SIADS 699 - Capstone - Team Alpha (#11)

Summer 2023

Greg Holden (gsholden), Jason Peloquin (peloquin), Emmanuel Sengendo (esenge)

Can NLP Techniques Be Used To Optimize Investment Portfolios?

Background

Maximizing investment returns has long been the primary goal among investors. Historically, a combination of fundamental and technical analysis has been used to make purchase/sale decisions. Initially, these decisions were focused on individual stocks. However, as the benefits of holding a diversified portfolio of investments became more apparent, investors shifted their focus to creating optimized portfolios. The theory behind this shift was to include multiple securities which responded differently to changing market conditions over time. This allowed for returns to be maximized while minimizing risk. A common benchmark portfolio utilizing this approach might consist of 60% equity and 40% debt securities.

How can NLP Help Investors?

Advances in NLP techniques have enabled investors to look beyond traditional fundamental and technical analysis and incorporate signals gleaned from the financial press into their investment decisions. One much-scrutinized source of financial information is the disclosure from meetings of the Federal Open Market Committee ("FOMC"). This government body, which typically meets eight times each year, provides a short Transcript at the conclusion of each meeting and a lengthier copy of the meeting Minutes ("FOMC Minutes") three weeks after the event. "Fed-watching" investors carefully analyze these releases looking for subtle changes in tone that could signal near-term changes in interest rate policy, which creates opportunities for investors to reposition their portfolios.

Objectives

Given this background we had two primary goals:

1. **Enhanced Portfolio Optimization Using NLP**: Analyze whether signals derived from various NLP algorithms (and run on the FOMC Minutes) could be used to construct an investment portfolio that outperforms a base portfolio relying solely on traditional technical indicators.

2. **Security-Level Impact**: Analyze whether subsegments of the equity and debt markets respond differently based on the tone of FOMC Minutes documents. This more granular analysis involves looking at (a) equity securities focused on specific industries such as banks, and (b) debt securities across the credit spectrum including U.S. Treasury bonds, and high ("Investment Grade") or low ("High Yield") quality Corporate bonds.

Literature Review

Our research revealed numerous precedents focused on the use of NLP to predict individual stock prices. Fortunately, we also found examples that addressed a wider array of investments.

Tadle (2022) measures sentiment analysis across multiple asset classes – Fed Futures, Foreign Exchange and select Equities – and finds that the FOMC Minutes provide stronger signals for Equities than the much-shorter FOMC Transcripts released immediately after the Fed Meeting. Handlan (2022) also highlights an important distinction between the FOMC's discussion of the current state of the economy versus the future economic outlook (known as "forward guidance"). She demonstrates that forward guidance has a much greater impact on the ability to predict Fed Funds rates. Importantly, Handlan also highlights potential issues neural networks face when relying on relatively scarce amounts of FOMC data.

We also found other research that helped with our model selection process. Nadendla (2020) highlights how LSTM models have been eclipsed by Transformers - which are able to better preserve document context over longer periods of time - for many NLP tasks.

Methodology

Our analytical process was broken down into three broad steps – Data Collection, NLP Modeling, and Portfolio Optimization. At a high level, our premise was that the FOMC Minutes should be additive to the investing process given their high visibility among investors looking to reposition portfolios between equity and debt securities.

However, we did not initially know which NLP algorithm would yield the best results, so we decided to conduct A/B tests on every NLP metric in search of the most promising approach.

Phase 1 - Data Collection

Our two primary data inputs were (1) FOMC Minutes documents, which we accessed using a python package called FedTools, and (2) historical stock prices and related technical indicators, which we accessed using the yFinance and Ta-Lib python packages.

Key Data Facts:

Dates: 2008 – 2023 YTD
Total Documents: 123
Avg. Length: ~10,000 words

Figure 1: FedTools Data Sample

	Federal_Reserve_Mins
2008-01-30	March 18, 2008 Present:Mr. Bernanke, Chairman...
2008-03-18	The Federal Reserve, the central bank of the U...
2008-04-30	April 29-30, 2008 PRESENT:Mr. Bernanke, Chair...
2008-08-05	August 5, 2008 PRESENT: Mr. Bernanke, Chairma...
2008-09-16	September 16, 2008 PRESENT:Mr. Bernanke, Chai...

Phase 2 - NLP Modeling

Our next step was to apply multiple NLP models to each FOMC Minutes document. The goal was to create vector representations of the text data that could be used as input features in our classification-based portfolio optimization model. We used multiple families of NLP models to (a) isolate the impact of each NLP strategy on our optimized portfolio, and (b) assess which types of models best addressed our relatively small dataset. These included:

Topic Modeling: LDA, Doc2Vec and BERTopic
Named Entity Recognition: Focused on named FOMC Board Members
Transformer Sentiment: FinBert
Document Similarity: Universal Sentence Encoder and Cosine Similarity

Collectively, this analysis allowed us to (a) identify key topics and entities that potentially signaled bullish or bearish market tone, and (b) understand how different models separate documents into topics.

For instance, Figure 2 shows which words BERTopic identified as prominent and grouped together into its first topic, while Figure 3 shows the frequency with which these words appear across all documents as per LDA. Interestingly, many words a trader would look for -- such as “risk” (2,512 occurrences), “credit” (2,125) and “employment”

(1,768) - do not appear among the top 15 most common words. This suggests that NLP models may prioritize different items than investors.

