Classroom Collaboration Analytics: Designing and Building Automated Systems for Collaboration Monitoring in Classroom Settings

Pankaj Chejara
pankajch@tlu.ee

Researcher
Center of Education Technology
Tallinn University, Estonia

Data Scientist
Applied Health Data Division
Metrosert AS, Estonia

AGENDA

Introduction
Modeling collaboration
Generalizability
Multimodal data & Collaboration monitoring
Guidelines for building models
Challenges/Future directions
Conclusion

INTRODUCTION

COLLABORATION

Collaboration is a complex construct **(Rummel et al., 2011).**

image: Freepik.com

INTRODUCTION
COLLABORATION

Collaboration is a complex construct **(Rummel et al., 2011).**

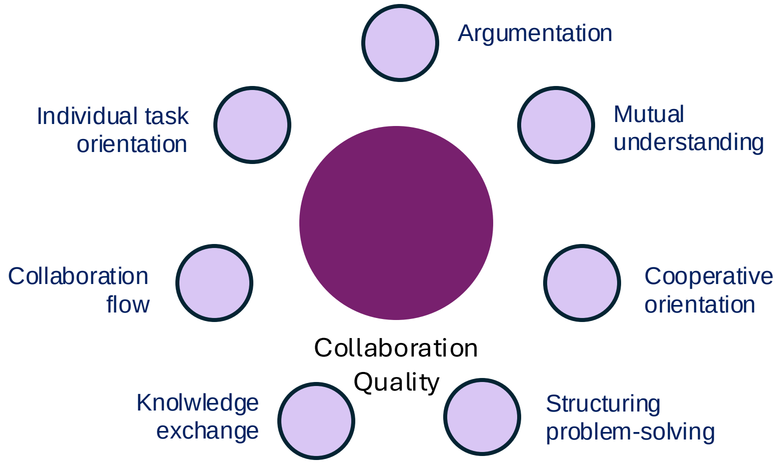
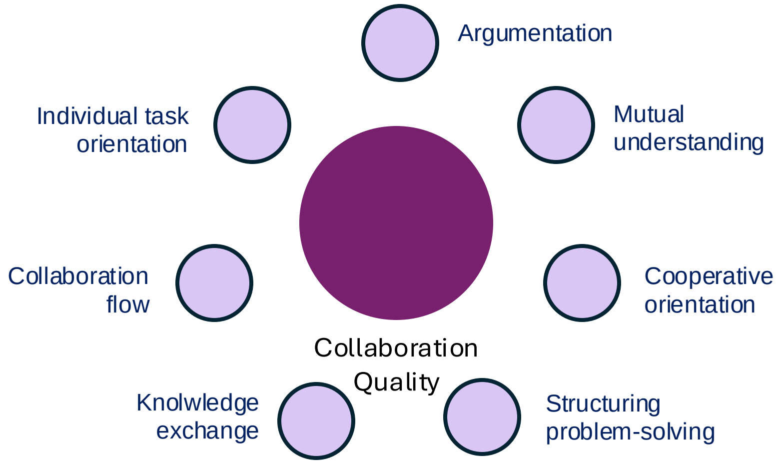


image: Freepik.com

INTRODUCTION

COLLABORATION

Collaboration is a complex construct **(Rummel et al., 2011).**



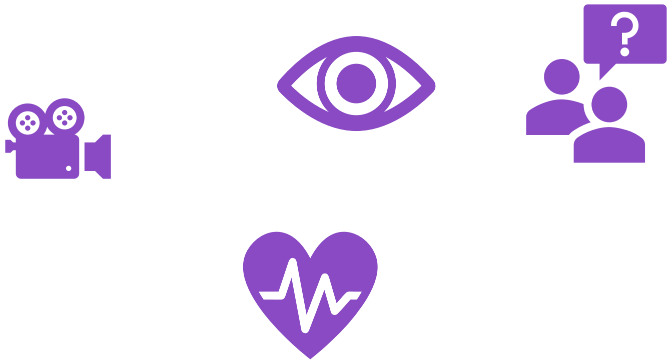
Difficult for teachers to monitor and detect problems **(Chounta & Avouris, 2016)**

image: Freepik.com

INTRODUCTION

MULTIMODAL LEARNING ANALYTICS

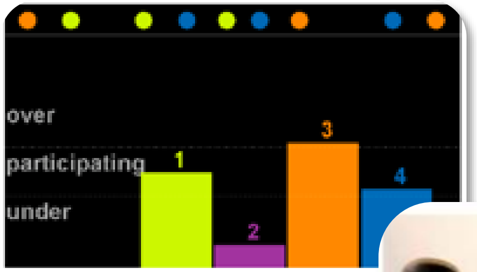
Uses sensors along with log data
(Ochoa et al., 2017).



Captures multimodality of students' interactions.

INTRODUCTION

MULTIMODAL LEARNING ANALYTICS



Speaking participation
(DiMicco et al., 2004)

LED matrix display
(Bachour et al., 2010)

Visualization

Speaking time
(collaboration quality)

(Martinez-Maldonado et al., 2011)

Distance between hands
(Performance)

(Spikol et al., 2018)

Pattern

Collaboration quality

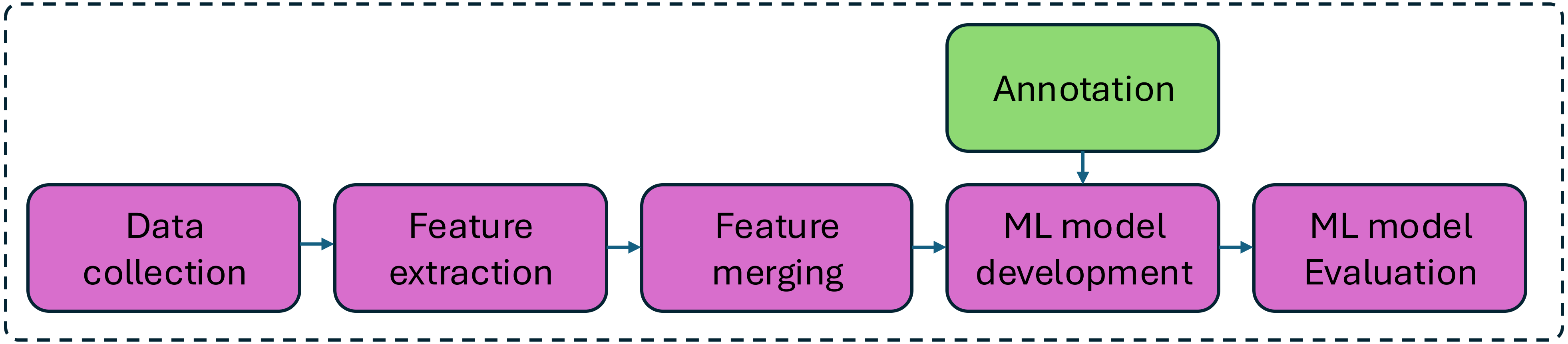
(Martinez-Maldonado et al., 2011)

Rapport

(Lubold et al., 2014)

Modeling

MODELING COLLABORATION



Data collection → Feature extraction → Feature merging → ML model development → ML model Evaluation

Multimodal data collection systems

1. iMotions Lab
   Commercialized
2. Multimodal Learning Hub
   Schneider et al., 2018
3. Social Signal Framework
   Wagner et al., 2013
4. Lab Streaming Layer
   Kothe et al., 2014
5. EZ-MMLA toolkit
   Schneider et al., 2022
6. CoTrack
   Chejara et al., 2024

Data collection → **Feature extraction** → Feature merging → ML model development → ML model Evaluation

Feature extraction

Audio data

Voice activity detection (VAD, py-webrtcvad)	Speaking time Turn-taking
Acoustic feature extraction (OpenSmile tool)	Pitch Energy
Speech-to-Text (Google Speech-to-Text)	Speech data

Video data

Facial exrepssions (OpenFace, Py-Feat)	Facial action units Emotions
Body movement (OpenPose)	Head movement Hand movement

Data collection → Feature extraction → **Feature merging** → ML model development → ML model Evaluation

Feature merging

Group

Facial expressions (**25 data points per second**)

Voice activity(**5 data points per second**)

→ Summary stats (10 seconds)

Facial expressions (**25 data points per second**)

Voice activity(**5 data points per second**)

