Enriched Feedback of Classroom Dynamics Using AI

Federico Pardo[1,*,†], Oscar Canovas[1,†] and Felix Garcia[1,†]

[1]*Universidad de Murcia (UM), 30100, Murcia, Spain*

Abstract

This doctoral thesis investigates the interactions between students and teachers by analyzing audio from classroom sessions through multi-modal learning analytics. The objective is to enhance professional development opportunities for teachers by leveraging a variety of features extracted from classroom audio recordings. The research aims to provide insights such as teaching profiles, interaction statistics, automatic classification, and speech analysis. To achieve this, we are developing a software solution that processes audio recordings to extract various features, including signal-related attributes, speaker diarization, and transcriptions, utilizing state-of-the-art technologies. We hypothesize that combining these features in a multi-modal approach will help identify new types of features, enabling us to design experiences that address our research questions.

Keywords

Machine Learning, Education Feedback, Speech Analysis, Classroom Dynamics, Teaching Methodologies, Multi-Modal Learning Analytics

1. Introduction

1.1. Background and Research Problems

Teachers require reliable feedback to enhance their teaching methods and assess their classroom approaches. However, monitoring the dynamics of classroom activities during instruction poses a significant challenge. The traditional path to expertise, which involves extensive practice under a mentor's guidance, is impractical due to the substantial time required to train mentors and the logistical difficulties in observing and analyzing classroom interactions.

Automated solutions powered by Artificial Intelligence (AI) offer a viable alternative, providing consistency, affordability, and the ability to uncover novel data correlations that can deliver actionable feedback to teachers, thereby improving their pedagogical techniques and supporting student success. Nevertheless, it is crucial to define what constitutes "meaningful feedback." Current machine learning algorithms are adept at identifying patterns and classifying data, but their "black box" nature renders the insights they generate opaque to both engineers and non-technical users.

To address these challenges, we propose the development of a unified software platform that delivers clear, practical feedback to educators in an easily understandable format using classroom audio recordings. By 'clear and practical feedback', we mean information that is

LASI Europe 2024 DC: Doctoral Consortium of the Learning Analytics Summer Institute Europe 2024, May 29-31 2024, Jerez de la Frontera, Spain

*Corresponding author.

†These authors contributed equally.

federico.pardog@um.es (F. Pardo); ocanovas@um.es (O. Canovas); fgarcia@um.es (F. Garcia)

directly applicable to teaching practices, specific enough to guide actionable changes, and presented in a way that teachers can quickly grasp and implement without requiring technical expertise. Our objective is to design experiments that elucidate the impact of teacher feedback on the teaching experience.

1.2. Research Goals and questions

The research aims to develop a comprehensive software solution capable of processing audio to perform diarization, transcribe text, and identify audio features. Additionally, it seeks to design experiments to evaluate the software's usefulness for educators. By utilizing this platform, the system will analyze these characteristics to offer insights to teachers, categorize the teaching methods observed in the audio recordings using machine learning techniques, provide profiles for both teachers and learners, and enable teachers to focus on improving specific aspects of their performance. The software will also visualize this information in multiple formats tailored to teachers' needs.

With this advanced software platform, we plan to develop multiple experiments or experiences in collaboration with educators and teaching experts to understand what works for teachers and how they approach classes using this tool. This research can be summarized into three main questions:

- **RQ1:** Is it possible to extract useful features for teachers to improve their teaching from audio recordings?
- **RQ2:** Can these extracted features, after being processed with state-of-the-art techniques such as machine learning, provide feedback to teachers?
- **RQ3:** How might this feedback influence teachers' instructional strategies?

2. Current knowledge and existing solutions

In recent years, numerous studies have analyzed classroom environments [1]. While recording classroom sessions is not a new practice, the shift toward automatic analysis of these recordings has gained momentum recently [2]. Most current proposals employ classical machine learning or deep neural network techniques to examine various teaching practices and styles or to measure the overall classroom climate [3].

Regarding teacher classroom discourse, several research teams have developed and validated automated systems for classifying teaching practices. For instance, [4] trained supervised machine learning models to classify audio recordings. Other studies, based on automatic speech recognition, have focused on segmenting classroom speech between teachers and students [5], leveraging low-level acoustic features. These proposals pay particular attention to the teacher's role [6] in classifying active learning tasks [7, 8, 9].

