Kelly's ETL project for financial data

ETL   Machine learning

ETL pipeline

Enter ticker symbols (separated by space). FOr e.g: AAPL GOOG NVIA

AAPL MSFT GOOG

Start date

2023/01/01

End date

2025/08/11

Select chunk size

200   −   +

Run ETL

Expand to see logs		⌃

	Level	Message
0	INFO	Starting ETL process
1	INFO	Completed fetching data for AAPL
2	INFO	Completed fetching data for MSFT
3	INFO	Completed fetching data for GOOG
4	INFO	Completed feature engineering for AAPL
5	INFO	Completed feature engineering for MSFT
6	INFO	Completed feature engineering for GOOG
7	INFO	Loading strategy data for AAPL, 122 records
8	INFO	Batch 1 loaded: 122 rows, total 122, elapsed 0.32s
9	INFO	Finished loading 122 rows in 1 batches, total time 0.32s
10	INFO	Loading strategy data for AAPL, 122 records
11	INFO	Batch 1 loaded: 122 rows, total 122, elapsed 0.3s
12	INFO	Finished loading 122 rows in 1 batches, total time 0.3s
13	INFO	Loading strategy data for MSFT, 122 records
14	INFO	Batch 1 loaded: 122 rows, total 122, elapsed 0.38s
15	INFO	Finished loading 122 rows in 1 batches, total time 0.38s
16	INFO	Loading strategy data for MSFT, 122 records
17	INFO	Batch 1 loaded: 122 rows, total 122, elapsed 0.46s
18	INFO	Finished loading 122 rows in 1 batches, total time 0.47s
19	INFO	Loading strategy data for GOOG, 122 records
20	INFO	Batch 1 loaded: 122 rows, total 122, elapsed 0.55s
21	INFO	Finished loading 122 rows in 1 batches, total time 0.55s
22	INFO	Loading strategy data for GOOG, 122 records
23	INFO	Batch 1 loaded: 122 rows, total 122, elapsed 0.39s
24	INFO	Finished loading 122 rows in 1 batches, total time 0.39s

	Level	Message
25	INFO	Loading data for AAPL, 122 records
26	INFO	Batch 1 loaded: 122 rows, total 122, elapsed 0.35s
27	INFO	Finished loading 122 rows in 1 batches, total time 0.35s
28	INFO	Loading data for MSFT, 122 records
29	INFO	Batch 1 loaded: 122 rows, total 122, elapsed 0.23s
30	INFO	Finished loading 122 rows in 1 batches, total time 0.23s
31	INFO	Loading data for GOOG, 122 records
32	INFO	Batch 1 loaded: 122 rows, total 122, elapsed 0.26s
33	INFO	Finished loading 122 rows in 1 batches, total time 0.26s
34	INFO	ETL process completed
35	INFO	Starting ETL process
36	INFO	Completed fetching data for AAPL
37	INFO	Completed fetching data for MSFT
38	INFO	Completed fetching data for GOOG
39	INFO	Completed feature engineering for AAPL
40	INFO	Completed feature engineering for MSFT
41	INFO	Completed feature engineering for GOOG
42	INFO	Loading strategy data for AAPL, 624 records
43	INFO	Batch 1 loaded: 200 rows, total 200, elapsed 0.34s
44	INFO	Batch 2 loaded: 200 rows, total 400, elapsed 0.8s
45	INFO	Batch 3 loaded: 200 rows, total 600, elapsed 1.18s
46	INFO	Batch 4 loaded: 24 rows, total 624, elapsed 1.24s
47	INFO	Finished loading 624 rows in 4 batches, total time 1.24s
48	INFO	Loading strategy data for AAPL, 624 records
49	INFO	Batch 1 loaded: 200 rows, total 200, elapsed 0.47s
50	INFO	Batch 2 loaded: 200 rows, total 400, elapsed 0.79s
51	INFO	Batch 3 loaded: 200 rows, total 600, elapsed 1.1s
52	INFO	Batch 4 loaded: 24 rows, total 624, elapsed 1.15s
53	INFO	Finished loading 624 rows in 4 batches, total time 1.15s
54	INFO	Loading strategy data for MSFT, 624 records
55	INFO	Batch 1 loaded: 200 rows, total 200, elapsed 0.75s
56	INFO	Batch 2 loaded: 200 rows, total 400, elapsed 1.22s
57	INFO	Batch 3 loaded: 200 rows, total 600, elapsed 1.52s
58	INFO	Batch 4 loaded: 24 rows, total 624, elapsed 1.62s
59	INFO	Finished loading 624 rows in 4 batches, total time 1.62s
60	INFO	Loading strategy data for MSFT, 624 records
61	INFO	Batch 1 loaded: 200 rows, total 200, elapsed 0.35s
62	INFO	Batch 2 loaded: 200 rows, total 400, elapsed 0.76s
63	INFO	Batch 3 loaded: 200 rows, total 600, elapsed 1.06s
64	INFO	Batch 4 loaded: 24 rows, total 624, elapsed 1.11s
65	INFO	Finished loading 624 rows in 4 batches, total time 1.11s
66	INFO	Loading strategy data for GOOG, 624 records
67	INFO	Batch 1 loaded: 200 rows, total 200, elapsed 0.45s
68	INFO	Batch 2 loaded: 200 rows, total 400, elapsed 0.73s
69	INFO	Batch 3 loaded: 200 rows, total 600, elapsed 1.06s
70	INFO	Batch 4 loaded: 24 rows, total 624, elapsed 1.1s