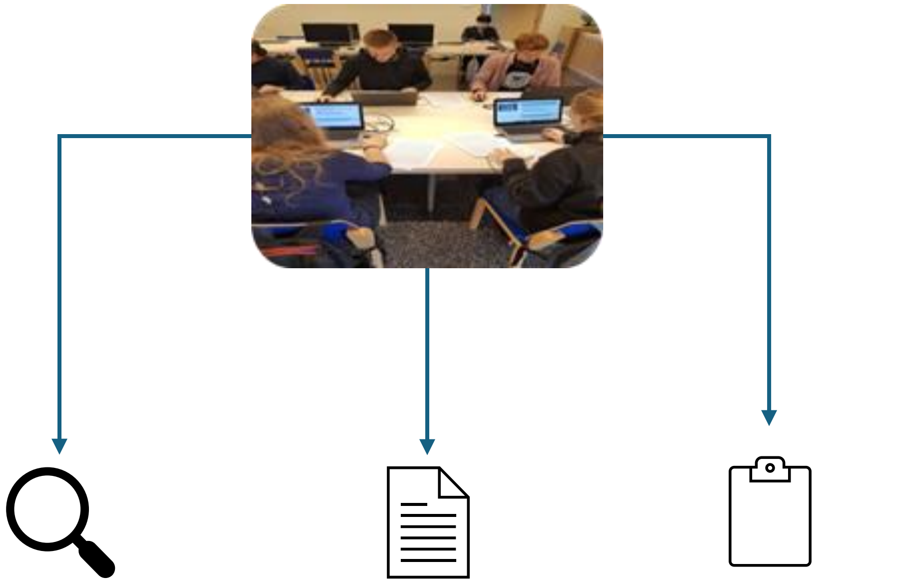
→ Summary stats (10 seconds)

→ Group level features

Annotation

Data collection → Feature extraction → Feature merging → ML model development → ML model Evaluation

Annotation



Expert evaluation

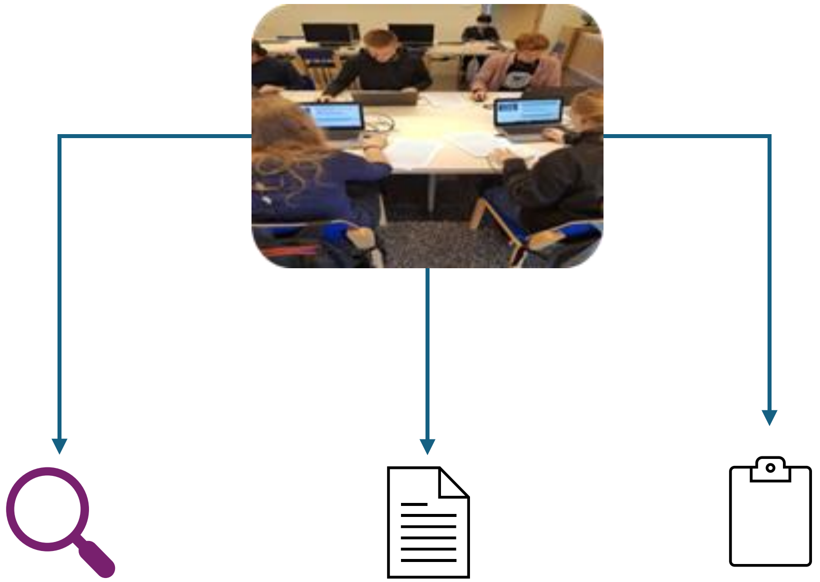
Artifact assessment

Self-reported survey

Annotation

Data collection → Feature extraction → Feature merging → ML model development → ML model Evaluation

Annotation

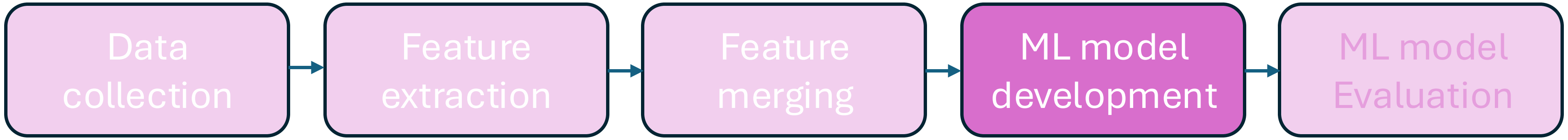


Expert evaluation

Artifact assessment

Self-reported survey

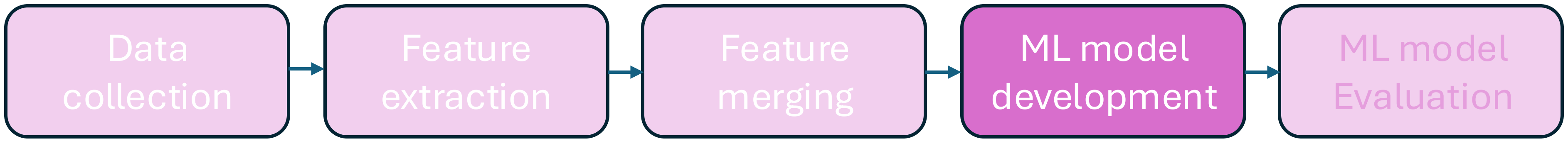
[Rating handbook](#)
(Rummel et al., 2011)