Figure 2: Illustrative Word Cloud

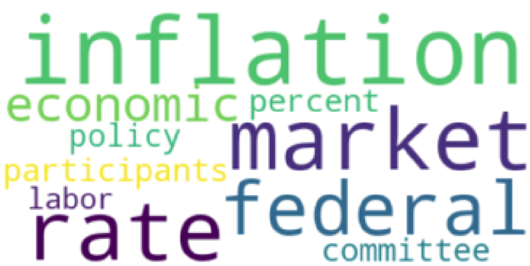
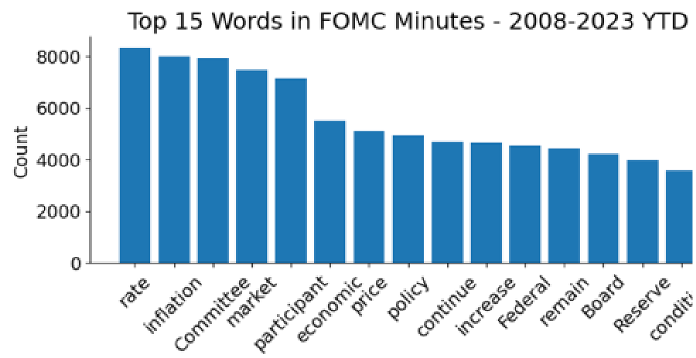
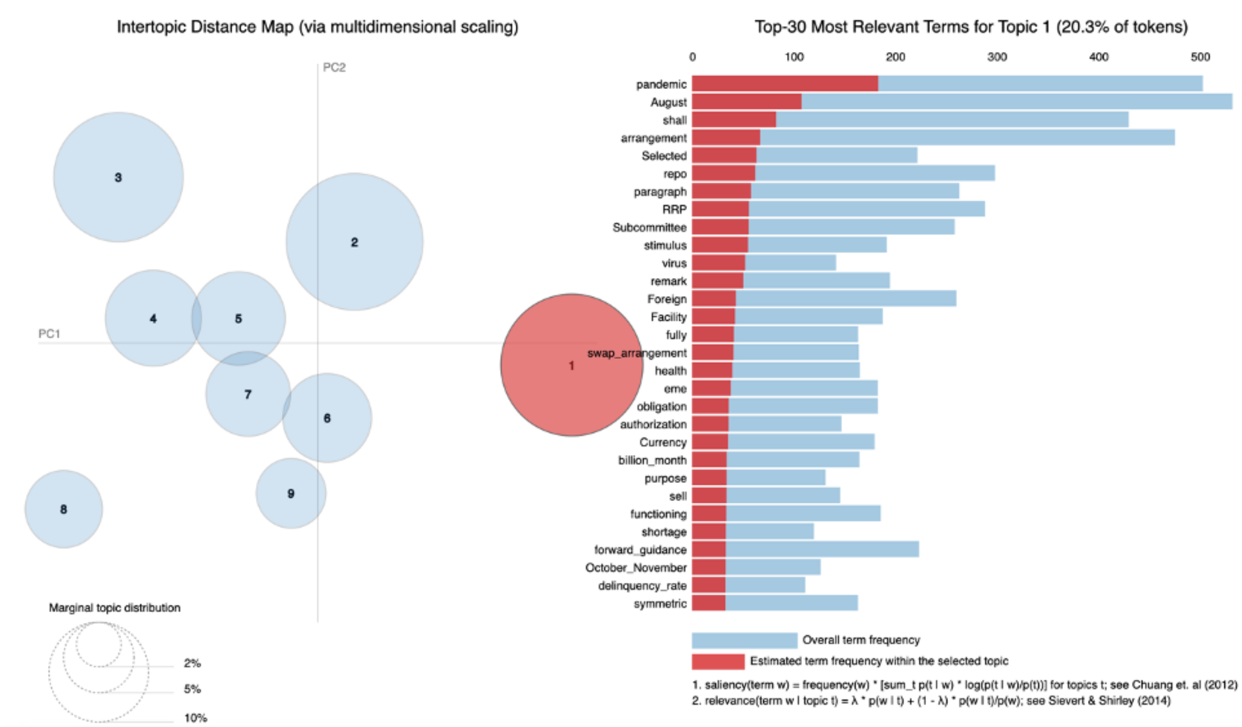


Figure 3: Most Common Words



This process also provided insights as to how different models group terms and form topics. As an example, Figure 4 shows how the LDA model generates topics, which differs from BERTopic in terms of the number of optimal topics and the composition of each topic.

Figure 4: LDA Topics – FOMC Minutes (2008 – 2023)

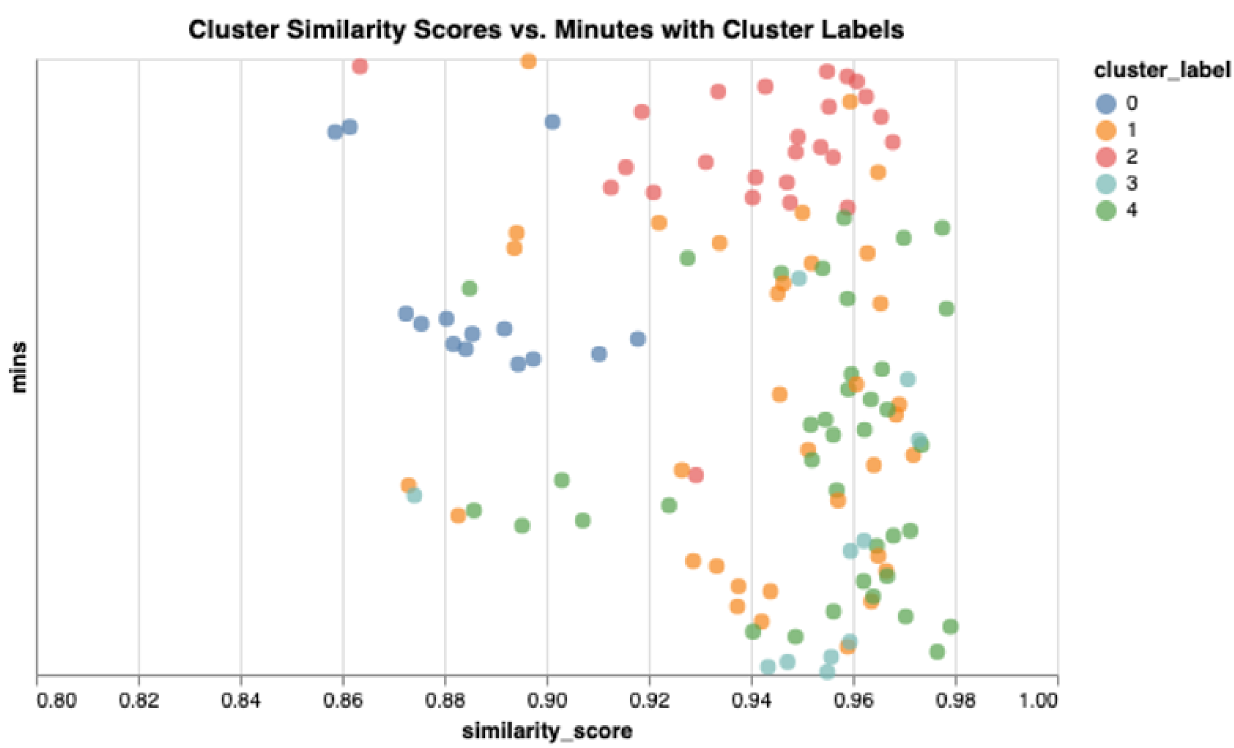


Beyond Topic Modeling, we also (a) calculated Document Similarity scores using Cosine Similarity and Universal Sentence Encoders to measure changes in tone between subsequent documents, and (b) used Named Entity Recognition to count the number of times each FOMC Board Member was mentioned.

Universal Sentence Encoder (USE):

Pre-trained models can be enhanced with domain-specific knowledge to achieve more accurate topic modeling results, particularly in the context of finance. We encountered challenges as we could not identify pre-trained models that had been specifically trained on precisely relevant data. As a result, we explored alternative approaches, such as building a custom topic modeling solution from scratch or fine-tuning existing pre-trained models using a relevant dataset. Our data source did not provide a sizable enough corpus for these tasks. We therefore relied on the Tensorflow Universal Sentence Encoder for document classification and similarity scores (Figure 5). USE captures the overall content of the document while using dimensionality reduction to be more computationally efficient.

