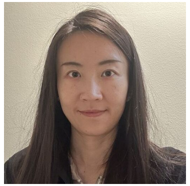
# Predicting The Impact of FOMC Communications on Market Returns

Team: Irene Na, William Lei



# Introduction and Background

- **The central bank plays a crucial role** in economic policy making, setting the tone.
- Federal Open Market Committee (FOMC) communication is **known for making a notable impact on the financial market**.
- Most analysis: focus **solely on numerical or a blend of numerical and non-numerical** (e.g. tones, sentiment) FOMC information to analyze the impact on the financial market.

<ins>**In this analysis,**</ins> we use transformer based text encoders to generate the representation of FOMC documents, and predict the impact on asset return in following days <ins>**directly**</ins>


**"Stock market will react to Fed meeting by scanning for views of the economy and future rate path."**

— CNBC

**"Dollar, yields rise ahead of speech by Fed's Powell | Reuters"**

**"CNBC Daily Open: Jerome Powell's speech was hawkish. Investors' mood was bullish"**

# Data Sources and Processing

## FOMC Documents

- FOMC Statement
- Conf Call Transcript
- Meeting Transcript
- Press Conf Transcript
- Meeting Minutes
- Economic Forecast
- Other documents

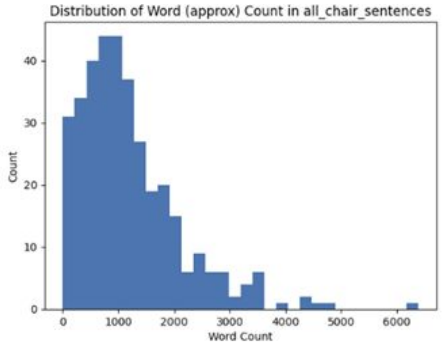
Period: 1980-2024
Train: 1980-2012
Val: 2013-2018
Test: 2019-2024
Caveat: practice changes over time

Off-script, timely release

## Main Processes

- Extract text by page
- Strip Unusable symbols
- Extract Useful Content using Regex
- Standardize content format
- Zoom in selective content within reason - Chair's talk

EDA to check token properties




| | all_chair_words |
|---|---|
| count | 356.000000 |
| mean | 1197.730337 |
| std | 917.775731 |
| min | 1.000000 |
| 25% | 565.000000 |
| 50% | 998.500000 |
| 75% | 1562.500000 |
| max | 6387.000000 |

Takeaways: median number of words from 'chairman' per document is ~1000 in sample, but with large std;

Caveat: we know that the main documents shortened since 2011

# Benchmark - BERT Classification and Regression Models

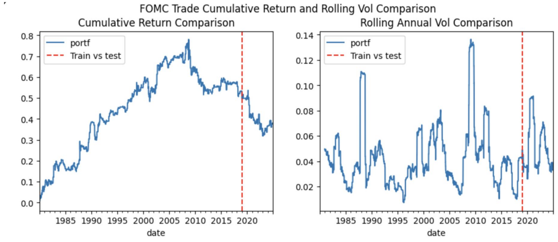
## Classification Task

**Goal:** Use BERT to predict the directional change in S&P 2-day return using FOMC statements.

**Key finding:** BERT achieves very modest improvements in directional accuracy, but deeper models struggle with generalization.

**BERT Base case - Classification**

| Classification | S&P as sole target | Layers | Batch Size | Epoch | Seed | Input token size | Keep Best Model (by val loss) | Train Accuracy | Val Accuracy | Backtest Sharpe Train-val | Backtest Sharpe Val |
|---|---|---|---|---|---|---|---|---|---|---|---|
| #1 | Yes | 3 | 8 | 10 | 42 | 512 | Y | 0.618768 | 0.458333 | 0.600497 | 0.414616 |
| #2 | Yes | 12 | 8 | 10 | 42 | 512 | Y | 0.59824 | 0.479167 | 0.400117 | 0.16101 |
| #3 | Yes | 12 | 5 | 10 | 42 | 512 | Y | 0.463343 | 0.458333 | 0.598815 | 0.143673 |
| #4 | Yes | 3 | 5 | 10 | 42 | 512 | Y | 0.583871 | 0.520833 | 0.772874 | 0.232797 |




| | sp_rtn_sign |
|---|---|
| **Train_accuracy** | 0.618768 |
| **Val_accuracy** | 0.458333 |
| **In-sample_accuracy** | 0.596144 |
| **Test_accuracy** | 0.583333 |

```
                 portf  Unscaled Market
avg_rtn_ann     0.008165        0.101741
vol_ann         0.049477        0.175313
sharpe_ann      0.165018        0.580337
max_drawdown   -0.383514       -0.567754
sharpe_ann_test -0.347016       0.794068
```

