TMUNLP at the NTCIR-17 FinArg-1 Task

Heng-Yu Lin, Eugene Sy, Tzu-Cheng Peng, Shih-Hsuan Huang, Yung-Chun Chang*
Graduate Institute of Data Science, Taipei Medical University

Introduction

In FinArg-1, the novel concept of identifying arguments within financial narratives was introduced, subsequently applied to argument-based sentiment analysis. Simultaneously, dedicated efforts were directed towards another task: discerning argumentative relationships within social media discussions. The amalgamation of these methodologies empowers us to delve deeply into the diversity and intricacies of argument structures within the financial domain.

Method

In ECC ARGUMENT UNIT SUBTASK, we applied conventional preprocessing, suitable for pretrained language models, to the unified ECC Dataset. Six widely employed pretrained language models were trained and evaluated. In ECC ARGUMENT RELATION SUBTASK, We use pretrained models and the Log Likelihood Ratio (LLR) to measure word relationships. To address imbalanced data challenges, we employ Random Sampling, Class Weighting, and SMOTE. In Social Media ARGUMENT RELATION SUBTASK, We employed a simple duplication method for both Post1 and Post2 to enhance data diversity, strengthening the model's ability to capture textual correlations and contextual features. Additionally, using a Large Language Model, we extracted five relevant finance-related keywords from each post, combining them into a string to form a novel training feature. These strings and features were then incorporated into a large language model for training.

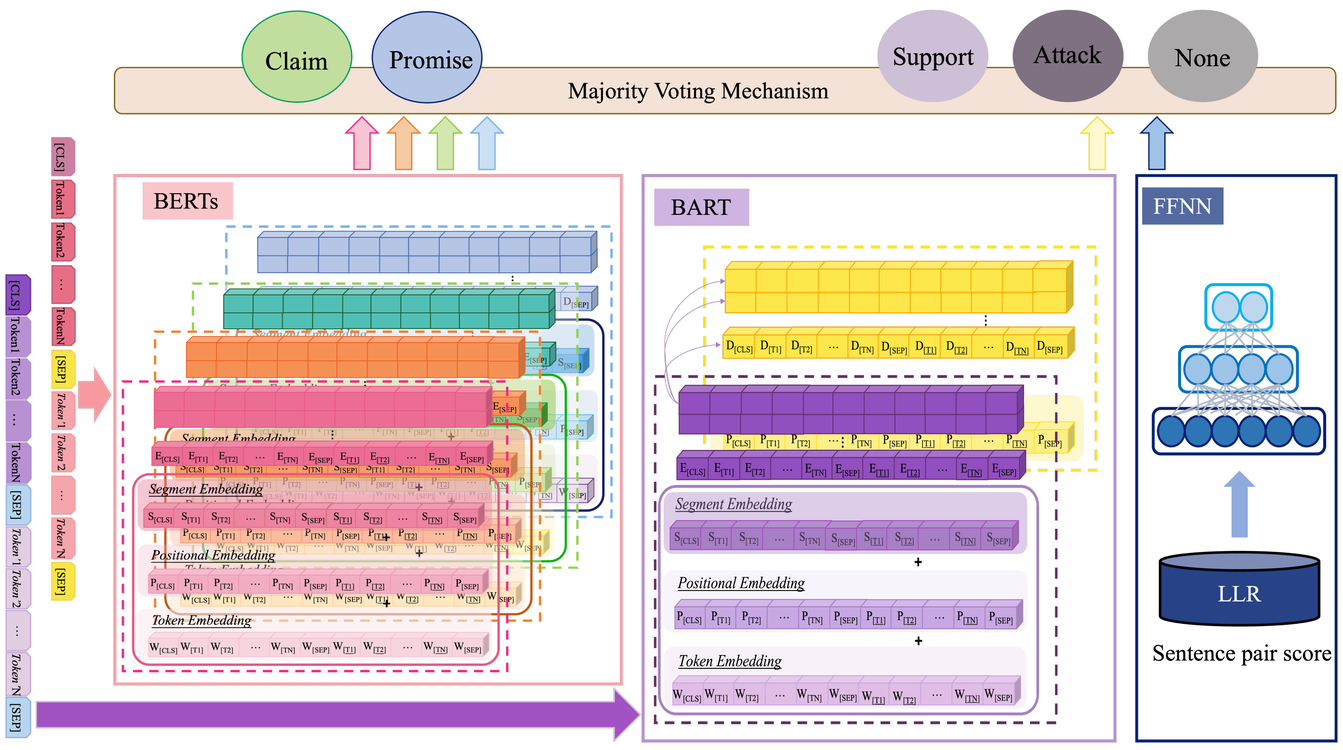

Figure 1. Architecture of ECC ARGUMENT UNIT (Left) and ECC ARGUMENT RELATION (Right)