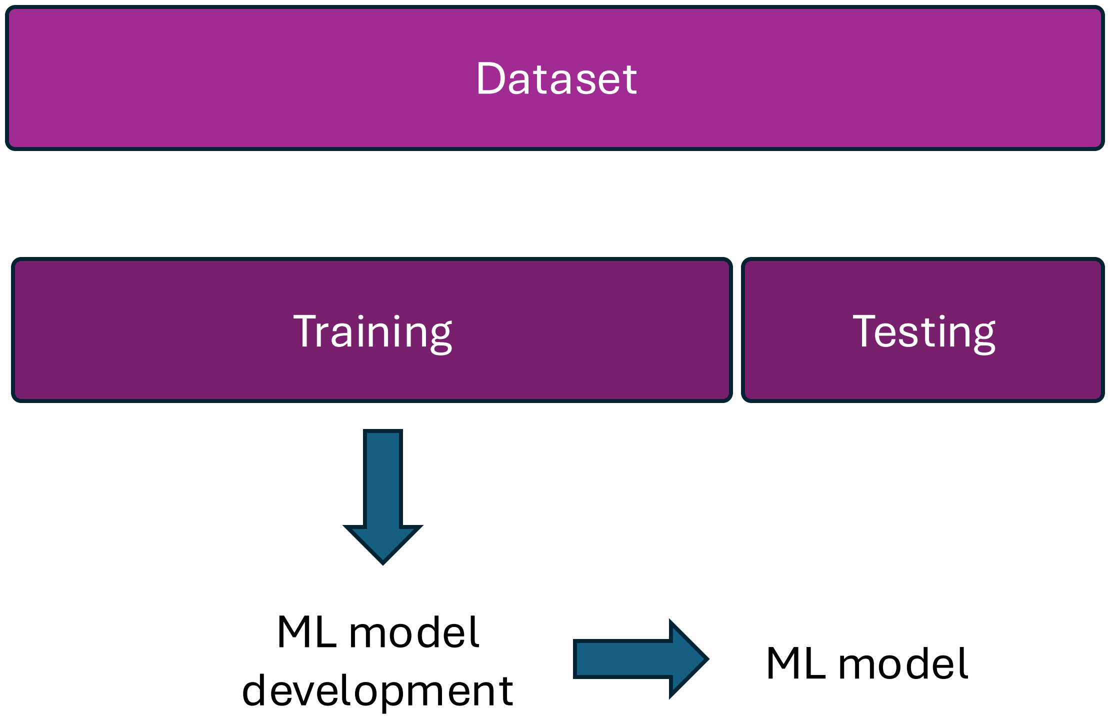
ML model development

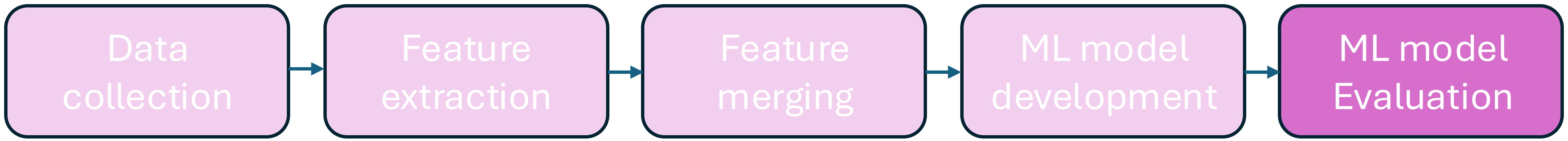
index	frame	group	user_speak_mean	user_speak_sd	user_turns_mean	user_turns_sd	write_text	user_wh_mean	user_self_mear	user_u
0	1	87_1	3.960666666666667	1.095701398901888	2	0.4714045207910317		0	1	0
1	2	87_1	2.578333333333333	1.208881117213580	1	0.4714045207910	*The Growth	0	0	0
2	3	87_1	5.400333333333332	3.269392638124429	2	1.2472191289246	*The Growth	0	1	0
3	4	87_1	Feb.56	2.222851022148508	1	0.9428090415820	*The Growth	0	0	0
4	5	87_1	3.168333333333333	2.292290024310963	1	0.4714045207910	*The Growth	0	0	0
5	6	87_1	4.741333333333333	1.288999956900266	2	0.4714045207910	*The Growth	0	0	0
6	7	87_1	3.965333333333333	3.169162069415546	1	0.9428090415820	*The Growth	0	0	0
7	8	87_1	4.828	2.568189634742730	1	0.4714045207910	*The Growth	0	1	0
8	9	87_1	3.239000000000000	2.962571293094339	1	0.8164965809277	*The Growth	0	0	0
9	10	87_1	5.597	0.444648925183303	2	0.0	*The Growth	0	2	0
10	11	87_1	3.986666666666667	1.102041842319166	2	0.8164965809277	*The Growth	0	2	0
11	12	87_1	7.578	2.561475746518011	2	0.4714045207910	*The Growth	0	4	0
12	13	87_1	7.608000000000000	2.735462422821169	2	0.0	*The Growth	0	1	0
13	14	87_1	1.529666666666666	2.163275345910044	0	0.4714045207910	*The Growth	0	0	0
14	15	87_1	8.931333333333333	0.962386732152009	5	0.816496580927726		0	0	0
15	16	87_1	5.556333333333334	2.261789900843037	2	0.4714045207910	*The Growth	0	3	1
16	17	87_1	6.727	1.017153216908183	3	0.0	*The Growth	1	3	1
17	18	87_1	10.11666666666666	0.788052592717573	2	0.4714045207910	*The Growth	0	0	0
18	19	87_1	4.003333333333333	2.393908148242580	1	0.4714045207910	*The Growth	1	2	0
19	20	87_1	2.688333333333333	1.207213412045369	1	0.4714045207910	*The Growth	0	0	0
20	21	87_1	3.849666666666666	2.990847854520334	2	1.6996731711975	*The Growth	1	1	0
21	22	87_1	3.414333333333333	2.844651160023355	1	1.2472191289246	*The Growth	0	0	0
22	23	87_1	2.529666666666667	2.335577349512440	1	0.4714045207910	*The Growth	0	0	0
23	24	87_1	3.084333333333333	2.186550759123195	2	1.4142135623730	*The Growth	0	0	0
24	25	87_1	5.055000000000001	4.149196629067688	1	0.9428090415820	*The Growth	0	0	0
25	26	87_1	5.559333333333332	4.346422385773793	1	0.8164965809277	*The Growth	0	0	0
26	27	87_1	7.597333333333332	3.063529808715575	2	0.8164965809277	*The Growth	0	1	0
27	28	87_1	3.507333333333333	3.013227984884132	1	0.4714045207910	*The Growth	2	2	0
28	29	87_1	4.020000000000000	2.012179581117616	1	0.4714045207910	*The Growth	1	1	0
29	30	87_1	7.004666666666666	4.640425650974511	2	1.2472191289246	*The Growth	0	0	0
30	31	87_1	6.007666666666666	1.914539863489107	2	0.4714045207910	*The Growth	0	2	0
31	32	87_1	6.849	0.687692276142945	2	0.0	*The Growth	0	0	0
32	33	87_1	2.880000000000000	2.347438320098457	1	0.8164965809277	*The Growth	0	0	0
33	34	87_1	3.532	2.464377135640295	2	0.8164965809277	*The Growth	0	0	1
34	35	87_1	2.981666666666667	3.437058076643777	0	0.4714045207910	*The Growth	0	1	0

Dataset

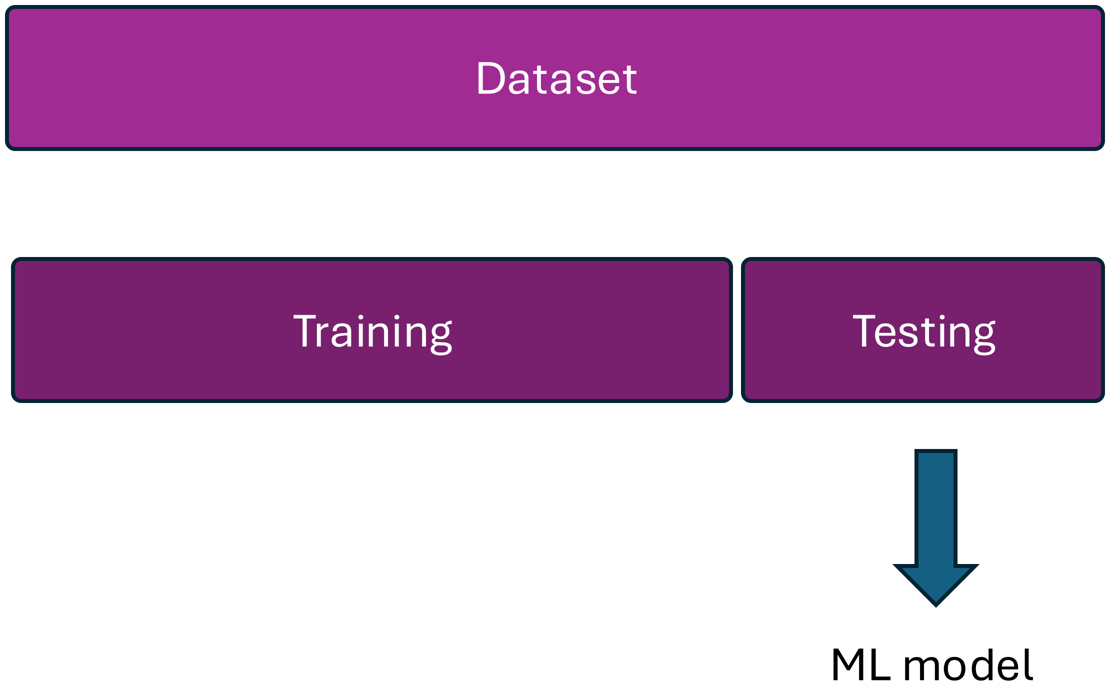


ML model development





ML model Evaluation



WHAT NEXT?

Would our model work well in authentic classroom settings?

NOISE ISSUE

VARYING CONTEXT

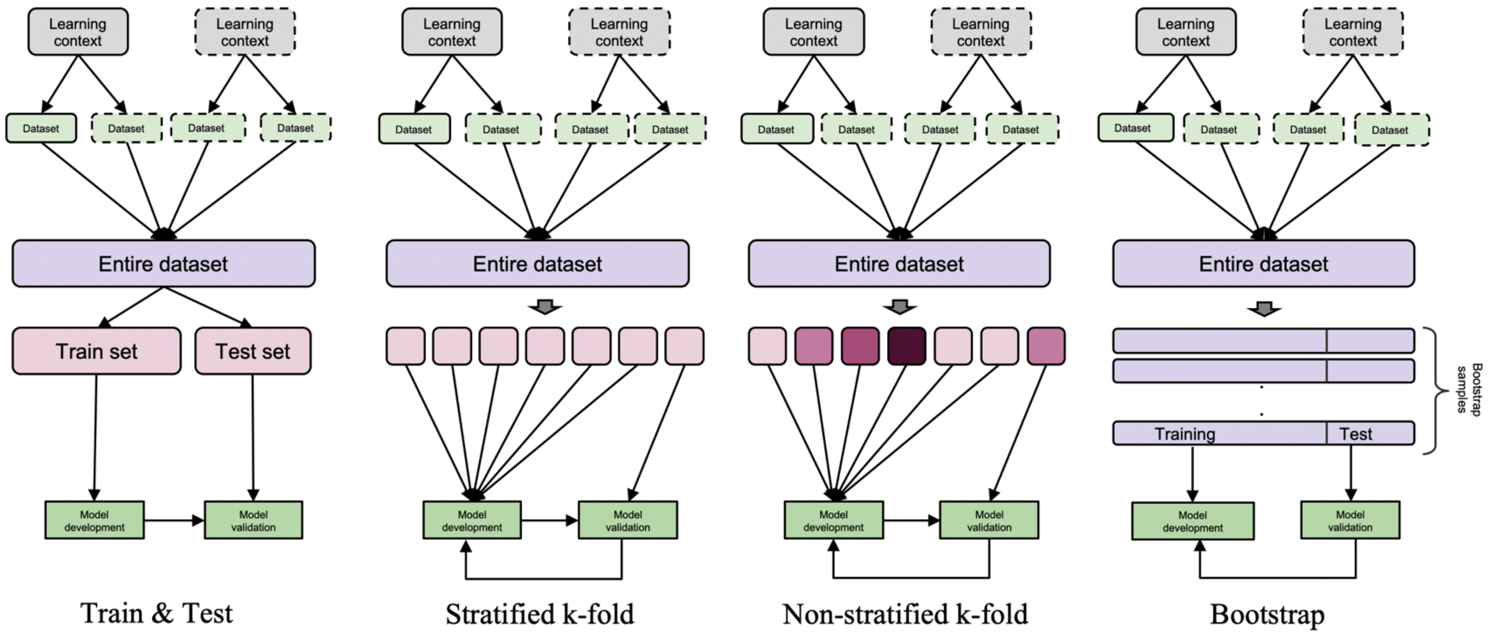
Collaboration quality model →

Transitioning from research to practice

GENERALIZABILITY

1

Ability of machine learning models to perform well on unseen data (**Raschka, 2018**).

