Majumdar, R. et al. (Eds.) (2025). Proceedings of the 1st International Conference on Learning Evidence and Analytics. Asia-Pacific Society for Computers in Education

A Distilled Model for Collaborative Problem Solving Skill Classification on Resource-Limited Devices

Shuqing LIU[a*], Li CHEN[b], Haiqiao LIU[a], Cheng TANG[c], Fumiya OKUBO[c] & Atsushi SHIMADA[c]

[a]*The Graduate School of Information Science and Electrical Engineering, Kyushu University, Japan*

[b]*Division of Math, Sciences, and Information Technology in Education, Osaka Kyoiku University, Japan*

[c]*Faculty of Information Science and Electrical Engineering, Kyushu University, Japan*

*liu.shuqing.718@s.kyushu-u.ac.jp

Abstract: Collaborative Problem Solving (CPS), a critical 21st-century skill, requires real-time classification to optimize classroom instruction. While mobile technologies offer unprecedented opportunities for such evaluation, existing methods fail to meet real-time demands in dynamic mobile environments due to their requirement for heavy computational resources. This paper introduces a novel knowledge distillation framework to enable lightweight, accurate CPS classification on resource-limited devices. Our approach distills knowledge from large-scale teacher models into compact student networks, reducing model size by 32% compared to baseline model while maintaining 82.6% accuracy on the target dataset. This work bridges the gap between AI-driven classification and mobile educational ecosystems, thereby offering educators a scalable tool for continuous formative evaluation and adaptive pedagogy.

Keywords: Collaborative Problem-Solving, Knowledge Distillation, Natural Language Processing

1. Introduction

Collaborative Problem Solving (CPS) has emerged as a pivotal 21st-century competency (Griffin et al., 2012), proven to enhance learning outcomes, foster engagement, and cultivate critical social-cognitive skills in educational settings (Hwang & Chen, 2023). The proliferation of artificial intelligence and mobile technologies presents unprecedented opportunities for real - time CPS evaluation (Latif et al., 2024), enabling educators to dynamically refine instructional strategies through granular skill analytics (Okubo et al., 2023) . However, existing CPS classification methods, relying on resource - intensive AI models, are impractical for mobile deployment due to latency and energy demands.

This study bridges this gap by developing a knowledge distillation - driven framework for mobile - optimized CPS classification. Knowledge distillation (Hinton et al., 2015), is a widely adopted technique for transferring knowledge from a large, computationally intensive teacher model to a smaller, efficient student model. This process minimizes a composite loss function that balances the student's adherence to ground - truth labels and its alignment with the teacher's predictive distribution. By distilling expertise from large teacher models into lightweight student networks, we achieve a paradigmatic shift in computational efficiency while barely sacrificing accuracy. Our threefold contributions aim to empower educators to conduct continuous formative CPS classifications directly on edge devices, enabling real-time pedagogical interventions:

- This paper proposes a novel mobile - optimized CPS classification model. Our approach addresses the gap in mobile CPS skill evaluation by leveraging mobile devices' unique capabilities and evaluation needs.

- This study is based on a new strategy for sequence building and representation, which explicitly encodes speaker roles and transition dynamics, thereby overcoming the limitations of traditional sequence - based models in CPS dialogue classification.
- Through careful model design, we have significantly reduced the model size by approximately 32% and decreased the computational load, thereby allowing the model to run smoothly on mobile devices with limited computational resources.

2. Methods and Experiments

The study utilizes a dataset comprising transcripts from 40 middle school chemistry classroom discussions, totaling 1,310 annotated conversational sentences by a professional teacher. Considering the evaluation, facilitation, and improvement of collaborative processes, this study follows the framework outlined by Hesse et al. and categorizes CPS into five essential dimensions: participation, perspective - taking, social regulation, task regulation, and learning and knowledge building. Approximately 10% of the samples were selected as the test set, while the remaining samples served as the training set.

In our model, DiaRoBERTa and DistilBERT (Sanh et al., 2020) are employed as the teacher model and the student model, respectively. Additionally, we introduce two custom markers, [SPK] and [THN], to enhance dialogue context and speaker dynamics. The [SPK] marker marks the start of an utterance, helping the model distinguish speakers and identify role - specific patterns. The [THN] marker indicates speaker transitions, improving the model's understanding of sequential relationships.

By inputting the training data and the soft and hard labels generated by the teacher model into the student model, the student model can acquire the classification ability of the teacher model. Subsequently, the student model can be utilized for skill monitoring on smaller devices.