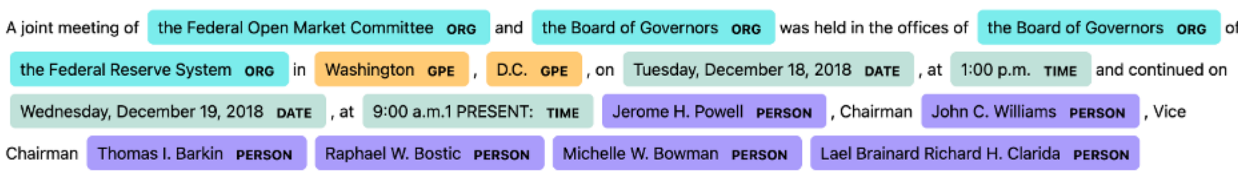
Figure 5: Universal Sentence Encoder Document Clusters & Similarity Scores



Using USE, the 123 documents (earliest examples from 2008 on the lower part of the y-axis) were clustered into 5 groups and had a mean similarity score of 0.93.

For the NER analysis, our goal was to see if mentions of certain FOMC Board Members -- many of whom have reputations for favoring either more (known as "hawks") or less ("doves") restrictive monetary policy – could signal future interest rate increases/decreases, respectively.

Figure 6: Sample Named Entity Recognition Analysis



Part 3 - Portfolio Allocation

Armed with our NLP analysis, we developed a "back trading" strategy which incorporated ETF daily price and technical indicator values as model inputs. These values were averaged around the 5-day trading window of each FOMC release to reflect typical market repositioning activity.

This data - along with the NLP vectors – was then used to conduct A/B portfolio testing. Our goal was to determine if including NLP data in a portfolio rebalancing scenario creates more accurate predictions and excess returns versus a baseline portfolio. To do this, two portfolios were created:

Each portfolio initially consisted of the following two securities:

A. 60% equity. Uses an Exchange-Traded Fund (ETF) as a proxy for the S&P 500 equity index (Ticker Symbol: SPY)
B. 40% debt. Uses an ETF proxy for long-dated U.S. Treasury Bonds (Ticker Symbol: TLT)

Adjustments to the portfolios were only made at the time of each FOMC meeting (8x per year), but based on slightly different criteria:

Portfolio #1 (Base Case - Technical Indicators Only, No NLP): Relies solely on three traditional stock market indicators to make portfolio rebalancing decisions:

a. Exponential Moving Average (“EMA”). Compares the current price of an ETF to its longer-term average.
b. Relative Strength Index (‘RSI”). Momentum-based indicator scaled from 0 to 100. Typically, RSI levels below 30 suggest an oversold security, while RSI levels above 70 suggest an overvalued security.
c. Moving Average Convergence / Divergence (“MACD”). Comparison of short and longer-term EMAs.

Portfolio #2 (NLP Case): Same technical indicators as Portfolio #1, but also individually layers in each of the NLP-based metrics described above.

Modeling Process

We treated this exercise as a classification task, where at each FOMC Meeting the model classified the economy as “bullish” (favorable) or “bearish” (unfavorable). Our training period was from 2008 to 2019, while our test period was from 2020 through June of 2023.

Each of the 92 FOMC meetings in the training period was treated as a potential rebalancing event, where allocations to equity/debt could increase/decrease based on positive/negative economic conditions determined through our derived bullish/bearish rebalancing signals.

A market event was determined to be bullish if the S&P 500 return was positive for the following market cycle leading to the next FOMC release. Portfolio allocations between equity and debt could change by 20% immediately after each FOMC meeting. For example, if portfolio indicators predicted a favorable (bullish) market environment, the equity/debt portion of the portfolio would increase/decrease by 20%, respectively.

A/B Testing NLP Portfolios

To determine which NLP metrics to include in the final portfolio, significant A/B testing was conducted. Each NLP metric was utilized in a custom NLP portfolio and was compared versus a baseline portfolio with no NLP vector input. Multiple classifiers were utilized in conjunction with each custom NLP portfolio to determine the best NLP input/model combination.

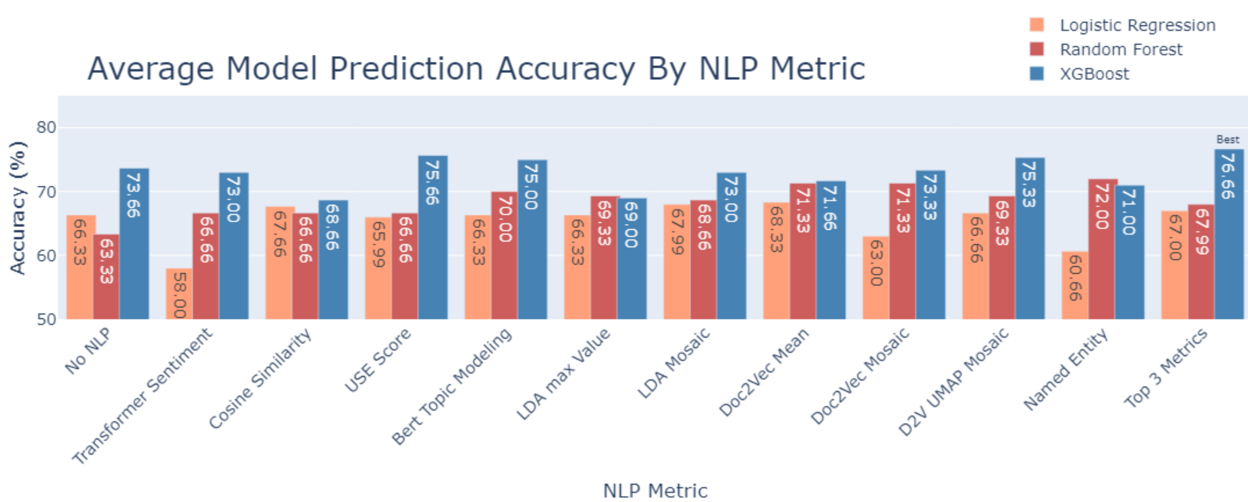
The three classifiers chosen for our analysis were: Logistic Regression, Random Forest and XGBoost. After each NLP metric was run for each classification model, the top 3 metrics for each classifier were then combined as a mosaic to determine if the aggregation of the top NLP performers increased model accuracy.

Model Optimization A/B Testing

In order to optimize each classification model, a random grid search was performed inside a Monte Carlo analysis, allowing for 10 simulations to be run for each model with 10 cross validations and 100 iterations for each simulation. This allowed us to find the most consistent performers with respect to accuracy. Figure 7 displays A/B testing results based on model and classifier pairings.

The XGBoost mosaic approach (far right blue bar) produces the highest overall accuracy score of 76.66% on the training data and includes Use Score, a Doc2vec composite and Bert Topic Modeling as its NLP feature inputs. This combination was down selected for use in our final portfolio rebalancing strategy.

Figure 7: Model Prediction Accuracy By NLP Metric



Key Observations:

1. Random Forest benefitted the most from the inclusion of various NLP metrics, while XGBoost recorded the highest accuracy overall at 76.66%.
2. Among Topic Modeling models, the Doc2Vec mosaic using UMAP improved accuracy of all three models.
3. Logistic Regression performed poorly across the board, potentially due to the model's inability to adjust itself in a nonlinear fashion, whereas XGBoost

capitalized on gradient boosting and produced more accurate results with limited training data.
4. Transformer sentiment only increased accuracy for Random Forest.
5. Identifying shifts in document tone using Cosine Similarity had a small positive effect on both Logistic Regression & Random Forest but a negative impact on XGBoost.
6. Use Score improved both the Random Forest & XGBoost models
7. USE Score alone produces similar results to the top scoring mosaic model using XGBoost.

