IMNTPU at the NTCIR-17 FinArg-1 Argument-based Sentiment Analysis and Identifying Attack and Support Argumentative Relations in Social Media Discussion Threads







Chia-Tung Tsai[1], Wen-Hsuan Liao[1], Hsiao-Chuan Liu[1], Vidhya Nataraj[2], Tzu-Yu Liu[3], Mike Tian-Jian Jiang[4] and Min-Yuh Day[1,*]

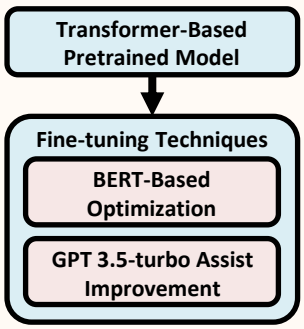
*myday@gm.ntpu.edu.tw

[1]Information Management, [2]Smart Healthcare Management, [3]Business Administration,
National Taipei University, New Taipei City, Taiwan

[4]Zeals Co., Ltd. Tokyo, Japan

In recent years, there has been a surge of interest in argument-based sentiment analysis and the identification of argumentative relationships in social media. These tasks encompass sentiment analysis of premises and claims, as well as the classification of argumentative relationships. Within these tasks, we have developed a fine-tuning method for transformer models. To evaluate and showcase this concept, we established a comprehensive framework to test and display the performance of BERT, RoBERTa, FinBERT, ALBERT, and GPT 3.5-turbo models on financial data and social media texts. Ultimately, the experimental results of these sub-tasks validate the effectiveness of our strategies. The primary contribution of our research is our proposal of two key elements: fine-tuning predominantly with BERT models and employing GPT for generative classification, aiming to enhance the identification of argumentative classifications. Through fine-tuning techniques, the state-of-the-art models can achieve better performance than the baseline.

IMNTPU Research Architecture



Fine-tuning Techniques

- Our research in Natural Language Processing (NLP) explores deep learning models like BERT, ALBERT, and RoBERTa for sentence classification. RoBERTa, in particular, shows superior performance in NLP tasks due to more data and extended training, refining BERT's original training approach. The study used RobertaTokenizer for tokenization and RobertaForSequenceClassification for training and evaluation.
- A 5-fold cross-validation technique was employed to fine-tune and assess model performance, involving dividing the dataset into five parts and using each in turn for validation. This ensures a stable and reliable performance evaluation. The study also adjusted hyperparameters such as sentence length, batch size, and training epochs to improve learning efficiency. For fair comparison, the same settings were applied to both RoBERTa and BERT models during fine-tuning.

GPT Generation Strategies and Optimization

- In our study, we demonstrate the application of OpenAI's ChatGPT API, integrating deep learning with Natural Language Processing (NLP) for detailed text analysis. The technology is finely tuned to classify sentences accurately as either "claim" or "premise", aiding researchers in identifying core arguments and their supporting reasons. This classification is part of a multi-step process, with specific sentiment labels providing clear targets for the model.
- The distinction between "claim" and "premise" is vital for understanding arguments and their justifications. Moreover, the method's scalability and adaptability make it versatile, suitable for not only basic sentiment analysis but also for more complex text analysis with additional classification labels.

Hyperparameter Settings

NTCIR-17 FinArg-1 Hyperparameter Settings

Hyperparameter	Value
Learning Rate	1e-5, 5e-5
Max Length	128, 256
Batch Size	8, 16
Epochs	3, 4, 5

Performance

NTCIR-17 FinArg-1 Argument Unit Classification

Model	Micro-F1	Macro-F1	Weight-F1	Accuracy
IMNTPU-1 (BERT-base)	75.44%	75.31%	75.40%	74.82%
IMNTPU-2 (RoBERTa-base)	**76.06%**	**76.05%**	**76.07%**	**75.64%**
IMNTPU-3 (GPT 3.5-turbo)	56.97%	56.82%	56.70%	55.08%

NTCIR-17 FinArg-1 Argument Relation Detection and Classification

Model	Micro-F1	Macro-F1	Weight-F1	Accuracy
IMNTPU-1 (RoBERTa-base)	78.99%	47.36%	76.54%	78.55%
IMNTPU-2 (FinBERT)	**82.61%**	**52.97%**	**82.14%**	**79.13%**
IMNTPU-3 (BERT-uncased)	80.72%	50.73%	79.67%	78.55%

NTCIR-17 FinArg-1 Identifying Attack and Support Argumentative Relations in Social Media Discussion Threads

Model	Micro-F1	Macro-F1	Weight-F1
IMNTPU-1 (Finetuned-Albert)	**52.88%**	**34.77%**	**48.73%**
IMNTPU-2 (RoBERTa-Large)	48.71%	24.64%	40.50%

Conclusions and Contributions

- We combined fine-tuning BERT and RoBERTa with the innovative use of GPT 3.5 Turbo, effectively capturing subtle nuances in conversational texts while demonstrating significant performance in generative tasks.
- Our study offers a comprehensive solution to the Argument Unit Classification challenge, thoroughly evaluating various methods' pros and cons. Additionally, in the multi-class classification task of financial sentiment analysis, we've revealed deeper semantic aspects of texts by analyzing inter-sentential relationships.

ACKNOWLEDGMENTS

This research was supported in part by the National Science and Technology Council (NSTC), Taiwan, under grants MOST 110-2410-H-305-013-MY2, NSTC 112-2425-H-305-002-, and NSTC 112-2627-M-038-001-, and National Taipei University (NTPU), Taiwan under grants 112-NTPU-ORDA-F-003, 112- NTPU-ORDA-F-004, USTP-NTPU-TMU-112-01, NTPU-112A413E01, and NTPU-112A513E01.