

Portfolio Management Executive Summary

Team Members: Li Zhu, Jingheng Wang

Github: https://github.com/HiggsP10/Portfolio_Optimization

Goal:

Use various deep learning models to do short-term and long-term portfolio management.

Data:

- Sentiment analysis CSV data from ~5000 tweets
- AlphaVantage Stock Data
- Hugging Face News Data
- Yahoo Finance

Structure:

- Seeking for both long and short term prediction for the stock with optimal allocation
- Short term: use NLP sentiment analysis combined with Black-Litterman model
- Long term: use LSTM network to do prediction based on stock history
- Portfolio Allocation: Markowitz v.s. Genetic

Evaluation Methods:

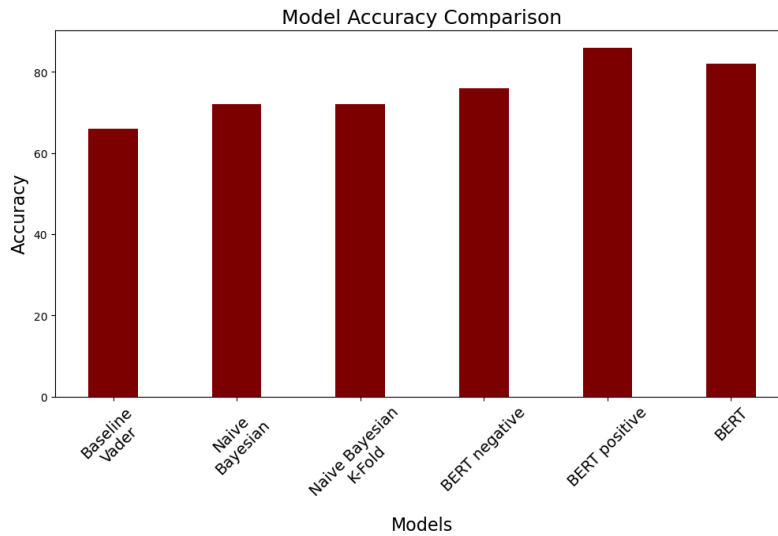
- Sentiment analysis
 - Generally, all models were compared to each other using **accuracy** (correct guesses/total data)
 - Specifically
 - For the best model, BERT, negative and positive sentiment was also computed to see any bias in the algorithm
- Stock prediction
 - We look at the return rate in both the short term and long term.

Models:

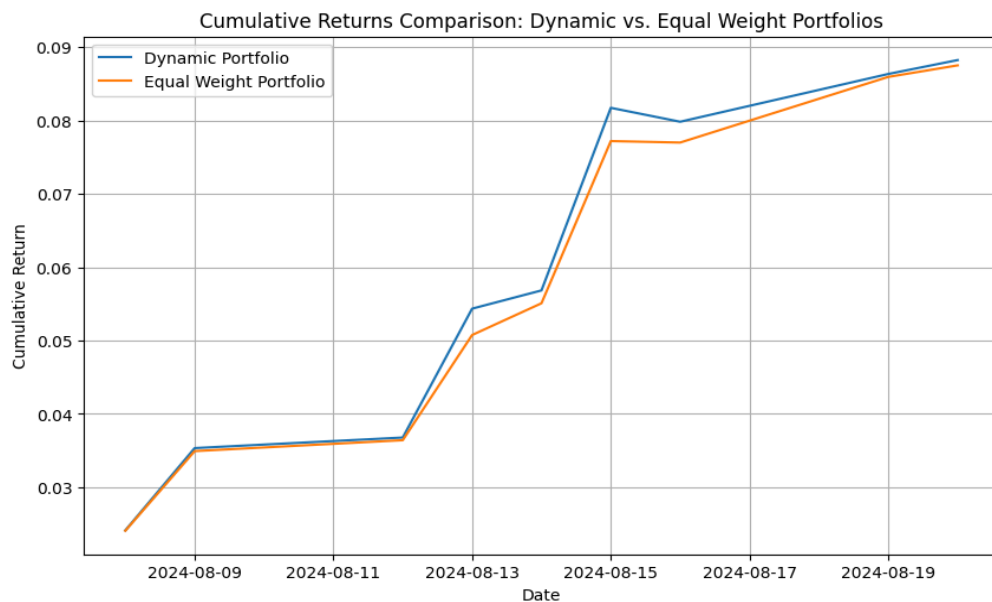
- Sentiment analysis
 - **BERT** - Model selected for final sentiment analysis
 - Vader
 - Naive-Bayesian
- Stock prediction
 - Long-Short-Term-Memories Network (Long-term)
 - The Black-Litterman model with sentiment analysis (Short-term)
 - Markowitz Model v.s Genetic Algorithm

Comparison of Models:

