) (USD) (variable numérica continua).
- Perc_diff_aluminium: diferencial porcentual del precio del aluminio en la fecha [x+1] con respecto a la fecha [x] (float) (variable numérica continua)
- Price_platinum: indica el precio de cierre del platino en el día de la fecha (float) (USD) (variable numérica continua)
- Perc_diff_platinum: diferencial porcentual del precio del platino en la fecha [x+1] con respecto a la fecha [x] (float) (variable numérica continua)
- Price_palladium: indica el precio de cierre del paladio en el día de la fecha (float) (USD) (variable numérica continua)
- Perc_diff_palladium: diferencial porcentual del precio del paladio en la fecha [x+1] con respecto a la fecha [x] (float) (variable numérica continua)

Ubicación:

<https://drive.google.com/file/d/1Mw8RrLFMLLgnRTQ0gK3XAO4kgR98orwY/view?usp=sharing>

Notebook: BTC_metals.ipynb

Ubicación notebook:

https://drive.google.com/file/d/1JfSeGoWkLvyQJnd2Db8gAlUMYgh_mUmI/view?usp=sharing

Base de datos (sin tratamiento):

- Aluminium Historical Data.csv
<https://drive.google.com/file/d/1tu8mU5pgmEP9zWcPQxml4VMytlaY96ip/view?usp=sharing>
- Copper Futures Historical Data.csv
https://drive.google.com/file/d/1fdM_-MAnfWDGqICs13G89QqdTY0TWvah/view?usp=sharing
- Gold Futures Historical Data.csv
<https://drive.google.com/file/d/1BokDIq8epLSMk2g88yCSnIrsbJ0dHel/view?usp=sharing>
- Silver Futures Historical Data.csv
https://drive.google.com/file/d/1qY9j8f8drd1eOL_MC0fntg5AOCi-qNYb/view?usp=sharing
- Palladium Futures Historical Data.csv
https://drive.google.com/file/d/1ZS_YeZVSBmBmZqKMwPYzYj7Dasa0Rmuy/view?usp=sharing
- Platinum Futures Historical Data.csv
https://drive.google.com/file/d/1h_k_JleG1_quEu9gZqolz52YHN6TvPA/view?usp=sharing

Fuente de datos (sin tratamiento): <https://www.investing.com/commodities/metals>

Fuente de BTC:

- BTC_final.csv
https://drive.google.com/file/d/1M38n5So0-6r_Q0vhs1i5nIW-cpigDlww/view?usp=sharing

1.3. *BTC_ind_trend*

Fecha de creación: 23 de julio de 2022

Diccionario y tipos de datos:

- Date: fecha en YYYY-MM-DD (object)
- Price: precio de cierre de BTC en el día de la fecha (USD) (float) (Variable numérica continua)
- Open: precio de apertura de BTC en el día de la fecha (USD) (float) (Variable numérica continua)
- High: precio más alto de BTC en el día de la fecha (USD) (float) (Variable numérica continua)
- Low: precio más bajo de BTC en el día de la fecha (USD) (float) (Variable numérica continua)

- Vol.: volumen de BTC (Número de intercambios) en el día de la fecha (float) (Variable numérica continua)
- Percentage_diff: diferencia porcentual del precio de BTC en la fecha [x+1] con respecto a la fecha [x] (float) (Variable numérica continua)
- Target: 1 indica que en el día de la fecha el precio de BTC subió, y 0 que el precio bajó (float) (Se la tratará como variable categórica)
- SMA_7d, SMA_15d, SMA_30d, SMA_50d, SMA_100d, SMA_200d**: media móvil simple (diferentes períodos) (USD) (float) (Variable numérica continua)
- EMA_7d, EMA_15d, EMA_30d, EMA_50d, EMA_100d, EMA_200d**: media móvil exponencial (diferentes períodos) (USD) (float) (Variable numérica continua)
- HT_TRENDLINE: indicador de Hilbert Transform - Instantaneous Trendline (USD) (float) (Variable numérica continua)
- BBANDS_upper_band: banda superior del indicador Bandas de Bollinger (USD) (float) (Variable numérica continua)
- BBANDS_middle_band: banda media del indicador Bandas de Bollinger (USD) (float) (Variable numérica continua)
- BBANDS_lower_band: banda inferior del indicador Bandas de Bollinger (USD) (float) (Variable numérica continua)
- PSAR: indicador Parabolic SAR (USD) (float) (Variable numérica continua)
- Price_silver: indica el precio de cierre de la plata en el día de la fecha (float) (USD) (variable numérica continua)

Ubicación: <https://drive.google.com/file/d/1Gj8BFSjnOvHpWoS9IQ4LN7DU-kxs0a3/view?usp=sharing>

Notebook: BTC_ind_trading.ipynb

Ubicación notebook:

<https://drive.google.com/file/d/1wgCY6oD52ZBlr8iKYy972MDTDi0v5ab2/view?usp=sharing>

Librería utilizada: ta-lib -> https://mrjbq7.github.io/ta-lib/doc_index.html

Fuente de BTC:

- BTC_final.csv
https://drive.google.com/file/d/1M38n5So0-6r_Q0vhs1i5nIW-cpigDlww/view?usp=sharing

1.4. BTC_ind_trend_ET

Fecha de creación: 23 de julio de 2022

Diccionario y tipos de datos:

- Date: fecha en YYYY-MM-DD (object)
- Price: precio de cierre de BTC en el día de la fecha (USD) (float) (Variable numérica continua)
- Open: precio de apertura de BTC en el día de la fecha (USD) (float) (Variable numérica continua)
- High: precio más alto de BTC en el día de la fecha (USD) (float) (Variable numérica continua)
- Low: precio más bajo de BTC en el día de la fecha (USD) (float) (Variable numérica continua)
- Vol.: volumen de BTC (Número de intercambios) en el día de la fecha (float) (Variable numérica continua)
- Percentage_diff: diferencia porcentual del precio de BTC en la fecha [x+1] con respecto a la fecha [x] (float) (Variable numérica continua)
- Target: 1 indica que en el día de la fecha el precio de BTC subió, y 0 que el precio bajó (float) (Se la tratará como variable categórica)
- SMA_7d, SMA_15d, SMA_30d, SMA_50d, SMA_100d, SMA_200d**: media móvil simple (diferentes períodos) (USD) (float) (Variable numérica continua)

- EMA_7d, EMA_15d, EMA_30d, EMA_50d, EMA_100d, EMA_200d*: media móvil exponencial (diferentes períodos) (USD) (float) (Variable numérica continua)
- HT_TRENDLINE: indicador de Hilbert Transform - Instantaneous Trendline (USD) (float) (Variable numérica continua)
- BBANDS_upper_band: banda superior del indicador Bandas de Bollinger (USD) (float) (Variable numérica continua)
- BBANDS_middle_band: banda media del indicador Bandas de Bollinger (USD) (float) (Variable numérica continua)
- BBANDS_lower_band: banda inferior del indicador Bandas de Bollinger (USD) (float) (Variable numérica continua)
- PSAR: indicador Parabolic SAR (USD) (float) (Variable numérica continua) Price_silver: indica el precio de cierre de la plata en el día de la fecha (float) (USD) (variable numérica continua)
- SMA_7d_ET, SMA_15d_ET, SMA_30d_ET, SMA_50d_ET, SMA_100d_ET, SMA_200d_ET: valores categóricos de la aplicación de la estrategia de trading SMA (comprar, vender, mantener posición) (object) (Variable categórica)
- EMA_7d_ET, EMA_15d_ET, EMA_30d_ET, EMA_50d_ET, EMA_100d_ET, EMA_200d_ET: valores categóricos de la aplicación de la estrategia de trading EMA (comprar, vender, mantener posición) (object) (Variable categórica)
- HT_TRENDLINE_ET: valores categóricos de la aplicación de la estrategia de trading HT Trendline (comprar, vender, mantener posición) (object) (Variable categórica)
- PSAR_ET: valores categóricos de la aplicación de la estrategia de trading PSAR (comprar, vender, mantener posición) (object) (Variable categórica)

Ubicación:

https://drive.google.com/file/d/14lnPpEThlA6gVoc0CZPXVJ4PPZM0PpH_/view?usp=sharing

Notebook: BTC_ind_trading.ipynb

Ubicación notebook:

<https://drive.google.com/file/d/1wgCY6oD52ZBlr8iKYy972MDTDi0v5ab2/view?usp=sharing>

Librería utilizada: ta-lib -> https://mrjbq7.github.io/ta-lib/doc_index.html

Fuente de BTC:

- BTC_final.csv
https://drive.google.com/file/d/1M38n5So0-6r_Q0vhs1i5nIW-cpigDlww/view?usp=sharing

1.5. BTC_ind_mom

Fecha de creación: 23 de julio de 2022

Diccionario y tipos de datos:

- Date: fecha en YYYY-MM-DD (object)
- Price: precio de cierre de BTC en el día de la fecha (USD) (float) (Variable numérica continua)
- Open: precio de apertura de BTC en el día de la fecha (USD) (float) (Variable numérica continua)
- High: precio más alto de BTC en el día de la fecha (USD) (float) (Variable numérica continua)
- Low: precio más bajo de BTC en el día de la fecha (USD) (float) (Variable numérica continua)
- Vol.: volumen de BTC (Número de intercambios) en el día de la fecha (float) (Variable numérica continua)
- Percentage_diff: diferencia porcentual del precio de BTC en la fecha [x+1] con respecto a la fecha [x] (float) (Variable numérica continua)
- Target: 1 indica que en el día de la fecha el precio de BTC subió, y 0 que el precio bajó (float) (Se la tratará como variable categórica)

- ADX: valores del índice de movimiento direccional medio (float) (Variable numérica continua)
- ADX_MDI: valores del Corriente Menos Movimiento Direccional (MDI) (float) (Variable numérica continua)
- ADX_MDM: valores del Mínimo Actual Menos el Mínimo Anterior (-DM) (float) (Variable numérica continua)
- ADX_DX: valores del índice de movimiento direccional (DX) (float) (Variable numérica continua)
- ADX_PDI: valores del Corriente Plus Movimiento Direccional (PDM) (float) (Variable numérica continua)
- ADX_PDM: valores del Máximo Actual Menos el Máximo Anterior (+DM) (float) (Variable numérica continua)
- Aroondown: valores del número de períodos desde los últimos 25 períodos máximos (float) (Variable numérica continua)
- Aroonup: valores del número de períodos desde los últimos 25 períodos máximos (float) (Variable numérica continua)
- Aroon_Oscillator: valores del Oscilador Aroon (float) (Variable numérica continua)
- MACD: valores del MACD (float) (Variable numérica continua)
- MACDSIGNAL: valores de la señal de MACD (float) (Variable numérica continua)
- MACDHIST: valores del histograma del MACD (float) (Variable numérica continua)
- MOM: valores del indicador Momentum (float) (Variable numérica continua)
- ROC: valores de la tasa de cambio de una variable durante un período específico de tiempo (float) (Variable numérica continua)
- RSI: valores del índice de fuerza relativa (float) (Variable numérica continua)
- STOCHRSI_fastk: valores de la aplicación del método estocástico al indicador RSI (con variable d) (más lento) (float) (Variable numérica continua)
- STOCHRSI_fastd: valores de la aplicación del método estocástico al indicador RSI (con variable k) (más rápido) (float) (Variable numérica continua)
- ULTOSC: valores del indicador Ultimate Oscillator (float) (Variable numérica continua)
- WILLR: valores del indicador Williams %R (float) (Variable numérica continua)