Evaluation Strategy

We measured the success of our portfolio optimization strategy by comparing the returns of our Base Portfolio (with no NLP) to our NLP Portfolio over time. Our primary metric was to track how often our model correctly predicted a bull or bear market between the time of a FOMC meeting and the next FOMC meeting, and most importantly, whether the inclusion of NLP metrics increased the model's predictive ability.

Translation / Interpretation of Results

Figure 8 shows the investment performance of both our Base and NLP portfolios over time. The orange line shows the performance of our Base portfolio, while the blue line shows the performance of our optimized NLP portfolio.

Each portfolio was presented with the opportunity to rebalance up or down by 20% with respect to its equity and debt allocations after each FOMC release.

Prediction errors were also plotted to indicate when the NLP and Base models incorrectly predicted a bull or bear market. The dark red shading represents when both models incorrectly predicted a bull market while the lighter red shading signals show when both models incorrectly predicted a bear market. Additionally, the two gray shaded regions represent the 2 instances where the NLP model correctly guessed bull or bear market and the baseline portfolio did not.

It should be noted several unpredictable systematic events occurred during this time span, specifically Covid in the spring of 2020 and the war in Ukraine in 2022, and neither model accurately predicted either black swan event. Figure 9 summarizes the occurrences where the NLP or base model produces prediction errors versus actual market conditions.

Figure 8: Return Comparison – Base vs NLP Portfolios

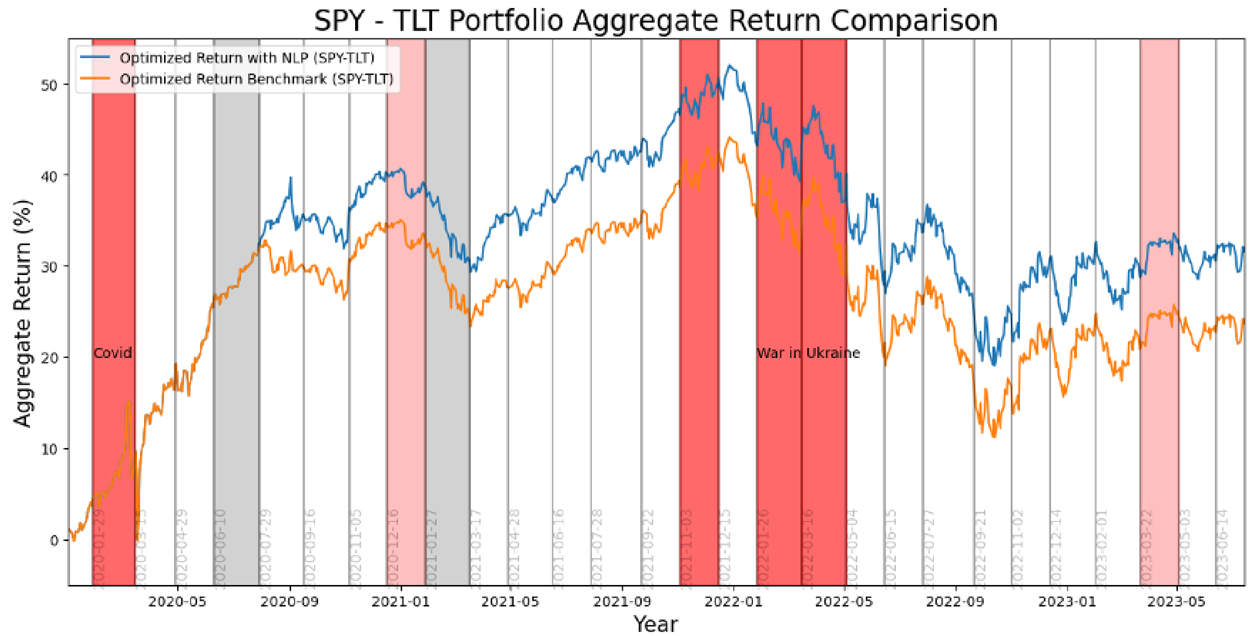


Figure 9: Summary Of Prediction Errors

FOMC Release Date	Market Performance	NLP Prediction	Baseline Prediction
2020-03-15	Bear	Bull	Bull
2020-07-29	Bull	Bull	Bear
2021-01-27	Bull	Bear	Bear
2021-03-17	Bull	Bull	Bear
2021-12-15	Bear	Bull	Bull
2022-03-16	Bear	Bull	Bull
2022-05-04	Bear	Bull	Bull
2023-05-03	Bull	Bear	Bear

Measuring Risk-Adjusted Performance

One popular method of comparing the performance of multiple portfolios is to use a metric called the Sharpe Ratio. This statistic calculates the excess return of a portfolio above a risk-free rate of return (typically a U.S government security), then divides that result by the volatility of the returns. This metric rewards portfolios that achieve superior returns without taking excessive risk.

For comparison purposes, both portfolios were analyzed against each other and against a passive 60% SPY / 40% TLT portfolio. Figure 10 shows that the NLP portfolio produced superior Sharpe Ratios versus the baseline for 2020 and 2021 and tied in 2022. However, the three-year aggregate Sharpe Ratio for the NLP Enhanced portfolio was approximately 36% higher versus the Baseline and 77% higher compared to the Passive investing approach. These findings imply the NLP Enhanced portfolio may offer superior risk-adjusted returns as an investment strategy.

Figure 10: SPY-TLT Portfolio Sharpe Ratio Comparison

Portfolio	2020	2021	2022	3 Year Average
NLP Enhanced	2.45	0.93	-1.46	0.64
Baseline Non NLP	2.15	0.71	-1.46	0.47
Passive 60/40	1.07	1.47	-1.48	0.36

Analysis of Results

Our key takeaway is that NLP metrics can provide an incremental benefit compared to relying solely on technical indicators. We suspect having a larger dataset – or extending beyond the FOMC Minutes to other financial press - would have provided our model with a stronger base for making predictions.

Security-Level Causal Impact

Our second question revolved around whether subsegments of the equity and debt markets respond differently to the bull/bear signals generated from our models based on FOMC releases. The extra movement of a subsector versus the market is referred to as “Beta”.

An analysis of average subsector returns was performed on the training data from 2008 to 2019 to determine if the NLP portfolio could further be enhanced (based on Sharpe Ratios) by exchanging the default SPY and TLT ETFs for other subsector ETFs - such as XLF (Financials) or SJNK (Junk Bonds) - described in the Figure 11.