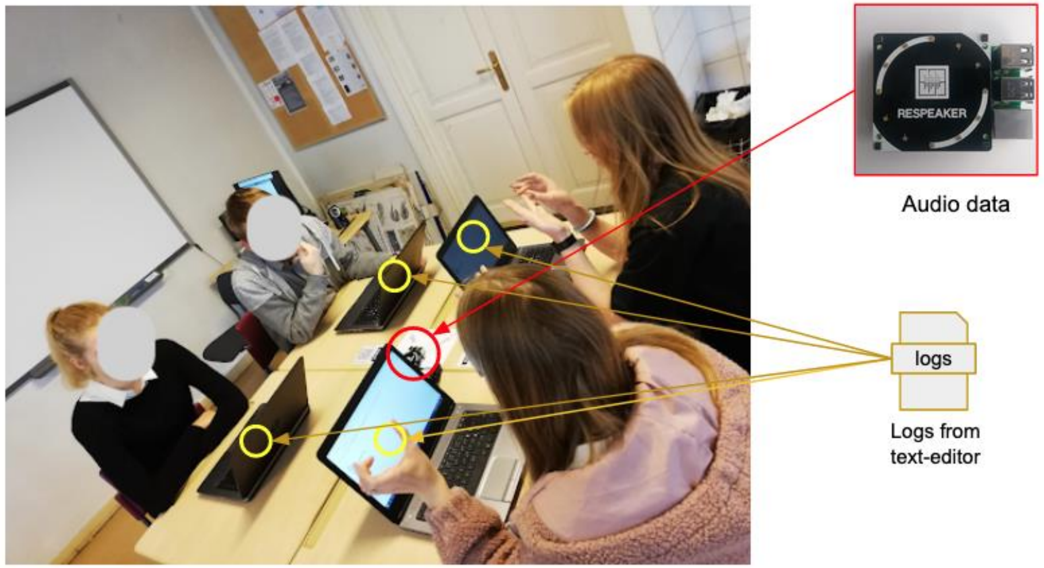


Model Evaluation Methods

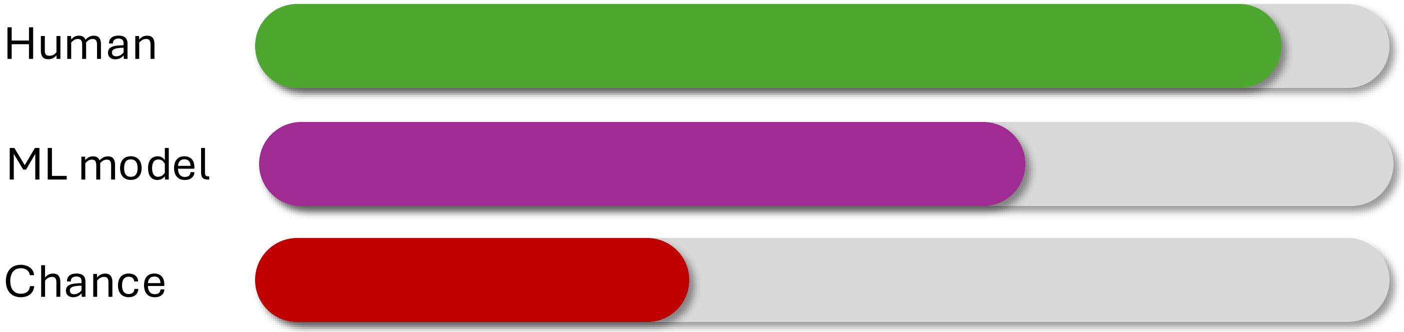
1 Lack of systematization for generalizability evaluation & reporting

2 Do not consider educational nature of MMLA

EXAMPLE



Features (30 sec)
Speaking time
number of char added
number of char deleted



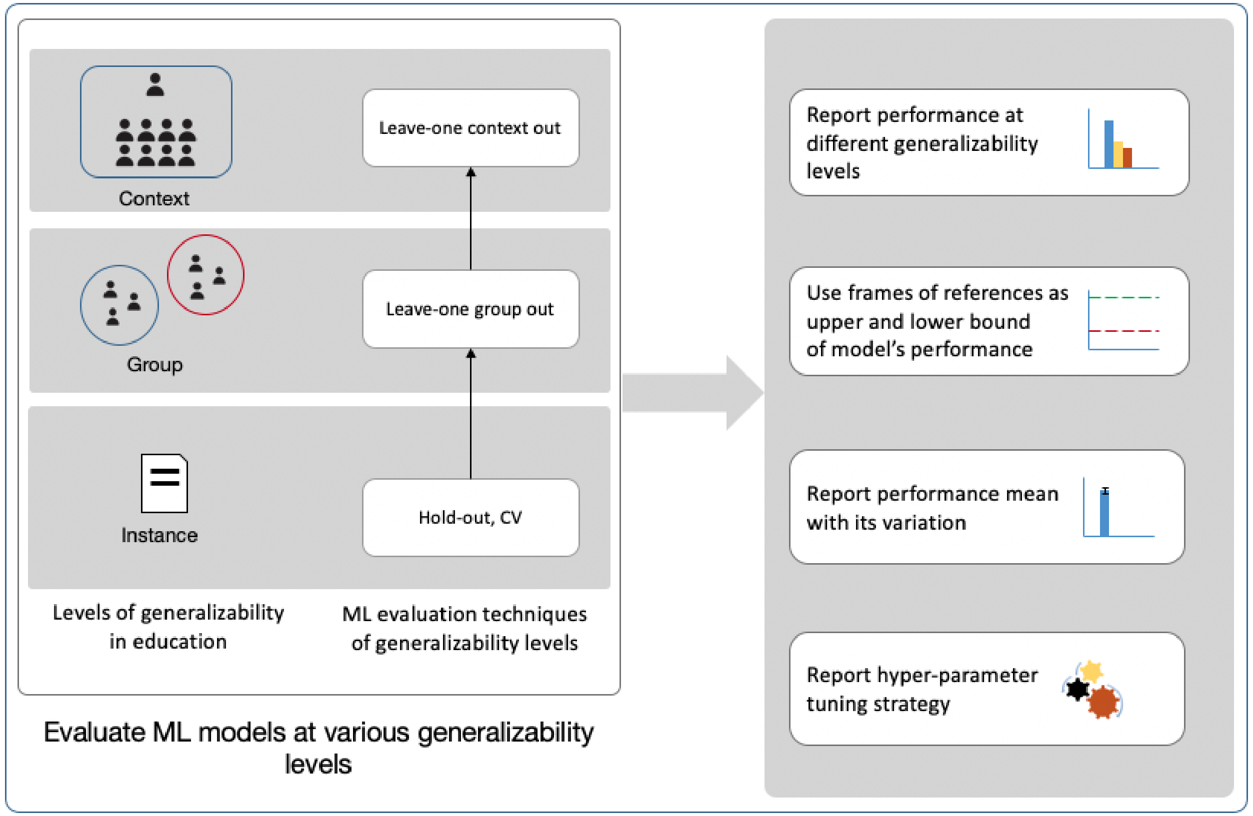
Group	2
Group-size	4
Duration	30 mins
Instances	121

Chejara, P., Prieto, L. P., Ruiz-Calleja, A., Rodríguez-Triana, M. J., Shankar, S. K., & Kasepalu, R. (2020). Quantifying collaboration quality in face-to-face classroom settings using MMLA. In *International Conference on Collaboration Technologies and Social Computing (CollabTech)* (pp. 159-166). Springer, Cham. https://doi.org/10.1007/978-3-030-58157-2_11

How to systematically assess and report generalizability in MMLA?

