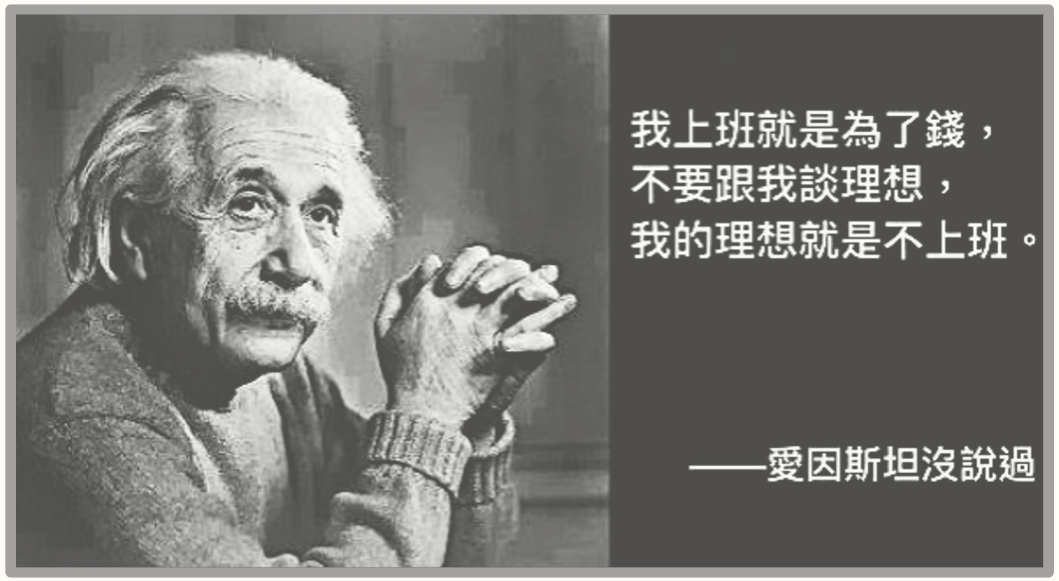
Stock price prediction by using ML algorithms

(symbol: VOO)

