MULTI-TASK LEARNING ON FINBERT FOR FINANCIAL FINE-GRAINED SENTIMENT ANALYSIS

Christelle George , David Lemieux , Marc-Antoine Provost

INTRODUCTION

Context

- Era of fast and continuous information flow
- Importance of capturing and extracting the right information in finance as it can determine future stock prices
- Sentiment analysis used by traders, portfolio managers and investors

Challenges

- Lack of labeled data
- Domain-specific vocabulary
- Ineffectiveness of traditional models such as BERT

Hypothesis

- Multi-task learning approach offers better performance compared to original FinBERT model in the financial field
- The unfreezing strategy should be part of the hyper-parameter tuning step as it is expected to boost the performance

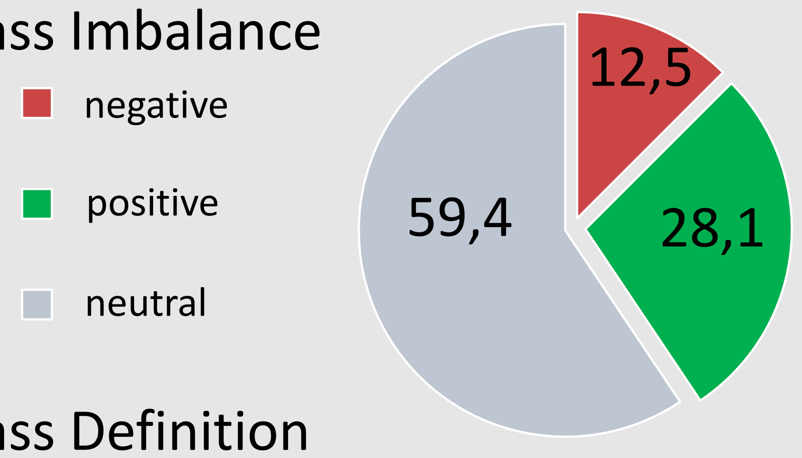
KEY CONTRIBUTIONS

- Multitask Learning approach with FinBERT
- Impact of different strategies of de-freezing layers
- Weight & biases platform for hyper parameter tuning

DATASET

FINANCIAL PHRASE BANK

- Class Imbalance



- Class Definition

Net sales increased by 5.2 % to EUR 205.5 mn, and operating profit by 34.9 % to EUR 23.5 mn → Profit

Deliveries are to start later in 2010, and the volume will increase in the years 2011-2012

Compared with the FTSE 100 index, this was a relative price change of -0.4 % → Loss