Ubicación:

<https://drive.google.com/file/d/1mloaGRsEdh4GU32jD5e9SQMc7Qhr08y1/view?usp=sharing>

Notebook: BTC_ind_trading.ipynb

Ubicación notebook:

<https://drive.google.com/file/d/1wgCY6oD52ZBlr8iKYy972MDTDi0v5ab2/view?usp=sharing>

Librería utilizada: ta-lib -> https://mrjibq7.github.io/ta-lib/doc_index.html

Fuente de BTC:

- BTC_final.csv
https://drive.google.com/file/d/1M38n5So0-6r_Q0vhs1i5nIW-cpigDlww/view?usp=sharing

1.6. *BTC_ind_mom_ET*

Fecha de creación: 23 de julio de 2022

Diccionario y tipos de datos:

- Date: fecha en YYYY-MM-DD (object)
- Price: precio de cierre de BTC en el día de la fecha (USD) (float) (Variable numérica continua)
- Open: precio de apertura de BTC en el día de la fecha (USD) (float) (Variable numérica continua)
- High: precio más alto de BTC en el día de la fecha (USD) (float) (Variable numérica continua)
- Low: precio más bajo de BTC en el día de la fecha (USD) (float) (Variable numérica continua)

- Vol.: volumen de BTC (Número de intercambios) en el día de la fecha (float) (Variable numérica continua)
- Percentage_diff: diferencia porcentual del precio de BTC en la fecha [x+1] con respecto a la fecha [x] (float) (Variable numérica continua)
- Target: 1 indica que en el día de la fecha el precio de BTC subió, y 0 que el precio bajó (float) (Se la tratará como variable categórica)
- ADX: valores del índice de movimiento direccional medio (float) (Variable numérica continua)
- ADX_MDI: valores del Corriente Menos Movimiento Direccional (MDI) (float) (Variable numérica continua)
- ADX_MDM: valores del Mínimo Actual Menos el Mínimo Anterior (-DM) (float) (Variable numérica continua)
- ADX_DX: valores del índice de movimiento direccional (DX) (float) (Variable numérica continua)
- ADX_PDI: valores del Corriente Plus Movimiento Direccional (PDM) (float) (Variable numérica continua)
- ADX_PDM: valores del Máximo Actual Menos el Máximo Anterior (+DM) (float) (Variable numérica continua)
- Aroondown: valores del número de períodos desde los últimos 25 períodos máximos (float) (Variable numérica continua)
- Aroonup: valores del número de períodos desde los últimos 25 períodos mínimos (float) (Variable numérica continua)
- Aroon_Oscillator: valores del Oscilador Aroon (float) (Variable numérica continua)
- MACD: valores del MACD (float) (Variable numérica continua)
- MACDSIGNAL: valores de la señal de MACD (float) (Variable numérica continua)
- MACDHIST: valores del histograma del MACD (float) (Variable numérica continua)
- MOM: valores del indicador Momentum (float) (Variable numérica continua)
- ROC: valores de la tasa de cambio de una variable durante un período específico de tiempo (float) (Variable numérica continua)
- RSI: valores del índice de fuerza relativa (float) (Variable numérica continua)
- STOCHRSI_fastk: valores de la aplicación del método estocástico al indicador RSI (con variable d) (más lento) (float) (Variable numérica continua)
- STOCHRSI_fastd: valores de la aplicación del método estocástico al indicador RSI (con variable k) (más rápido) (float) (Variable numérica continua)
- ULTOSC: valores del indicador Ultimate Oscillator (float) (Variable numérica continua)
- WILLR: valores del indicador Williams %R (float) (Variable numérica continua)
- ADX_ET: valores categóricos de la aplicación de la estrategia ADX (comprar, vender, mantener posición) (object)
- ADXR_ET: valores categóricos de la aplicación de la estrategia ADXR (comprar, vender, mantener posición) (object)
- MACD_ET: valores categóricos de la aplicación de la estrategia MACD (comprar, vender, mantener posición) (object)
- ROC_ET: valores categóricos de la aplicación de la estrategia ROC (comprar, vender, mantener posición) (object)
- STOCHRSI_fastk_ET: valores categóricos de la aplicación de la estrategia STOCHRSI_fastk (comprar, vender, mantener posición) (object)
- STOCHRSI_fastd_ET: valores categóricos de la aplicación de la estrategia STOCHRSI_fastd (comprar, vender, mantener posición) (object)

Ubicación: <https://drive.google.com/file/d/1cUBbMAbGg1iELh1pdg-dOJmrcBwbFmzp/view?usp=sharing>

Notebook: BTC_ind_trading.ipynb

Ubicación notebook:

<https://drive.google.com/file/d/1wgCY6oD52ZBlr8iKYy972MDTDi0v5ab2/view?usp=sharing>

Librería utilizada: ta-lib -> https://mrjbq7.github.io/ta-lib/doc_index.html

Fuente de BTC:

- BTC_final.csv
https://drive.google.com/file/d/1M38n5So0-6r_Q0vhs1i5nIW-cpigDlww/view?usp=sharing

1.7. *BTC_ind_vol*

Fecha de creación: 23 de julio de 2022

Diccionario y tipos de datos:

- Date: fecha en YYYY-MM-DD (object)
- Price: precio de cierre de BTC en el día de la fecha (USD) (float) (Variable numérica continua)
- Open: precio de apertura de BTC en el día de la fecha (USD) (float) (Variable numérica continua)
- High: precio más alto de BTC en el día de la fecha (USD) (float) (Variable numérica continua)
- Low: precio más bajo de BTC en el día de la fecha (USD) (float) (Variable numérica continua)
- Vol.: volumen de BTC (Número de intercambios) en el día de la fecha (float) (Variable numérica continua)
- Percentage_diff: diferencia porcentual del precio de BTC en la fecha [x+1] con respecto a la fecha [x] (float) (Variable numérica continua)
- Target: 1 indica que en el día de la fecha el precio de BTC subió, y 0 que el precio bajó (float) (Se la tratará como variable categórica)
- ADOSC: indica la fuerza de los movimientos de precios y la presión subyacente de compra y venta (float) (Variable numérica continua)

Ubicación: https://drive.google.com/file/d/1d-Vtnz_mvVHCucOt4VkTfhYcsvsyVI7/view?usp=sharing

Notebook: BTC_ind_trading.ipynb

Ubicación notebook:

<https://drive.google.com/file/d/1wgCY6oD52ZBlr8iKYy972MDTDi0v5ab2/view?usp=sharing>

Librería utilizada: ta-lib -> https://mrjbq7.github.io/ta-lib/doc_index.html

Fuente de BTC:

- BTC_final.csv
https://drive.google.com/file/d/1M38n5So0-6r_Q0vhs1i5nIW-cpigDlww/view?usp=sharing

1.8. *BTC_ind_cycle_indicators*

Fecha de creación: 23 de julio de 2022

Diccionario y tipos de datos:

- Date: fecha en YYYY-MM-DD (object)
- Price: precio de cierre de BTC en el día de la fecha (USD) (float) (Variable numérica continua)
- Open: precio de apertura de BTC en el día de la fecha (USD) (float) (Variable numérica continua)
- High: precio más alto de BTC en el día de la fecha (USD) (float) (Variable numérica continua)
- Low: precio más bajo de BTC en el día de la fecha (USD) (float) (Variable numérica continua)
- Vol.: volumen de BTC (Número de intercambios) en el día de la fecha (float) (Variable numérica continua)
- Percentage_diff: diferencia porcentual del precio de BTC en la fecha [x+1] con respecto a la fecha [x] (float) (Variable numérica continua)
- Target: 1 indica que en el día de la fecha el precio de BTC subió, y 0 que el precio bajó (float) (Se la tratará como variable categórica)

- HT_DCPERIOD: Hilbert Transform Dominant Cycle Period (período de tiempo del ciclo dominante actual en la serie de Datos, en barras; rango válido = 10.60 barras, retraso = -10 barras) (float) (Variable numérica continua).
- HT_DCPHASE: Hilbert Transform - Dominant Cycle Phase (fase del ciclo dominante en una barra específica entre 0 a 360 grados dentro del período de transformada de Hilbert actual medido instantáneamente en esa barra) (grados) (float) (Variable numérica continua)
- HT_PHASOR_inphase: el elemento inphase de la amplitud de la senoide modulada (float) (Variable numérica continua)
- HT_PHASOR_quadrature: el elemento quadrature de la amplitud de la senoide modulada (float) (Variable numérica continua)
- HT_SINE_sine: Hilbert Transform - SineWave (-1 a 1) (float) (Variable numérica continua)
- HT_PHASOR_leadsine: Hilbert Transform seno del lead (-1 a 1) (float) (Variable numérica continua)
- HT_SINE_integer: Hilbert Transform trend indicator (retorna 1 para Trend Mode, 0 para Cycle Mode) (int) (Funciona como variable categórica)

Ubicación: <https://drive.google.com/file/d/133y-TBLhD6JFJxXA2xGvVnVTFY1hww06/view?usp=sharing>

Notebook: BTC_ind_trading.ipynb

Ubicación notebook:

<https://drive.google.com/file/d/1wgCY6oD52ZBlr8iKYy972MDTDi0v5ab2/view?usp=sharing>

Librería utilizada: ta-lib -> https://mrjbq7.github.io/ta-lib/doc_index.html

Fuente de BTC:

- BTC_final.csv
https://drive.google.com/file/d/1M38n5So0-6r_Q0vhs1i5nIW-cpigDlww/view?usp=sharing

1.9. BTC_ind_volat

Fecha de creación: 23 de julio de 2022

Diccionario y tipos de datos:

- Date: fecha en YYYY-MM-DD (object)
- Price: precio de cierre de BTC en el día de la fecha (USD) (float) (Variable numérica continua)
- Open: precio de apertura de BTC en el día de la fecha (USD) (float) (Variable numérica continua)
- High: precio más alto de BTC en el día de la fecha (USD) (float) (Variable numérica continua)
- Low: precio más bajo de BTC en el día de la fecha (USD) (float) (Variable numérica continua)
- Vol.: volumen de BTC (Número de intercambios) en el día de la fecha (float) (Variable numérica continua)
- Percentage_diff: diferencia porcentual del precio de BTC en la fecha [x+1] con respecto a la fecha [x] (float) (Variable numérica continua)
- Target: 1 indica que en el día de la fecha el precio de BTC subió, y 0 que el precio bajó (float) (Se la tratará como variable categórica)
- ATR: valores del indicador Average True Range (Promedio de Rango Verdadero) (float) (Variable numérica continua)

Ubicación: https://drive.google.com/file/d/18kg9TzmQjfDgCA1Lca2czMPSN5JdRx_d/view?usp=sharing

Notebook: BTC_ind_trading.ipynb

Ubicación notebook:

<https://drive.google.com/file/d/1wgCY6oD52ZBlr8iKYy972MDTDi0v5ab2/view?usp=sharing>

Librería utilizada: ta-lib -> https://mrjbq7.github.io/ta-lib/doc_index.html

Fuente de BTC:

- BTC_final.csv
https://drive.google.com/file/d/1M38n5So0-6r_Q0vhs1i5nIW-cpigDlww/view?usp=sharing

1.10. *BTC_statistic_functions*

Fecha de creación: 23 de julio de 2022

Diccionario y tipos de datos:

- Date: fecha en YYYY-MM-DD (object)
- Price: precio de cierre de BTC en el día de la fecha (USD) (float) (Variable numérica continua)
- Open: precio de apertura de BTC en el día de la fecha (USD) (float) (Variable numérica continua)
- High: precio más alto de BTC en el día de la fecha (USD) (float) (Variable numérica continua)
- Low: precio más bajo de BTC en el día de la fecha (USD) (float) (Variable numérica continua)
- Vol.: volumen de BTC (Número de intercambios) en el día de la fecha (float) (Variable numérica continua)
- Percentage_diff: diferencia porcentual del precio de BTC en la fecha [x+1] con respecto a la fecha [x] (float) (Variable numérica continua)
- Target: 1 indica que en el día de la fecha el precio de BTC subió, y 0 que el precio bajó (float) (Se la tratará como variable categórica)
- Beta: valores beta de correlación (-1 a 1) (float) (Variable numérica continua)
- Pearson: valores de correlación de Pearson (-1 a 1) (float) (Variable numérica continua)
- Linearregression: precio estimado por regresión lineal (USD) (float) (Variable numérica continua)
- Linearangle: ángulo de la mejor línea de regresión (ángulos) (float) (Variable numérica continua)
- Linearintercept: intercepción de la regresión lineal (USD) (float) (Variable numérica continua)
- Linearslope: curvatura de la regresión lineal (grados) (float) (Variable numérica continua)
- VAR: varianza (medida de dispersión definida como la esperanza del cuadrado de la desviación de dicha variable respecto a su media) (float) (Variable numérica continua)
- STDDEV: desviación estándar (medición de la dispersión de los datos) (float) (Variable numérica continua)
- TSF: Time Series Forecasting (predicción de precio) (USD) (float) (Variable numérica continua)

Ubicación:

<https://drive.google.com/file/d/1Zu4pmlYQKy6jrAHRR2ojjLREQ4MQywJ/view?usp=sharing>

Notebook: BTC_ind_trading.ipynb

Ubicación notebook:

<https://drive.google.com/file/d/1wgCY6oD52ZBlr8iKYy972MDTDi0v5ab2/view?usp=sharing>

Librería utilizada: ta-lib -> https://mrjbq7.github.io/ta-lib/doc_index.html

Fuente de BTC:

- BTC_final.csv
https://drive.google.com/file/d/1M38n5So0-6r_Q0vhs1i5nIW-cpigDlww/view?usp=sharing

1.11. *BTC_patterns_indicator*

Fecha de creación: 23 de julio de 2022

Diccionario y tipos de datos:

- Date: fecha en YYYY-MM-DD (object)
- Close: precio de cierre de BTC en el día de la fecha (USD) (float) (Variable numérica continua)
- Open: precio de apertura de BTC en el día de la fecha (USD) (float) (Variable numérica continua)
- High: precio más alto de BTC en el día de la fecha (USD) (float) (Variable numérica continua)
- Low: precio más bajo de BTC en el día de la fecha (USD) (float) (Variable numérica continua)
- Vol.: volumen de BTC (Número de intercambios) en el día de la fecha (float) (Variable numérica continua)

- Percentage_diff: diferencia porcentual del precio de BTC en la fecha [x+1] con respecto a la fecha [x] (float) (Variable numérica continua)
- Target: 1 indica que en el día de la fecha el precio de BTC subió, y 0 que el precio bajó (float) (Se la tratará como variable categórica)
- CDL2CROWS: patrón 2 crows (hacia abajo/arriba) (-100, 0, 100) (int) (Funcionan como variable categórica)
- CDL3BLACKCROWS: patrón 3 black crows (hacia abajo/arriba) (-100, 0, 100) (int) (Funcionan como variable categórica)
- CDL3INSIDE: patrón 3 inside (hacia abajo/arriba) (-100, 0, 100) (int) (Funcionan como variable categórica)
- CDL3LINESTRIKE: patrón three lines strike (hacia abajo/arriba) (-100, 0, 100) (int) (Funcionan como variable categórica)
- CDL3OUTSIDE: patrón three outside (hacia abajo/arriba) (-100, 0, 100) (int) (Funcionan como variable categórica)
- CDL3STARSINSOUTH: patrón 3 stars in south (hacia abajo/arriba) (-100, 0, 100) (int) (Funcionan como variable categórica)
- CDL3WHITESOLDIERS: patrón 3 white soldiers (hacia abajo/arriba) (-100, 0, 100) (int) (Funcionan como variable categórica)
- CDLABANDONEDBABY: patrón abandone baby (hacia abajo/arriba) (-100, 0, 100) (int) (Funcionan como variable categórica)
- CDLADVANCEBLOCK: patrón advance block (hacia abajo/arriba) (-100, 0, 100) (int) (Funcionan como variable categórica)
- CDLBELTHOLD: patrón belt hold (hacia abajo/arriba) (-100, 0, 100) (int) (Funcionan como variable categórica)
- CDLBREAKAWAY: patrón break away (hacia abajo/arriba) (-100, 0, 100) (int) (Funcionan como variable categórica)
- CDLCLOSINGMARUBOZU: patrón closing Marubozu (hacia abajo/arriba) (-100, 0, 100) (int) (Funcionan como variable categórica)
- CDLCONCEALBABYSWALL: patrón concealing baby swallow (hacia abajo/arriba) (-100, 0, 100) (int) (Funcionan como variable categórica)
- CDLCOUNTERATTACK: patrón counter attack (hacia abajo/arriba) (-100, 0, 100) (int) (Funcionan como variable categórica)
- CDLDARKCLOUDCOVER: patrón dark cloud cover (hacia abajo/arriba) (-100, 0, 100) (int) (Funcionan como variable categórica)
- CDLDOJI: patrón doji (hacia abajo/arriba) (-100, 0, 100) (int) (Funcionan como variable categórica)
- CDLDOJISTAR: patrón doji star (hacia abajo/arriba) (-100, 0, 100) (int) (Funcionan como variable categórica)
- CDLDRAGONFLYDOJI: patrón dragonfly doji (hacia abajo/arriba) (-100, 0, 100) (int) (Funcionan como variable categórica)
- CDLENGULFING: patrón engulfing (hacia abajo/arriba) (-100, 0, 100) (int) (Funcionan como variable categórica)
- CDLEVENINGDOJISTAR: patrón evening doji star (hacia abajo/arriba) (-100, 0, 100) (int) (Funcionan como variable categórica)
- CDLEVENINGSTAR: patrón evening star (hacia abajo/arriba) (-100, 0, 100) (int) (Funcionan como variable categórica)
- CDLGAPSIDESIDEWHITE: patrón gap side-by-side (hacia abajo/arriba) (-100, 0, 100) (int) (Funcionan como variable categórica)

- CDLGRAVESTONEDOJI: patrón gravestone doji (hacia abajo/arriba) (-100, 0, 100) (int) (Funcionan como variable categórica)
- CDLHAMMER: patrón hammer (hacia abajo/arriba) (0, 100) (int) (Funcionan como variable categórica)
- CDLHANGINGMAN: patrón hanging man (hacia abajo/arriba) (-100, 0, 100) (int) (Funcionan como variable categórica)
- CDLHARAMI: patrón harami (hacia abajo/arriba) (-100, 0, 100) (int) (Funcionan como variable categórica)
- CDLHARAMICROSS: patrón harami cross (hacia abajo/arriba) (-100, 0, 100) (int) (Funcionan como variable categórica)
- CDLHIGHWAVE: patrón high wave (hacia abajo/arriba) (-100, 0, 100) (int) (Funcionan como variable categórica)
- CDLHIKKAKE: patrón Hikkake (hacia abajo/arriba) (-200, 0, 200) (int) (Funcionan como variable categórica)
- CDLHIKKAKEMOD: patrón Hikkake moderated (hacia abajo/arriba) (-200, 0, 200) (int) (Funcionan como variable categórica)
- CDLHOMINGPIGEON: patrón homing pigeon (hacia abajo/arriba) (-200, 0, 200) (int) (Funcionan como variable categórica)
- CDLIDENTICAL3CROWS: patrón identical three crows (hacia abajo/arriba) (-200, 0, 200) (int) (Funcionan como variable categórica)
- CDLINNECK: patrón in neck (hacia abajo/arriba) (-100, 0, 100) (int) (Funcionan como variable categórica)
- CDLINVERTEDHAMMER: patrón inverted hammer (hacia abajo/arriba) (0, 100) (int) (Funcionan como variable categórica)
- CDLKICKING: patrón kicking (hacia abajo/arriba) (-100, 0, 100) (int) (Funcionan como variable categórica)
- CDLKICKINGBYLENGTH: patrón kicking by length (hacia abajo/arriba) (-100, 0, 100) (int) (Funcionan como variable categórica)
- CDLLADDERBOTTOM: patrón ladder bottom (hacia abajo/arriba) (-100, 0, 100) (int) (Funcionan como variable categórica)
- CDLLONGLEGGEDDOJI: patrón long-legged doji (hacia abajo/arriba) (-100, 0, 100) (int) (Funcionan como variable categórica)
- CDLLONGLINE: patrón long line (hacia abajo/arriba) (-100, 0, 100) (int) (Funcionan como variable categórica)
- CDLMARUBOZU: patrón Marubozu (hacia abajo/arriba) (-100, 0, 100) (int) (Funcionan como variable categórica)
- CDLMATCHINGLOW: patrón matching lows (hacia abajo/arriba) (-100, 0, 100) (int) (Funcionan como variable categórica)
- CDLMATHOLD: patrón Matt hold (hacia abajo/arriba) (-100, 0, 100) (int) (Funcionan como variable categórica)
- CDLMORNINGDOJISTAR: patrón morning doji star (hacia abajo/arriba) (-100, 0, 100) (int) (Funcionan como variable categórica)
- CDLMORNINGSTAR: patrón morning star (hacia abajo/arriba) (-100, 0, 100) (int) (Funcionan como variable categórica)
- CDLONNECK: patrón on-neck (hacia abajo/arriba) (-100, 0, 100) (int) (Funcionan como variable categórica)
- CDLPIERCING: patrón piercing (hacia abajo/arriba) (-100, 0, 100) (int) (Funcionan como variable categórica)
- CDLRICKSHAWMAN: patrón Rickshaw man (hacia abajo/arriba) (-100, 0, 100) (int) (Funcionan como variable categórica)

