Finmily Final Presentation

ZHAOTONG LIU, ZHANBO LAN, TIAN DAI

Data (from Yahoo Finance)

1. Historical Data–Model

Date	Open	High	Low	Close	Adj Close	Volume
2021-09-16	108.239998	108.750000	106.529999	108.080002	103.826820	5821800
2021-09-17	107.500000	108.250000	107.099998	107.730003	103.490585	14157200
2021-09-20	106.099998	107.489998	105.559998	106.400002	102.212921	7103400
2021-09-21	106.250000	107.410004	106.199997	107.150002	102.933418	6400400
2021-09-22	107.010002	107.580002	105.930000	106.410004	102.222542	6319100
...	...	...	...	...	...	...
2022-09-09	140.399994	142.169998	140.100006	141.419998	141.419998	4427000
2022-09-12	141.009995	142.869995	140.320007	142.240005	142.240005	4938400
2022-09-13	140.550003	141.929993	137.490005	138.529999	138.529999	6195900
2022-09-14	139.169998	140.440002	138.270004	139.550003	139.550003	4673600
2022-09-15	139.990005	143.899994	139.389999	142.509995	142.509995	6416900

Web scraping package: urllib

2. Beta–Decision Making

Previous Close	**141.17**	Market Cap	**2.334T**
Open	**141.40**	Beta (5Y Monthly)	**1.25**
Bid	**141.13 x 800**	PE Ratio (TTM)	**23.13**
Ask	**141.17 x 800**	EPS (TTM)	**6.11**
Day's Range	**140.55 - 146.88**	Earnings Date	**Jan 25, 2023 - Jan 30, 2023**
52 Week Range	**129.04 - 182.94**	Forward Dividend & Yield	**0.92 (0.65%)**
Volume	**62,204,943**	Ex-Dividend Date	**Nov 04, 2022**
Avg. Volume	**89,746,460**	1y Target Est	**178.15**

3. Industrial Sector–Decision Making

Sector(s): **Technology**

Industry: **Consumer Electronics**

Full Time Employees: **164,000**

Web scraping package: beautiful soap

Model (LSTM)

1. Long short-term memory (LSTM) is an artificial neural network used in the fields of artificial intelligence and deep learning.

2. Strong prediction power especially with a long-term data sequence.

3. From ARIMA to LSTM: LSTM model can train time series data with multiple features with more accuracy.

Parameters:

(1) Same parameters for all assets at first;

(2) Find best parameters for assets with large error.

```
mem_days=[5]
lstm_layers=[3]
dense_layers=[1]
units=[32]
from tensorflow.keras.callbacks import ReduceLROnPlateau
for the_mem_days in mem_days:
    for the_lstm_layers in lstm_layers:
        for the_dense_layers in dense_layers:
            for the_units in units:
```

Model (LSTM)

Inputs:

Date	Open	High	Low	Close	Adj Close	Volume
2021-09-16	108.239998	108.750000	106.529999	108.080002	103.826820	5821800
2021-09-17	107.500000	108.250000	107.099998	107.730003	103.490585	14157200
2021-09-20	106.099998	107.489998	105.559998	106.400002	102.212921	7103400
2021-09-21	106.250000	107.410004	106.199997	107.150002	102.933418	6400400
2021-09-22	107.010002	107.580002	105.930000	106.410004	102.222542	6319100
...	...	...	...	...	...	...
2022-09-09	140.399994	142.169998	140.100006	141.419998	141.419998	4427000
2022-09-12	141.009995	142.869995	140.320007	142.240005	142.240005	4938400
2022-09-13	140.550003	141.929993	137.490005	138.529999	138.529999	6195900
2022-09-14	139.169998	140.440002	138.270004	139.550003	139.550003	4673600
2022-09-15	139.990005	143.899994	139.389999	142.509995	142.509995	6416900

Transform data into same scale:

```
from sklearn.preprocessing import StandardScaler
scaler = StandardScaler()
sca_X = scaler.fit_transform(df.iloc[:,:-1])
```

Outputs:

	name	price_fri	price_pre	mape
0	ABBV	154.979996	148.252579	3.241197
1	ACN	286.500000	278.057434	4.854886
2	AEP	91.279999	89.071167	4.628783
3	AIZ	124.949997	122.626015	4.482192
4	ALLE	112.959999	113.564033	5.207478
...	...	...	...	...
105	SOL-USD	12.787567	10.944075	14.981492
106	DOGE-USD	0.081392	0.064880	11.362064
107	DOT-USD	5.449032	6.293902	5.621378
108	SHIB-USD	90.000000	91.382843	8.761893
109	AVAX-USD	12.667869	15.426949	6.925270

110 rows × 4 columns

std_6_months
0.015112
0.020487
0.016962
0.021431
0.022311
...
0.065592
0.065225
0.041512
0.054520
0.049970

The model is not good at predicting cryptocurrency.