Figure 11: Summary of Sector ETFs

ETF	Sector	Type
XLF	Financials	Equity
XLY	Consumer Discretionary	Equity
XLI	Industrials	Equity
XLB	Materials	Equity
XLK	Technology	Equity
XLE	Energy	Equity
XLC	Communications	Equity
XLV	Healthcare	Equity
XLRE	Real Estate	Equity
XLP	Consumer Staples	Equity
XLU	Utilities	Equity
SJNK	Short Term High Yield	Debt
LQD	Investment Grade Corporate Bonds	Debt
SPTI	Intermediate Treasuries	Debt
TLT	Long Term Treasuries	Debt

Figures 12 and 13 show average ETF price performance in bullish and bearish periods. As expected, equities outperform in bullish market environments, while investors prefer the safety of U.S Treasury bonds in bearish environments. This relationship is highlighted in both plots below with XLF (Financials), XLY (Consumer Discretionary) and XLI (Industrials) having the greatest sensitivity to bullish and bearish sentiment. Fixed Income remains relatively muted under both scenarios and serves as a stabilizing presence in periods of volatility. In bear markets, long term (TLT) and intermediate (SPTI) Treasuries outperform their counterparts but underperform in bullish cycles. Not surprisingly, the Junk Bond ETF (SJNK) tends to behave more like equity than debt due to the speculative nature and lower average credit quality of its constituents.

Figure 12: Subsector Performance - Bull Market

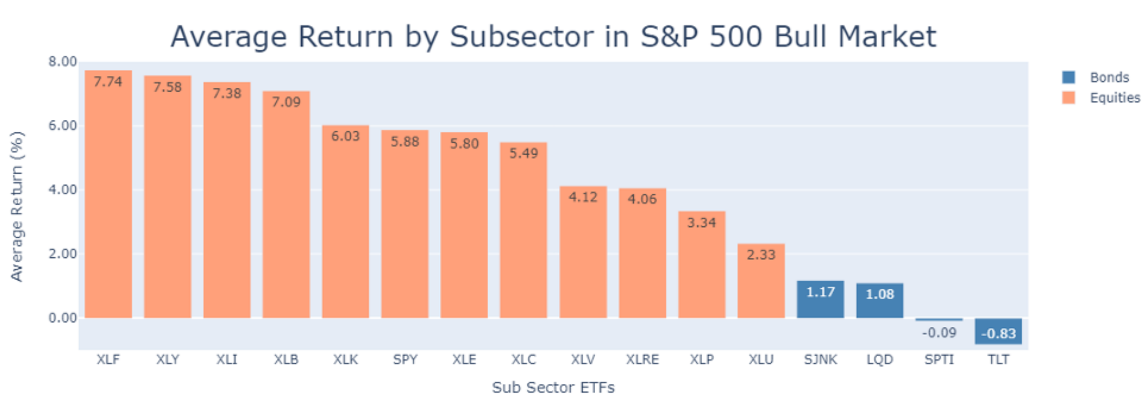
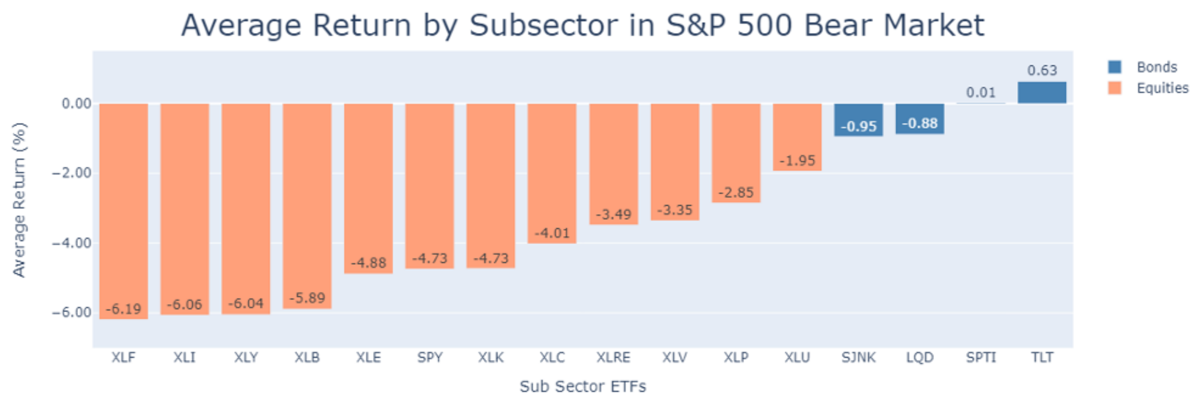


Figure 13: Subsector Performance - Bear Market



Further Portfolio Optimization:

To expand upon both the Beta analysis above and the previous findings from the NLP portfolio optimization, a few additional rebalancing strategies were developed to create portfolios consisting of the top combinations of equity and debt ETFs.

For brevity, only a few scenarios are presented in Figures 14 - 18 - with the top Sharpe Ratio combinations and time series plots located below:

Figure 14: Select ETF Subsector Portfolio Returns

Scenario	Equity ETF	Debt ETF	3 Year Avg. NLP Sharpe	3 Year Avg. Non NLP Passive	3 Year Avg. Passive Sharpe	NLP % Improvement Vs. Non NLP	NLP % Improvement Vs. Passive
Portfolio 1	XLF	TLT	0.53	0.43	0.29	+23.25%	+82.75%
Portfolio 2	XLF	SJNK	0.63	0.59	0.47	+6.78%	+34.04%
Portfolio 3	XLI	TLT	0.68	0.48	0.24	+41.66%	+183.33%
Portfolio 4	XLY	TLT	0.70	0.48	0.24	+45.83%	+191.67%