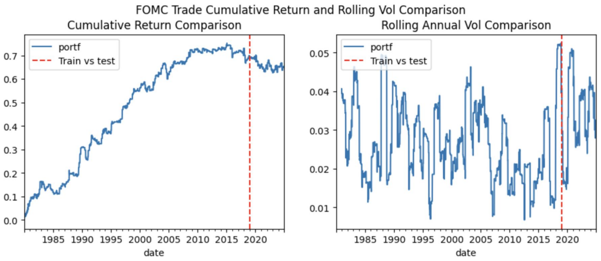
## Regression Task

**Goal:** Use BERT to predict the magnitude of S&P 2-day return using FOMC statements.

**Key finding:** BERT's most effective setup (3 layers, batch size 5, 10 epochs) amd with smaller models outperforming deeper ones in robustness.

**BERT Base case - Regression**

| Regression | S&P as sole target | Layers | Batch Size | Epoch | Seed | Input token size | Train Loss | Val Loss | Backtest Sharpe Train-val | Backtest Sharpe Val |
|---|---|---|---|---|---|---|---|---|---|---|
| #1 | Yes | 3 | 8 | 10 | 42 | 512 | 0.000700 | 0.000251 | 0.736001 | 0.480337 |
| #2 | Yes | 12 | 8 | 10 | 42 | 512 | 0.002400 | 0.000214 | 0.351572 | 0.002500 |
| #3 | Yes | 12 | 5 | 10 | 42 | 512 | 0.00059 | 0.000310 | 0.339012 | 0.269006 |
| #4 | Yes | 3 | 5 | 10 | 42 | 512 | 0.000253 | 0.000398 | 1.106904 | 0.354984 |




```
                 portf  Unscaled Market
avg_rtn_ann     0.014031        0.101741
vol_ann         0.029847        0.175313
sharpe_ann      0.470106        0.580337
max_drawdown   -0.122336       -0.567754
sharpe_ann_test -0.141078       0.794068
```

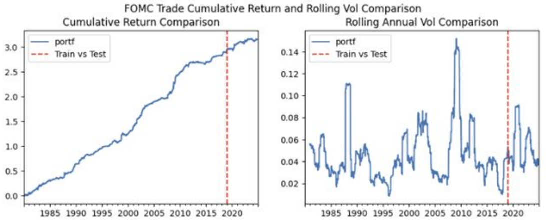
# Proposed - MordernBERT Classification and Regression

## Classification Task

**Goal:** see if MordernBERT could improve the directional prediction of S&P next 2-day return using FOMC transcripts post release. Various parameter / structural searching experiments were done.

**Key finding:** ModernBERT is able to improve both the accuracy and backtest Sharpe over BERT (in and out of sample). BUT, the improvement is not from extra input length.

| Classification Experiments | S&P as sole target | Trainable-layers | Batch size | Epochs | Seed | Input token size | Keep Best Model (by val loss)? | Train Accuracy | Val Accuracy | Backtest Sharpe-Train-Val | Backtest Sharpe-Val |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Y | 12 | 5 | 10 | 42 | 512 | Y | 0.8506 | 0.6458 | 1.3788 | 1.0270 |
| 2 | Y | 3 | 5 | 10 | 42 | 1000 | Y | 0.6408 | 0.5625 | 0.6273 | 0.2079 |
| 3 | Y | 3 | 5 | 10 | 42 | 512 | Y | 0.6782 | 0.5208 | 0.5900 | 0.1981 |
| 4 | Y | 3 | 5 | 10 | 42 | 2500 | Y | 0.6322 | 0.5208 | 0.6526 | 0.1842 |
| 5 | Y | 12 | 5 | 10 | 42 | 1000 | Y | 0.7414 | 0.5208 | 0.8794 | 0.1755 |
| 6 | Y | 12 | 5 | 10 | 42 | 2500 | Y | 0.5172 | 0.5417 | 0.1801 | (0.1397) |
| 7 | Y | 3 | 5 | 10 | 42 | 800 | Y | 0.5891 | 0.5833 | 0.5249 | (0.2354) |
| 8 | Y | 12 | 5 | 10 | 42 | 800 | Y | 0.5489 | 0.5417 | 0.3796 | (0.3140) |



