Deep Learning for Portfolio Optimization

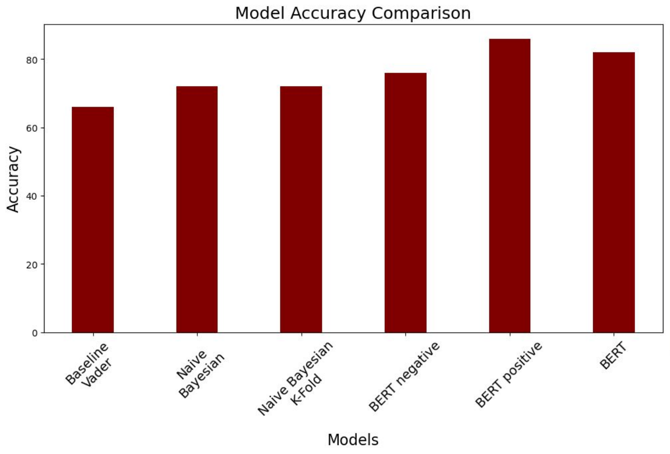
Deep Learning Boot Camp May-Summer 2024
The Erdös institute
Li Zhu, Jingheng Wang, Arvind Suresh

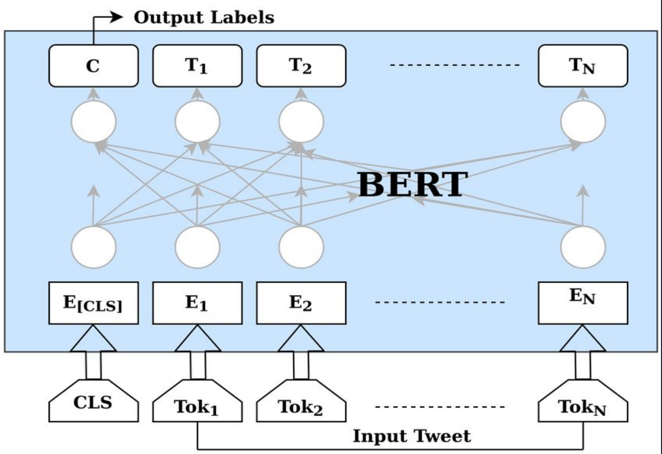
Problem description

- Selecting the best possible mix of stocks to maximize returns while minimizing risk, based on an investor's risk tolerance, investment goals, and market conditions
- Short-term: NLP sentiment analysis combined with the Black-Litterman model
- Long-term
  - Long-Short-Term-Memories Network using stock history
  - Genetic Algorithm
  - Markowitz Mean-Variance Model
- Stakeholders: Individual Investors, Institutional Investors, Portfolio Managers and Financial Advisors, Fund Managers, Regulators and Compliance Officers, Shareholders, etc.

Sentiment Analysis

- BERT is a transformer based encoder trained to read in a full sentence to understand context
- We preprocess the news to standardize the text
- BERT model perform better for positive messages than negative
- Overall, BERT performs better than other NLP models