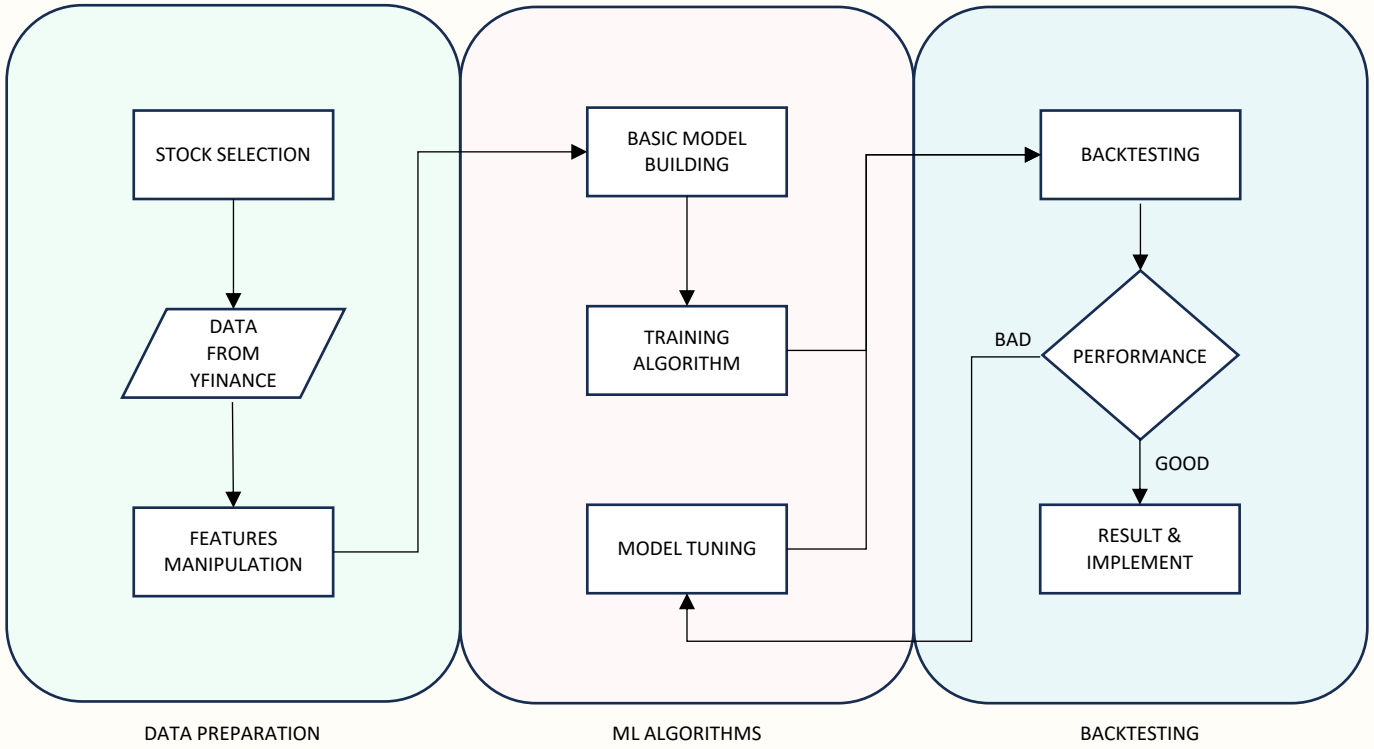
Jin He 何建進

Photo by @riverhetw

Motivation



Overview


STOCK SELECTION
DATA FROM YFINANCE
FEATURES MANIPULATION
BASIC MODEL BUILDING
TRAINING ALGORITHM
MODEL TUNING
BACKTESTING
BAD
PERFORMANCE
GOOD
RESULT & IMPLEMENT
DATA PREPARATION
ML ALGORITHMS
BACKTESTING

Features

- OHLC + VOLUME
- [x1,x2,x3,…] days MA ratio + TREND
  TREND: up/down percentage during a period, e.g. only 3 up over the past 10 days scores 3/10
- [x1,x2,x3,…] days RSI, >50 means gaining more during a period

$$RSI_{\text{step one}} = 100 - \left[\frac{100}{1+\frac{\text{Average gain}}{\text{Average loss}}}\right]$$

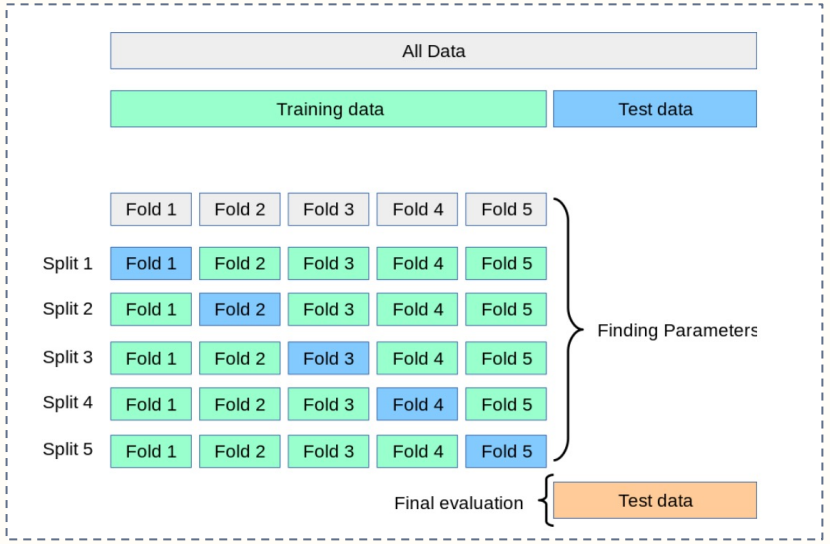
- Deterministic features:
  Time series with 1, 2, 3,…to len(); or squared time series with 1^2, 2^2, 3^2,…to $\text{len}()^2$
- CPI / FED FUND RATES (optional due to almost a month time delay)