However, it is important to acknowledge that while these efforts aim for high classification accuracy [8, 10], they often neglect defining discourse features that could provide descriptive and informative data. Advances in discourse technology, such as diarization techniques, should be applied in this field [11]. Informative data are essential for capturing the nuances of teaching practices and providing meaningful insights.

Currently, there are commercial products available, such as the start-up TeachFX[1], which primarily focuses on quantifying the proportion of teacher and student speech by providing graphs related to participation and interaction. However, their approach is limited in terms of the number of analyzed features, hindering a detailed and comprehensive temporal analysis.

In addition to the non-verbal approach, traditional methods of automated teacher discourse analysis have relied on automatic speech recognition (ASR) transcripts. This widely adopted method involves extracting high-level linguistic features that span words, sentences, and discourse levels. Some studies have employed deep learning techniques to identify specific dialogic strategies within mathematics classrooms [12] or to measure how teachers consider students' contributions and engagement [13]. We are currently developing a hybrid system for classroom analysis based on non-verbal features and Natural Language Processing (NLP) techniques, although this work is still in progress.

In alignment with the hybrid approach concept, our methodology fits within the Multi-Modal Learning Analytics (MMLA) framework, as we intend to utilize various data sources to deliver feedback to educators. Recent studies in MMLA predominantly emphasize student collaboration, employing video and audio data as primary information sources [14, 15]. Our research will concentrate on diverse audio processing techniques to extract multiple forms of information, which will be integrated to provide actionable feedback for teachers.

3. Proposed Solution's Novelty

The proposed solution is distinguished by its foundation on a unified multi-modal software platform, which aims to offer educators feedback through straightforward methods and easy to understand metrics, such as the percentage of participation of students, the amount of silence in class or the words per minute of the teacher. The application of these metrics to categorize teaching methodologies within specific intervals demonstrates the potential for accurately identifying teaching strategies based on teacher-student interactions. This capability not only aids in gaining deeper insights into the unique dynamic characteristics of each method but also significantly enhances the quality of feedback. This improvement stems from the feedback being more than just statistical data about the class; it includes contextualized insights tailored to the adopted teaching approach and customized by the educator's desired areas of improvements.

Recently, we have integrated audio signal features, such as tone, to assess speech monotony and enhance the accuracy of previously extracted metrics, such as speaking time for each identified speaker. Additionally, natural language processing (NLP) capabilities have been incorporated through state-of-the-art audio transcription tools. This supplementary information allows for the analysis of speech content and its association with tone and existing speaker diarization features, enabling the identification of filler words, sentiment analysis, and the evaluation of vocabulary complexity, among other aspects.

With these features extracted, the plan is to combine them in various ways and propose novel approaches using MMLA. These include identifying teaching profiles and providing constructive feedback for teachers aspiring to adopt those profiles, automatically identifying learning strategies within teaching methods, and examining how ongoing feedback influences

[1] https://teachfx.com

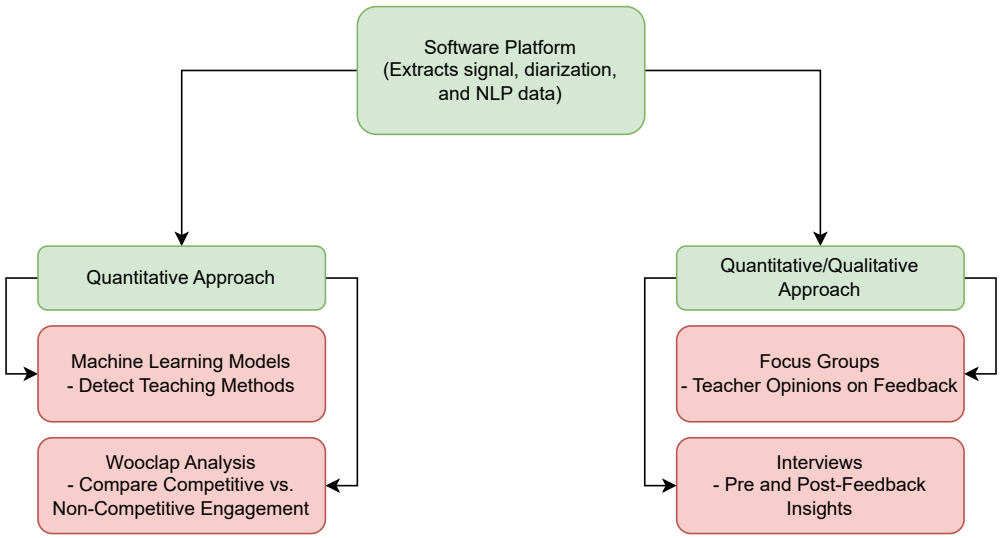

Figure 1: Methodology schema followed in this research with examples.

the dynamics of teachers over time. By using these features collectively in a multi-modal fashion, there is potential to improve both teachers' abilities and students' results. All of this is achieved using exclusively audio information processed with different techniques that extracts the different kinds of information previously mentioned.

