Leveraging Speech Data and Text Classification for Identifying Socially Shared Regulation in Collaborative Problem

Presenter: Wenting Sun

Wenting Sun, Jiangyue Liu

Humboldt-Universität zu Berlin, Germany
Suzhou University, China

Introduction

- With the shift in Computer-Supported Collaborative Learning (CSCL) from the perspective of individual learners to groups of learners (Lämsä et al., 2021), the transition of metacognition from individualistic models to socially situated models has become increasingly prevalent in research (Vaughan et al., 2020).

- To deeply explore essential activities, accumulating research in engineering education has stressed the value of learners' conversation during CPS. For example, through epistemic network analysis, Lyu et al. (2023) found that different engineering design behavioural patterns under two instructional approaches. These studies demonstrated the great value of a fine-grained analysis of learners' collaborative learning process when using CPS in engineering courses.

- The popularity of ChatGPT has significantly increased researchers' interest in exploring text classification techniques in online discussions. For example, Flor and Andrews-Todd (2022) combined n-grams and various machine learning (ML) algorithms to automatically classify Collaborative Problem Solving (CPS) events based on text-chat messages. To the best of our knowledge, research on using text classification techniques to classify SSR phases remains scarce, particularly using speech data collected in the context of authentic CPS practice courses.

Research questions

As a classical problem in natural language processing (NLP), text classification targets at assigning labels or tags to textual units (Minaee et al., 2021). Most researchers using text classification in CSCL involved large datasets from online discussion. For example, in English language, using 19,105 sentences from online inquiry-based discussion, Ba et al. (2023) employed different models to classify cognitive presence, social presence, and teaching presence, like Logistic Regression (LR), Naïve Bayes (NB), Random Forest (RF), and Bidirectional Encoder Representations from Transformers (BERT).

It is understudied that how these models perform in a small speech transcription dataset from authentic face-to-face engineer practice course. Neither about in the SSR phase identification.

In summary, several research gaps still exist in the research about sequential SSR behaviours in authentic face-to-face engineering practice settings. Based on these, some research questions were proposed:

- RQ1: Compared to other CPS environments, what other activities can be observed in authentic face-to-face practice courses?
- RQ2: To what extent can text classification techniques identify different SSR phases?

Methods

- **Participants:** 36 undergraduates participated from a Chinese public university formed 18 dyads groups. This lesson was the practical course (one week) after their theoretical course (four weeks).

- **Context**: In the CPS environment, the learning activities were IP sending and receiving using the Internet Control Message Protocol (ICMP) package. All groups had simple CSCL scripts to support them monitor their progress. Group members were required to use their own computers to take turns being the IP sender and receiver.

Methods

- **Data collection:** The session lasted nearly 3 h 50 minutes and the operation session lasted approximately two hours with an instructor present throughout. Speech dialogues were recorded during the operation session.

- **Data analysis**:

Thematic analysis: the speech recordings data were transcribed manually, and then content analysis and qualitatively coded transcription were conducted (Braun & Clarke, 2006). Two researchers independently coded the data, with one coding 20% of the data and the other coding all of them. The agreement of the coding results achieved 80%.

Text classification: 5 SSR phases were used as labelling in our dataset, namely Orientation, Planning, Support strategies, Monitoring, Evaluation and reflection. The dataset was split into training and test datasets at 8:2 ratio. Several common text classifiers were built, including Logistic Regression (LR), Naïve Bayes (NB), Random Forest (RF), k-nearest-neighbours (KNN), Support Vector Machine (SVM), XGBoost, BERT.

Methods

Feature extraction used two steps, including "Jieba" Chinese text segmentation and Term Frequency-Inverse Document Frequency (TF-IDF). Then the outputs were put into the text classifier. This study performed all text classification processing with Python and its available package except BERT model used TensorFlow. For evaluation metrics, accuracy, precision, recall and F1 score were included.