- CDLRISEFALL3METHODS: falling and rising 3 method (hacia abajo/arriba) (-100, 0, 100) (int) (Funcionan como variable categórica)
- CDLSEPARATINGLINES: patrón separating lines (hacia abajo/arriba) (-100, 0, 100) (int) (Funcionan como variable categórica)
- CDLSHOOTINGSTAR: patrón shooting star (hacia abajo/arriba) (-100, 0, 100) (int) (Funcionan como variable categórica)
- CDLSHORTLINE: patrón shortline (hacia abajo/arriba) (-100, 0, 100) (int) (Funcionan como variable categórica)
- CDLSPINNINGTOP: patrón spinning top (hacia abajo/arriba) (-100, 0, 100) (int) (Funcionan como variable categórica)
- CDLSTALLEDPATTERN: patrón stalled (hacia abajo/arriba) (-100, 0, 100) (int) (Funcionan como variable categórica)
- CDLSTICKSANDWICH: patrón sandwich (hacia abajo/arriba) (-100, 0, 100) (int) (Funcionan como variable categórica)
- CDLTAKURI: patrón Takuri (hacia abajo/arriba) (-100, 0, 100) (int) (Funcionan como variable categórica)
- CDLTASUKIGAP: patrón Tasuki gap (hacia abajo/arriba) (-100, 0, 100) (int) (Funcionan como variable categórica)
- CDLTHRUSTING: patrón Thrusting lines (hacia abajo/arriba) (-100, 0, 100) (int) (Funcionan como variable categórica)
- CDLTRISTAR: patrón tri-star (hacia abajo/arriba) (-100, 0, 100) (int) (Funcionan como variable categórica)
- CDLUNIQUE3RIVER: patrón unique three river (hacia abajo/arriba) (-100, 0, 100) (int) (Funcionan como variable categórica)
- CDLUPSIDE GAP2CROWS: patrón downside/upside two crows (hacia abajo/arriba) (-100, 0, 100) (int) (Funcionan como variable categórica)
- CDLXSIDEGAP3METHODS: downside/upside three methods (hacia abajo/arriba) (-100, 0, 100) (int) (Funcionan como variable categórica)

Ubicación: <https://drive.google.com/file/d/1QMbvIXkIhHZsGySv5LD7ltY7Kzwc2Ra/view?usp=sharing>

Notebook: BTC_ind_trading.ipynb

Ubicación notebook:

<https://drive.google.com/file/d/1wgCY6oD52ZBlr8iKYy972MDTDi0v5ab2/view?usp=sharing>

Librería utilizada: ta-lib -> https://mrjbq7.github.io/ta-lib/doc_index.html

Fuente de BTC:

- BTC_final.csv
https://drive.google.com/file/d/1M38n5So0-6r_Q0vhs1i5nIW-cpigDlww/view?usp=sharing

1.12. *BTC_Forex*

Fecha de creación: 23 de julio de 2022

Diccionario y tipos de datos:

- Date: fecha en YYYY-MM-DD (datetime64 [ns])
- Price: precio de cierre de BTC en el día de la fecha (USD) (float) (Variable numérica continua)
- Open: precio de apertura de BTC en el día de la fecha (USD) (float) (Variable numérica continua)
- High: precio más alto de BTC en el día de la fecha (USD) (float) (Variable numérica continua)
- Low: precio más bajo de BTC en el día de la fecha (USD) (float) (Variable numérica continua)
- Vol.: volumen de BTC (Número de intercambios) en el día de la fecha (float) (Variable numérica continua)

- **Percentage_diff**: diferencia porcentual del precio de BTC en la fecha [x+1] con respecto a la fecha [x] (float) (Variable numérica continua)
- **Target**: 1 indica que en el día de la fecha el precio de BTC subió, y 0 que el precio bajó (float) (Se la tratará como variable categórica)
- **Price_EUR_USD**: precio de la relación Euro y Dolar norteamericano (float) (Variable numérica continua)
- **Percentage_diff_EUR_USD**: diferencia porcentual de la relación Euro y Dolar norteamericano (float) (Variable numérica continua)
- **Price_USD_JPY**: precio de la relación Dolar norteamericano y el Yen japonés (float) (Variable numérica continua)
- **Percentage_diff_USD_JPY**: diferencia porcentual de la relación Dolar norteamericano y Yen japonés (float) (Variable numérica continua)
- **Price_EUR_USD**: precio de la relación Dolar norteamericano y el Yuan chino (float) (Variable numérica continua)
- **Percentage_diff_EUR_USD**: diferencia porcentual de la relación Dolar norteamericano y Yuan chino (float) (Variable numérica continua)

Ubicación:

https://drive.google.com/file/d/1I8kN3PwyVamLIHcXqz0XUFuWmwK13fF_/view?usp=sharing

Notebook: BTC_Forex.ipynb

Ubicación notebook: https://drive.google.com/file/d/111_5tWsAXYZVdwrPTg_qa5zpR-A8QB7-/view?usp=sharing

Base de datos (sin tratamiento):

- **BTC_final.csv**
https://drive.google.com/file/d/1M38n5So0-6r_Q0vhs1i5nIW-cpigDlww/view?usp=sharing
- **EUR_USD Historical Data.csv**
<https://drive.google.com/file/d/16iaYVmlv5vtNDN4WZKaQmeO3MsI9QG12/view?usp=sharing>
- **USD_JPY Historical Data.csv**
https://drive.google.com/file/d/1xGO8ejLusQGmQ8_cNLhhX5Mvh0yk2YRO/view?usp=sharing
- **USD_CNY Historical Data.csv**
https://drive.google.com/file/d/1tyfAwJctyEE1Amvy9C7B3n_AaJmqw4av/view?usp=sharing

Fuente de datos (sin tratamiento): <https://www.investing.com/currencies/streaming-forex-rates-majors>

1.13. *BTC_US_Finance*

Fecha de creación: 23 de julio de 2022

Diccionario y tipos de datos:

- **Date**: fecha en YYYY-MM-DD (datetime64 [ns])
- **Price**: precio de cierre de BTC en el día de la fecha (USD) (float) (Variable numérica continua)
- **Open**: precio de apertura de BTC en el día de la fecha (USD) (float) (Variable numérica continua)
- **High**: precio más alto de BTC en el día de la fecha (USD) (float) (Variable numérica continua)
- **Low**: precio más bajo de BTC en el día de la fecha (USD) (float) (Variable numérica continua)
- **Vol.**: volumen de BTC (Número de intercambios) en el día de la fecha (float) (Variable numérica continua)
- **Percentage_diff**: diferencia porcentual del precio de BTC en la fecha [x+1] con respecto a la fecha [x] (float) (Variable numérica continua)
- **Target**: 1 indica que en el día de la fecha el precio de BTC subió, y 0 que el precio bajó (float) (Se la tratará como variable categórica)
- **Price_SP500**: indica el valor de cierre de la bolsa de valores S%P500 (USD) (Variable numérica continua)

- Percentage_diff_SP500: indica el porcentaje de diferencia del cierre con respecto al cierre anterior de S%P500 (float)(Variable numérica continua)
- Price_NASDAQ: indica el valor de cierre de la bolsa de valores NASDAQ (USD) (float) (Variable numérica continua)
- Percentage_diff_NASDAQ: indica el porcentaje de diferencia del cierre con respecto al cierre anterior de NASDAQ (float)(Variable numérica continua)
- Price_DJ: indica el valor de cierre de la bolsa de valores Dow Jones (USD) (float) (Variable numérica continua)
- Percentage_diff_DJ: indica el porcentaje de diferencia del cierre con respecto al cierre anterior de Dow Jones (float) (Variable numérica continua)
- Price_US_Bond: indica el valor del bono a 10 años de EEUU (es uno de los bonos más relevantes a nivel macro) (float) (Variable numérica continua)
- Percentage_diff_US_Bond: indica la diferencia porcentual entre valores del bono a 10 años de EEUU (float) (Variable numérica continua)
- I_Percentage_diff_ACPI: indica la diferencia porcentual interanual entre los valores del Consumer Price Index (float) (inflación) (Variable numérica continua)
- M_Percentage_diff_MCPI: indica la diferencia porcentual mensual entre los valores del Consumer Price Index (inflación) (float) (Variable numérica continua)
- FEDfunds: indica la tasa de inflación de EEUU (float) (Variable numérica continua)

Ubicación:

<https://drive.google.com/file/d/1dg4hIUheDZok9NMXpnycEcZ5HtwrDX/view?usp=sharing>

Notebook: BTC_US_Finance.ipynb

Ubicación notebook:

<https://drive.google.com/file/d/1iz7KPNT7V1uvDPexkNvPLAiZuRWGevgT/view?usp=sharing>

Base de datos (sin tratamiento):