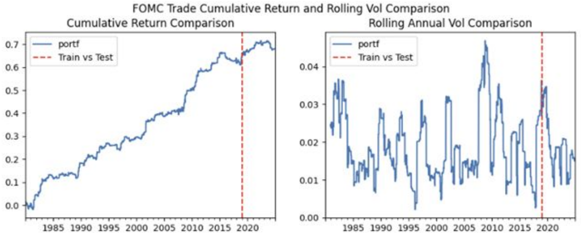

| | portf | Unscaled Market |
|---|---|---|
| avg_rtn_ann | 0.067623 | 0.101168 |
| vol_ann | 0.052944 | 0.175342 |
| sharpe_ann | 1.277252 | 0.576976 |
| max_drawdown | -0.056241 | -0.567754 |
| sharpe_ann_test | 0.569479 | 0.795467 |

## Regression Task

**Goal:** see if MordenBERT could improve the regression prediction of S&P next 2-day return using the FOMC transcripts post release. 40+ parameter / structural searching experiments were done.

**Key finding:** ModernBERT can improve out-of-sample loss and backtest Sharpe over BERT, but struggles to surpass the classification backtest. The change of Fed communication practice makes the more nuanced prediction harder to perform.

| Experiments | S&P + Treasury as Targets | Max Pooling as Seq Rep | CLS as Seq Rep | Trainable-layers | Batch size | Input token size | attention_dropout | attention_bias | embedding_dropout | num_hidden_layers (above base) | dropout ( above base) | Epochs (keep best) | Seed | Swapped val and test? | Train Loss | Val Loss | Backtest Sharpe-Train-Val | Backtest Sharpe-Val |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Y | Y | N | 3 | 5 | 2500 | - | FALSE | - | 2 | 0.500 | 20 | 42 | N | 0.000224 | 0.000207 | 0.7493 | 0.4150 |
| 2 | Y | Y | N | 12 | 5 | 2500 | 0.150 | TRUE | 0.351 | 3 | 0.233 | 20 | 123 | N | 0.000235 | 0.000213 | 0.4721 | 0.4987 |
| 3 | Y | Y | N | 12 | 5 | 512 | - | FALSE | - | 2 | 0.500 | 20 | 42 | N | 0.000247 | 0.000213 | 0.2781 | 0.3839 |
| 4 | Y | Y | N | 6 | 5 | 2500 | 0.242 | TRUE | 0.103 | 4 | 0.153 | 20 | 123 | N | 0.000222 | 0.000218 | 0.8576 | 0.3847 |
| 5 | Y | Y | N | 6 | 5 | 2500 | 0.127 | TRUE | 0.110 | 4 | 0.185 | 20 | 123 | N | 0.000226 | 0.000221 | 0.6913 | 0.5612 |
| 6 | Y | Y | N | 6 | 5 | 2500 | 0.111 | TRUE | 0.254 | 4 | 0.182 | 20 | 123 | N | 0.000229 | 0.000225 | 0.6306 | 0.2745 |
| 7 | Y | Y | N | 3 | 5 | 512 | - | FALSE | - | 2 | 0.500 | 20 | 42 | N | 0.000285 | 0.000225 | 0.1280 | (0.0215) |
| 8 | Y | Y | N | 6 | 5 | 2500 | 0.204 | TRUE | 0.305 | 4 | 0.106 | 20 | 123 | N | 0.000215 | 0.000226 | 1.0180 | 0.0838 |




| | portf | Unscaled Market |
|---|---|---|
| avg_rtn_ann | 0.014594 | 0.101168 |
| vol_ann | 0.020841 | 0.175342 |
| sharpe_ann | 0.70024 | 0.576976 |
| max_drawdown | -0.053615 | -0.567754 |
| sharpe_ann_test | 0.169096 | 0.795467 |

# Conclusion and Main Takeaways

## MAIN PROCESS

Experimented with both BERT and ModernBERT to conduct **directional and magnitude prediction** tasks for next 2-day S&P return, using FOMC transcripts around FOMC meetings.