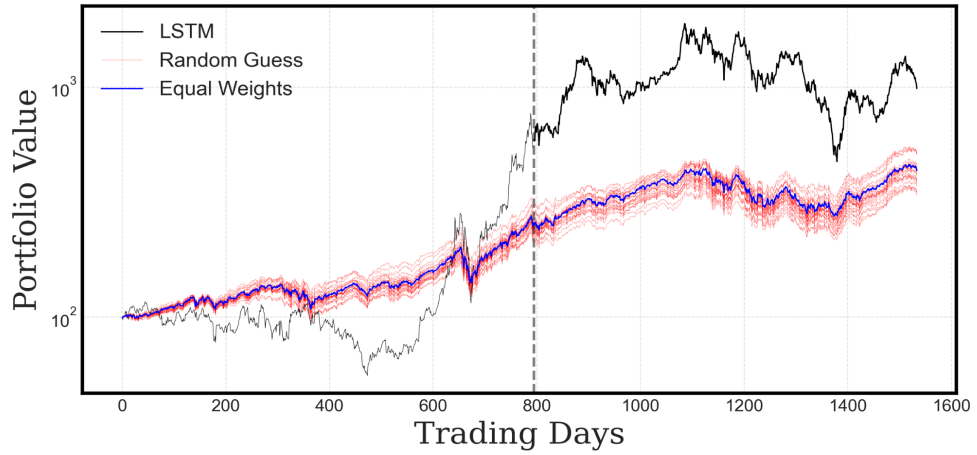
- Sentiment analysis
 - Amongst all models compared, BERT performed better than the Vader and Naive Bayes models in overall accuracy
 - We see that BERT's positive and negative sentiment accuracy performs above all models as well but there is a bias toward better prediction for positive sentiment.
 - We choose data generated by better trained model (with more training data) over self-generated data for better prediction.



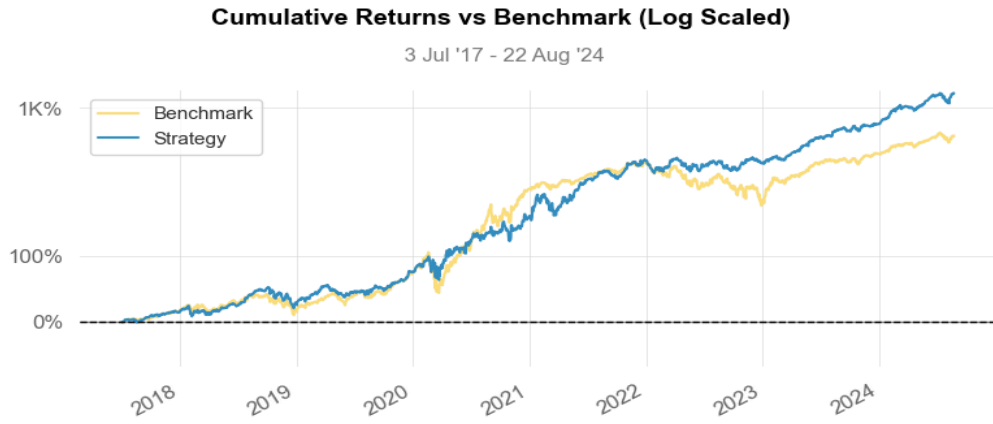
- Stock prediction
 - The Black-Litterman model dominates the baseline model the whole time.



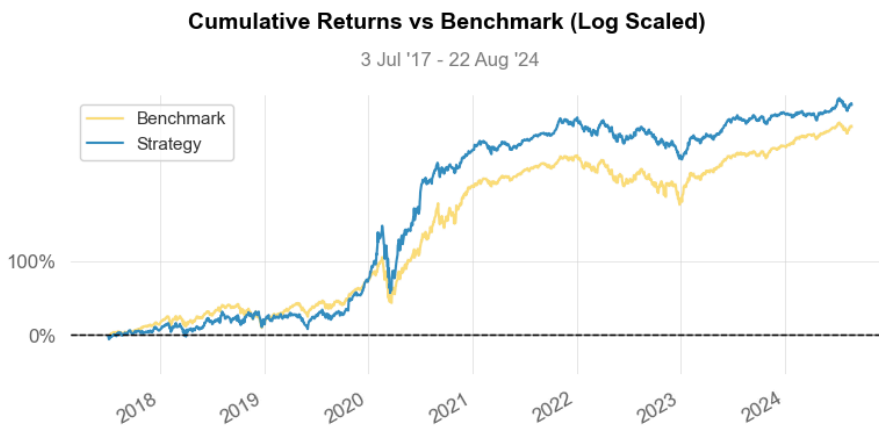
- LSTM works good in long-term holding



- Portfolio Allocation
 - Markowitz Model



- Genetic Model



Future Directions / Conclusion:

- We need a better and larger training data set for the sentiment classification. Due to the nature of the NLP model, negative classification needs more data to train.
- Work better on enhancing profit when the market is good than preventing loss when the market is bad.
- LSTM model behaves the best
- Genetic Algorithms need more tuning of parameters to get better results
- Use sentiment analysis doesn't have significant improvement
- Each optimization method has its strengths and weaknesses, and the choice of method often depends on the specific objectives, constraints, and preferences of the investor or portfolio manager