- S&P 500 Historical Data.csv
<https://drive.google.com/file/d/16qG9goRySV2HBlrqWPgpXxKcDRp49WPY/view?usp=sharing>
- NASDAQ Composite Historical Data.csv
<https://drive.google.com/file/d/1S7dwau6JZTIK96pleWbzG-09ftoh5ZNO/view?usp=sharing>
- Dow Jones Industrial Average Historical Data.csv
<https://drive.google.com/file/d/1zNZcUcLHk6lMgEZJC-aUg4xutezSelQ/view?usp=sharing>
- United States 10-Year Bond Yield Historical Data.csv
<https://drive.google.com/file/d/1sm9FndfZDWEbYAA9p7-RaGh3w1BSxyvZ/view?usp=sharing>
- FEDFUNDS.csv
https://drive.google.com/file/d/1LxMAldLl0cDhgYl0WOysiq3XJf28_BY/view?usp=sharing
- CPI (mensual).csv
<https://drive.google.com/file/d/1pKv5OtHps0XI0PvxF-f2avUinrrcFRyr/view?usp=sharing>
- CPI (acumulado interanual).csv
<https://drive.google.com/file/d/1Su1uRSrwPFp2O-6wAJny1SWg2mPucHeS/view?usp=sharing>

Fuente de datos (sin tratamiento):

- FEDFunds: <https://fred.stlouisfed.org/series/FEDFUNDS>
- CPI: <https://fred.stlouisfed.org/series/CPIAUCSL>
- United States 10-Year Bond Yield Historical Data: <https://www.investing.com/rates-bonds/u.s.-10-year-bond-yield-historical-data>

Fuente de BTC:

- BTC_final.csv
https://drive.google.com/file/d/1M38n5So0-6r_Q0vhs1i5nIW-cpigDlww/view?usp=sharing

2. Análisis univariado, bivariado y multivariado

2.1. Análisis univariado, bivariado y multivariado de BTC_metals

Fecha de creación: 13 de agosto de 2022

Notebook: Análisis univariado, bivariado y multivariado de BTC_metals.ipynb

Ubicación notebook:

https://drive.google.com/file/d/1WckUnjXI_D_9h3WoZTJ6upSo0a_WTKt4/view?usp=sharing

Bases de datos analizadas:

- BTC_final.csv:
https://drive.google.com/file/d/1M38n5So0-6r_Q0vhs1i5nIW-cpigDlww/view?usp=sharing
- BTC_metals.csv:
<https://drive.google.com/file/d/1Mw8RrLFMLLgnRTQ0gK3XAO4kgR98orwY/view?usp=sharing>
- Variables de análisis de serie de tiempo: se emplearon las librerías *statsmodels.tsa.filters.hpfilter* y *statsmodels.tsa.seasonal_decompose*

2.2. Análisis univariado, bivariado y multivariado de BTC_ind_trend_ET

Fecha de creación: 13 de agosto de 2022

Notebook: Análisis univariado, bivariado y multivariado de BTC_ind_trend_ET.ipynb

Ubicación notebook: <https://drive.google.com/file/d/1aPunCHAsSelnI8CNng-tL4851HfBosh/view?usp=sharing>

Bases de datos analizadas:

- BTC_final.csv:
https://drive.google.com/file/d/1M38n5So0-6r_Q0vhs1i5nIW-cpigDlww/view?usp=sharing
- BTC_ind_trend_ET.csv:
https://drive.google.com/file/d/14lnPpEThIA6gVoc0CZPXVJ4PPZM0PpH_/view?usp=sharing
- Variables de análisis de serie de tiempo: se emplearon las librerías *statsmodels.tsa.filters.hpfilter* y *statsmodels.tsa.seasonal_decompose*

2.3. Análisis univariado, bivariado y multivariado de BTC_ind_mom_ET

Fecha de creación: 13 de agosto de 2022

Notebook: Análisis univariado, bivariado y multivariado de BTC_ind_mom_ET.ipynb

Ubicación notebook: <https://drive.google.com/file/d/1UjT9j0xA92oGFKuyt-BrfvjQ3v3Labyt/view?usp=sharing>

Bases de datos analizadas:

- BTC_final.csv:
https://drive.google.com/file/d/1M38n5So0-6r_Q0vhs1i5nIW-cpigDlww/view?usp=sharing
- BTC_ind_trend_ET.csv:
https://drive.google.com/file/d/14lnPpEThIA6gVoc0CZPXVJ4PPZM0PpH_/view?usp=sharing
- Variables de análisis de serie de tiempo: se emplearon las librerías *statsmodels.tsa.filters.hpfilter* y *statsmodels.tsa.seasonal_decompose*

2.4. Análisis univariado, bivariado y multivariado de BTC_patterns_indicator

Fecha de creación: 13 de agosto de 2022

Notebook: Análisis univariado, bivariado y multivariado de BTC_patterns_indicator.ipynb

Ubicación notebook: <https://drive.google.com/file/d/1khtle21f6eQ8nvyP-WpvV2bJB2jIXMrS/view?usp=sharing>

Bases de datos analizadas:

- BTC_final.csv:
https://drive.google.com/file/d/1M38n5So0-6r_Q0vhs1i5nIW-cpigDlww/view?usp=sharing
- BTC_ind_trend_ET.csv:
https://drive.google.com/file/d/14lnPpEThIA6gVoc0CZPXVJ4PPZM0PpH_/view?usp=sharing
- Variables de análisis de serie de tiempo: se emplearon las librerías *statsmodels.tsa.filters.hpfilter* y *statsmodels.tsa.seasonal_decompose*

2.5. Análisis univariado, bivariado y multivariado de BTC_multiple_indicators

Fecha de creación: 13 de agosto de 2022

Notebook: Análisis univariado, bivariado y multivariado de BTC_multiple_indicators.ipynb

Nota especial: es la combinación de diversas bases de datos

- BTC_ind_volat
- BTC_statistic_functions
- BTC_ind_cycle_indicators
- Variables de análisis de serie de tiempo (Seasonality, Trend, etc) que se aplicaron al momento del estudio

Ubicación notebook:

<https://drive.google.com/file/d/1F47XOkpuTaKqFhNSnxTNYVrwO6XcfNb/view?usp=sharing>

Bases de datos analizadas:

- BTC_ind_vol_ET.csv:
<https://drive.google.com/file/d/1I3UTC3uAkV00BKDoAWPn06mj2j-ZCf7c/view?usp=sharing>
- BTC_statistic_functions.csv:
<https://drive.google.com/file/d/1Zu4pmlYQKya6jrAHRR2ojjLREQ4MQywJ/view?usp=sharing>
- BTC_ind_cycle_indicators.csv:
<https://drive.google.com/file/d/133y-TBLhD6JFjXxA2xGvVnVTFY1hww06/view?usp=sharing>
- Variables de análisis de serie de tiempo: se emplearon las librerías *statsmodels.tsa.filters.hpfilter* y *statsmodels.tsa.seasonal_decompose*

2.6. Análisis univariado, bivariado y multivariado de BTC_US

Fecha de creación: 13 de agosto de 2022

Notebook: Análisis univariado, bivariado y multivariado de BTC_US.ipynb

Nota especial: es la combinación de diversas bases de datos

- BTC_Forex
- BTC_US_Finance
- Variables de análisis de serie de tiempo (Seasonality, Trend, etc) que se aplicaron al momento del estudio

Ubicación notebook:

<https://drive.google.com/file/d/1fS5cH0pPSsENjCo37HE5avU1bpMAxkrv/view?usp=sharing>

Bases de datos analizadas:

- BTC_US_Finance.csv:
<https://drive.google.com/file/d/1dg4hIUheDZok9NMXpnycEcfZE5HtwrDX/view?usp=sharing>
- BTC_Forex.csv:
https://drive.google.com/file/d/1I8kN3PwyVamLIHcXqz0XUFuWmwK13fF_/view?usp=sharing

- Variables de análisis de serie de tiempo: se emplearon las librerías *statsmodels.tsa.filters.hpfilter* y *statsmodels.tsa.seasonal_decompose*

3. Modelado

3.1.BTC_metals_fv

Fecha de creación: 27 de agosto de 2022

Notebook: BTC_metals_fv.ipynb

Nota especial: se manipuló BTC_metals.csv según las conclusiones del análisis del data set.

- Se eliminaron las variables 'Price_silver', 'Perc_diff_silver', 'Price_aluminium', 'Perc_diff_aluminium', 'Price_platinum', 'Perc_diff_platinum', 'Price_palladium', 'Perc_diff_palladium', 'Day', 'Month', 'Year'.
- Se añadieron las variables 'Trend' y 'Residuals'.

Ubicación notebook: <https://drive.google.com/file/d/1IU9G2Kptk1-QmijP6mHNERoICudwgPDg/view?usp=sharing>

Base de datos: https://drive.google.com/file/d/1IU7RCERkjbbrHv8OiQb2EI4PIpk_3BW0/view?usp=sharing

Mejor modelo: Decision Tree ()

- <https://drive.google.com/file/d/1EvIkQaLTYu5362Mw16rLjn7S4mnQeS8Z/view?usp=sharing>
- % de aciertos sobre el set de entrenamiento: 0.6736081597960051
- % de aciertos sobre el set de evaluación (accuracy): 0.6723259762308998
- Precision Score of the classifier is: 0.6911764705882353
- Recall Score of the classifier is: 0.632996632996633
- F1 Score of the classifier is: 0.6608084358523726
- AUC for our classifier is: 0.7045166274618329

Notebooks de modelos:

- *Algoritmo de clasificación (Decision Tree + BTC_metals_fv).ipynb:*
 - <https://drive.google.com/file/d/1EvIkQaLTYu5362Mw16rLjn7S4mnQeS8Z/view?usp=sharing>
- *Algoritmos de clasificación (Random forest + BTC_metals_fv).ipynb:*
 - <https://drive.google.com/file/d/1HY7oMJwti-1ScVB03ttw363PbZO16DKU/view?usp=sharing>
- *Algoritmos de clasificación (KNN + BTC_metals_fv).ipynb:*
 - <https://drive.google.com/file/d/15VpLgDRtGuLKRYlbdKfKWUtnOPlhM-P3q/view?usp=sharing>
- *Algoritmos de clasificación (Logistic Regression + BTC_metals_fv).ipynb:*
 - https://drive.google.com/file/d/1yWYMxh8zCi2Mp3epRFMc_Tci4L9cZQPb/view?usp=sharing
- *Algoritmos de clasificación (SVM + BTC_metals_fv).ipynb:*
 - <https://drive.google.com/file/d/1q7dDJBu1pYKwBnO4fTD8EJ26nSPb7n6o/view?usp=sharing>
- *Algoritmos de clasificación (Boosting models + BTC_metals_fv).ipynb:*
 - https://drive.google.com/file/d/13uUfBMDIQiWvFIH_jvzZzX3LYwXzMKKP/view?usp=sharing

3.2.BTC_ind_trend_ET_fv

Fecha de creación: 27 de agosto de 2022

Notebook: BTC_ind_trend_ET_fv.ipynb

Nota especial: se manipuló BTC_ind_trend_ET.csv según las conclusiones del análisis del data set.