1

## MAIN FINDINGS

Judging by data science and investing metrics, both types of models **demonstrated the ability to make meaningful predictions**.

ModernBert is able to improve BERT results.

2

## ANSWERS TO KEY RESEARCH QUESTIONS

There is meaningful information in FOMC communication, which can be extracted to predict the market return **directly**, with proper processing using advanced LLM technologies.

3

## FUTURE CONSIDERATIONS

The changes of Fed's communication could introduce challenges in using past data for future modeling.

Thus it is recommended to **stay informed** about these changes.

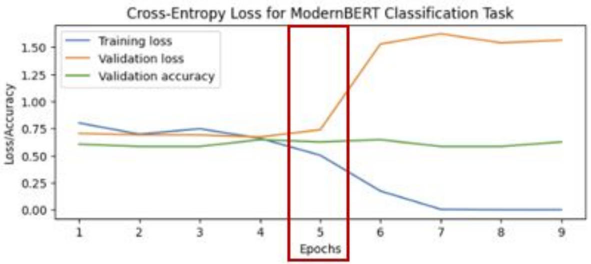
4

# Reference


1. Gorodnichenko, Yuriy and Pham, Tho and Talavera, Oleksandr, The Voice of Monetary Policy (March 2021).
2. Michael Ehrmann, Jonathan Talmi, Starting from a blank page? Semantic similarity in central bank communication and market volatility (May 2020).
3. Bilal Taskin, Fuat Akal, Tales of Turbulence: BERT-based Multimodal Analysis of FED Communication Dynamics Amidst COVID-19 Through FOMC Minutes (March 2024)
4. Asif M. Ruman, A Comparative Textual Study of FOMC Transcripts Through Inflation Peaks, Journal of International Financial Markets, Institutions & Money (August 2023)
5. Federal Reserve government website: https://www.federalreserve.gov/monetarypolicy/fomc_historical_year.htm

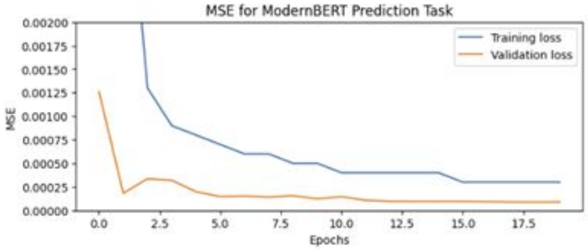
# Appendix I Chosen ModernBert Models' Metrics

## Metrics of Best Performing ModernBert Classification Model




| Metrics | sp_rtn_sign |
|---|---|
| Train_accuracy | 0.8506 |
| Val_accuracy | 0.6458 |
| In-sample_accuracy | 0.8258 |
| Test_accuracy | 0.4694 |

## Metrics of Best Performing ModernBert Prediction/Regression Model




| | next2d_sp_return | next2d_tsy10yr_diff | next2d_tsy5yr_diff |
|---|---|---|---|
| In-sample_direction_accuracy | 0.575758 | 0.494949 | 0.441919 |
| Test_direction_accuracy | 0.55102 | 0.367347 | 0.55102 |
| In-sample_MSE | 0.000222 | 0.000055 | 0.000036 |
| Test_MSE | 0.000331 | 0.000037 | 0.000028 |