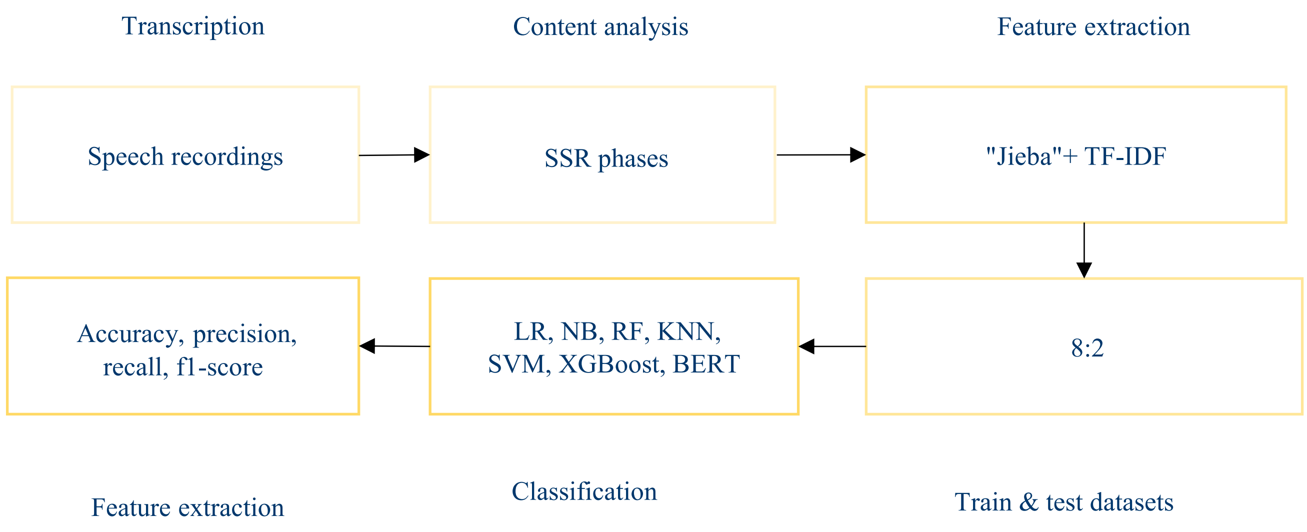

Figure 1. Speech recordings analysis procedure.

Results

Considering the research questions, the features of learning materials and peer interaction, the coding instrument from De Backer et al. (2016) was employed as the initial version of code schemes. After multiple rounds of listening to the speech recordings and discussion with the course instructor and coauthors, based on De Backer et al. (2016) framework, we developed the main categories of the coding scheme of the SSR activities in CPS engineer practice course with several subcodes respectively.

Table 1 demonstrate the coding schemes that include the Orientation, Planning, Support strategies, Monitoring, as well as Evaluation and reflection phases and further activities of SSR (the content in italics is new content added in this study). Here are examples:

SSR phases	Event	Activities
Orientation	Task Analysis	Exploring task demands
		Processing task demands/learning objectives
	Content Orientation	Generating hypotheses
		Activating prior knowledge
		Becoming aware of task perceptions
Planning	Planning in advance	Formulating problem solving plan (planning in advance)
		Selecting problem solving plan (planning in advance)
	Interim Planning	***Formulating problem solving plan repeated (interim planning)***
		Formulating problem solving plan new (interim planning)
		Peers' formulating problem solving plan new (interim planning)
		Teacher's formulating problem solving plan new
		Selecting problem solving plan
		Questioning the problem solving plan
Support strategies	***Peer Interaction***	***Asking for peers' support***
		Replying to peers' help asking
		Asking for teacher's support
	Online Searching	***Searching Online solutions***

Results

Among the 7 classifiers (Table 2), SVM model overall produced better performance (Confusion matrix heatmap can be seen in Figure 2). The model trained with TF-IDF features outputs the accuracy and recall of 0.67, precision of 0.66.

Classifier	Accuracy	Precision	Recall	F1-score
LR	0.66	0.65	0.66	0.59
NB	0.64	0.5	0.64	0.5
RF	0.66	0.62	0.66	0.59
KNN	0.64	0.61	0.64	0.57
SVM	0.67	0.66	0.67	0.59
XGBoost	0.66	0.62	0.66	0.62
BERT	0.64	0.48	0.64	0.5

Table 2. Summaries of model performance.

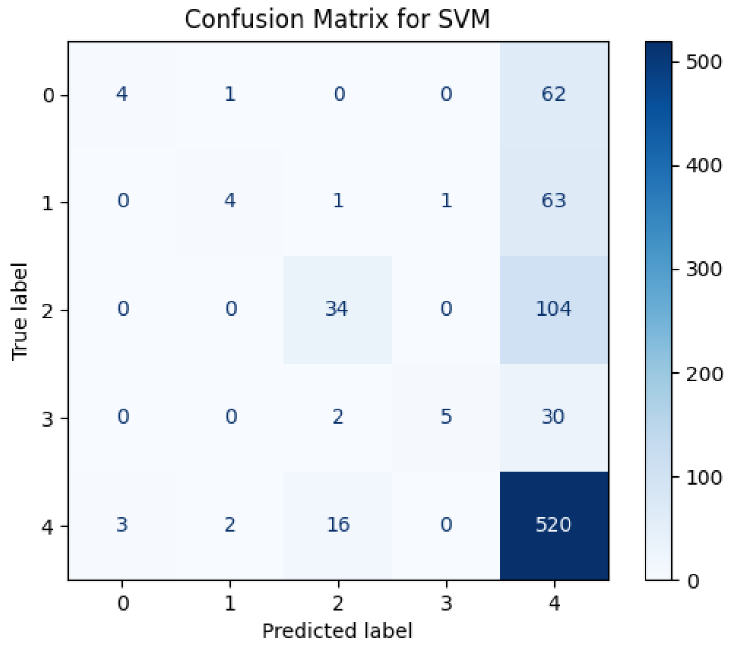

Figure 2. Confusion matrix heatmap of (SVM).

