Introduction

* Background and motivation
* Importance of finance-related tweets in influencing financial markets and investor behavior
* Objective of the study: Developing an effective classification model for finance-related tweets by utilizing state of the art Bert models to accurately categorize these tweets based on their topic.
* Our thesis: We believe that we can produce greater accuracy by 1) using domain-specific pre-trained models (i.e. FinBert) and 2) Fine-tuning these models utilizing well-studied methodologies for optimizing transformer models

Literature Review

* Overview of related research on sentiment analysis and topic classification in finance
* Introduction to language models and their role in text classification
* How previous works utilizing BERT-based models for text classification tasks have come up short
* Introduction to FinBERT and its advantages in finance domain

Data Collection and Preprocessing

* Description of the Twitter Financial News dataset from Kaggle
  + Contains 21,107 records with 20 unique labels that classify their topic.
  + This dataset has been labeled with finance-related topics, such as earnings, currencies, macro, and company news.
  + The training dataset provided contains 16,990 records and the validation (test) set contains 4,118.
* Preprocessing steps
  + Tweet text contains hyperlinks, emojis, &, along with other symbols = $
  + Removed only hyperlinks
  + Shortest tweet = 2, longest tweet = 316
  + Tweet length has a normal distribution.
* Label distribution analysis and data splitting for training and testing
  + Data is unbalanced across the 20 labels
  + Left data unbalanced during split to represent real-world imbalance and gain a realistic evaluation
  + We will be treating the validation set provided as our test set and creating another validation set from the training dataset.

Methodology

* Detailed explanation of BERT & FinBERT pretrain language model (PLM)
  + Description of the BERT-base model and comparison with FinBERT
  + Optimization strategies will be conducted on both models for comparison
* Explanation of fine-tuning and inter-training strategies from Sun et al. and Shnarch et al.
* Overview of the experiments combining and individually implementing these strategies

Experimental Setup

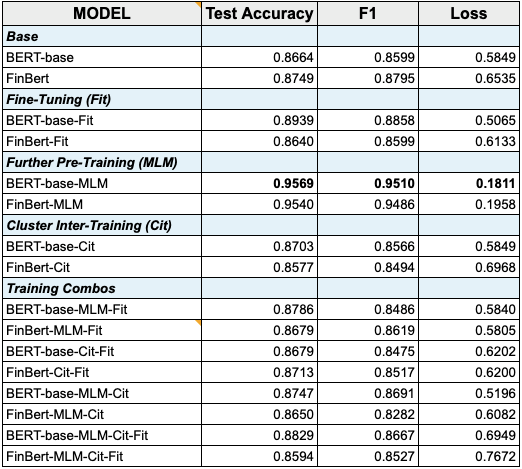
* Evaluation metrics for the classification model
  + Accuracy, F1 Score,Confusion Matrix
* Description of checkpoint: We’d build the individual models and run their evaluations, then we would export the checkpoint as a pre-trained model then load that model for the combo models

Models

* + Baseline:
  + BertBase:
  + FinBert:
  + Fine-Tuning:
  + Further Pre-Training:
  + Cluster Inter-Training:
  + Training Combos: Discuss the experiments that combine different training strategies and compare the results with individual approaches.
* Implementation details and hyperparameters used in the experiments

Results and Analysis

* Presentation of the experimental results (Need to update numbers)



* Comparison of FinBERT and BERT-base models
* Discussion of the performance improvements achieved by fine-tuning and inter-training strategies
* Comparison with the results of linear classifiers as presented in Lin et al.

Discussion

* Interpretation of the findings from the different models and strategies
* Analysis of the results and performance comparisons
* Discussion of the strengths and limitations of each model and strategy
* Implications of the results for the objective of accurately classifying finance-related tweets.

Conclusion

* Recapitulation of the objective and contributions of the paper
* Summary of the achieved results and their significance
* Future directions for further improvement and research in this domain
  + GPU /RAM, Time, Access to a larger Dataset

References

* Yang, Yi & UY, Mark & Huang, Allen. (2020). *FinBERT: A Pretrained Language Model for Financial Communications.* https://www.researchgate.net/publication/342198406\_FinBERT\_A\_Pretrained\_Language\_Model\_for\_Financial\_Communications
* Lin, Y.-C., Chen, S.-A., Liu, J.-J., & Lin, C.-J. (2023, June 12). *Linear Classifier: An Often-Forgotten Baseline for Text Classification.* https://arxiv.org/abs/2306.07111
* Sun, C., Qiu, X., Xu, Y., & Huang, X. (2019, May 14). *How to Fine-Tune BERT for Text Classification?*<https://arxiv.org/abs/1905.05583>
* Shnarch, E., Gera, A., Halfon, A., Dankin, L., Choshen, L., Aharonov, R., & Slonim, N. (2022, March 20). *Cluster & Tune: Boost Cold Start Performance in Text Classification.* <https://arxiv.org/abs/2203.10581>

Appendix

* Hardware and software specifications
  + Google Colab with a paid credits to increase Google Compute Engine (GPU) performance.
  + Python 3, huggingface transformers, pytorch, keras (not sure we need this)