This approach has significant implications for the affordability and practicality of the system, making it easier to integrate into classrooms compared to systems that rely on more complex data types such as images or video. Audio data requires less computational power to process, which reduces the overall cost and technical barriers associated with implementation. Moreover, the reliance on audio alone simplifies the setup and minimizes the need for additional equipment, making it a scalable solution for diverse educational settings. The simplicity of audio-based analysis ensures that the system can be easily adopted by educational institutions without requiring significant changes to existing classroom infrastructure. This ease of use, combined with the robust analytical capabilities of the platform, positions it as an accessible and effective tool for enhancing educational practices across a wide range of environments.

While the primary focus is on enhancing teaching practices, this tool holds promise for broader applications across diverse settings. Considering that the features extracted can originate from any audio source and are not limited to academic contexts, the tool could be beneficial for providing feedback in oral presentations, meetings, or interviews. However, this is out of scope of this thesis.

4. Research Methodology

The research methodology of this doctoral thesis focuses on the design of both qualitative and quantitative experiments, with software serving as a key tool. Initially, the emphasis was on

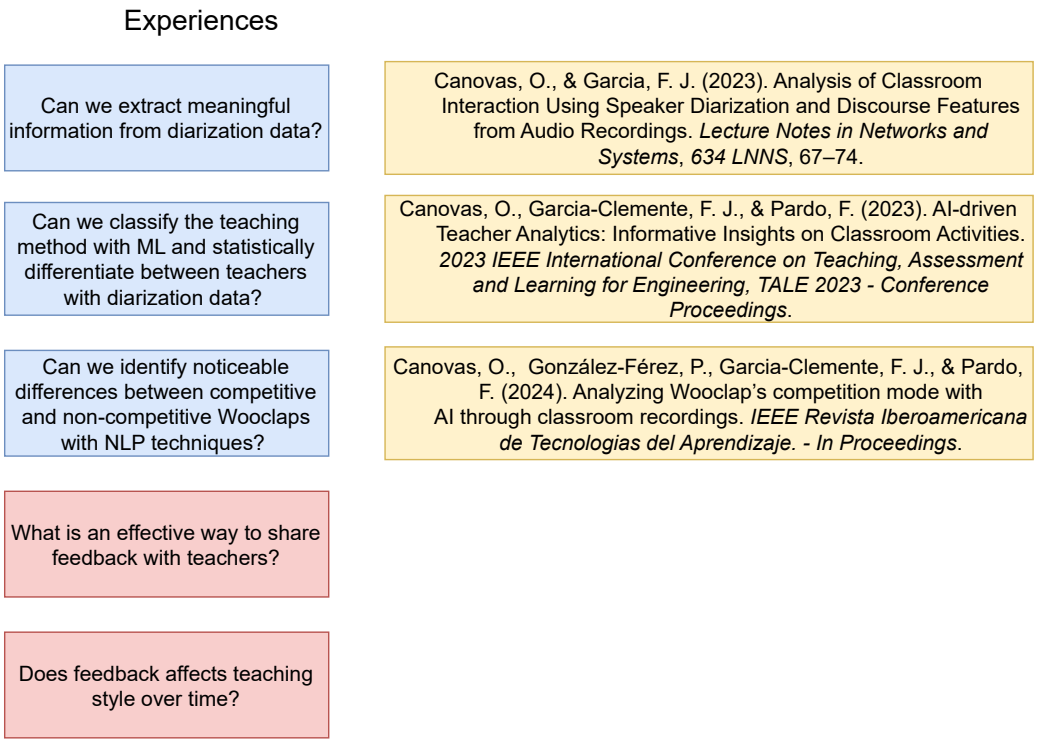

Figure 2: Diagram of the designed experiences and their related publications (blue) with examples of potential future experiences (red).

extracting features from audio diarization to provide meaningful feedback and classify teaching methodologies within specific time frames—an objective that has been achieved. Recently, additional features, such as audio signal properties and transcriptions, have been integrated to enhance the tool's capabilities.

