



Analyzing Forex Market Trends and L/S strategy on GBP/USD and USD/JPY

Data Engineering Final project

MG-GY 8411

Instructor: Carlos De Oliveira

Xinjing Guo

May 3, 2024

Data Collection and Procession

1. Data information

Currency pair: EUR/USD, GBP/CHF, USD/CAD, EUR/CHF, EUR/CAD, GBP/EUR, GBP/USD, GBP/CAD , USD/CHF, USD/JPY

Data Fetching Frequency: Fetch every second, and set 6 minutes as a loop, repeat 50 loops (5 hours) to classify 8 currency pairs. Based on GBP/USD and USD/JPY, simulate long and short trends to evaluate the performance and profitability in next 3 hours.

1. Data Storage

Collections: Each currency pair has its own collection in MongoDB.

Classification currency pairs in `classification_results.csv`

1. Data Analysis

Using function 'insert_forex_data_and_calculate_stats' get Mean, Max, Min, and Volatility

Using 'calculate_keltner_bands' and 'calc_fd' function get upper and lower Keltner bands and Fractal Dimension

Regression and classification part:

Using 'create_synth_pair' function get average value of mean, max, min, vol and fd for 40th and 50th loop

Using 'run_pycaret_experiment' function get best model at 40th and 50th loop

Using 'calculate_mae_for_currency_pairs' get the MAE

Long/Short strategy on GBP/USD and USD/JPY:

Using 'analyze_slope_for_pairs' function get the slope of GBP/USD and USD/JPY

Use real-time data from Polygon and PyCaret for the regression & classification

1. Load and AggregateData:
After fetch 40th and 50th loops of real time currency pairs, get latest 20 data points.
Get the average value of mean, max, min, vol and fd.
Store into MongoDB 'synth_pair_stats'
2. Model Building with PyCaret:
Using the aggregate data, set target is mean value.
Automatically filter the best models with PyCaret.
3. Get value MAE to classify CPs:
Use selected model get predicted mean value.
Using predicted mean value and real mean value get MAE.
Sort the mae, the three smallest are forecastable. Next 2 currency pairs are undefined. Others are non forecastable.
4. Conclusion for classification:
GBP/USD: forecastable
USD/JPY: non forecastable

Cycle	Pair	Classification	MAE
40	GBPUSD	FORECASTABLE	0.027598971756395564
40	GBPEUR	FORECASTABLE	0.061271485813394214
40	GBPCHF	FORECASTABLE	0.09114846
40	USDCAD	UNDEFINED	0.14114206890515274
40	EURUSD	UNDEFINED	0.14972953068148323
40	EURCAD	UNDEFINED	0.24612382383220482
40	EURCHF	UNDEFINED	0.2520184351143727
40	USDJPY	NON FORECASTABLE	0.30142017556344153
40	USDCHF	NON FORECASTABLE	0.32154039025652403
40	GBPCAD	NON FORECASTABLE	0.4895352646328501
50	GBPUSD	FORECASTABLE	0.027445113451778337
50	GBPEUR	FORECASTABLE	0.061141758
50	GBPCHF	FORECASTABLE	0.090748229
50	USDCAD	UNDEFINED	0.14118767626055692
50	EURUSD	UNDEFINED	0.1499936626732511
50	EURCAD	UNDEFINED	0.2458434357540556
50	EURCHF	UNDEFINED	0.2518211856937091
50	USDJPY	NON FORECASTABLE	0.30175565244465974
50	USDCHF	NON FORECASTABLE	0.3211552075205593
50	GBPCAD	NON FORECASTABLE	0.489438513

L/S strategy on GBPUSD and USDJPY

1. Choose the Long and Short CPs:

Using last 20 data points (2 hours) and get slope based on PyCaret

Slope for GBPUSD: -0.000014, Slope for USDJPY: -0.00027

Long position: GBPUSD, Short position: USDJPY

2. Adjusted initial trade to balance the ratio to 1:

\$50 for long and short, divide by 100 on USDJPY to simplify

Initial rate for GBPUSD: 1.255010, USDJPY: 152.9180

Long 39.840320 units of GBPUSD and short 2683.4899 units of USDJPY

3. Hour 6, 7 and 8 ratio:

Hour 6: GBPUSD to USDJPY ratio: 1.2187216742, Profit

Hour 7: GBPUSD to USDJPY ratio: 1.2169976626, Profit

Hour 8: GBPUSD to USDJPY ratio: 1.2169976626, Profit

	initial	6th hour	7th hour	8th hour	Total
P&L	0	0.2187%	0.2169%	0.2169%	0.6525%
P&L			0.2169%	0.2169%	0.4338%
P&L				0.2169%	0.2169%
Total		0.2187%	0.4338%	0.65%	1.3032%

	MAE	MSE	RMSE	R2	RMSLE	MAPE
Fold						
0	0.0001	0.0	0.0001	-0.3808	0.0001	0.0001
1	0.0004	0.0	0.0004	-4539.7637	0.0002	0.0003
2	0.0001	0.0	0.0002	-6.7930	0.0001	0.0001
3	0.0002	0.0	0.0002	-6.3499	0.0001	0.0002
4	0.0005	0.0	0.0005	NaN	0.0002	0.0004
5	0.0001	0.0	0.0001	NaN	0.0001	0.0001
6	0.0003	0.0	0.0003	NaN	0.0001	0.0002
7	0.0001	0.0	0.0001	NaN	0.0001	0.0001
8	0.0002	0.0	0.0002	NaN	0.0001	0.0001
9	0.0004	0.0	0.0004	NaN	0.0002	0.0003
Mean	0.0002	0.0	0.0002	NaN	0.0001	0.0002
Std	0.0001	0.0	0.0001	NaN	0.0001	0.0001
Slope for GBPUSD: -0.00001463294029235836						

	MAE	MSE	RMSE	R2	RMSLE	MAPE
Fold						
0	0.0004	0.0	0.0004	-3.5785	0.0002	0.0003
1	0.0006	0.0	0.0006	-17250.6952	0.0002	0.0004
2	0.0007	0.0	0.0009	-1.0296	0.0004	0.0004
3	0.0002	0.0	0.0002	-0.0103	0.0001	0.0002
4	0.0002	0.0	0.0002	NaN	0.0001	0.0001
5	0.0004	0.0	0.0004	NaN	0.0002	0.0003
6	0.0003	0.0	0.0003	NaN	0.0001	0.0002
7	0.0002	0.0	0.0002	NaN	0.0001	0.0001
8	0.0003	0.0	0.0003	NaN	0.0001	0.0002
9	0.0004	0.0	0.0004	NaN	0.0001	0.0002
Mean	0.0004	0.0	0.0004	NaN	0.0002	0.0002
Std	0.0002	0.0	0.0002	NaN	0.0001	0.0001
Slope for USDJPY: -0.00027282039324442526						
Long position: GBPUSD, Short position: USDJPY						