Basic model building

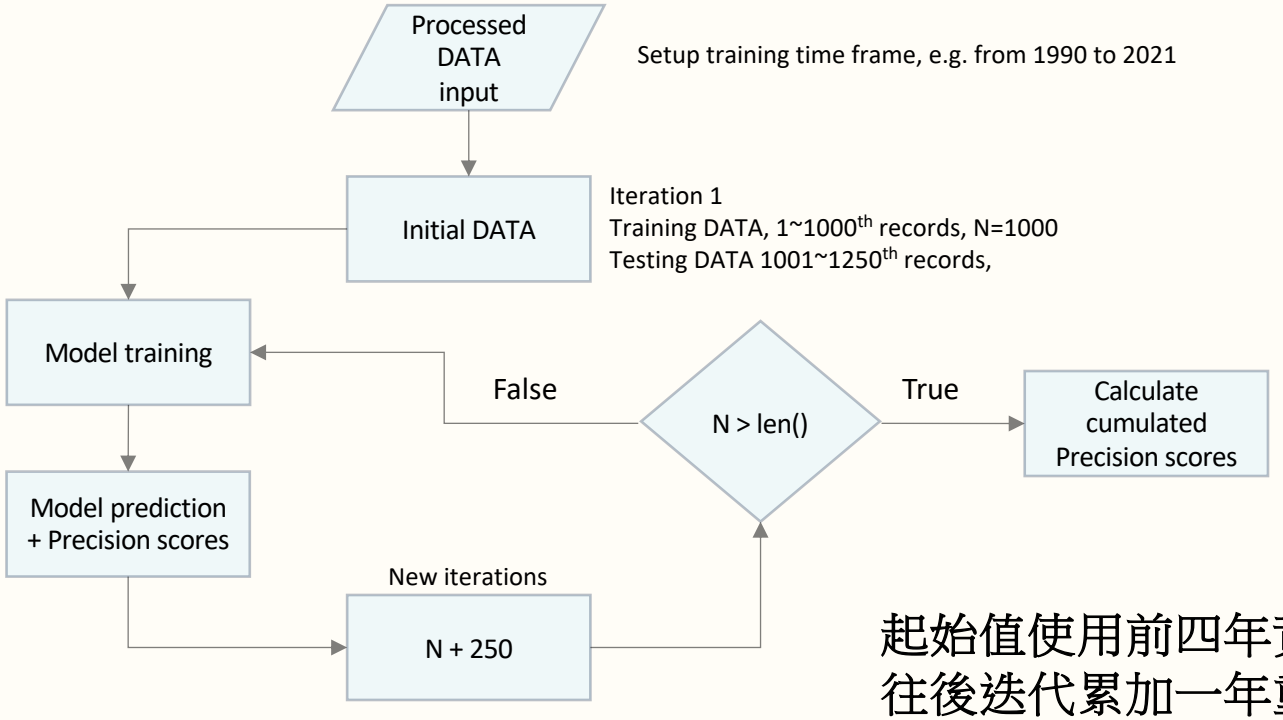
- Using Sklearn cross_val_score, CV groups = 5

	rf	gnb	lr	dt	knn	xgb	voting_clf_soft	voting_clf_hard
baseline	0.475789	0.528353	0.541199	0.48188	0.512018	0.481238	0.505275	0.5197

Logistic regression shows the best score from others



Training algorithm



起始值使用前四年資料訓練, 預測第五年
往後迭代累加一年重新訓練, 預測後一年
使用累計準確率評估模型表現

Results

	RandomForestClassifier	GaussianNB	LogisticRegression	DecisionTreeClassifier	KNeighborsClassifier	XGBClassifier	VotingClassifier_soft	VotingClassifier_hard
cumulated_precision	0.61	0.56	0.55	0.58	0.56	0.58	0.59	0.56