# Appendix II Additional Experimented ModernBert Prediction Models

| Experiments | S&P + Treasury as Targets | Max Pooling as Seq Rep | CLS as Seq Rep | Trainable-layers | Batch size | Input token size | attention_dropout | attention_bias | embedding_dropout | num_hidden_layers (above base) | dropout (above base) | Epochs (keep best) | Seed | Swapped val and test? | Train Loss | Val Loss | Backtest Sharpe-Train-Val | Backtest Sharpe-Val |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Y | Y | N | 3 | 5 | 2500 | - | FALSE | - | 2 | 0.500 | 20 | 42 | N | 0.000224 | 0.000207 | 0.7493 | 0.4150 |
| 2 | Y | Y | N | 12 | 5 | 2500 | 0.150 | TRUE | 0.351 | 3 | 0.233 | 20 | 123 | N | 0.000235 | 0.000213 | 0.4721 | 0.4987 |
| 3 | Y | Y | N | 12 | 5 | 512 | - | FALSE | - | 2 | 0.500 | 20 | 42 | N | 0.000247 | 0.000213 | 0.2781 | 0.3839 |
| 4 | Y | Y | N | 6 | 5 | 2500 | 0.242 | TRUE | 0.103 | 4 | 0.153 | 20 | 123 | N | 0.000222 | 0.000218 | 0.8576 | 0.3847 |
| 5 | Y | Y | N | 6 | 5 | 2500 | 0.127 | TRUE | 0.110 | 4 | 0.185 | 20 | 123 | N | 0.000226 | 0.000221 | 0.6913 | 0.5612 |
| 6 | Y | Y | N | 6 | 5 | 2500 | 0.111 | TRUE | 0.254 | 4 | 0.182 | 20 | 123 | N | 0.000229 | 0.000225 | 0.6306 | 0.2745 |
| 7 | Y | Y | N | 3 | 5 | 512 | - | FALSE | - | 2 | 0.500 | 20 | 42 | N | 0.000285 | 0.000225 | 0.1280 | (0.0215) |
| 8 | Y | Y | N | 6 | 5 | 2500 | 0.204 | TRUE | 0.305 | 4 | 0.106 | 20 | 123 | N | 0.000215 | 0.000226 | 1.0180 | 0.0838 |
| 9 | Y | Y | N | 10 | 5 | 2500 | - | FALSE | - | 2 | 0.500 | 20 | 42 | N | 0.000248 | 0.000227 | 0.3660 | (0.1103) |
| 10 | Y | Y | N | 15 | 5 | 2500 | 0.148 | TRUE | 0.349 | 3 | 0.147 | 20 | 123 | N | 0.000233 | 0.000229 | 0.5781 | (0.1181) |
| 11 | Y | Y | N | 3 | 5 | 800 | - | FALSE | - | 2 | 0.500 | 20 | 42 | N | 0.000225 | 0.000234 | 0.6321 | (0.0560) |
| 12 | Y | Y | N | 12 | 5 | 1000 | - | FALSE | - | 2 | 0.500 | 20 | 42 | N | 0.000254 | 0.000239 | 0.5105 | (0.1190) |
| 13 | Y | Y | N | 12 | 5 | 800 | - | FALSE | - | 2 | 0.500 | 20 | 42 | N | 0.000357 | 0.000241 | 0.0461 | 0.0376 |
| 14 | Y | Y | N | 12 | 5 | 2500 | - | FALSE | - | 2 | 0.500 | 20 | 42 | N | 0.000272 | 0.000252 | 0.0831 | (0.4195) |
| 15 | Y | Y | N | 3 | 5 | 1000 | - | FALSE | - | 2 | 0.500 | 20 | 42 | N | 0.000217 | 0.000257 | 0.6212 | (0.6069) |
| 16 | Y | N | Y | 3 | 5 | 2500 | - | FALSE | - | 2 | 0.500 | 10 | 42 | N | 0.000272 | 0.000259 | 0.5208 | 0.4386 |
| 17 | Y | N | Y | 3 | 5 | 2500 | - | FALSE | - | 2 | 0.500 | 30 | 42 | N | 0.000234 | 0.000260 | 0.5516 | (0.0416) |
| 18 | N | Y | N | 3 | 5 | 2500 | - | FALSE | - | 2 | 0.500 | 30 | 42 | N | 0.000232 | 0.000266 | 0.4550 | (0.7145) |
| 19 | N | Y | N | 3 | 10 | 2500 | - | FALSE | - | 2 | 0.500 | 30 | 42 | N | 0.000245 | 0.000278 | 0.3798 | 0.0467 |
| 20 | N | N | Y | 3 | 5 | 2500 | - | FALSE | - | 2 | 0.500 | 10 | 42 | N | 0.000252 | 0.000328 | 0.3590 | (0.0248) |
| 21 | Y | Y | N | 3 | 5 | 2500 | - | FALSE | - | 2 | 0.500 | 20 | 42 | Y | 0.000230 | 0.000353 | 0.5285 | (0.1482) |
| 22 | N | Y | N | 3 | 5 | 2500 | - | FALSE | - | 2 | 0.500 | 20 | 42 | N | 0.000263 | 0.000368 | 0.4170 | 0.2392 |