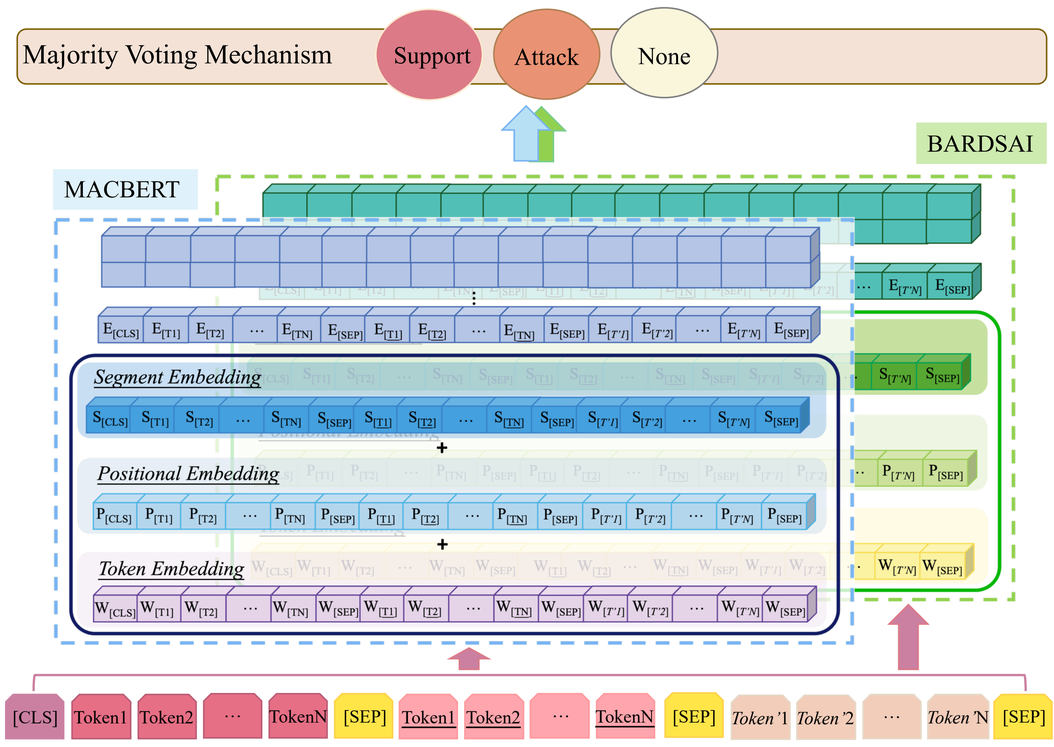

Figure 2. Architecture of Social Media Argument Relation

Result

For the ECC ARGUMENT UNIT SUBTASK, the model employing hard voting on a combination of top 4 model achieved the best performance, with a macro F1-score of 76.55% and a micro F1-score of 76.57%. In the ECC ARGUMENT RELATION SUBTASK, the top-performing model utilized soft voting on a combination of BART (CL) & LLR (RS), yielding a macro F1-score of 57.90% and a micro F1-score of 82.03%. Additionally, for the ECC Social Media ARGUMENT RELATION SUBTASK, the leading model employed soft voting on a combination of Macbert added feature & Bardsai added feature, achieving a macro F1-score of 73.12% and a micro F1-score of 73.13%.

TASK	ECC ARGUMENT UNIT		ECC ARGUMENT RELATION		Social Media Argument Relation	
Submission ID	Macro-F1 (%)	Micro-F1 (%)	Macro-F1 (%)	Micro-F1 (%)	Macro-F1 (%)	Micro-F1 (%)
[TMUNLP1]	1 **76.551**	**76.574**	81.88	57.36	71.10	67.61
[TMUNLP2]	75.826	75.851	3 **82.03**	**82.03**	**73.39**	**73.13**
[TMUNLP3]	76.036	76.058	57.90	57.90	73.12	70.18

*Bolded Fonts are the achieve a high rank on its respective task

Conclusion

Compared to other competing teams, our approach demonstrates superior performance, validating that (1) the voting algorithm contributes to enhancing and balancing the effectiveness across integrated models, (2) random sampling and class weighting effectively address data imbalance issues, (3) the Log Likelihood Ratio (LLR) method improves the capture of relationships between sentences, and (4) the integration of keyword features generated by Large Language Model enhances the performance of all models, emphasizing the broad potential of LLM in classification problems.

Acknowledgement

This study was supported by the National Science and Technology Council under grant NSTC-112-2922-I-038-086 and 112-2622-E-038-001- and 112-2410-H-038-007-, as well as the National Health Research Institutes NHRI-12A1-PHCO-1823244.

Scan the QR-code to know us more