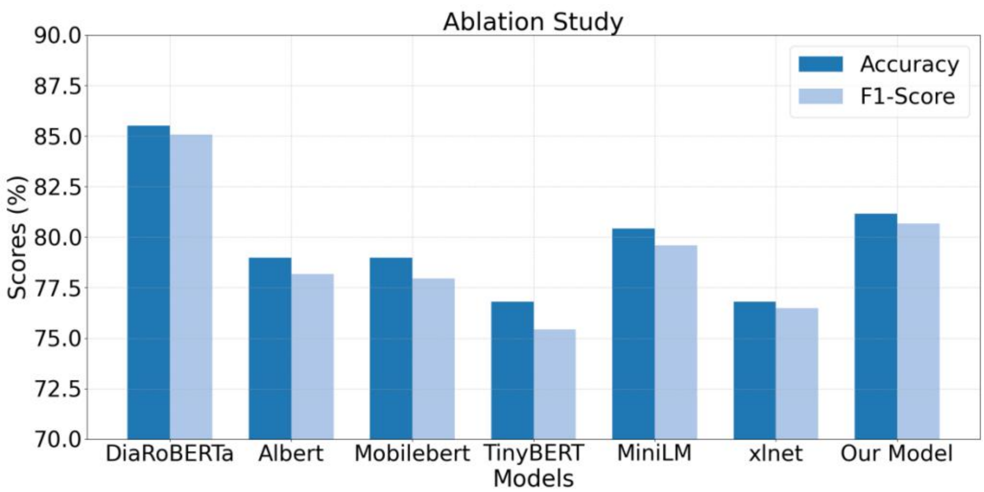

*Figure 1.*Performance Comparison of Models on the CPS Dataset.

The comparative analysis of model accuracy on the CPS dataset is presented in Figure 1. Results demonstrate the effectiveness of our student model in achieving competitive performance relative to the teacher model, DiaRoBERTa. Specifically, our model attained 81.16% accuracy and 80.66% F1 - score. Our model significantly outperformed other models such as MobileBERT(Sun et al., 2020) and ALBERT(Lan et al., 2019) in both accuracy and F1-score. The integration of [SPK] and [THN] tokens likely enhanced contextual awareness, improving classification of overlapping categories. [SPK] enables the model to distinguish between a speaker and their utterances, while [THN] helps the model separate text from different speakers.

Our model has a significantly smaller size, being only two-thirds the size of the teacher model, while still demonstrating effectiveness in handling the CPS dataset. In terms of speed, during the testing phase, each model was able to complete the classification of more than 1000 lines of dialogue within 10 seconds, which can meet the needs of actual real-time dialogue detection.

3. Conclusion

In this study, we have successfully addressed the challenges of CPS skill evaluation in mobile scenarios. Our proposed mobile - specific CPS classification method has effectively filled the research gap in mobile - based CPS skill classification. The model, with its reduced size and overhead, has demonstrated its excellent adaptability to mobile devices with limited resources. The proposed model holds significant potential for application in education. It can be utilized in remote and in - person collaborative learning settings to assess students' CPS skills more accurately. Our future work includes comparing more datasets for better model generalization and conducting real-world scenario experiments to test its effectiveness, which is expected to boost the development and application of mobile CPS skill evaluation technology.

Acknowledgements

This work was supported by JST CREST Grant Number JPMJCR22D1 and JSPS KAKENHI Grant Number JP22H00551, JP24K16759, Japan.

References

Chen, L., Taniguchi, Y., Shimada, A., & Yamada, M. (2024). Exploring behavioral and strategic factors affecting secondary students’ learning performance in collaborative problem-solving-based STEM lessons. *Sage Open,* 14(2), 21582440241251641. https://doi.org/10.1177/21582440241251641

Griffin, P., McGaw, B., & Care, E. (2012). *Assessment and teaching of 21st century skills* (Vol. 10, pp. 978-994). Dordrecht: Springer.

Hesse, F., Care, E., Buder, J., Sassenberg, K., & Griffin, P. (2015). A framework for teachable collaborative problem-solving skills. *Artificial intelligence in education*. Springer.

Hinton, G. E., Vinyals, O., & Dean, J. (2015). Distilling the knowledge in a neural network. *CoRR, abs/1503.02531*. http://arxiv.org/abs/1503.02531

Hwang, G.-J., & Chen, P.-Y. (2023). Effects of a collective problem-solving promotion-based flipped classroom on students’ learning performances and interactive patterns. *Interactive Learning Environments,* 31(5), 2513–2528. https://doi.org/10.1080/10494820.2019.1568263

Lan, Z., Chen, M., Goodman, S., Gimpel, K., Sharma, P., & Soricut, R. (2019). ALBERT: A lite BERT for self-supervised learning of language representations. arXiv preprint arXiv:1909.11942.

Latif, E., Fang, L., Ma, P., & Zhai, X. (2024). Knowledge distillation of LLMs for automatic scoring of science assessments. In A. M. Olney, I.-A. Chounta, Z. Liu, O. C. Santos, & I. I. Bittencourt (Eds.), *Artificial intelligence in education: Posters and late-breaking results, workshops and tutorials* (pp. 166–174). Springer Nature Switzerland.

Okubo, F, Shiino, T, Minematsu, T, Taniguchi, Y & Shimada, A. (2023). Adaptive Learning Support System Based on Automatic Recommendation of Personalized Review Materials, in *IEEE Transactions on Learning Technologies*, vol. 16, no. 1, pp. 92-105, 1 Feb. 2023.

Sanh, V., Debut, L., Chaumond, J., & Wolf, T. (2020). DistilBERT, a distilled version of BERT: Smaller, faster, cheaper and lighter. arXiv preprint arXiv:1910.01108. https://arxiv.org/abs/1910.01108

Sun, Z., Yu, H., Song, X., Liu, R., Yang, Y., & Zhou, D. (2020). MobileBERT: A compact task-agnostic BERT for resource-limited devices. In *Proceedings of the 58th annual meeting of the Association for Computational Linguistics* (pp. 2158–2170). Association for Computational Linguistics.