Each new set of features added to the software will be explored for their impact on various dimensions. This exploration aims to understand the diversity of language used by different educators across various teaching methods within the same subject. It will also involve performing verbal analyses, which are valuable both independently and in generating teaching profiles from the extracted features. Furthermore, this approach aims to provide targeted feedback to assist teachers in aligning with established, desired teaching profiles. Through this multifaceted approach, the utility of each newly integrated group of features in enhancing educational effectiveness can be evaluated.

The research will focus on both qualitative and quantitative experience design, as both approaches are beneficial. In the quantitative approach, machine learning models could be used to automatically classify teaching methodologies or compare engagement differences between competitive and non-competitive response systems. A mixed-methods approach could involve creating focus groups to gather teachers' opinions on the feedback. Different disciplines

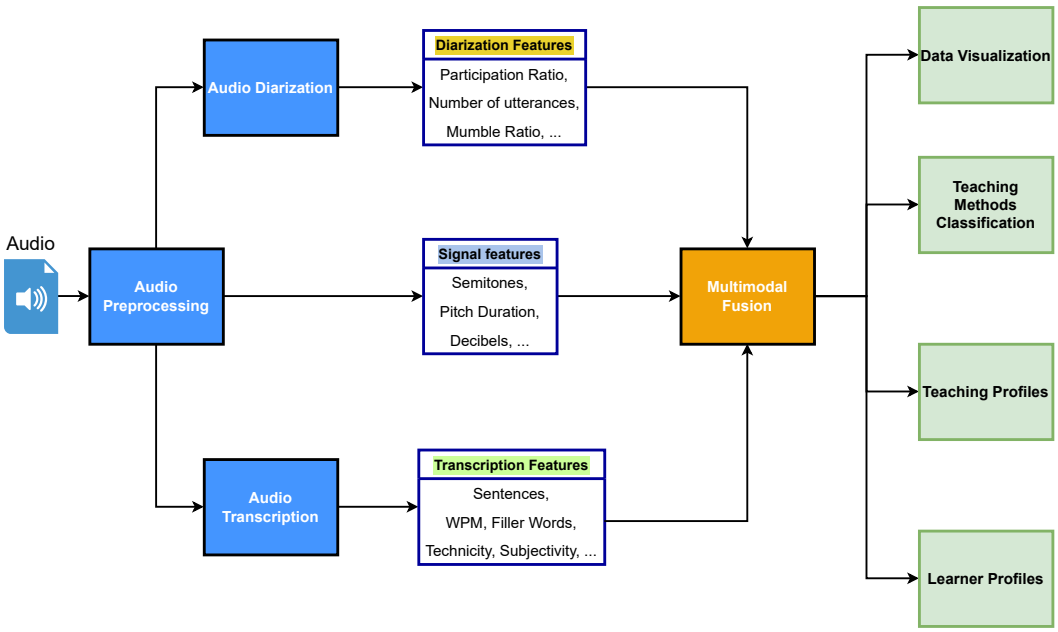

Figure 3: Schema of the software platform development. Blue boxes indicate already developed modules, orange boxes represent planned modules, and green boxes denote the software's capabilities. Examples of extracted features to be processed are shown next to their corresponding sources.

might require different formats or emphasize different aspects of the features provided, needing direct collaboration with teachers. An example of this methodology could involve designing an experiment where teachers are first asked for their opinions about feedback on their activities. After a semester of providing feedback, the teachers could be interviewed again to determine if they changed their approach to class preparation and performance. This qualitative data could then be compared with the quantitative data collected for feedback, offering a comprehensive view of how feedback affects teaching. This more mixed approach would greatly benefit from the assistance of teachers and education experts that could provide a new point of view more focused on the pedagogical part of this work. A visual representation of this methodology is shown in Figure 1, while a visual representation of the experiences already developed and their respective publications is available in Figure 2.

All recordings used for the development and testing of the software strictly adhere to current data protection regulations, ensuring that all participants have provided informed consent for their participation in this research.