- Se eliminaron las variables 'SMA_15d', 'SMA_30d', 'SMA_50d', 'SMA_100d', 'SMA_200d', 'EMA_15d', 'EMA_30d', 'EMA_50d', 'EMA_100d', 'EMA_200d', 'SMA_7d_ET', 'SMA_15d_ET', 'SMA_30d_ET', 'SMA_50d_ET', 'SMA_100d_ET', 'SMA_200d_ET', 'EMA_15d_ET', 'EMA_30d_ET', 'EMA_50d_ET', 'EMA_100d_ET', y 'EMA_200d_ET', 'Day', 'Month', 'Year'.

- Se añadieron las variables [Trend] y [Residuals].

Ubicación notebook: https://drive.google.com/file/d/1ySy-Tn15av_yOA6Te2on8F0Cs2YbpY23/view?usp=sharing

Base de datos: https://drive.google.com/file/d/14lnPpEThlA6gVoc0CZPXVJ4PPZM0PpH_/view?usp=sharing

Mejor modelo: Decision Tree

- <https://drive.google.com/file/d/1wVhE-awW1R7cxbhiLNMyz5ISTKUWYbt7/view?usp=sharing>
- % de aciertos sobre el set de entrenamiento: 0.6715856481481481
- % de aciertos sobre el set de evaluación (accuracy): 0.6701388888888888
- Precision Score of the classifier is: 0.699724517906336
- Recall Score of the classifier is: 0.5906976744186047
- F1 Score of the classifier is: 0.6406052963430012
- AUC for our classifier is: 0.7084583645911477

Notebooks de modelos:

- *Algoritmo de clasificación (Decision Tree + BTC_ind_trend_ET_fv).ipynb:*
 - <https://drive.google.com/file/d/1wVhE-awW1R7cxbhiLNMyz5ISTKUWYbt7/view?usp=sharing>
- *Algoritmos de clasificación (Random forest + BTC_ind_trend_ET_fv).ipynb:*
 - <https://drive.google.com/file/d/1UYNo9-nbtjwhx-jl--Mb-7EGz61okFUv/view?usp=sharing>
- *Algoritmos de clasificación (KNN + BTC_ind_trend_ET_fv).ipynb:*
 - https://drive.google.com/file/d/1HFiWT_vXYUGQaWb5-bMF5PLzzZOKi5Pj/view?usp=sharing
- *Algoritmos de clasificación (Logistic Regression + BTC_ind_trend_ET_fv).ipynb:*
 - https://drive.google.com/file/d/1_p0LmLBNkQrTybgKyWfzJdDaEGspFe3/view?usp=sharing
- *Algoritmos de clasificación (SVM + BTC_ind_trend_ET_fv).ipynb:*
 - <https://drive.google.com/file/d/1oKZQgnb4mSZvERKNVL-yHXt4ceFwsFC2/view?usp=sharing>
- *Algoritmos de clasificación (Boosting models + BTC_ind_trend_ET_fv).ipynb:*
 - <https://drive.google.com/file/d/1oBQCM4clft2NnR7fzLyz19a9b5n1jqzW/view?usp=sharing>

3.3.BTC_ind_mom_ET_fv

Fecha de creación: 27 de agosto de 2022

Notebook: BTC_ind_mom_ET_fv.ipynb

Nota especial: se manipuló BTC_ind_mom_ET.csv según las conclusiones del análisis del data set.

- Se eliminaron las variables 'Aroondown', 'Aroonup', 'STOCHRSI_fastd', 'STOCHRSI_fastd_ET', 'Day', 'Month', 'Year'.
- Se añadieron las variables [Trend] y [Residuals].

Ubicación notebook: https://drive.google.com/file/d/15HbibZt8znr_au_QVwkOgGQHkAth3wHJ/view?usp=sharing

Base de datos: <https://drive.google.com/file/d/1r970f10CJdhcCX8rt9Ld1pfbw6DjxLiB/view?usp=sharing>

Mejor modelo: DecisionTreeClassifier ()

- https://drive.google.com/file/d/1AkjI_QfMi_yG6rXYOih-mYAMmpGj6aSR/view?usp=sharing
- % de aciertos sobre el set de entrenamiento: 0.6720905553590378
- % de aciertos sobre el set de evaluación: 0.6710439921208142
- Precision Score of the classifier is: 0.6841294298921418
- Recall Score of the classifier is: 0.6
- F1 Score of the classifier is: 0.6393088552915767
- AUC for our classifier is: 0.7009733871802837

Notebooks de modelos:

- *Algoritmo de clasificación (Decision Tree + BTC_ind_mom_ET_fv).ipynb:*
 - https://drive.google.com/file/d/1AkjI_QfMi_yG6rXYOih-mYAMmpGj6aSR/view?usp=sharing

- *Algoritmos de clasificación (Random forest + BTC_ind_mom_ET_fv).ipynb:*
 - <https://drive.google.com/file/d/1F-IByjyh-OuOZW9IB9bAS2DDs95nxSM3/view?usp=sharing>
- *Algoritmos de clasificación (KNN + BTC_ind_mom_ET_fv).ipynb:*
 - <https://drive.google.com/file/d/1F-IByjyh-OuOZW9IB9bAS2DDs95nxSM3/view?usp=sharing>
- *Algoritmos de clasificación (Logistic Regression + BTC_ind_mom_ET_fv).ipynb:*
 - <https://drive.google.com/file/d/1F-IByjyh-OuOZW9IB9bAS2DDs95nxSM3/view?usp=sharing>
- *Algoritmos de clasificación (SVM + BTC_ind_mom_ET_fv).ipynb:*
 - <https://drive.google.com/file/d/1oKZQgnb4mSZvERKNVL-yHXt4ceFwsFC2/view?usp=sharing>

3.4.BTC_US_fv

Fecha de creación: 13 de agosto de 2022

Notebook: Análisis univariado, bivariado y multivariado de BTC_US_Finance.ipynb

Nota especial: se manipuló BTC_Forex.csv y BTC_US_Finance.csv según las conclusiones de los diversos tipos de análisis del data set.

- Se eliminaron las variables
'Percentage_diff_SP500', 'Percentage_diff_NASDAQ', 'Price_US_Bond', 'Percentage_diff_US_Bond', ,
'Price_EUR_USD', 'Percentage_diff_USD_JPY', 'Price_USD_CNY', 'Percentage_diff_USD_CNY', 'Year',
'Month', 'Day'.

Ubicación notebook: <https://drive.google.com/file/d/1fS5cH0pPSsENjCo37HE5avU1bpMAxkrv/view?usp=sharing>

Base de datos: <https://drive.google.com/file/d/1khOq6X7TVPah5Gh-HHCyEr382MLPINvS/view?usp=sharing>

Mejor modelo: Decision Tree ()

- https://drive.google.com/file/d/1PnHW89DzookC6syIwu4144JR_3toQZ6e/view?usp=sharing
- % de aciertos sobre el set de entrenamiento: 0.6894977168949772
- % de aciertos sobre el set de evaluación (accuracy): 0.6971080669710806
- Precision Score of the classifier is: 0.683076923076923
- Recall Score of the classifier is: 0.6981132075471698
- F1 Score of the classifier is: 0.6905132192846034
- AUC for our classifier is: 0.7540154635349993

Notebooks de modelos:

- *Algoritmo de clasificación (Decision Tree + BTC_US_fv).ipynb:*
 - https://drive.google.com/file/d/1PnHW89DzookC6syIwu4144JR_3toQZ6e/view?usp=sharing
- *Algoritmos de clasificación (Random forest + BTC_US_fv).ipynb:*
 - <https://drive.google.com/file/d/10X04KT9HuSxiCRZxpq60dOSp0SvJN8-O/view?usp=sharing>
- *Algoritmos de clasificación (KNN + BTC_US_fv).ipynb:*
 - <https://drive.google.com/file/d/1AKvrv5qKrvfnaYtRjUefZsY5JlyHisD/view?usp=sharing>
- *Algoritmos de clasificación (Logistic Regression + BTC_US_fv).ipynb:*
 - https://drive.google.com/file/d/11Yqd6_CRYNbQ7GPUTpo0_V7kHPGZbIBz/view?usp=sharing
- *Algoritmos de clasificación (SVM + BTC_US_fv).ipynb:*
 - https://drive.google.com/file/d/10uDOaQB6aiU5X8fzLyiTsl_T_MSbG0KhI/view?usp=sharing

3.5.BTC_multiple_indicators_fv

Fecha de creación: 13 de agosto de 2022

Notebook: Análisis univariado, bivariado y multivariado de BTC_multiple_indicators.ipynb

Nota especial: se manipuló BTC_ind_vol_ET.csv, BTC_ind_volat.csv, BTC_statistic_functions.csv y BTC_ind_cycle_indicators.csv según las conclusiones de los diversos tipos de análisis del data set.

- Se eliminaron las variables
'HT_DCPERIOD', 'HT_PHASOR_inphase', 'HT_PHASOR_quadrature', 'HT_SINE_integer', 'Beta'
'Pearson', 'TSF', 'Year', 'Month', 'Day'

Ubicación notebook: <https://drive.google.com/file/d/1F47XOkpuTaKqFhNSnxTNYVrwO6XcfNb/view?usp=sharing>

Base de datos: <https://drive.google.com/file/d/17dwqcGhoYCqLE5HWcu7TVT8lWB4lAjjT/view?usp=sharing>

Mejor modelo: HistGradientBoostingClassifier ()

- <https://drive.google.com/file/d/1WpDOY8KjQNBbhJCYCSzBQ7Lfe8J9nATN/view?usp=sharing>
- % de aciertos sobre el set de entrenamiento: 0.9485174323884001
- % de aciertos sobre el set de evaluación: 0.7120060790273556
- Accuracy of the classifier is: 0.7120060790273556
- Precision Score of the classifier is: 0.6859375
- Recall Score of the classifier is: 0.7115072933549432
- F1 Score of the classifier is: 0.6984884645982498
- AUC for our classifier is: 0.7969569864798752

Notebooks de modelos:

- *Algoritmo de clasificación (Decision Tree + BTC_multiple_indicators_fv).ipynb:*
 - <https://drive.google.com/file/d/1WpDOY8KjQNBbhJCYCSzBQ7Lfe8J9nATN/view?usp=sharing>
- *Algoritmos de clasificación (Random forest + BTC_multiple_indicators_fv).ipynb:*
 - https://drive.google.com/file/d/1c8WWoBa5_ZaE182XNhGh2UkrGXVIZdi7/view?usp=sharing
- *Algoritmos de clasificación (KNN + BTC_multiple_indicators_fv).ipynb:*
 - <https://drive.google.com/file/d/1dZt62XrHkJG2XOHMxJB2Wg4XG8FqvB9C/view?usp=sharing>
- *Algoritmos de clasificación (Logistic Regression + BTC_multiple_indicators_fv).ipynb:*
 - <https://drive.google.com/file/d/1QVjvRvFLxN7Y-SoUzkVdIIZ2Vtp8DVk/view?usp=sharing>
- *Algoritmos de clasificación (SVM + BTC_multiple_indicators_fv).ipynb:*
 - https://drive.google.com/file/d/1XWSX6V3YO2PMWY1hkG_mjLVMrS-A4Z37/view?usp=sharing

3.6.BTC_pattern_indicators_fv

Fecha de creación: 13 de agosto de 2022

Notebook: BTC_pattern_indicators_fv.ipynb

Nota especial: se manipuló BTC_pattern_indicator.csv según las conclusiones de los diversos tipos de análisis del data set.