- Train-Val-Test Split : 80 /10 /10

METHODOLOGY

Benchmark

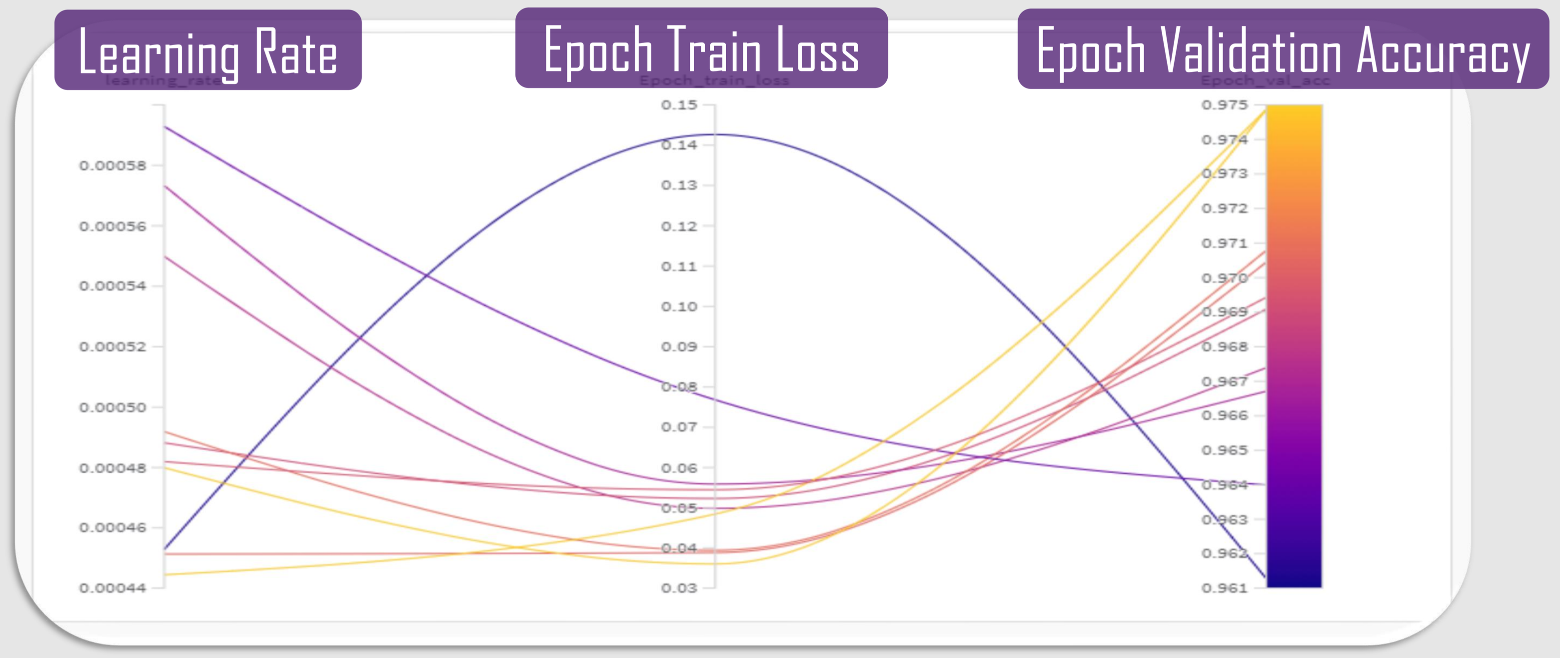
RNN GRU

- Pre-trained word embeddings GloVe
- Embedding's dimension 300
- 25 epochs

Hyper Parameter Tuning

FinBERT

- Optimizer Scheduler
- Learning rate
- Batch Size



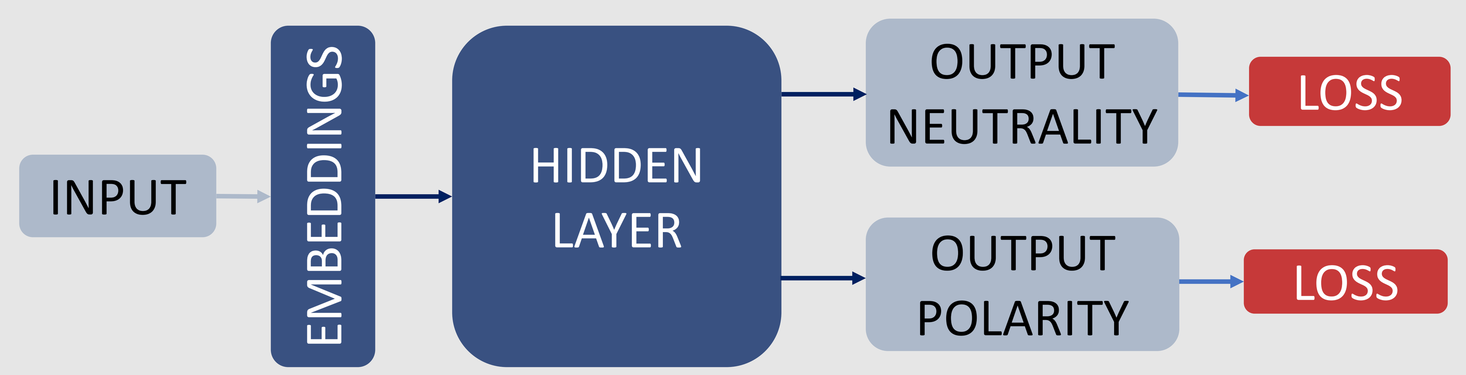
De-Freezing

FinBERT

- The best learning rates from the previous original FinBERT model is used
- Analyze multiple unfreezing strategies as a good structure can provide extra performance :
  no freezing / all freezing / ascending freezing / descending freezing / random layer freezing

Multi-Task

- Drop the output layer of the regular FinBERT
- Add a multi head with the following approach :
  1- Determine the neutrality of the data (presence of sentiment)
  2- Determine the polarity of the data (positive or negative)
- Non-deterministic approach for the train data