EFAR-MMLA

GENERALIZABILITY EVALUATION FRAMEWORK



Chejara, P., Prieto, L. P., Ruiz-Calleja, A., Rodríguez-Triana, M. J., Shankar, S. K., & Kasepalu, R. (2021). EFAR-MMLA: An evaluation framework to assess and report generalizability of machine learning models in MMLA. *Sensors* (21), 2863. https://doi.org/10.3390/s21082863

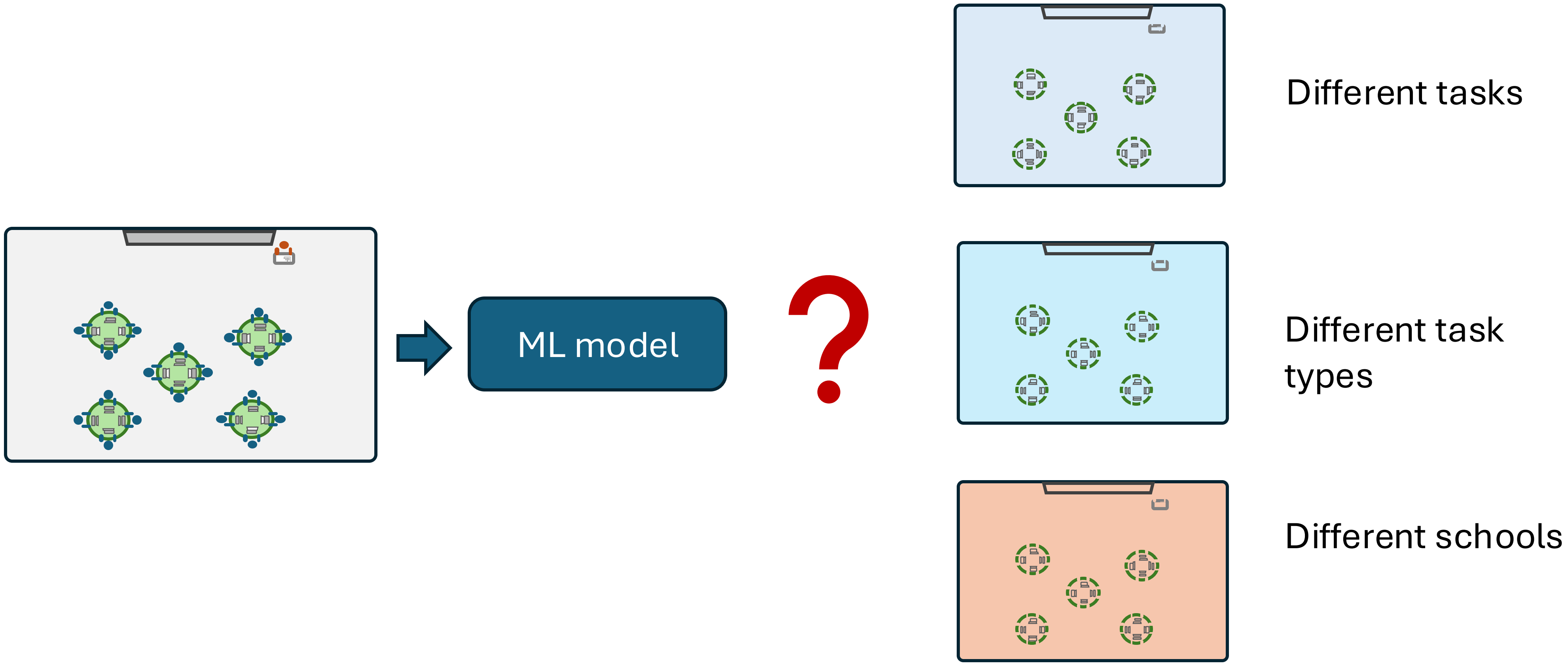
EFAR-MMLA

GENERALIZABILITY EVALUATION FRAMEWORK

1. Assessment of generalizability relavant to MMLA
2. Bias identification

How well automated collaboration estimation models perform across different contexts varying on task, task type and school?

CONTEXT GENERALIZABILITY



CONTEXT GENERALIZABILITY

CoTrack

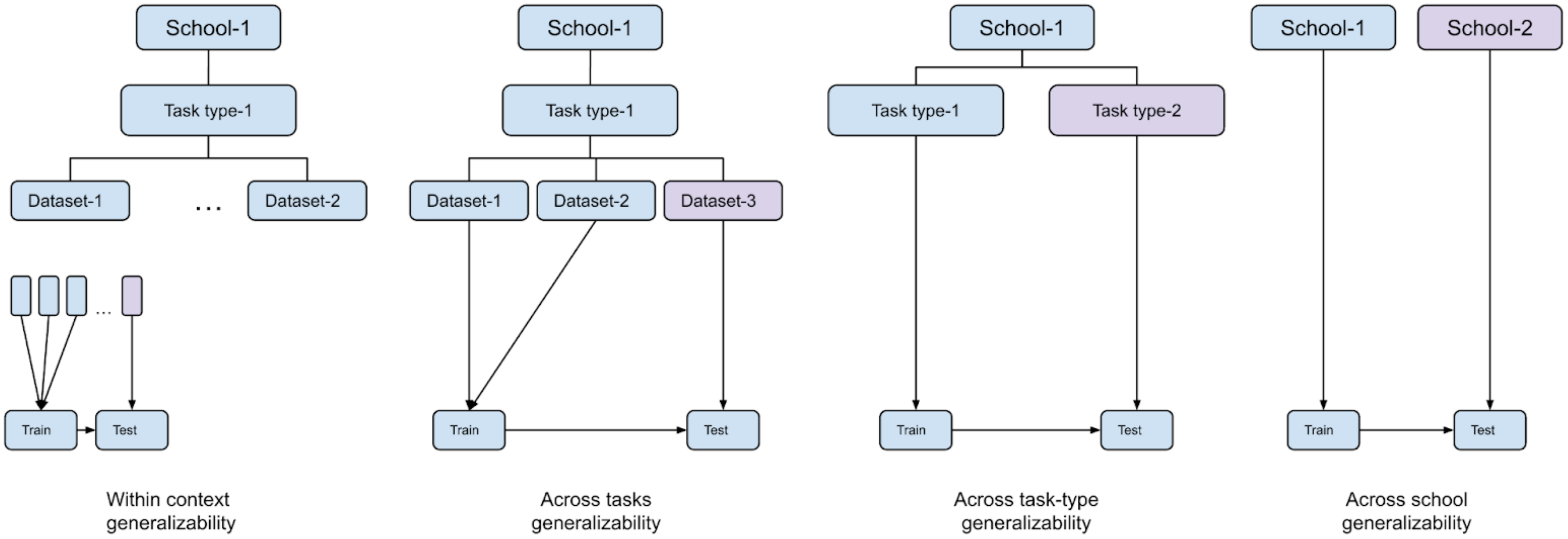
Feature extraction & merging

Supervised machine learning

Estonian Vocational school

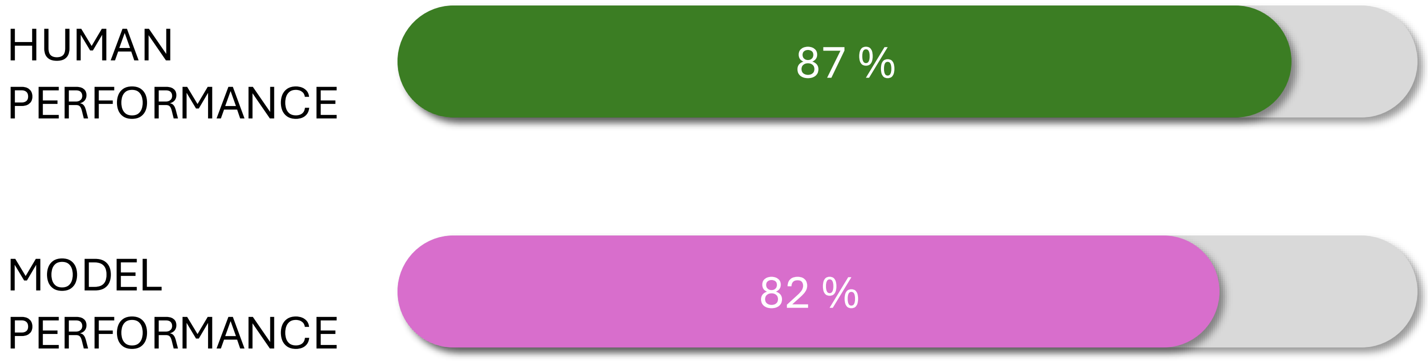
Chejara, P., Kasepalu, R., Prieto, L., P., Rodríguez-Triana, M. J., Ruiz-Calleja, A., & Schneider, B. (2023). How well do collaboration quality estimation models generalize across authentic school contexts. *British Journal of Educational Technology*, 00, 1–23. https://doi.org/10.1111/bjet.13402

EVALUATION


School-1
Task type-1
Dataset-1
...
Dataset-2
Train
Test
Within context generalizability
School-1
Task type-1
Dataset-1
Dataset-2
Dataset-3
Train
Test
Across tasks generalizability
School-1
Task type-1
Task type-2
Train
Test
Across task-type generalizability
School-1
School-2
Train
Test
Across school generalizability