- Se eliminaron las variables
'CDL2CROWS', 'CDL3BLACKCROWS', 'CDL3LINESTRIKE', 'CDL3OUTSIDE',
'CDL3STARSINSOUTH', 'CDL3WHITESOLDIERS', 'CDLABANDONEDBABY',
'CDLBREAKAWAY', 'CDLCONCEALBABYSWALL', 'CDLCOUNTERATTACK',
'CDLDARKCLOUDCOVER', 'CDLDOJI', 'CDLDRAGONFLYDOJI', 'CDLEVENINGDOJISTAR',
'CDLEVENINGSTAR', 'CDLGAPSIDESIDEWHITE', 'CDLGRAVESTONEDOJI', 'CDLHAMMER',
'CDLHANGINGMAN', 'CDLHIKKAKEMOD', 'CDLHOMINGPIGEON',
'CDLIDENTICAL3CROWS', 'CDLINNECK', 'CDLINVERTEDHAMMER', 'CDLKICKING',
'CDLKICKINGBYLENGTH', 'CDLLADDERBOTTOM', 'CDLLONGLEGGEDDOJI',
'CDLMATCHINGLOW', 'CDLMATHOLD', 'CDLMORNINGSTAR',
'CDLONNECK', 'CDLPIERCING', 'CDLRICKSHAWMAN', 'CDLRISEFALL3METHODS',
'CDLSTALLEDPATTERN', 'CDLSTICKSANDWICH', 'CDLTAKURI',
'CDLTASUKIGAP', 'CDLTHRUSTING', 'CDLTRISTAR', 'CDLUNIQUE3RIVER',

'CDLUPSIDEGAP2CROWS',

Ubicación notebook: <https://drive.google.com/file/d/1GpCszTpNtAQ36cCc4VZee9ctdcdX96jS/view?usp=sharing>

Base de datos: <https://drive.google.com/file/d/1n4UTRL4N6ooD9tLl-fSFhuWMVA-OXxRK/view?usp=sharing>

Mejor modelo: HistGradientBoostingClassifier ()

- https://drive.google.com/file/d/1DEy4xmwZaIpeIrrlY_Kk5bG82fg_cIEC/view?usp=sharing
- % de aciertos sobre el set de entrenamiento: 0.8963831867057673
- % de aciertos sobre el set de evaluación (accuracy): 0.6808510638297872
- Precision Score of the classifier is: 0.6513056835637481
- Recall Score of the classifier is: 0.6871961102106969
- F1 Score of the classifier is: 0.668769716088328
- AUC for our classifier is: 0.7587477364050983

Notebooks de modelos:

- Algoritmo de clasificación (Decision Tree + BTC_pattern_indicators_fv).ipynb:
 - https://drive.google.com/file/d/1DEy4xmwZaIpeIrrlY_Kk5bG82fg_cIEC/view?usp=sharing
- Algoritmos de clasificación (Random forest + BTC_pattern_indicators_fv).ipynb:
 - <https://drive.google.com/file/d/1yZjVypgfdlNt26bYRXZvFL7IVnI5u0nG/view?usp=sharing>
- Algoritmos de clasificación (KNN + BTC_pattern_indicators_fv).ipynb:
 - https://drive.google.com/file/d/1m-YfNDUL4XvvtZN0_zoK9mASp4hWz0H3/view?usp=sharing
- Algoritmos de clasificación (Logistic Regression + BTC_pattern_indicators_fv).ipynb:
 - https://drive.google.com/file/d/1uzf6wJR0c7uwcJiFx5zrIhKOIrOvwq_7/view?usp=sharing
- Algoritmos de clasificación (SVM + BTC_pattern_indicators_fv).ipynb:
 - https://drive.google.com/file/d/1-DINfl_OD1Z-49GEtw-vqrPjyZn8pni-/view?usp=sharing

3.7.BTC_various_indicators_fv

Fecha de creación: 27 de agosto de 2022

Notebook: BTC_various_indicators_fv.ipynb

Nota especial: se manipuló BTC_metals_fv.csv, BTC_ind_trend_ET.csv, BTC_ind_mom_ET_fv.csv, BTC_multiple_indicators_fv.csv, y BTC_US_fv.csv.

- Se adoptaron las variables:
'Date', 'Price', 'Open', 'High', 'Low', 'Vol.', 'Percentage_diff', 'Target', 'Price_gold', 'Price_copper', 'Price_aluminium', 'Trend', 'Residuals', 'EMA_7d', 'HT_TRENDLINE', 'BBANDS_upper_band', 'BBANDS_middle_band', 'BBANDS_lower_band', 'PSAR', 'EMA_7d_ET', 'HTTL_ET', 'PSAR_ET', 'ADX_PDM', 'ATR', 'Linearregression', 'Linearintercept', 'STDDEV', 'Price_SP500', 'Price_NASDAQ', 'Price_DJ'.

Ubicación notebook: <https://drive.google.com/file/d/1XtTGO29xfmqHkHSlb6izhRMggEwxVbYO/view?usp=sharing>

Base de datos: <https://drive.google.com/file/d/12zbCtfvIQsTQdpbrVnjlhKgsBc5yxzwn/view?usp=sharing>

Mejor modelo: Decision Tree ()

- <https://drive.google.com/file/d/1UwltvpsN46Ev9V-rtqn8WyzsfTRBwrz0/view?usp=sharing>
- % de aciertos sobre el set de entrenamiento: 0.6872521246458924
- % de aciertos sobre el set de evaluación: 0.6867572156196944
- Accuracy of the classifier is: 0.6867572156196944
- Precision Score of the classifier is: 0.6949924127465857
- Recall Score of the classifier is: 0.731629392971246
- F1 Score of the classifier is: 0.7128404669260701
- AUC for our classifier is: 0.7052093462054915

Notebooks de modelos:

- *Algoritmo de clasificación (Decision Tree + BTC_various_indicators_fv).ipynb:*
 - <https://drive.google.com/file/d/1UwltvpsN46Ev9V-rtqn8WyzsfTRBwrz0/view?usp=sharing>
- *Algoritmos de clasificación (Random forest + BTC_various_indicators_fv).ipynb:*
 - <https://drive.google.com/file/d/1UwltvpsN46Ev9V-rtqn8WyzsfTRBwrz0/view?usp=sharing>
- *Algoritmos de clasificación (KNN + BTC_various_indicators_fv).ipynb:*
 - <https://drive.google.com/file/d/1tn7iQQCTH-LFcNbYRgkvCVq2eP0myaP9/view?usp=sharing>
- *Algoritmos de clasificación (Logistic Regression + BTC_various_indicators_fv).ipynb:*
 - https://drive.google.com/file/d/1AYj479p2rIEgc4Iyp6xIelv-wXs_L9Gv/view?usp=sharing
- *Algoritmos de clasificación (SVM + BTC_various_indicators_fv).ipynb:*
 - <https://drive.google.com/file/d/1ZSEA8hKkTEa9lQ0G0uS7dYjCna7utIto/view?usp=sharing>
- *Algoritmos de clasificación (Boosting models + BTC_various_indicators_fv).ipynb:*
 - <https://drive.google.com/file/d/1hhh40WhJq9rupF2JQrzR13XzTZXRoezk/view?usp=sharing>

3.8. MEJOR MODELO (BTC_multiple_indicators_fv)

Fecha de creación: 10 de octubre de 2022

Notebook: BTC_multiple_indicators_fv.ipynb

Ubicación notebook: <https://drive.google.com/file/d/1XfTGO29xfmqHkHSLb6izhRMggEwxVbYO/view?usp=sharing>

Base de datos: <https://drive.google.com/file/d/17dwqcGhoYCqLE5HWcu7TVT8lWB4lAjT/view?usp=sharing>

Mejor modelo: HistGradientBoostingClassifier ()

- <https://drive.google.com/file/d/1-xpmxsBEuQhFRxMC3eQUX8tIgyDvhZjS/view?usp=sharing>
- Accuracy of the classifier is: 0.729483282674772
- Precision Score of the classifier is: 0.7115072933549432
- Recall Score of the classifier is: 0.7115072933549432
- F1 Score of the classifier is: 0.7115072933549431
- AUC for our classifier is: 0.8009775483847034