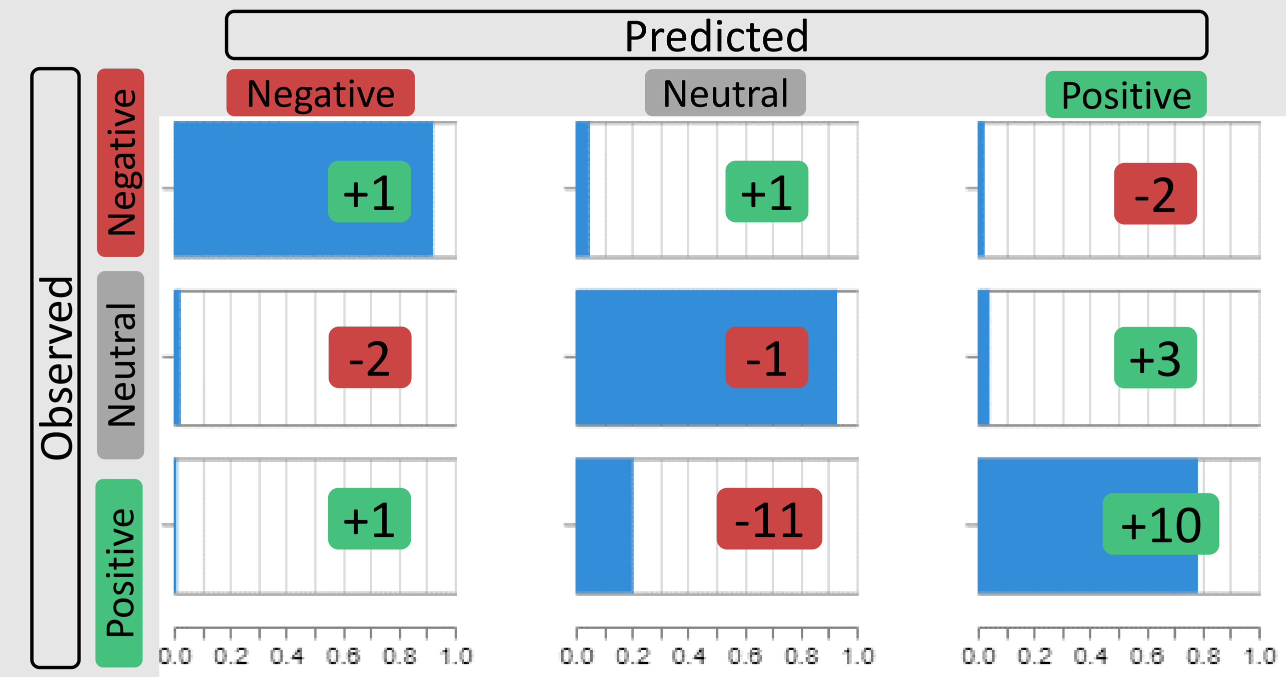
RESULTS

- **Experimental results on the Financial PhraseBank dataset**

Model	Accuracy	F1 score		
		Negative	Neutral	Positive
GRU - Benchmark	62,30	1,20	45,30	3,40
FinBERT - Paper	86	84 (Global)		
FinBERT - Single Task	87,92	80,89	88,62	90,74
FinBERT - Multi Task	88,96	83,07	89,43	91,42

Superiority of the Multitask Learning approach with the FinBERT with all performance metrics

- **Confusion matrix of the Multitask Learning FinBERT**



- **Experimental results of different freeze strategies**

Strategy	Accuracy	F1 Score
Freeze all	89,17	88,02
Descending freezing	87,60	86,88
Random layer freezing 1	88,33	85,21
Random layer freezing 2	86,46	87,11
Random layer freezing 3	87,29	85,89
No freezing	59,58	24,88
Ascending freezing	59,58	24,88

Freeze all → Optimal Strategy

No freezing, Ascending freezing → Catastrophical forgetting

RECOMMENDATIONS

- Broader hyper parameter search for model selection
- Data augmentation or loss weighting for the negative and neutral classes to tackle the class imbalance
- Hierarchical approach for the multi-task learning

REFERENCES

- https://github.com/dlemieux89/Deep_Learning2
- Yifan Peng, Qingyu Chen, and Zhiyong Lu. “An Empirical Study of Multi-Task Learning on BERT for Biomedical Text Mining”.2020.
- Georgios Balikas, Simon Moura, and Massih-Reza Amini. “Multitask learning for fine-grained twitter sentiment analysis”. 2017.
- Pekka Malo et al. “Good debt or bad debt: Detecting semantic orientations in eco-nomic texts”. 2014.