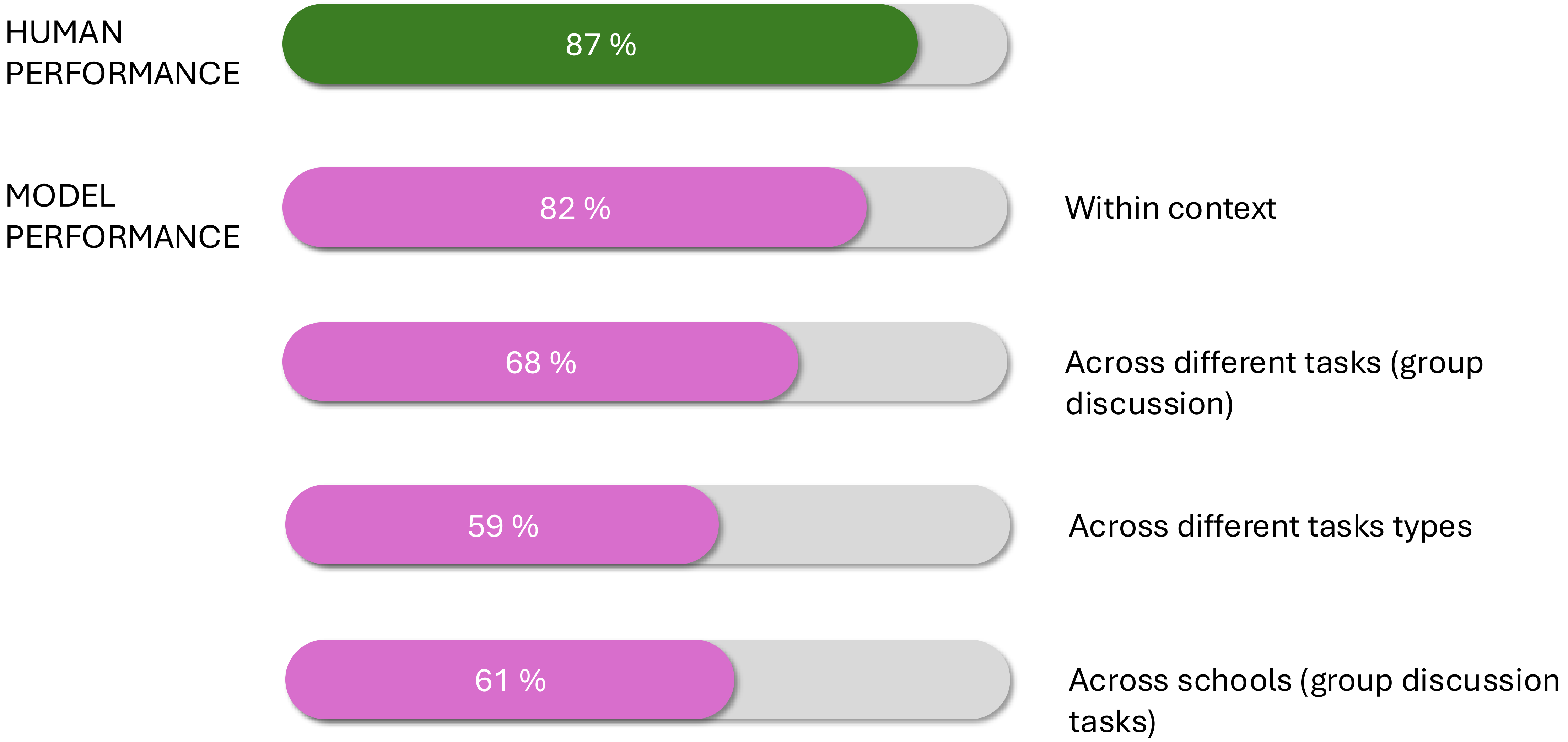
RESULTS

GENERALIZABILITY ACROSS CONTEXTS



RESULTS

GENERALIZABILITY ACROSS CONTEXTS

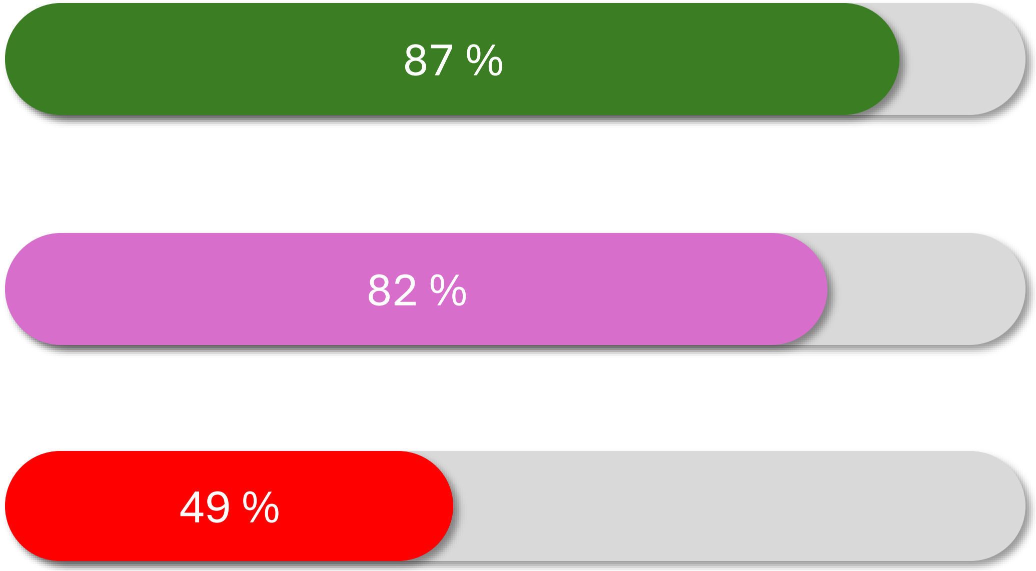


RESULTS

GENERALIZABILITY ACROSS CONTEXTS

HUMAN PERFORMANCE

MODEL PERFORMANCE

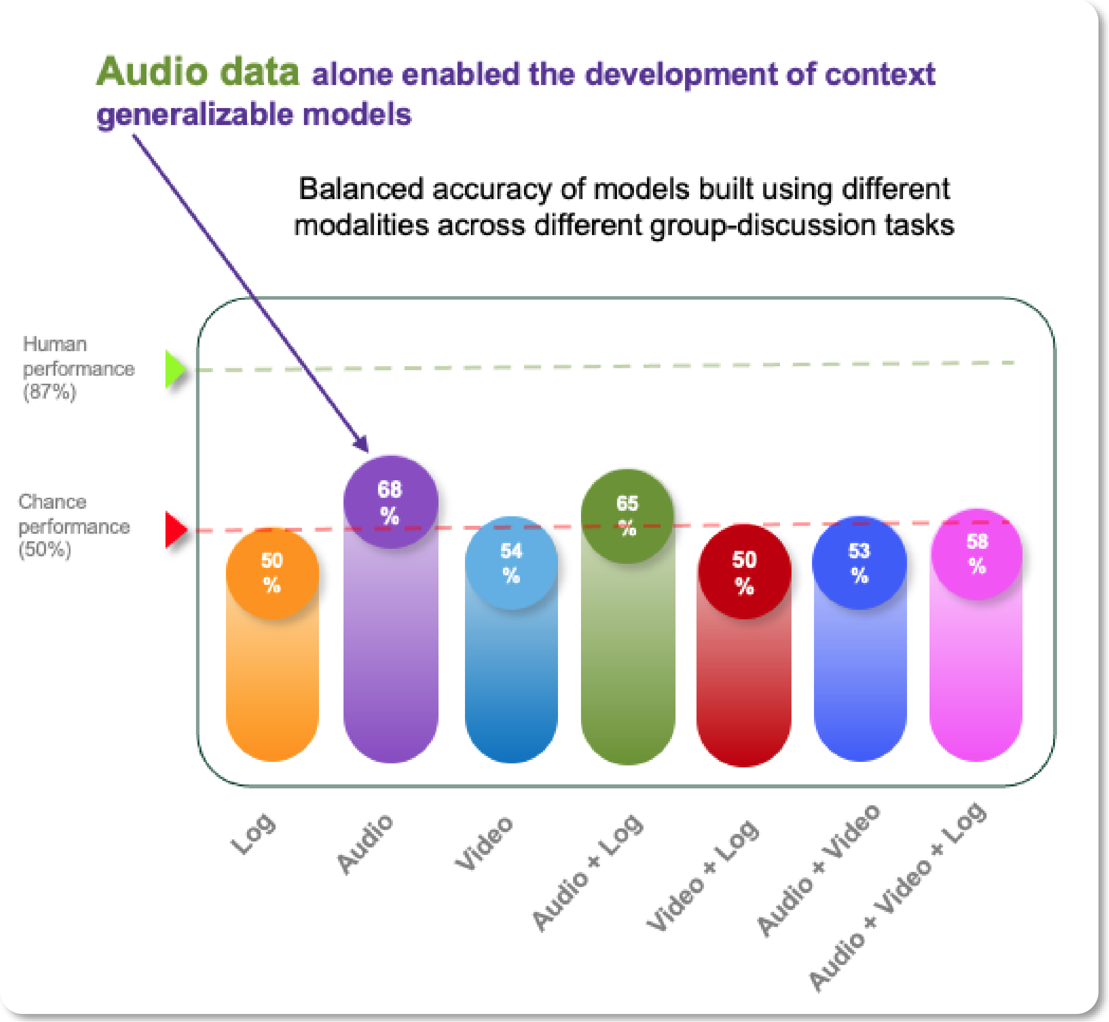


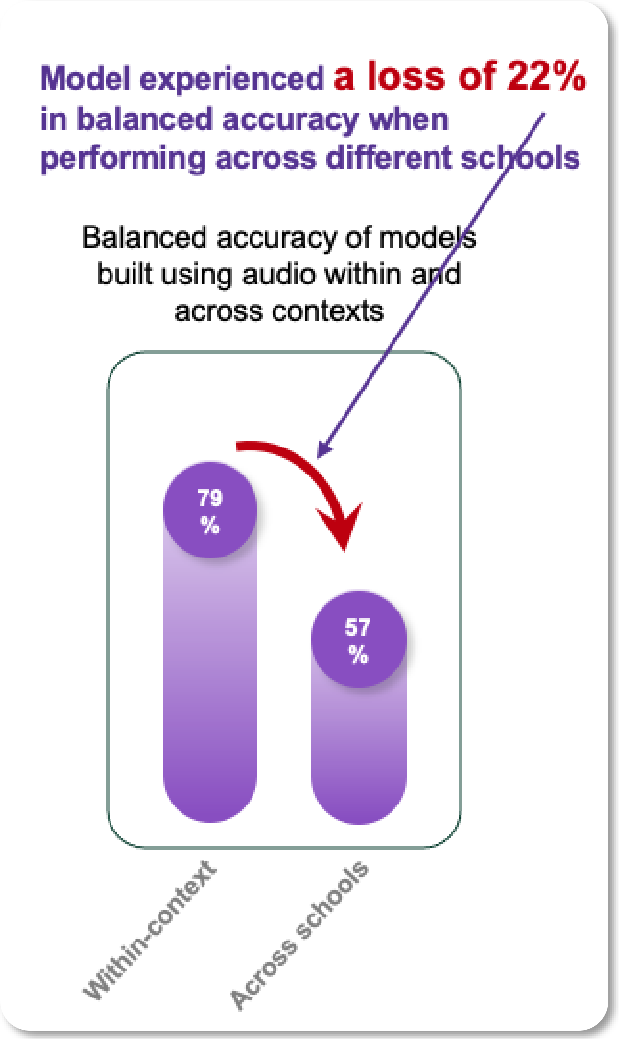
Within context

Across different tasks (collaborative writing)

RESULTS

DATA IMPORTANCE





Multimodal Data Collection & Collaboration Monitoring