Figure 15: Portfolio 1 XLF-TLT

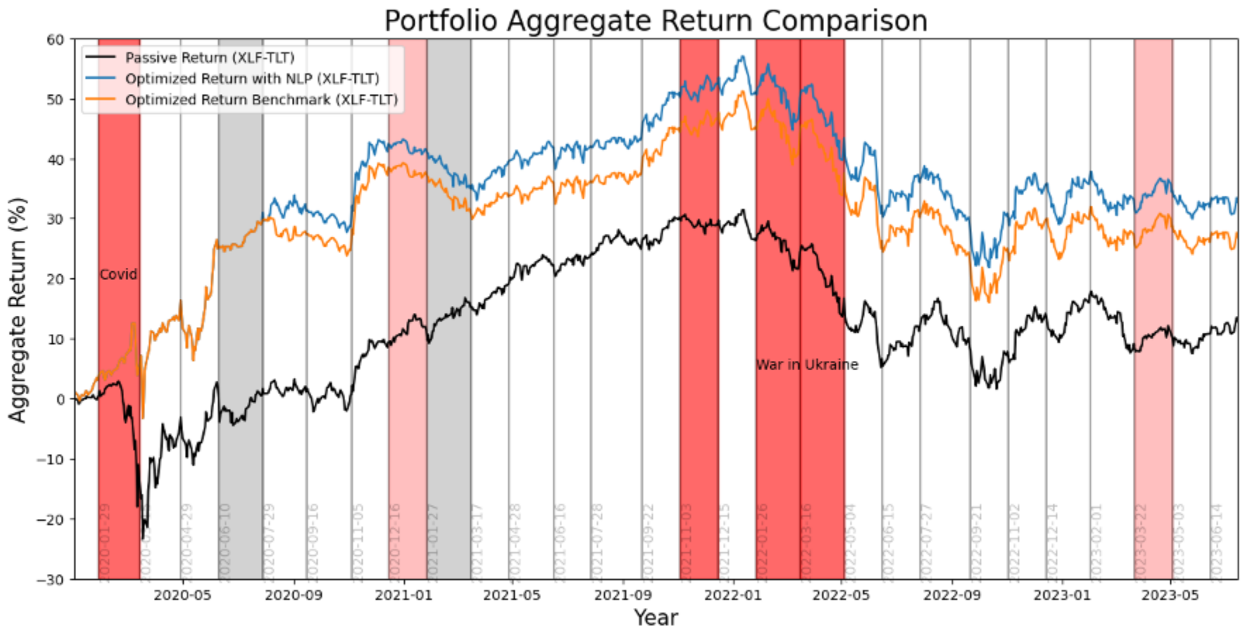


Figure 16: Portfolio 2 XLF- SJNK

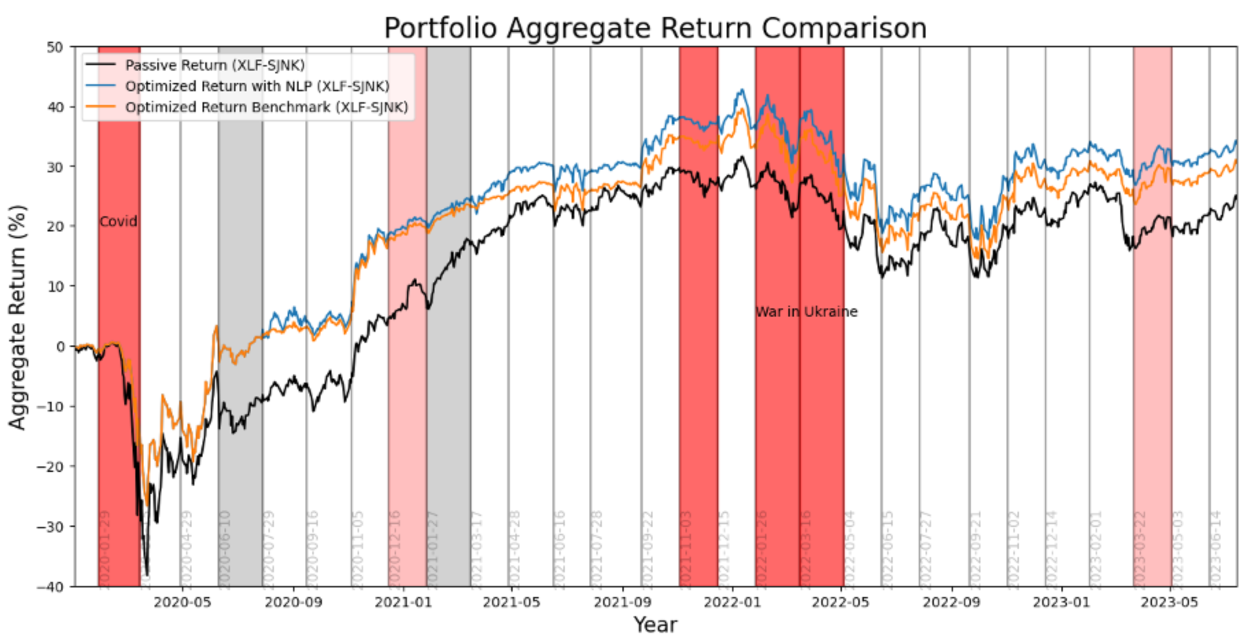


Figure 17: Portfolio 3 XLI-TLT

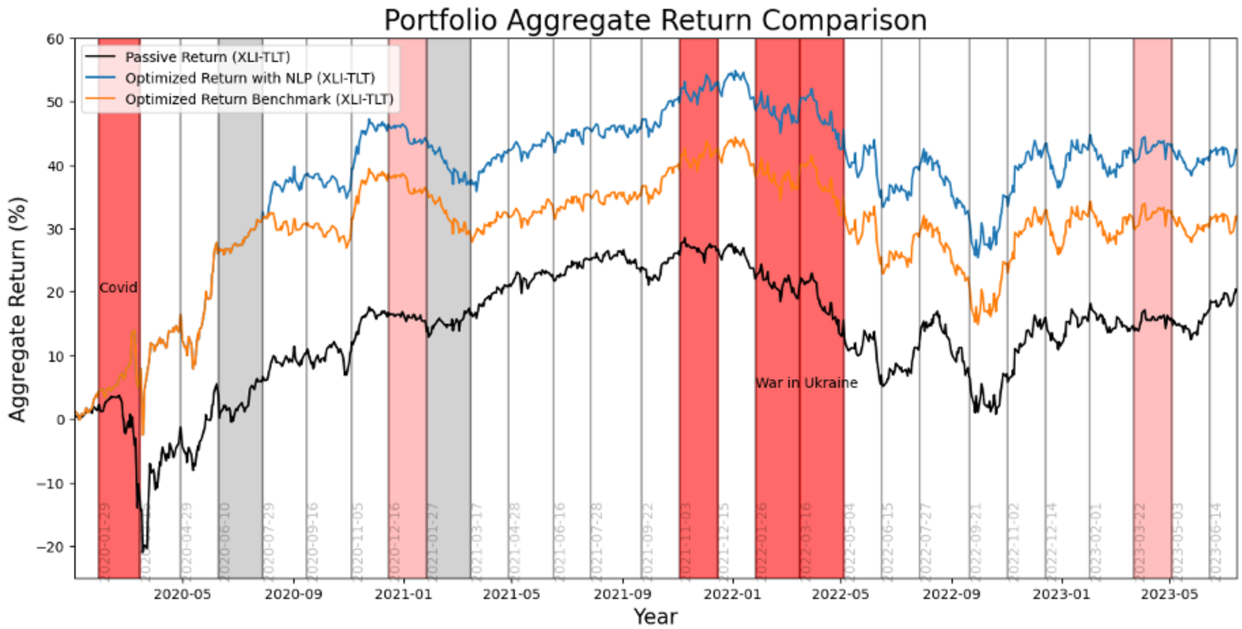
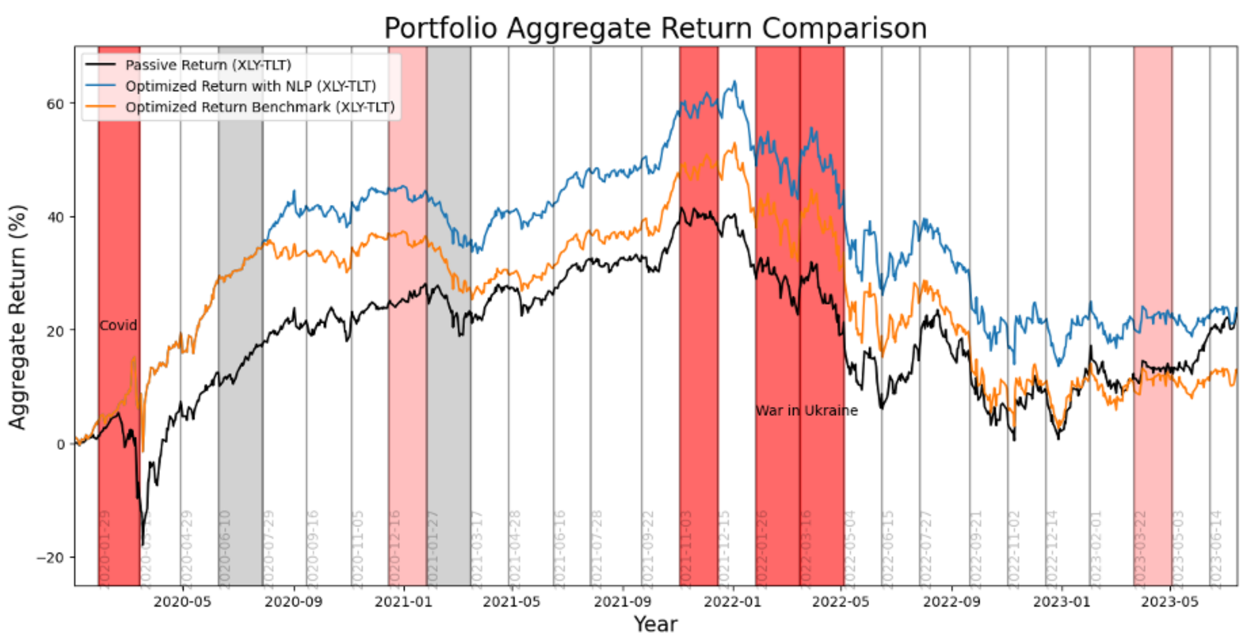