Randomforest shows the best score from others

Backtesting (using VectorBT)

- Model training data from 1990-01-01 ~ 2021-12-31
- VOO backtesting period: 2022-01-01 ~ 2023-09-05

	RandomForestClassifi	GaussianNB()	LogisticRegression()	DecisionTreeClassifi	KNeighborsClassifier	XGBClassifier(base_s	VotingClassifier(est
precision	0.5152	0.4881	0.4881	0.4716	0.4805	0.5031	0.4987
benchmark(%)	-3.6500	-3.6500	-3.6500	-3.6500	-3.6500	-3.6500	-3.6500
pred_return(%)	15.7100	-3.7400	-3.7400	-20.8100	-34.4800	9.0400	-3.2100

Randomforest shows the best ROI from others

Note 1: benchmark = last close / first close (%)
Note 2: pred_return is the ROI by following the models prediction, transaction fee 0.001425, initial cash $10,000

Performance (Randomforest classification)

Start	2022-01-03 05:00:00+00:00
End	2023-09-05 04:00:00+00:00
Period	420
Start Value	10000.0
End Value	11570.56069
Total Return [%]	15.705607
Benchmark Return [%]	-3.652898
Max Gross Exposure [%]	99.962871
Total Fees Paid	1223.355814
Max Drawdown [%]	14.855432
Max Drawdown Duration	211.0
Total Trades	43
Total Closed Trades	42
Total Open Trades	1
Open Trade PnL	301.777065
Win Rate [%]	61.904762
Best Trade [%]	5.872304
Worst Trade [%]	-6.025643
Avg Winning Trade [%]	1.830318
Avg Losing Trade [%]	-2.116501
Avg Winning Trade Duration	6.5
Avg Losing Trade Duration	7.5
Profit Factor	1.377333
Expectancy	30.209134

Summary

	Order Id	Column	Timestamp	Size	Price	Fees	Side
0	0	0	2022-01-03 05:00:00+00:00	23.0	428.554657	14.292298	Buy
1	1	0	2022-01-05 05:00:00+00:00	23.0	420.173828	14.012797	Sell
2	2	0	2022-01-10 05:00:00+00:00	23.0	417.607880	13.927223	Buy
3	3	0	2022-01-13 05:00:00+00:00	23.0	416.700500	13.896962	Sell
4	4	0	2022-01-14 05:00:00+00:00	23.0	416.827332	13.901192	Buy
...	...	...	...	...	...	...	...
80	80	0	2023-08-03 04:00:00+00:00	27.0	412.399994	16.145460	Buy
81	81	0	2023-08-10 04:00:00+00:00	27.0	409.790009	16.043279	Sell
82	82	0	2023-08-14 04:00:00+00:00	27.0	411.720001	16.118838	Buy
83	83	0	2023-08-23 04:00:00+00:00	27.0	406.950012	15.932093	Sell
84	84	0	2023-08-24 04:00:00+00:00	28.0	401.540009	16.302524	Buy

Trading records

Be able to avoid the down trend when in a bear market during 2022

Be able to gain the market as well when in a bull market in 2023

Predict if VOO will go up on 9/8

Input:
9/7 VOO DATA

Date	Open	High	Low	Close	Volume	Capital Gains	ratio_5	trend_5	ratio_15	trend_15
2023-09-07 04:00:00+00:00	407.170013	409.429993	406.859985	408.829987	2243800	0.0	0.992152	1.0	1.002849	8.0

Output:
If 9/8 will go up?

```
tmr = model.predict(stock_df_last_date[predictors])
print(tmr)
```

```
[1]
```
(Yes)

(基金投資有賺有賠 投資人申購前請詳閱基金公開說明書)

Thanks for listening

Photo by @riverhetw