CoTrack

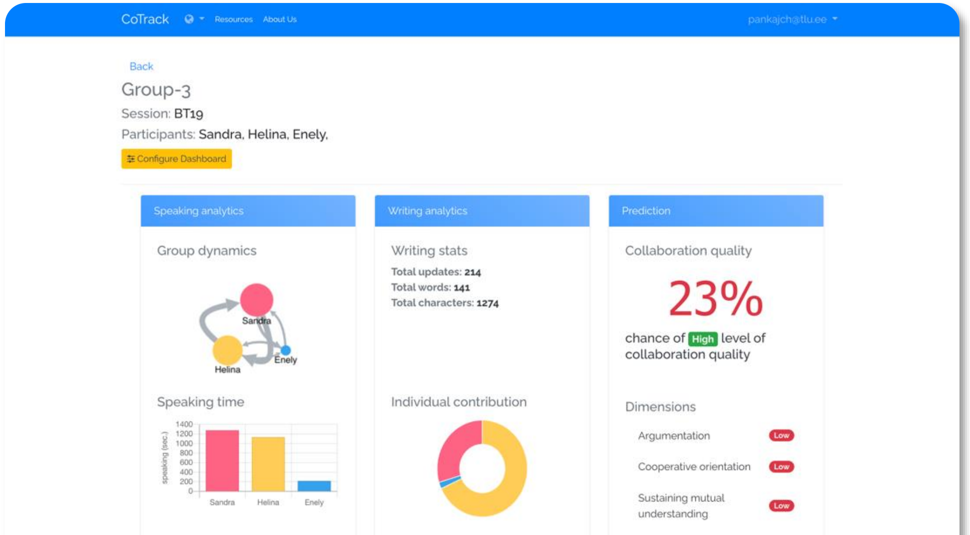
Tool Demo

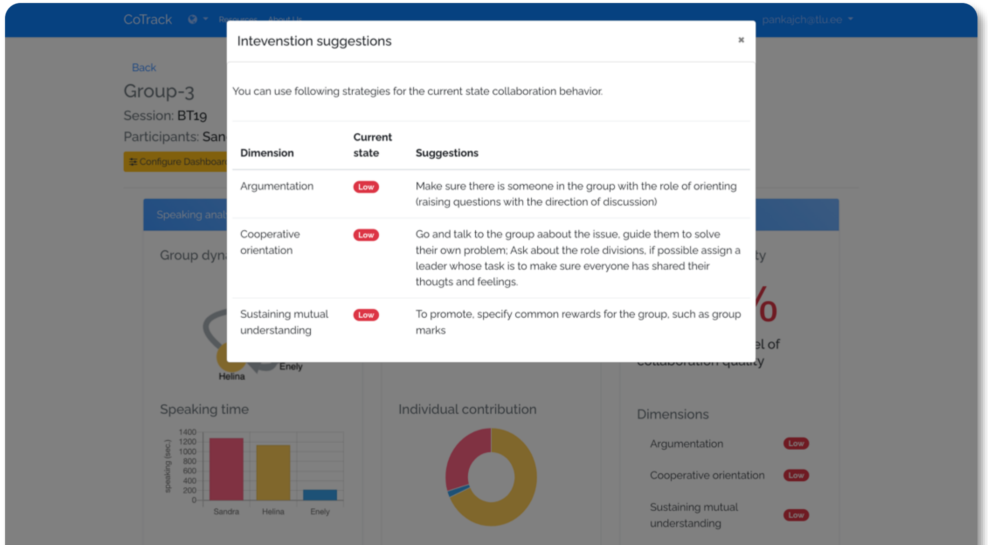
https://www.cotrack.website

6 Researchers

58 Teachers

600 Students





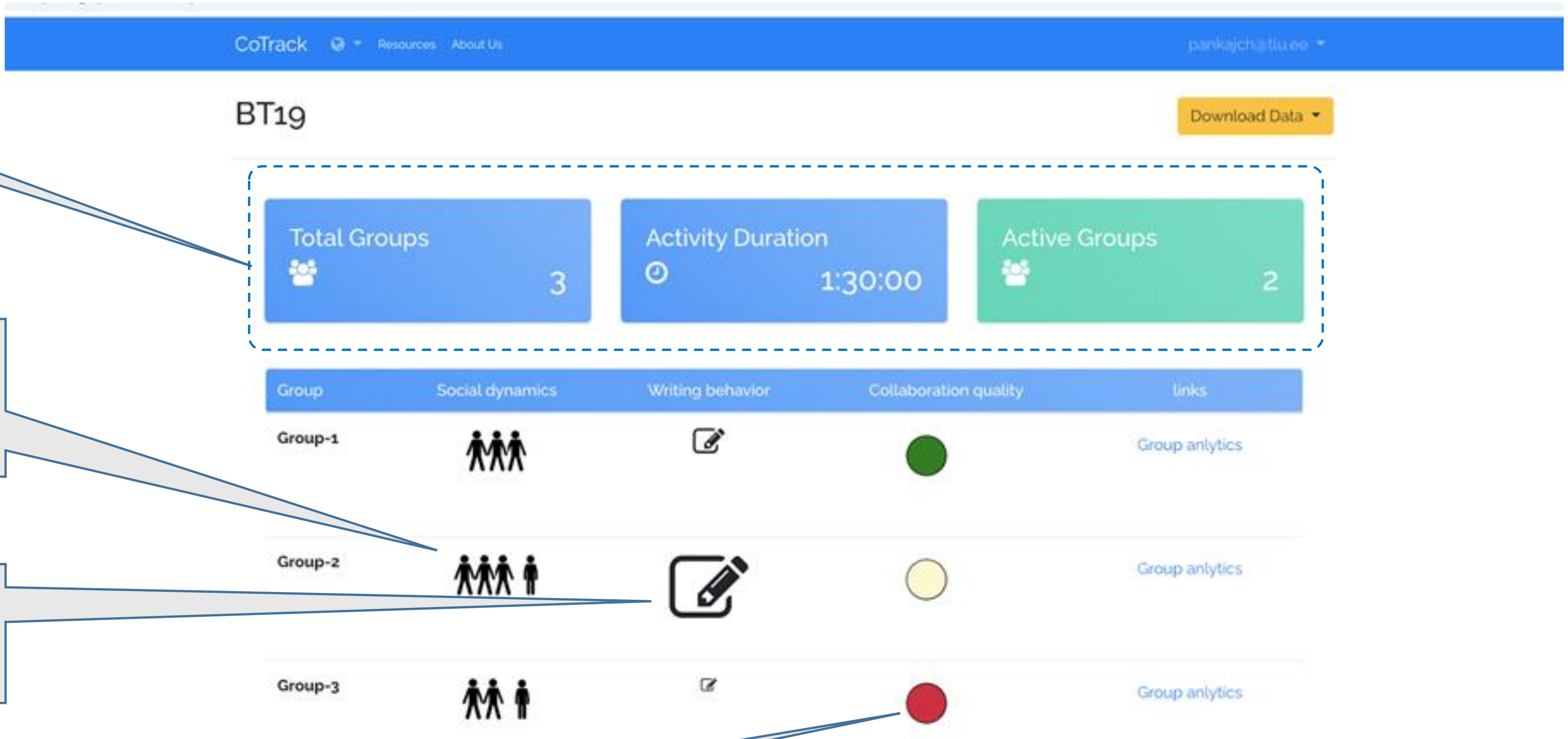
CLASSROOM VIEW

Basic details

Showing how equally group participants are speaking (in this example, only three students are talking to each other more often)

It shows the size of text each group has produced in comparison with other groups.