5. Current Status and Research

The software developed is now fully functional, as illustrated in Figure 3. Blue boxes indicate the modules already developed, with examples of features that can be extracted using these methods. The planned but not fully developed multi-modal fusion model is shown in orange,

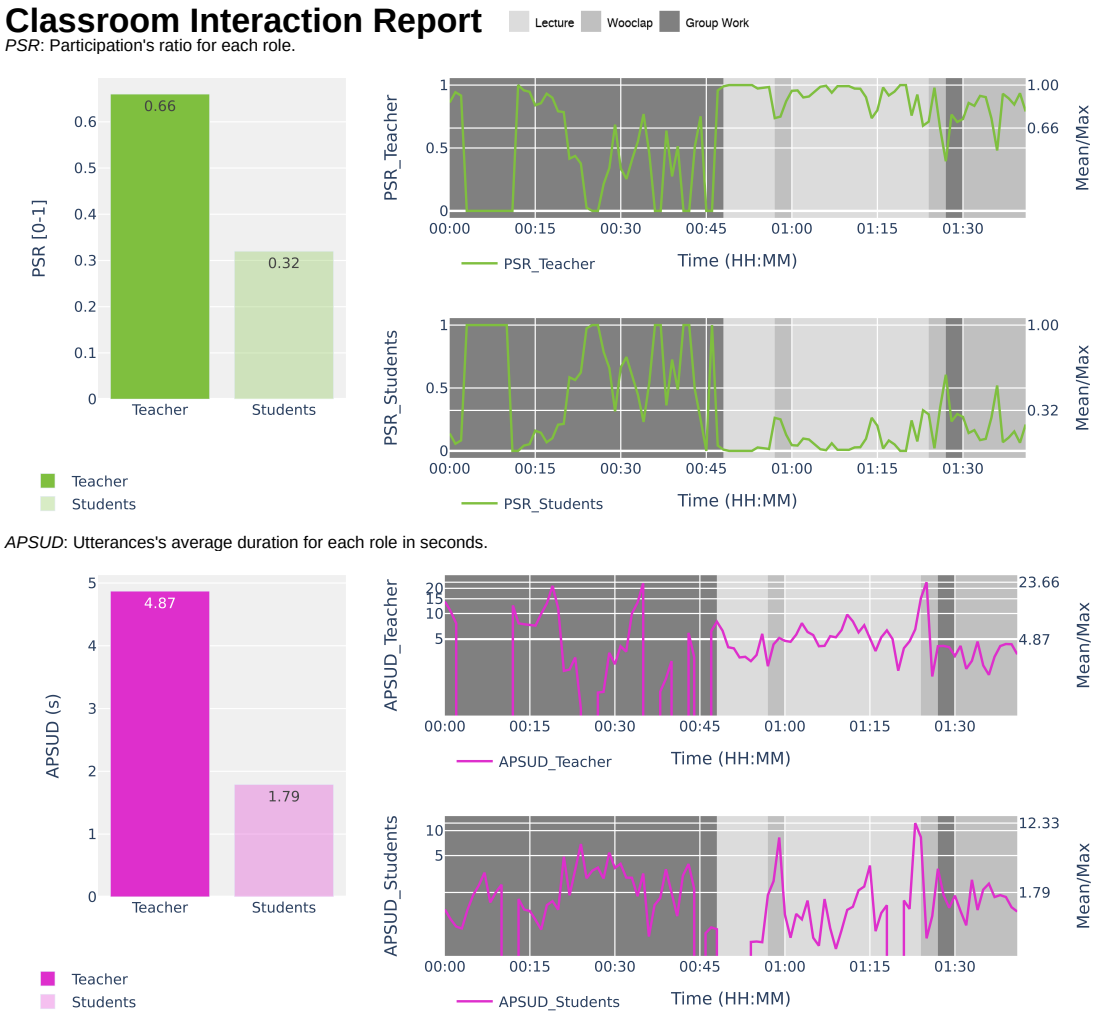

Figure 4: Extract of the information provided by the dashboard for a specific recording.

while the software's capabilities are marked in green. Using this data, we can compile and present a concise feedback report, as demonstrated in Figure 4. This report visually encapsulates various features, including a projection of the teaching methodology for each segment, layered over temporal graphics. The methodology's prediction leverages the collective suite of features extracted from the diarization process. Future versions will incorporate information extracted from natural language processing and signal analysis to further improve the results and enrich the presented information. Additionally, features extracted from these sources will be displayed in customizable graphs, allowing each teacher to focus on specific areas for improvement. It should be noted that this design is subject to significant changes. While visualization is a highly effective tool for teachers to understand their students' progress [16], its design must be carefully considered to ensure it is also beneficial from a learning perspective [17]. Therefore, the development of dashboards will be approached with caution and in consultation with

teachers and education experts to maximize their utility and effectiveness.