Discussions

- Based on the existing SSR phases, another 16 activities were found in the authentic face-to-face practical course in this study. In terms of code schemes used for specific tasks at the fine-grained levels in CPS, researchers need adjust original one based on their specific context.

- This study provides an example to use text classification to identify SSR phases using speech transcription from authentic face-to-face engineer practice course. The model achieved a comparatively good performance (accuracy of 0.67) considering our dataset is small and there were five categories. This shows that text classification techniques also work good at speech transcription data not only in text chat data (Flor M., Andrews-Todd, 2022; Ba et al., 2023; Zheng et al., 2023).

Conclusion and Impications

- the data models and research methods can be diverse.

The data in this study was only speech recordings collected to help analyse sequential patterns of SSR. Multimodal datasets are recommended to be collected to do a deep analysis of the collaboration process in CPS to provide a granular and comprehensive understanding of sequential SSR, such as through speech rates, eye tracking, and body movement (Spikol et al., 2018).

- one research direction about how to combine the CSCL script with large language models in CPS.

Peer feedback analysis in CPS using natural language processing display an approach to automated content analysis for extracting specific categories (Castro et al., 2023), including SSR behaviours.

References

Lämsä J., Hämäläinen R., Koskinen P., Viiri J., Lampi E., “What do we do when we analyse the temporal aspects of computer-supported collaborative learning? A systematic literature review”, Educational Research Review, City, Publishing House, 2021, 33, 100387.

Vaughan N., Wah J. L., “The Community of Inquiry framework: Future practical directions shared metacognition”, International Journal of E-Learning & Distance Education/Revue internationale du e-learning et la formation à distance, City, Publishing House, 2020, 35(1).

Flor M., Andrews-Todd J., “Towards automatic annotation of collaborative problem-solving skills in technology-enhanced environments”, Journal of Computer Assisted Learning, City, Publishing House, 2022, 38(5), 1434-1447.

Lyu Q., Chen W., Su J., Heng K. H., “Collaborate like expert designers: An exploratory study of the role of individual preparation activity on students' collaborative learning”, The Internet and Higher Education, City, Publishing House, 2023, 59, 100920.

Minaee S., Kalchbrenner N., Cambria E., Nikzad N., Chenaghlu M., Gao J., “Deep learning--based text classification: a comprehensive review”, ACM Computing Surveys (CSUR), City, Publishing House, 2021, 54(3), 1-40.

Ba S., Hu X., Stein D., Liu Q., “Assessing cognitive presence in online inquiry-based discussion through text classification and epistemic network analysis”, British Journal of Educational Technology, City, Publishing House, 2023, 54(1), 247-266.

Zheng L., Long M., Niu J., Zhong L., “An automated group learning engagement analysis and feedback approach to promoting collaborative knowledge building, group performance, and socially shared regulation in CSCL”, International Journal of Computer-Supported Collaborative Learning, City, Publishing House, 2023, 18(1), 101-133.

De Backer L., Van Keer H., Moerkerke B., Valcke M., “Examining evolutions in the adoption of metacognitive regulation in reciprocal peer tutoring groups”, Metacognition and Learning, City, Publishing House, 2016, 11, 187-213.

Braun V., Clarke V., “Using thematic analysis in psychology”, Qualitative Research in Psychology, City, Publishing House, 2006, 3(2), 77-101.

Spikol D., Ruffaldi E., Dabisias G., Cukurova M., “Supervised machine learning in multimodal learning analytics for estimating success in project-based learning”, Journal of Computer Assisted Learning, City, Publishing House, 2018, 34(4), 366-377.

Castro M.S.d.O., et al., “Understanding Peer Feedback Contributions Using Natural Language Processing”, In: Viberg O., Jivet I., Muñoz-Merino P., Perifanou M., Papathoma T. (eds) Responsive and Sustainable Educational Futures, EC-TEL 2023, Lecture Notes in Computer Science, Springer, Cham, 2023, 14200.