This traffic system shows the predicted collaboration quality level of the group. (Green: High, Yellow: Medium, Red: Low)



Chejara, P., Kasepalu, R., Prieto, L., P., Rodríguez-Triana, M. J., & Ruiz-Calleja, A. (2024). Bringing collaboration analytics using multimodal data to the masses: Evaluation and design guidelines for developing a mmla system for research and teaching practices in CSCL. *In the 14th International Learning Analytics and Knowledge Conference (LAK24)*. ACM. https://doi.org/10.1145/3636555.3636877

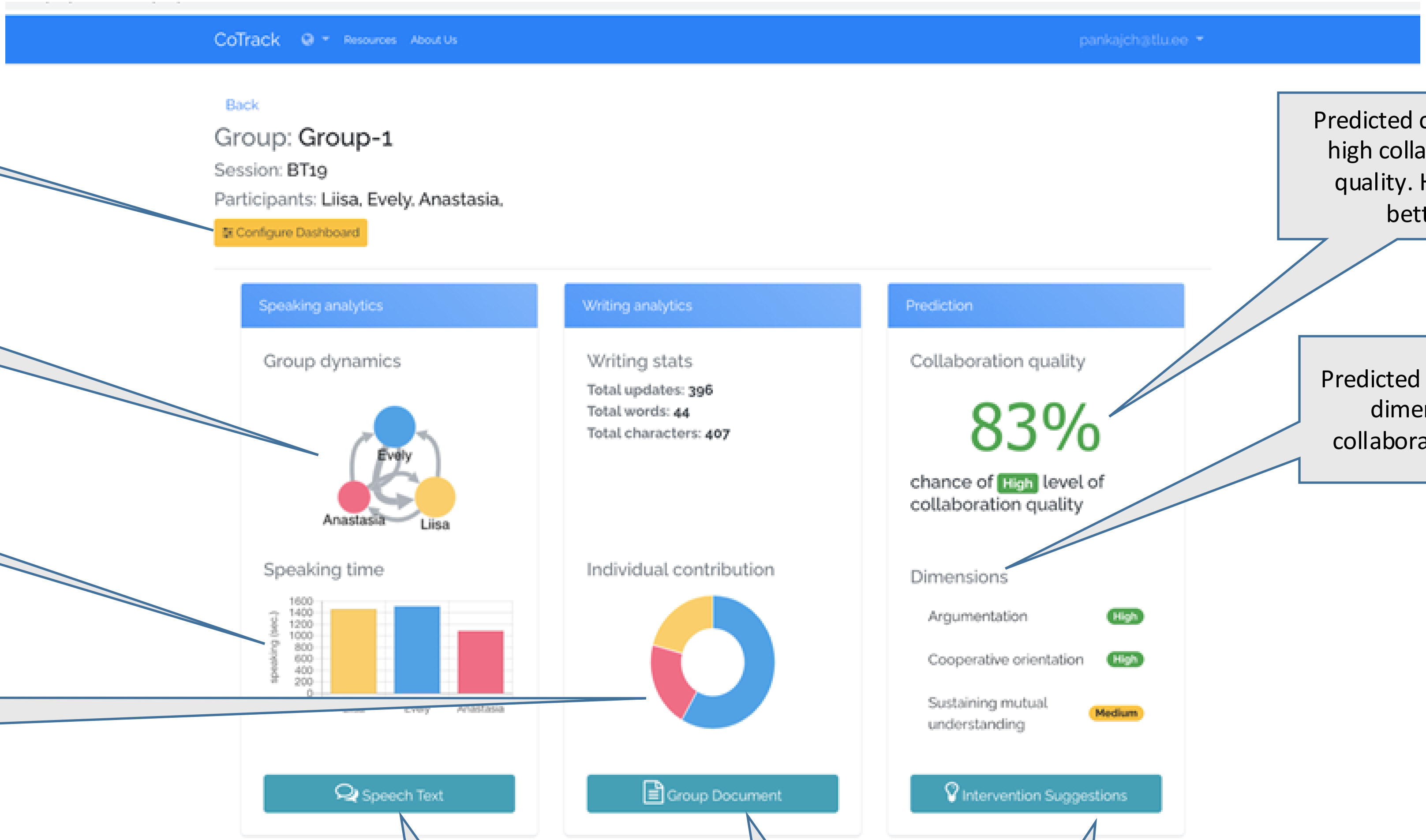
Group VIEW

Configure the dashboard.

Group dynamics in terms of who is talking to whom.

Speaking time of each group member.

It shows how much everyone in the group contributed to written text document.



Predicted chances of high collaboration quality. Higher is better.

Predicted level of three dimensions of collaboration quality.

To see the word cloud of group’s conversation.

To see group's written document.

To see suggestions on interventions.

Guidelines to build context generalizable collaboration estimation models

GUIDELINES

1. Use **60 seconds time window** for data segmentation for modeling collaboration quality using multimodal data

2. Use **Random Forest** for building robust ML models for collaboration quality

3. Use of **contextual data to** build context **generalizable models for estimating collaboration quality**

[Thesis link]

FUTURE DIRECTIONS

1. Small dataset size
2. Investigation using cross-modal features
3. Impact of choosing different choices of ML modeling step on generalizability
4. Privacy-preserving approaches for MMLA
5. Teacher's perception and response to AI-enabled systems

CONCLUSION

1. Time to move research from laboratories to practice.
2. Teacher-AI hybrid partnership

“ *Your **reasoning** for **WHY** you do WHAT YOU DO is more critical than WHAT YOU DO.*

Anonymous

Thank you

Pankaj Chejara
pankajch@tlu.ee

References

1. Chounta, I. A., & Avouris, N. (2016). Towards the real-time evaluation of collaborative activities: Integration of an automatic rater of collaboration quality in the classroom from the teacher's perspective. *Education and Information Technologies*, *21*(4), 815–835. https://doi.org/10.1007/s10639-014-9355-3

2. Rummel, N., Deiglmayr, A., Spada, H., Kahrimanis, G., & Avouris, N. (2011). Analyzing Interactions in CSCL. *Analyzing Interactions in CSCL*, 367–390. https://doi.org/10.1007/978-1-4419-7710-6

3. DiMicco, J. M., Pandolfo, A., & Bender, W. (2004). Influencing group participation with a shared display. *Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work - CSCW '04*, 614. https://doi.org/10.1145/1031607.1031713

4. Bachour, K., Kaplan, F., & Dillenbourg, P. (2010). An interactive table for supporting participation balance in face-to-face collaborative learning. *IEEE Transactions on Learning Technologies*, *3*(3), 203–213. https://doi.org/10.1109/TLT.2010.18

5. Lubold, N., & Pon-Barry, H. (2014). Acoustic-Prosodic Entrainment and Rapport in Collaborative Learning Dialogues. *Proceedings of the 2014 ACM Workshop on Multimodal Learning Analytics Workshop and Grand Challenge - MLA '14*, 5–12. https://doi.org/10.1145/2666633.2666635

5. Ochoa, X., & Worsley, M. (2016). Augmenting Learning Analytics with Multimodal Sensory Data. *Journal of Learning Analytics*, 3(2), 213–219. https://doi.org/10.18608/jla.2016.32.10

4. Raschka, S. Model Evaluation, Model Selection, and Algorithm Selection in Machine Learning. arXiv 2018, arXiv:1811.12808.

6. Martínez-maldonado, R. (2011). Modelling and Identifying Collaborative Situations in a Collocated Multi- display Groupware Setting. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *6738*(June), 39–46. https://doi.org/10.1007/978-3-642-21869-9

7. Spikol, D., Ruffaldi, E., Dabisias, G., & Cukurova, M. (2018). Supervised machine learning in multimodal learning analytics for estimating success in project-based learning. *Journal of Computer Assisted Learning*, *34*(4), 366–377. https://doi.org/10.1111/jcal.12263

8. Drachsler, H., & Greller, W. (2016, April). Privacy and analytics: it's a DELICATE issue a checklist for trusted learning analytics. In *Proceedings of the sixth international conference on learning analytics & knowledge* (pp. 89-98).