Our research objectives, detailed in subsection 1.2, are well-defined, and we possess a robust dataset for ongoing development and testing. We plan to continue augmenting this dataset with a broader variety of examples and scenarios. Additionally, and perhaps most crucially, we have delineated use cases for this research, which, while predominantly focused on academic settings, are not exclusively confined to such environments.

This thesis project has resulted in the publication of two papers [18, 19], with two more under development. The first of these focused on the classification of teaching methods based on diarization features, achieving nearly 0.94 F1 (0.97 in more recent versions). Notably, this study received the TALE Best Paper Award [18]. The second one explores how teacher's language differs using the Wooclaps methodology in competitive versus non-competitive settings, finding out that in a competitive environment, the teacher tends to take more time on the gamification part, commenting on the student's scores and the fights for the firsts positions in the ranking. However, on non-competitive wooclaps, much more time is dedicated to explainning the answers to the questions and making sure everyone understands the content. Moreover, a paper detailing the developed feedback system has been accepted for publication, including an analysis of the designed feedback reports (a fragment is shown in Figure 4, including data about participation ratio).

Our current work involves analyzing the transcriptions to derive new features that will be helpful in generating teaching and learner profiles. This process includes utilizing advanced natural language processing (NLP) techniques to explore the linguistic aspects of classroom interactions. Specifically, we are employing BERT-like models to classify the language used in different learning scenarios. These models, known for their ability to understand and generate human language, allow us to delve deeper into the contextual and semantic nuances of the transcriptions. By applying BERT-like models, we can identify patterns and categories in the language that correspond to various teaching methods and student engagement levels. Additionally, we are preparing a systematic review focused on the usage of audio in learning analytics. This review will concentrate on the diverse features extracted from audio data and the innovative approaches applied in recent studies. Our goal is to compile and analyze the state-of-the-art techniques in audio-based learning analytics to inform and enhance our own methodologies. By integrating insights from this systematic review, we aim to refine our approach and ensure that our tool remains at the forefront of educational technology innovation.

Acknowledgments

This work has been funded under grant TED2021-129300B-I00, by MCIN/AEI/10.13039/501100011033, NextGenerationEU/PRTR, UE, and grant PID2021-122466OB-I00, by MCIN/AEI/10.13039/501100011033/FEDER, UE.

References

[1] A. James, Y. H. V. Chua, T. Maszczyk, A. M. Núñez, R. Bull, K. Lee, J. Dauwels, Automated classification of classroom climate by audio analysis, Lecture Notes in Electrical

Engineering 579 (2019) 41–49. doi:10.1007/978-981-13-9443-0_4.

[2] M. E. Dale, A. J. Godley, S. A. Capello, P. J. Donnelly, S. K. D'Mello, S. P. Kelly, Toward the automated analysis of teacher talk in secondary ela classrooms, Teaching and Teacher Education 110 (2022). doi:10.1016/j.tate.2021.103584.

[3] A. Ramakrishnan, B. Zylich, E. Ottmar, J. Locasale-Crouch, J. Whitehill, Toward automated classroom observation: Multimodal machine learning to estimate class positive climate and negative climate, IEEE Transactions on Affective Computing 14 (2023) 664–679. doi:10.1109/TAFFC.2021.3059209.

[4] P. Donnelly, N. Blanchard, B. Samei, A. M. Olney, X. Sun, B. Ward, S. Kelly, M. Nystrand, S. K. D'Mello, Automatic teacher modeling from live classroom audio, UMAP 2016 - Proceedings of the 2016 Conference on User Modeling Adaptation and Personalization (2016) 45–53. doi:10.1145/2930238.2930250.

[5] S. K. D'Mello, A. M. Olney, N. Blanchard, X. Sun, B. Ward, B. Samei, S. Kelly, Multimodal capture of teacher-student interactions for automated dialogic analysis in live classrooms, Association for Computing Machinery, Inc, 2015, pp. 557–566. doi:10.1145/2818346.2830602.