Figure 18: Portfolio 4 XLY-TLT



A Final Comparison:

One final analysis was performed to compare the best performing Subsector ETF equity/debt combination to the benchmark SPY/TLT portfolio.

The XLY-TLT (Consumer Discretionary -Treasuries) combination was used as the Optimized NLP portfolio, as it produced the highest Sharpe Ratio versus its non-NLP peer. This was then compared to a passive SPY-TLT portfolio. Results are plotted in Figure 19 and show that the XLY-TLT NLP combination significantly outperforms the Passive approach.

Figure 20 indicates that the Optimized NLP portfolio produces a significantly higher 3-year average Sharpe Ratio, signifying it may provide better risk-adjusted returns versus the typical market approach.

Figure 19: Optimized Subsector Portfolio vs SPY/TLT Benchmark

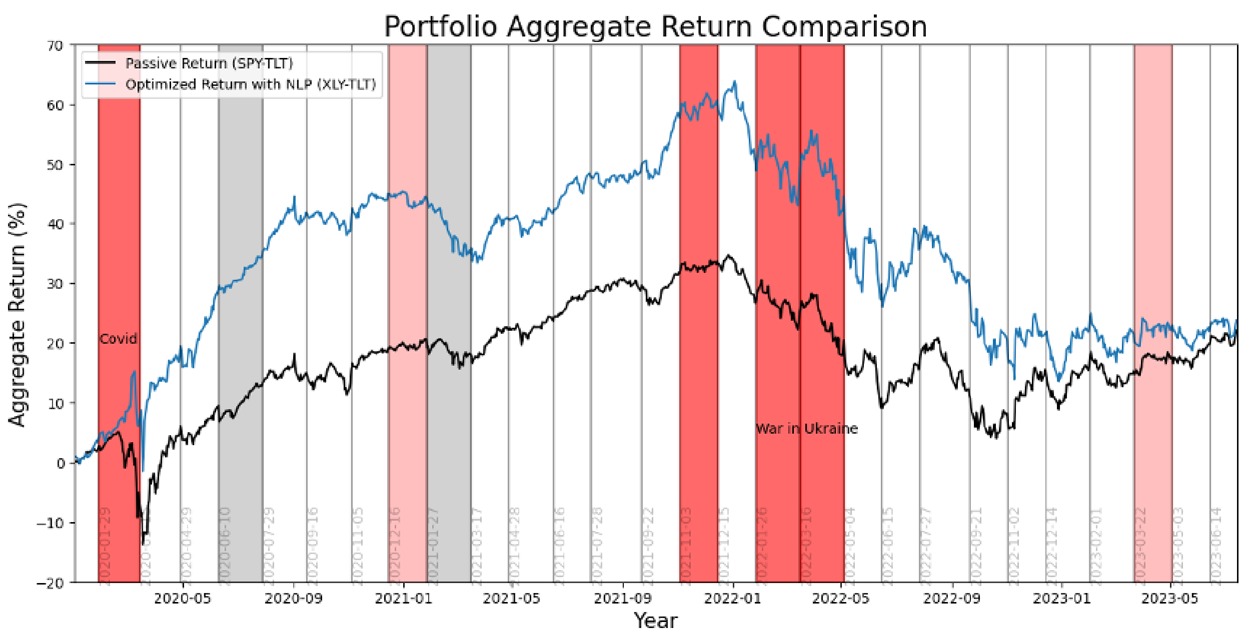


Figure 20: Risk-adjusted Performance - NLP Subsector (XLY-TLT) vs. Passive SPY-TLT

Portfolio	2020 Sharpe Ratio	2021 Sharpe Ratio	2022 Sharpe Ratio	3 Year Average
NLP (XLY-TLT)	2.64	1.29	-1.84	0.70
SPY-TLT	1.33	1.11	-1.71	0.24

Broader Impacts

Given that both the FOMC Minutes and stock price information are publicly available, there is no meaningful ethical risk in terms of privacy. However, it is possible that if we made our results available on Github or published an article that retail investors could make investment decisions that rely on the signals provided by our modestly sized dataset. This could lead to significant financial loss.

Statement of Work

Greg Holden	FOMC Document & Stock Price Data Collection, Portfolio Strategy Development, NLP Analysis, Report Writing & Editing
Jason Peloquin	Transformer Sentiment Analysis, Classifier/Prediction Models, Back testing & Performance Models, Report Writing & Editing.
Emmanuel Sengendo	NLP Analysis, Visualizations

References

Andersen, S. (2020, December 28). *Trading with MACD and RSI – Yfinance & Python*. Medium.com.
https://medium.com/analytics-vidhya/momentum-trading-with-macd-and-rsi-yfinance-python-e5203d2e1a8a

Anello, E. (2023, January 11). Topic Modeling Approached: Top2Vec vs BERTopic. KDnuggets.com.
https://www.kdnuggets.com/2023/01/topic-modeling-approaches-top2vec-bertopic.html

Aziz, Saqib & Dowling, Michael & Hammami, Helmi & Piepenbrink, Anke. (2019). Machine Learning in Finance: A Topic Modeling Approach. SSRN Electronic Journal. 10.2139/ssrn.3327277.