	Level	Message
71	INFO	Finished loading 624 rows in 4 batches, total time 1.1s
72	INFO	Loading strategy data for GOOG, 624 records
73	INFO	Batch 1 loaded: 200 rows, total 200, elapsed 0.37s
74	INFO	Batch 2 loaded: 200 rows, total 400, elapsed 0.77s
75	INFO	Batch 3 loaded: 200 rows, total 600, elapsed 1.19s
76	INFO	Batch 4 loaded: 24 rows, total 624, elapsed 1.24s
77	INFO	Finished loading 624 rows in 4 batches, total time 1.24s
78	INFO	Loading data for AAPL, 624 records
79	INFO	Batch 1 loaded: 200 rows, total 200, elapsed 0.36s
80	INFO	Batch 2 loaded: 200 rows, total 400, elapsed 0.71s
81	INFO	Batch 3 loaded: 200 rows, total 600, elapsed 1.06s
82	INFO	Batch 4 loaded: 24 rows, total 624, elapsed 1.1s
83	INFO	Finished loading 624 rows in 4 batches, total time 1.1s
84	INFO	Loading data for MSFT, 624 records
85	INFO	Batch 1 loaded: 200 rows, total 200, elapsed 0.22s
86	INFO	Batch 2 loaded: 200 rows, total 400, elapsed 0.42s
87	INFO	Batch 3 loaded: 200 rows, total 600, elapsed 0.65s
88	INFO	Batch 4 loaded: 24 rows, total 624, elapsed 0.69s
89	INFO	Finished loading 624 rows in 4 batches, total time 0.69s
90	INFO	Loading data for GOOG, 624 records
91	INFO	Batch 1 loaded: 200 rows, total 200, elapsed 0.33s
92	INFO	Batch 2 loaded: 200 rows, total 400, elapsed 0.57s
93	INFO	Batch 3 loaded: 200 rows, total 600, elapsed 0.88s
94	INFO	Batch 4 loaded: 24 rows, total 624, elapsed 0.93s
95	INFO	Finished loading 624 rows in 4 batches, total time 0.93s
96	INFO	ETL process completed

Features table preview

Number of feature columns retained for AAPL: 20

	ticker	date	0	4	5	7	8	12	23	24	25	34	35	36	37
								empty							

Strategy table (triple barrier labelling) preview

	ticker	date	close	vert_barriers	positions
493	AAPL	2023-01-03 00:00:00	123.4706	2023-01-04 00:00:00	
494	AAPL	2023-01-04 00:00:00	124.7441	2023-01-05 00:00:00	
495	AAPL	2023-01-05 00:00:00	123.4213	2023-01-06 00:00:00	
496	AAPL	2023-01-06 00:00:00	127.9624	2023-01-09 00:00:00	
497	AAPL	2023-01-09 00:00:00	128.4857	2023-01-10 00:00:00	
498	AAPL	2023-01-10 00:00:00	129.0582	2023-01-11 00:00:00	
499	AAPL	2023-01-11 00:00:00	131.7829	2023-01-12 00:00:00	
500	AAPL	2023-01-12 00:00:00	131.704	2023-01-13 00:00:00	
501	AAPL	2023-01-13 00:00:00	133.0367	2023-01-17 00:00:00	
502	AAPL	2023-01-17 00:00:00	134.2016	2023-01-18 00:00:00	

PnL

Target value counts (profitable vs. losing outcomes)

target	count
-1.0000	338
1.0000	311
0.0000	2

Sharpe ratio (annualised): 0.3417