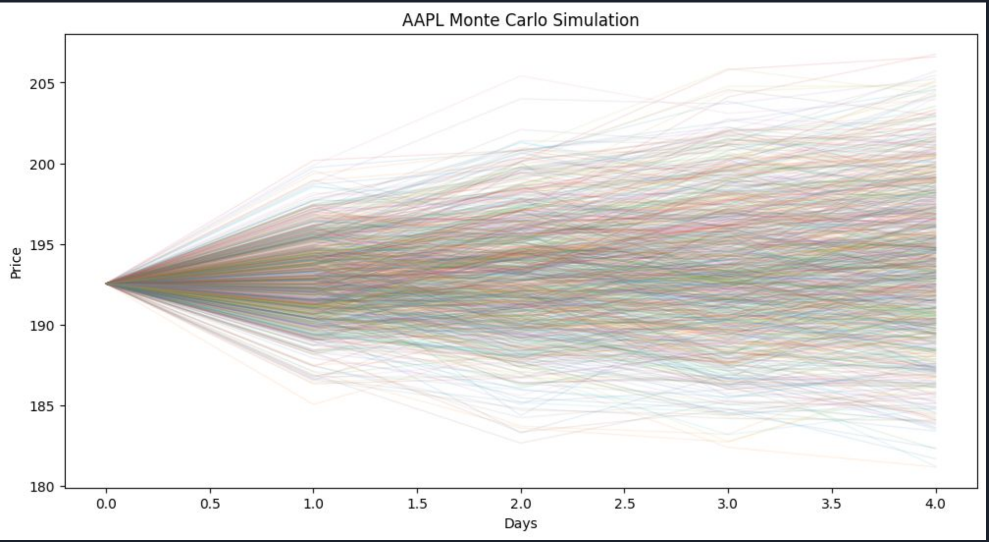


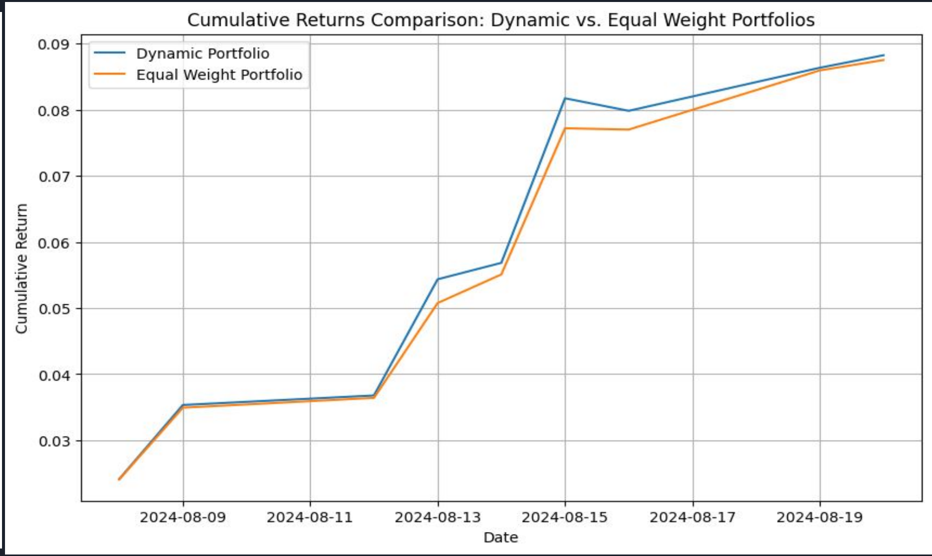
The Black-Litterman model – Short Term Prediction

- Posterior estimate of expected returns
- $E(R) = [(\tau\Sigma)^{-1} + P^T\Omega^{-1}P)]^{-1}[(\tau\Sigma)^{-1}\Pi + P^T\Omega^{-1}Q)]$
- Use prior estimate of returns with investor's views and confidence
- Use sentiment score $\gamma \in [-1,1]$ to calculate views

$$S_T = \begin{cases} S_0 + [(S_{MC}^{+} - S_0) \cdot \gamma] & \text{if } \gamma \text{ in } (0,1] \\ S_0 - [(S_0 - S_{MC}^{-}) \cdot \gamma] & \text{if } \gamma \text{ in } [-1,0) \end{cases}$$

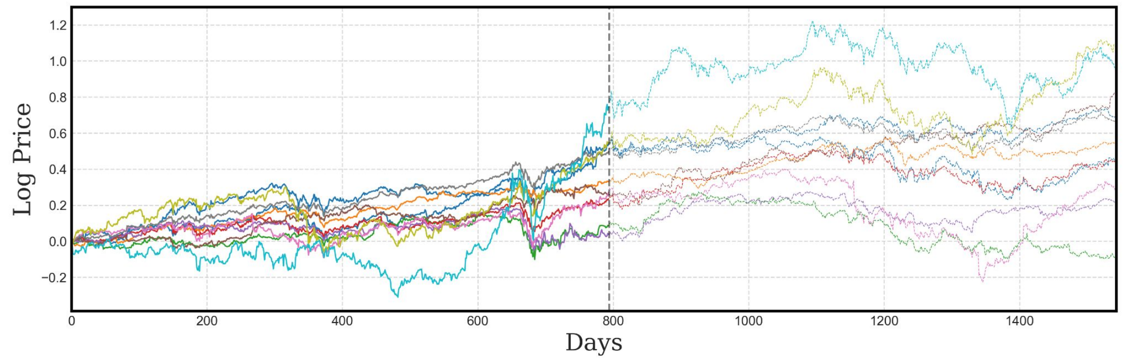
$$Views = \ln(\frac{S_T}{S_0})$$

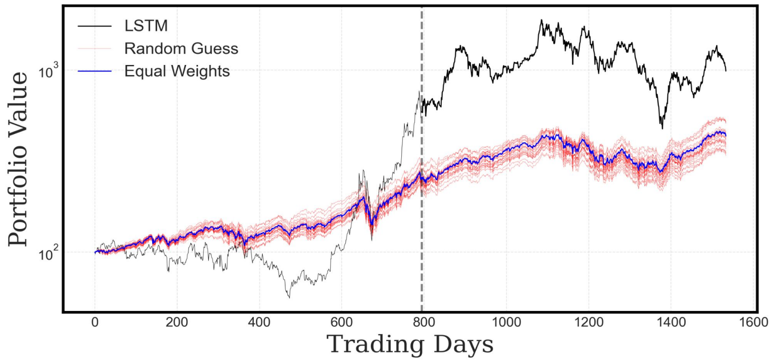

AAPL Monte Carlo Simulation
Price
Days
205
200
195
190
185
180
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0