Cer, D., Yang, Y., Kong, S.-Y., Hua, N., Limtiaco, N., St. John, R., Constant, N., Guajardo-Céspedes, M., Yuan, S., Tar, C., Sung, Y.-H., Strope, B., & Kurzweil, R. (2018). *Universal Sentence Encoder*.

https://storage.googleapis.com/pub-tools-public-publication-data/pdf/67e012bfb04aac8e1ed088d792a13bd234b885aa.pdf

CFA Institute. (2023). *Sharpe Ratio. Chartered Financial Analyst Institute*. https://corporatefinanceinstitute.com/resources/risk-management/sharpe-ratio-definition-formula/

Chen, Y. (2019, February 26). *How to build a LDA Topic Model using from text*. Medium. https://yanlinc.medium.com/how-to-build-a-lda-topic-model-using-from-text-601cdcbfd3a6

Fernandez Monsalve, I., Frank, S., & Vigliocco, G. (2012). *Lexical surprisal as a general predictor of reading time* (pp. 398–408). Association for Computational Linguistics. https://aclanthology.org/E12-1041.pdf

Gayen, A. (2019, August 20). *Python | Clustering, Connectivity and other Graph properties using Networkx*. GeeksforGeeks. https://www.geeksforgeeks.org/python-clustering-connectivity-and-other-graph-properties-using-networkx/

Google Colaboratory. (n.d.). Colab.research.google.com. Retrieved August 3, 2023, from https://colab.research.google.com/drive/18arPPe50szvcCp_Y6xS56H2tY0m-RLqv?usp=sharing#scrollTo=zGAmTJp7CDHe

Grootendorst, M. P. (n.d.). *4. Vectorizers - BERTopic*. Maartengr.github.io. https://maartengr.github.io/BERTopic/getting_started/vectorizers/vectorizers.html#countvectorizer

Guo, P. (2020, April 9). *Predict Stock Market Trend using News Headlines – Part 4 topic modelling as features*. Medium.com. https://medium.com/swlh/predict-stock-market-trend-using-news-headlines-part-3-topic-modelling-as-features-64ce30db955b

Handlan, A. (2022, February 15). *Text Shocks and Monetary Surprises: Text Analysis of FOMC Statements with Machine Learning.* Department of Economics. Brown University. https://handlanamy.github.io/MyFiles/Handlan_TextShocks.pdf

Hashmi, F. (n.d.). *How to classify text using Word2Vec.* Thinkingneuron.com. https://thinkingneuron.com/how-to-classify-text-using-word2vec/

Hozjan, M. (2021, October 8). *Calculate RSI with python and yahoo finance*. wire.insiderfinance.io. https://wire.insiderfinance.io/calculate-rsi-with-python-and-yahoo-finance-c8fb78b1c199

Jansen, S. (n.d). Topic Modeling For Earnings Calls and Financial News. In *Machine Learning For Trading*. https://stefan-jansen.github.io/machine-learning-for-trading/15_topic_modeling/.

Ji, X., Wang, J. & Yan, Z. (2021). A stock price prediction method based on deep learning technology. *International Journal of Crowd Science*, 5(1), 55-72. DOI 10.1108/IJCS-05-2020-0012

Keane et al. (2022, June 12). *NLP on FOMC Meetings*. Medium.com. https://medium.com/@nusfintech.ml/nlp-on-fomc-meetings-50b48c447fe1

Khadka, N. (2023, April 21). Gensim 101: *A Beginner's Guide For Understanding and Implementing Topic Modeling*. Dataaspirant. https://dataaspirant.com/gensim-guide-topic-modeling/

Li, S. (2018, August 17). *Named Entity Recognition with NLTK and Spacy*. Towards Data Science. https://towardsdatascience.com/named-entity-recognition-with-nltk-and-spacy-8c4a7d88e7da

Lokare, G. (2023). *Effortless Sentiment Analysis with Hugging Face Transformers For Beginners*. Medium.com https://medium.com/@lokaregns/effortless-sentiment-analysis-with-hugging-face-transformers-a-beginners-guide-359b0c8a1787

Mattingly, W. (2022). *Introduction to Python for Digital Humanities*. CRC Press. python-textbook.pythonhumanities.com.

Muhammad, T., Aftab, A., Ahsan, M., Muhu, M., Ibrahim, M., Khan, S. & Alam, M. (2022, August 18). *Transformer-Based Deep Learning Model for Stock Price Prediction: A Case Study on Bangladesh Stock Market*. DOI:10.48550/arXiv.2208.08300.

Nadendla, H. (2020, October 11). *Why are LSTMs struggling to matchup with Transformers?* Medium.com.
https://medium.com/analytics-vidhya/why-are-lstms-struggling-to-matchup-with-transformers-a1cc5b2557e3

NLP for Developers: Word Embeddings | Rasa. (2020, March 16). Www.youtube.com.
https://www.youtube.com/watch?v=oUpuABKoElw&list=PL75e0qA87dlFJiNMeKltWImhQxfFwaxvv&index=19

Pierre, S. (2021, October 5). *An Introduction to Portfolio Optimization in Python*. Builtin.com. https://builtin.com/data-science/portfolio-optimization-python

Roepke, B. (2022, January 9). *How to Build NLP Topic Models to Truly Understand What Customers Want*. Data Knows All.
https://www.dataknowsall.com/topicmodels.html

Sonkiya, P., Bajpai, V. & Bansal, A. (2021). Stock price prediction using BERT and GAN. In Proceedings of ACM Conference (Conference'17), New York, NY, USA.
https://doi.org/10.1145/

Sadananda, N. (2021). *The battle between logistic regression, random forest classifier, xg boost and support vector* machines. Medium.com
https://medium.com/@nischitasadananda/the-battle-between-logistic-regression-random-forest-classifier-xg-boost-and-support-vector-46d773c70f41

Tadle, R. (2022, January 6). FOMC minutes sentiment and their impact on financial markets. *Journal of Economics and Business.*
https://www.sciencedirect.com/science/article/pii/S0148619521000394

Tripathi, A. (2020, April 27). *Named Entity Recognition NER Using spaCy*. Data Science Duniya.
https://ashutoshtripathi.com/2020/04/27/named-entity-recognition-ner-using-spacy-nlp-part-4/

Van Otten, N. (2023, February 15). *Tutorial TF-IDF vs Word2Vec For Test Classification*. Spotintelligence.com.
https://spotintelligence.com/2023/02/15/word2vec-for-text-classification/

Vukglisovic. (2017). *Classification combining LDA and Word2Vec.* Kaggle.com. https://www.kaggle.com/code/vukglisovic/classification-combining-lda-and-word2vec#3)-Model-Training-

Woroniuk, D. (2020, September 19). *FOMC Named Entity Recognition – Has Information Content Evolved?* Towards Data Science. https://towardsdatascience.com/fomc-named-entity-recognition-has-information-content-evolved-2da6da8abd88

Zvornicanin, E. (2023, May 31). *When Coherence Score is Good or Bad in Topic Modeling.* Baeldung.com. https://www.baeldung.com/cs/topic-modeling-coherence-score