An Institute of

Teamwork Dimensions Classification Using BERT

Junyoung Lee[a] and Elizabeth Koh[b]

[a]Nanyang Technological University, Singapore
[b]National Institute of Education, Singapore

1 Introduction

To provide a formative assessment of student teamwork, an automated natural language processing approach was developed to identify teamwork dimensions of students' online team chat, using traditional machine learning methods [1, 2]. The teamwork dimensions are focused on the following areas in Figure 1.

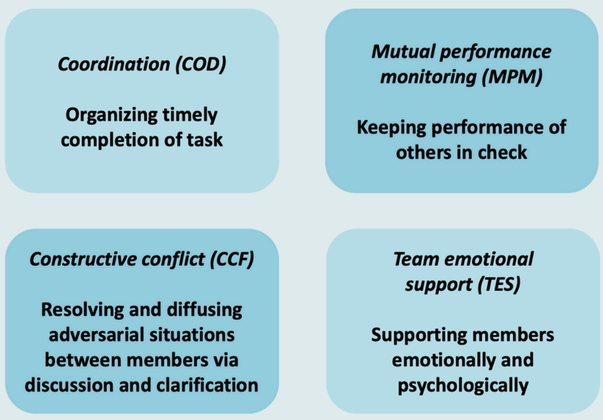

Figure 1. Framework for teamwork competency, derived from [3]

2 Current Challenges

The existing backend model and current training approach face the following challenges:

1. Resource-heavy data labelling
2. Generalizability of machine learning model
3. Improving classification performance

3 Objectives

To improve the existing backend model - random forest (RF) classifier with TF-IDF vectorizer - for teamwork classifications using deep transfer learning techniques, namely Bidirectional Encoder Representations from Transformers (BERT), in terms of classification performance and generalizability.

The approach considers the recent success of pre-trained language models in natural language processing tasks.

4 Dataset and Pre-processing

Chatlog of 19762 messages gathered from 76 teams of a total of 272 14-year-old students in a Singapore secondary school. Table 1 shows a sample of the human-annotated dataset from a past study for the four teamwork dimensions.

Table 1. Sample of annotated dataset

User	Message	COD	MPM	CCF	TES
Student A	Bob are you okay with it	1	1	0	0
Student B	ideal teacher as like um can work well with students and listen to students ideas	0	0	1	0
Student C	yes	0	0	0	1
Student C	So, caring, attentive to our needs and humourous?	0	0	1	0
Student A	humour yes	0	0	0	1
Student B	ok	0	0	0	1

Text pre-processing was carried out, with focus on textese and local terms replacement. Another approach previously taken to improve classification performance was feature engineering on the text data [2], hand-crafted from language cues related to the teamwork dimensions.

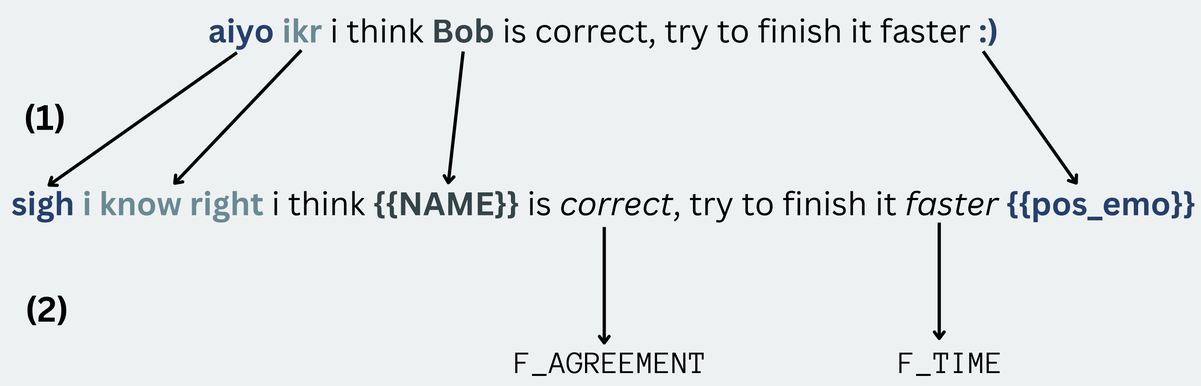

Figure 2. Example of (1) text pre-processing and (2) feature engineering

5 Results

A BERT-based classifier was developed, and its classification performance was compared on both the feature engineered and non-feature engineered versions of the corpus, as seen in Table 2.

Table 2. Classification performance comparison

	Without feature engineering		With feature engineering	
	RF	BERT	RF	BERT
Precision	**0.824**	0.810	**0.833**	0.801
Recall	0.406	**0.702**	0.594	**0.747**
F1 Score	0.527	**0.734**	0.684	**0.757**
Hamming distance	0.124	**0.076**	0.090	**0.070**

Inference on Unseen Data
Chatlog of 129 messages gathered from 2 groups of 15 students aged 19 to 42, manually annotated, pre-processed using the steps in Table 2 with no additional features. The trained BERT-based classifier achieved a Cohen's Kappa score of **0.640** compared to 0.467 achieved by the RF classifier, showing a higher inter-rater reliability.

6 Conclusion

The project demonstrates potential for **scalability** of such models **without need for extensive feature engineering**, as well as **generalizability** of the model with different demographics. This ultimately contributes towards an enhanced formative learning analytics tool for teamwork assessment, allowing students and teachers to become more aware of the teamwork competencies in the discussion process and receive more holistic feedback.

Future works include further in-depth analysis of task context, as well as more complex deep learning architecture or advanced language models.

Acknowledgements
This paper refers to data and analysis from OER62/12EK and OER09/15EK, funded by the Education Research Funding Programme, National Institute of Education (NIE), Nanyang Technological University, Singapore and the Incentivising ICT use Innovation Grant (I3G 9/19EK) from NIE. IRB: IRB-2020-08-006. This project was supported by Nanyang Technological University under the URECA Undergraduate Research Programme, and the CN Yang Scholars Programme.

References

1. Shibani, A., Koh, E., Lai, V., Shim, K.J.: Assessing the Language of Chat for Teamwork Dialogue. Journal of Educational Technology & Society 20(2), 224–237 (2017)
2. Suresh, D., Lek, H.H., Koh, E.: Identifying teamwork indicators in an online collaborative problem-solving task: A text-mining approach. In: Yang, J.C., Chang, M., Wong, L.H., Rodrigo, M.M.T. (eds.) Proceedings of the 26th International Conference on Computers in Education (ICCE) 2018. pp. 39–48. Asia-Pacific Society for Computers in Education (APSCE), Manila, Philippines (2018)
3. Koh, E., Hong, H., Tan, J.P.L.: Formatively assessing teamwork in technology-enabled twenty-first century classrooms: exploratory findings of a teamwork awareness programme in Singapore. Asia Pacific Journal of Education 38(1), 129–144 (2018)