4. Librerías utilizadas

El entorno ACJUP fue creado específicamente para este proyecto y cuenta con las siguientes librerías:

```
# $ conda create --name <env> --file <this file> # platform: win-64
```

```
_py-xgboost-mutex=2.0=cpu_0
amadeus=6.0.1=pypi_0
argon2-ffi=21.3.0=pyhd3eb1b0_0
argon2-ffi-bindings=21.2.0=py39h2bbff1b_0
attrs=20.2.0=py_0
aws-c-common=0.4.57=ha925a31_1
aws-c-event-stream=0.1.6=hd77b12b_5
aws-checksums=0.1.9=ha925a31_0
aws-sdk-cpp=1.8.185=hd77b12b_0
backcall=0.2.0=py_0
blas=1.0=mkl
bleach=3.2.1=py_0
bottleneck=1.3.4=py39h080aecd_0
brotli=1.0.9=ha925a31_2
brotlipy=0.7.0=py39h2bbff1b_1003
bzip2=1.0.8=he774522_0
ca-certificates=2022.9.24=h5b45459_0
cachetools=5.2.0=pypi_0
certifi=2022.9.24=pyhd8ed1ab_0
cffi=1.15.0=py39h2bbff1b_1
cfitsio=3.470=h2bbff1b_7
charset-normalizer=2.0.4=pyhd3eb1b0_0
click=7.1.2=pyhd3eb1b0_0
click-plugins=1.1.1=pyhd3eb1b0_0
cligj=0.7.2=py39haa95532_0
colorama=0.4.4=py_0
cryptography=37.0.1=py39h21b164f_0
curl=7.82.0=h2bbff1b_0
cyclers=0.11.0=pyhd3eb1b0_0
daal4py=2021.5.0=py39h8cb3d55_0
dal=2021.5.0=haa95532_796
debugpy=1.5.1=py39hd77b12b_0
decorator=4.4.2=py_0
defusedxml=0.6.0=py_0
```

entrypoints=0.3=py39haa95532_0
 et_xmlfile=1.0.1=py_1001
 expat=2.4.4=h6c2663c_0
 fiona=1.8.13.post1=py39h758c064_0
 fonttools=4.25.0=pyhd3eb1b0_0
 freetype=2.10.4=hd328e21_0
 freexl=1.0.6=h2bbff1b_0
 gdal=3.0.2=py39hb978731_1
 geographiclib=1.52=pyhd8ed1ab_0
 geopandas=0.9.0=py_1
 geopandas-base=0.9.0=py_1
 geopy=2.2.0=pyhd8ed1ab_0
 geos=3.8.0=h33f27b4_0
 geotiff=1.6.0=h5770a2b_0
 google-api-core=2.8.2=pypi_0
 google-api-python-client=2.54.0=pypi_0
 google-auth=2.9.1=pypi_0
 google-auth-httpplib2=0.1.0=pypi_0
 google-auth-oauthlib=0.5.2=pypi_0
 googleapis-common-protos=1.56.4=pypi_0
 hdbscan=0.8.28=py39h5d4886f_1
 hdf4=4.2.13=h712560f_2
 hdf5=1.10.6=h7ebc959_0
 htmlmin=0.1.12=pypi_0
 httpplib2=0.20.4=pypi_0
 icc_rt=2019.0.0=h0cc432a_1
 icu=58.2=vc14hc45fdbb_0
 idna=3.3=pyhd3eb1b0_0
 imagehash=4.2.1=pypi_0
 imbalanced-learn=0.9.1=pyhd8ed1ab_1
 importlib-metadata=2.0.0=py_1
 importlib-metadata=2.0.0=1
 intel-openmp=2021.4.0=haa95532_3556
 ipykernel=6.4.1=py39haa95532_1
 ipython=7.31.1=py39haa95532_0
 ipython_genutils=0.2.0=pyhd3eb1b0_1
 ipywidgets=7.5.1=py_1
 jdcals=1.4.1=py_0
 jedi=0.18.1=py39haa95532_1
 jinja2=2.11.2=py_0
 joblib=1.1.0=pyhd8ed1ab_0
 jpeg=9b=vc14h4d7706e_1
 jsonschema=3.2.0=py_2
 jupyter=1.0.0=py39haa95532_7
 jupyter-http-over-ws=0.0.8=pypi_0
 jupyter_client=6.1.7=py_0
 jupyter_console=6.2.0=py_0
 jupyter_core=4.9.2=py39haa95532_0
 kealib=1.4.14=hde4a422_1
 kiwisolver=1.3.2=py39hd77b12b_0
 krb5=1.19.2=h5b6d351_0
 libcurl=7.82.0=h86230a5_0
 libgdal=3.0.2=ha1b3edf_1
 libiconv=1.16=h2bbff1b_2
 libnetcdf=4.6.1=hf59b723_4
 libpng=1.6.37=h2a8f88b_0
 libpq=12.9=hb652d5d_1
 libspatialindex=1.9.3=h6c2663c_0
 libspatialite=4.3.0a=h14feca5_20
 libssh2=1.10.0=hcd4344a_0
 libta-lib=0.4.0=he774522_0
 libtiff=4.2.0=hd0e1b90_0
 libwebp=1.2.2=h2bbff1b_0
 libxgboost=1.5.0=hd77b12b_2
 libxml2=2.9.12=h0ad7f3c_0
 lightgbm=3.3.2=py39h415ef7b_0
 lz4-c=1.9.3=h2bbff1b_1
 m2w64-expat=2.1.1=2
 m2w64-gcc-libgfortran=5.3.0=6
 m2w64-gcc-libstdc++=5.3.0=7
 m2w64-gcc-libstdc++-core=5.3.0=7
 m2w64-gettext=0.19.7=2
 m2w64-gmp=6.1.0=2
 m2w64-libiconv=1.14=6
 m2w64-libwinpthread-git=5.0.0.4634.697f757=2
 m2w64-xz=5.2.2=2
 mapclassify=2.4.3=pyhd3eb1b0_0
 markupsafe=2.0.1=py39h2bbff1b_0
 matplotlib=3.5.1=py39haa95532_1
 matplotlib-base=3.5.1=py39hd77b12b_1
 matplotlib-inline=0.1.2=pyhd3eb1b0_2
 missingno=0.5.1=pypi_0
 mistune=0.8.4=py39h2bbff1b_1000
 mkl=2021.4.0=haa95532_640
 mkl-service=2.4.0=py39h2bbff1b_0
 mkl_fft=1.3.1=py39h277e83a_0
 mkl_random=1.2.2=py39hf11a4ad_0
 msys2-conda-epoch=20160418=1
 multimethod=1.8=pypi_0
 munch=2.5.0=pyhd3eb1b0_0
 munkres=1.1.4=py_0
 nbconvert=5.5.0=py_0
 nbformat=5.0.8=py_0
 networkx=2.8.5=pypi_0
 notebook=6.4.5=py39haa95532_0
 numexpr=2.8.1=py39hb80d3ca_0
 numpy=1.21.5=py39h7a0a035_1
 numpy-base=1.21.5=py39hca35cd5_1
 oauth2client=4.1.3=pypi_0
 oauthlib=3.2.0=pypi_0
 openjpeg=2.4.0=h4fc8c34_0
 openpyxl=3.0.5=py_0
 openssl=1.1.1q=h8ffe710_0
 packaging=20.4=py_0
 pandas=1.4.2=py39hd77b12b_0
 pandas-profiling=3.2.0=pypi_0
 pandoc=2.11=h9490d1a_0
 pandocfilters=1.5.0=pyhd3eb1b0_0
 parso=0.8.0=py_0
 patsy=0.5.2=py39haa95532_1

phik=0.12.2=pypi_0
 pickleshare=0.7.5=pyhd3eb1b0_1003
 pillow=9.0.1=py39hdc2b20a_0
 pip=21.2.4=py39haa95532_0
 plotly=5.7.0=py_0
 plotly_express=0.4.1=py_0
 postgresql=12.9=hb652d5d_1
 proj=6.2.1=h3758d61_0
 prometheus_client=0.8.0=py_0
 prompt-toolkit=3.0.8=py_0
 prompt_toolkit=3.0.8=0
 protobuf=4.21.3=pypi_0
 py-xgboost=1.5.0=py39haa95532_2
 pyasn1=0.4.8=pypi_0
 pyasn1-modules=0.2.8=pypi_0
 pycparser=2.20=py_2
 pydantic=1.9.1=pypi_0
 pydrive=1.3.1=pypi_0
 pygments=2.7.1=py_0
 pyopenssl=22.0.0=pyhd3eb1b0_0
 pyparsing=2.4.7=py_0
 pyproj=2.6.1.post1=py39h593ac45_1
 pyqt=5.9.2=py39hd77b12b_6
 pyrsistent=0.18.0=py39h196d8e1_0
 pysocks=1.7.1=py39haa95532_0
 python=3.9.12=h6244533_0
 python-dateutil=2.8.1=py_0
 python_abi=3.9=2_cp39
 pytz=2021.3=pyhd3eb1b0_0
 pywavelets=1.3.0=pypi_0
 pywin32=302=py39h827c3e9_1
 pywinpty=2.0.2=py39h5da7b33_0
 pyyaml=6.0=pypi_0
 pyzmq=22.3.0=py39hd77b12b_2
 qt=5.9.7=vc14h73c81de_0
 qtconsole=4.7.7=py_0
 qtpy=1.9.0=py_0
 requests=2.27.1=pyhd3eb1b0_0
 requests-oauthlib=1.3.1=pypi_0
 rsa=4.9=pypi_0
 rtree=0.9.7=py39h2eaa2aa_1
 scikit-learn=1.1.1=py39hd77b12b_0
 scikit-learn-intelex=2021.5.0=py39haa95532_0
 scikit-surprise=1.1.1=py39h5d4886f_2
 scipy=1.7.3=py39h0a974cb_0
 seaborn=0.11.2=pyhd3eb1b0_0
 send2trash=1.8.0=pyhd3eb1b0_1
 setuptools=61.2.0=py39haa95532_0
 shapely=1.7.1=py39h06580b3_0
 sip=4.19.13=py39hd77b12b_0
 six=1.12.0=pypi_0
 sqlite=3.38.2=h2bbff1b_0
 statsmodels=0.12.2=py39h2bbff1b_0
 ta-lib=0.4.19=py39h5d4886f_4
 tangled-up-in-unicode=0.2.0=pypi_0
 tbb=2021.5.0=h59b6b97_0
 tenacity=8.0.1=py39haa95532_0
 terminado=0.13.1=py39haa95532_0
 testpath=0.4.4=py_0
 threadpoolctl=3.1.0=pypi_0
 tiledb=2.2.9=hf7ce2e6_0
 tk=8.6.11=h2bbff1b_0
 tornado=6.2=pypi_0
 tqdm=4.64.0=pypi_0
 traitlets=5.0.5=py_0
 typing-extensions=3.7.4.3=0
 typing_extensions=3.7.4.3=py_0
 tzdata=2022a=hda174b7_0
 uritemplate=4.1.1=pypi_0
 urllib3=1.26.9=py39haa95532_0
 vc=14.2=h21ff451_1
 visions=0.7.5=pypi_0
 vs2015_runtime=14.27.29016=h5e58377_2
 wcwidth=0.2.5=py_0
 webencodings=0.5.1=py39haa95532_1
 wheel=0.37.1=pyhd3eb1b0_0
 widgetsnbextension=3.5.2=py39haa95532_0
 win_inet_pton=1.1.0=py39haa95532_0
 wincertstore=0.2=py39haa95532_2
 winpty=0.4.3=4
 xerces-c=3.2.3=ha925a31_0
 xgboost=1.5.0=py39haa95532_2
 xz=5.2.5=h62dcd97_0
 zipp=3.3.1=py_0
 zlib=1.2.11=vc14h1cdd9ab_1
 zstd=1.4.9=h19a0ad4_0