Optimization Objective function

Objective function (Choose from 3 of the following depends on our goals):

1) $\boxed{\text{Max } \sum_{i=1}^{110}(w_i * x_i) \text{ where } x_i \text{ is the sharp ratio for asset i}}$
2) Min $\sum_{i=1}^{110}(w_i * x_i)$ where x_i is the risk for asset i
3) Max $\sum_{i=1}^{110}(w_i * x_i)$ where x_i is the predicted return for asset i

Variables: w_i where i ∈ [1,110]

- *Finally choose maximizing based on sharpe ratio to take both risk and return into account.*
- *Due to limitation of free Gurobi package, convert x_i to positive value to reduce number of variables.*
- *Convert optimization result back to original sign based on predicted return rate.*

Optimization Constraints

Constraints:

1) $\sum_{i=1}^{110} w_i = 1$, the sum of the weights is 1.
2) $\sum_{i=101}^{110} w_i = 0$, we do not invest in cryptocurrency as it is too risky.
3) $w_i < 0.1$, we do not invest too much in one asset
4) $\sum_{i=1}^{110} r_i * w_i < 1/1.5/2$, we want to keep at a risk level, test for 1, 1.5, 2 to find reasonable solution
5) $w_{78} > 0$, when economy is down, the gold price will go up

We hedge between stock and ETF within the same sector.

6) $w_2 * \beta_2 + w_6 * \beta_6 + w_{15} * \beta_{15} + w_{17} * \beta_{17} + w_{28} * \beta_{28} + w_{30} * \beta_{30} - w_{82} * \beta_{82} - w_{90} * \beta_{90} = 0$, hedge between stock and ETF within the Tech sector.
7) $w_4 * \beta_4 + w_7 * \beta_7 + w_{11} * \beta_{11} + w_{33} * \beta_{33} + w_{35} * \beta_{35} + w_{40} * \beta_{40} + w_{41} * \beta_{41} + w_{42} * \beta_{42} + w_{47} * \beta_{47} - w_{91} * \beta_{91} = 0$, hedge between stock and ETF within the Finance sector.
8) $w_{22} * \beta_{22} + w_{50} * \beta_{50} + w_{93} * \beta_{93} = 0$, hedge between stock and ETF within the Energy sector.
9) $w_1 * \beta_1 + w_{12} * \beta_{12} + w_{14} * \beta_{14} + w_{20} * \beta_{20} + w_{37} * \beta_{37} - w_{92} * \beta_{92} = 0$, hedge between stock and ETF within the Healthcare sector.
10) $w_5 * \beta_5 + w_{10} * \beta_{10} + w_{16} * \beta_{16} + w_{23} * \beta_{23} + w_{46} * \beta_{46} + w_{48} * \beta_{48} - w_{95} * \beta_{95} = 0$, hedge between stock and ETF within the Industries sector.
11) $w_{19} * \beta_{19} + w_{29} * \beta_{29} + w_{31} * \beta_{31} - w_{96} * \beta_{96} = 0$, hedge between stock and ETF within the Communication sector.
12) $w_3 * \beta_3 + w_{21} * \beta_{21} + w_{39} * \beta_{39} - w_{97} * \beta_{97} = 0$, hedge between stock and ETF within the Utilities sector.
13) $w_{18} * \beta_{18} - w_{99} * \beta_{99} = 0$, hedge between stock and ETF within the Materials sector.

- *Make sure sum of weights equals to 1*
- *Limit maximum weight for each asset*
- *Constrain total risk level by standard deviation*
- *Ignore cryptocurrency to avoid too much volatility*
- *Hedging between stocks and ETF based on different sectors and beta values.*

Back test on return rate

We plan to conduct back test with the historical data of last 2 years.

For each week, we calculate both the return rate with our predicted portfolio, based on the LSTM and Optimization model, and the return rate with a random portfolio.

Decision Making

Input: Predicted Assets' Ranking and Return

- Method 1: Recent industrial trend + Predicted Return

Recent industrial trend	Predicted Return	Decision	Total Proportion
↑	↑	Large Long	40%
↑	↓	Small Short	10%
↓	↑	Small Long	10%
↓	↓	Large Short	40%

The total proportion is distributed evenly across the assets within each category

Decision Making

- Method 2: Recent industrial trend + Predicted Return + Market Trend Prediction

  An modification method of Method 1 that **adjust the total proportion of Long/short** with the market trend expectation. Market trend insights gained by actively read and sense from:

  - ➔ News on future Fed reserve announcement
  - ➔ News on Earning announcement
  - ➔ Yahoo finance Investment idea, stock market news
  - ➔ Various opinions from WeChat investment groups

- Method 3: Optimization + Market Trend Prediction

  - ➔ Limited by rounds of weeks, we got poor performance for this methods

Method 2 works best

Learning

- Stock market has too many undetermined factors which make it difficult to predict, entire or specific sectors may be affected by breaking news and jeopardize normal predictions.
- Predicting the market trends could be helpful in determining the trading strategies but it requires a closer trace of news.
- There are lots of existing factors (like beta value) and/or models (like Boosting or Deep Learning models) to help us create our portfolio, and we can build our own investment pipeline based on them.
- Trust your model even a kind of surprise result is showing, but also remind of the risk and avoid extreme decision.

Q&A

Thank you!