Cumulative Returns Comparison: Dynamic vs. Equal Weight Portfolios
Dynamic Portfolio
Equal Weight Portfolio
Cumulative Return
Date
0.09
0.08
0.07
0.06
0.05
0.04
0.03
2024-08-09
2024-08-11
2024-08-13
2024-08-15
2024-08-17
2024-08-19

LSTM – Long Term Prediction

- Recurrent neural network (RNN)
- Capture both short-term fluctuations and long-term trends in financial data
- Memory Cells, Forget Gates -> Adaptive learning





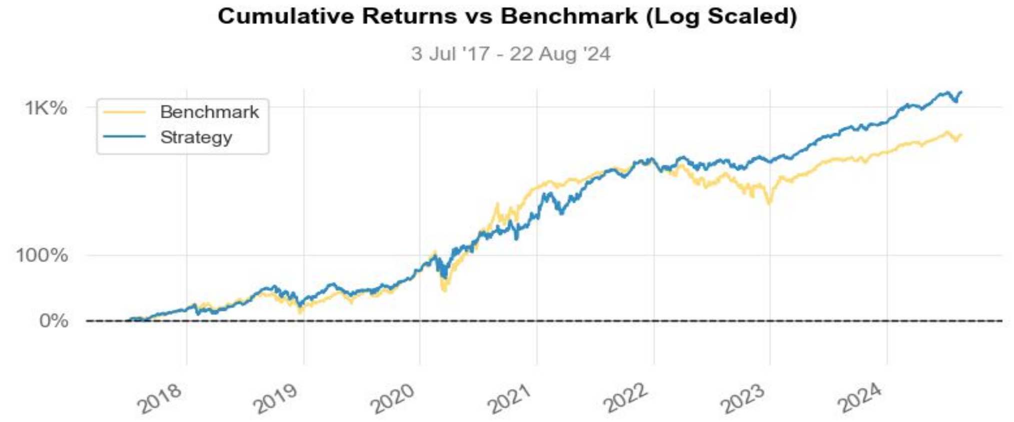
The Markowitz Model v.s GA

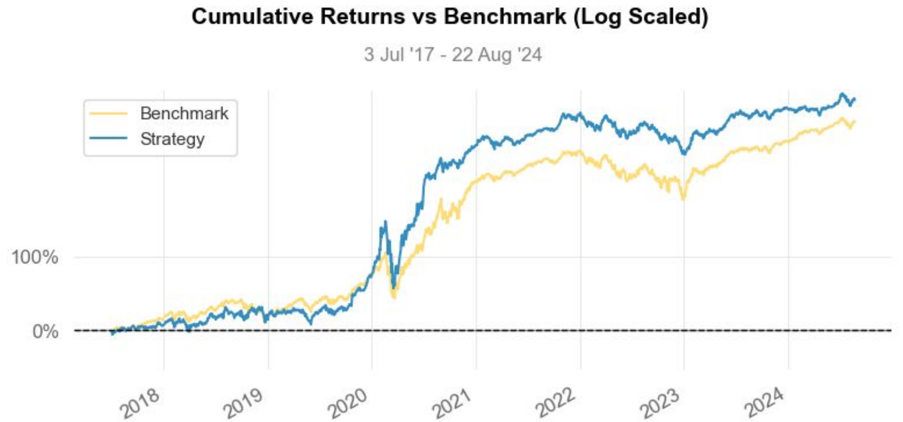
Performance Metrics

	Benchmark	Strategy
Start Period	2017-07-03	2017-07-03
End Period	2024-08-22	2024-08-22
Risk-Free Rate	0.0%	0.0%
Time in Market	100.0%	100.0%
Cumulative Return	572.26%	1,327.74%
CAGR﹪	20.22%	29.3%
Sharpe	1.21	1.67
Prob. Sharpe Ratio	99.92%	100.0%
Smart Sharpe	1.08	1.5
Sortino	1.73	2.62
Smart Sortino	1.55	2.35
Sortino/√2	1.22	1.85
Smart Sortino/√2	1.1	1.66
Omega	1.36	1.36

Performance Metrics

	Benchmark	Strategy
Start Period	2017-07-03	2017-07-03
End Period	2024-08-22	2024-08-22
Risk-Free Rate	0.0%	0.0%
Time in Market	100.0%	100.0%
Cumulative Return	572.26%	836.52%
CAGR﹪	20.22%	24.14%
Sharpe	1.21	1.17
Prob. Sharpe Ratio	99.92%	99.9%
Smart Sharpe	1.12	1.09
Sortino	1.73	1.72
Smart Sortino	1.6	1.59
Sortino/√2	1.22	1.22
Smart Sortino/√2	1.13	1.13
Omega	1.23	1.23





Conclusion and Future work

- 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