[6] D. Schlotterbeck, P. Uribe, R. Araya, A. Jimenez, D. Caballero, What classroom audio tells about teaching: A cost-effective approach for detection of teaching practices using spectral audio features, LAK21: LAK21: 11th International Learning Analytics and Knowledge Conference (2021) 132–140. URL: https://doi.org/10.1145/3448139.3448152. doi:10.1145/3448139.3448152.

[7] H. Li, Y. Kang, W. Ding, S. Yang, S. Yang, G. Y. Huang, Z. Liu, Multimodal learning for classroom activity detection, ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings 2020-May (2020) 9234–9238. doi:10.1109/ICASSP40776.2020.9054407.

[8] M. T. Owens, S. B. Seidel, M. Wong, T. E. Bejines, S. Lietz, et al., Classroom sound can be used to classify teaching practices in college science courses, Proceedings of the National Academy of Sciences of the United States of America 114 (2017) 3085–3090. doi:10.1073/pnas.1618693114.

[9] H. Su, B. Dzodzo, X. Wu, X. Liu, H. Meng, Unsupervised methods for audio classification from lecture discussion recordings, Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH 2019-Septe (2019) 3347–3351. doi:10.21437/Interspeech.2019-2384.

[10] Z. Wang, X. Pan, K. F. Miller, K. S. Cortina, Automatic classification of activities in classroom discourse, volume 78, 2014, pp. 115–123. doi:10.1016/j.compedu.2014.05.010.

[11] T. J. Park, N. Kanda, D. Dimitriadis, K. J. Han, S. Watanabe, S. Narayanan, A review of speaker diarization: Recent advances with deep learning, Computer Speech and Language 72 (2022). URL: http://arxiv.org/abs/2101.09624. doi:10.1016/j.csl.2021.101317.

[12] A. Suresh, T. Sumner, J. Jacobs, B. Foland, W. Ward, Automating analysis and feedback to improve mathematics teachers' classroom discourse, 33rd AAAI Conference on Artificial Intelligence, AAAI 2019, 31st Innovative Applications of Artificial Intelligence Conference, IAAI 2019 and the 9th AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2019 (2019) 9721–9728. doi:10.1609/aaai.v33i01.33019721.

[13] T. Nazaretsky, B. B. Klebanov, J. N. Mikeska, Empowering Teacher Learning with AI:

Automated Evaluation of Teacher Attention to Student Ideas during Argumentation-focused Discussion, volume 1, Association for Computing Machinery, 2023. doi:10.1145/3576050.3576067, explainable models. Usage of utterances to move discussion forward.

[14] P. Chejara, L. P. Prieto, M. J. Rodriguez-Triana, A. Ruiz-Calleja, M. Khalil, Impact of window size on the generalizability of collaboration quality estimation models developed using Multimodal Learning Analytics, in: ACM International Conference Proceeding Series, 2023, pp. 559–565. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85149265618&doi=10.1145%2F3576050.3576143&partnerID=40&md5=50dc16c50e67df37a96d6bdb02c49ad9. doi:10.1145/3576050.3576143.

[15] C. M. D'Angelo, R. J. Rajarathinam, Speech analysis of teaching assistant interventions in small group collaborative problem solving with undergraduate engineering students, British Journal of Educational Technology (2024). doi:10.1111/bjet.13449.

[16] R. Martinez Maldonado, J. Kay, K. Yacef, B. Schwendimann, An interactive teacher's dashboard for monitoring groups in a multi-tabletop learning environment, in: Intelligent Tutoring Systems: 11th International Conference, ITS 2012, Chania, Crete, Greece, June 14-18, 2012. Proceedings 11, Springer, 2012, pp. 482–492.

[17] V. Echeverria, R. Martinez-Maldonado, R. Granda, K. Chiluiza, C. Conati, S. Buckingham Shum, Driving data storytelling from learning design, in: Proceedings of the 8th international conference on learning analytics and knowledge, 2018, pp. 131–140.

[18] O. Canovas, F. J. Garcia-Clemente, F. Pardo, Ai-driven teacher analytics: Informative insights on classroom activities, in: 2023 IEEE International Conference on Teaching, Assessment and Learning for Engineering (TALE), IEEE, 2023, pp. 1–8.

[19] Óscar Cánovas Reverte, P. G. Férez, F. J. G. Clemente, F. P. García, Análisis con ia del modo competición de wooclap a través de las grabaciones de las clases, Revista Iberoamericana de Tecnología en Aprendizaje y Enseñanza de la Programación (2024). URL: https://vaep-rita.org/, submitted.