# AI based algorithms for teaching method selection: Using tandem learning in mathematics.

Bor Breganta \* orcid id 0009-0000-9331-5391, Darjo Feldaa, Daniel Doza

aFaculty of Education, University of Primorska, Koper, Slovenia

\*corresponding author

Bor Bregant

Faculty of Education, University of Primorska

Cankarjeva 5, 6000 – Koper

Slovenia

bor.bregant@gmail.com

# AI based algorithms for teaching method selection: Using tandem learning in mathematics.

## Abstract

### Background

The main objective of higher education institutions is to provide quality education to its students. One way to achieve this is is by introducing various teaching methods, one of which is tandem learning. Not everyone responds well to a one-size-fits-all method, and therefore, uncovering insights for predictive model selection tailored to individual students or classrooms becomes imperative for teaching institutions. The knowledge is hidden among the educational data set and is extractable through data mining techniques. The aim of this study was to evaluate the performance of machine learning algorithms for predicting student response to tandem learning.

### Methods

A sample of 89 high school students and 13 predictor variables has been used. The outcome of interest was a three state variable indicating whether the student responded well to implementation of tandem learning into education environment or not. In this paper, we implemented 9 classification machine learning algorithms that can be used to predict a target variable with three states and evaluated their performance with a 5 by 2-fold cross-validation with stratified folds.

### Results

Algorithm ... performed best with ... accuracy, f\_1 score ...

### Conclusion

The results imply that machine learning algorithms struggle to accurately predict students' responses to group learning in mathematics using the variables and sample size employed. As a result, they may not be appropriate for aiding teachers in making decisions about selecting teaching methods.

### Keywords

Assessment, education, machine learning, tandem learning, data mining, teching methods

### Math subject classification, MSC2020

97D40, 97D60, 62P99

## Introduction and theoretical framework

### Teaching methods and tandem learning

Critic of frontal teaching and new theoretical didactics, psychological, pedagogic, sociologic findings and positive experience in practical work have lead to the development of new indirect forms of education process (Blažič et al., 2003). Based on strong research literatures various forms of small-group learning are effective in promoting greater academic achievement, more favorable attitudes toward learning, and increased persistence through SMET courses and programs (Roschelle et al., 2010). Tandem learning is a special learning approach, where two students make an experiment together, formulate a report, solve a problem etc. (Tomić, 2002). It is a simple approach from organizational standpoint, as pair members have more chance for activity than in frontal teaching and group teaching, however they are not alone as in individual teaching method (Blažič et al., 2003).

Hundreds of studies have been conducted with main objective being being to determine the effects of cooperative learning on student achievement. We must keep in mind that this learning method is not only theoretical and a debate of research; it is used at some level by millions on teachers (Slavin et al., 2003) (v viru tudi dejanska cifra 81% v ameriki uciteljev to na dnevni ravni...). Many studies, which can be found in (Johnson & Johnson, 2011; Slavin, 1996; Webb, 1991) have found positive effect of for cooperative learning. (iz roschelle – tu notri tudi s kaksnim faktorjem posamezne raziskave)

Spremenljivke, ki vplivajo na delo v skupini – zakaj smo izbrali v raziskavi

By synthesizing these diverse factors, we can develop a more holistic framework for predicting the effects of tandem learning on student performance and tailor educational strategies accordingly.

Tole zgori je bolje napisano v Variable importance.... Bomo še spremenili.

From the above mentioned literature it might be understood that several factors impact the efficacy of working in small groups, such as tandem learning. For instance, students’ demographic characteristics (gender, age, class), teacher-related factors, and student-related factors (levels of mathematics anxiety, students’ personality, previous grade in mathematics) might have a non-negligible impact on predicting the efficacy of tandem learning in mathematics. Therefore, the aim of the present study is to investigate whether these factors can accurately predict the efficacy of tandem learning. To explore these influences, machine learning techniques might be applied.

### Machine learning and classification

Data mining is the process of uncovering hidden patterns, relationships, or insights within vast datasets through techniques from statistics and database management (Baradwaj & Pal, 2012). It involves data preprocessing to prepare information for analysis and utilizes methods such as clustering and association rule mining (Singhal & Jena, 2013). In contrast, machine learning, a subset of artificial intelligence, focuses on building predictive models by allowing computers to learn from data and make decisions or predictions (Candanedo et al., 2018). The sequences of steps identified in extracting knowledge from data is shown in Figure 2 below.

A black arrow pointing to a square

Description automatically generated

Figure 1: Knowledge discovery process

Machine learning encompasses various learning paradigms, including classification process and finds applications in areas like recommendation systems. Classification, a fundamental task in both data mining and machine learning, involves categorizing data into predefined classes, such as binary or three-state classification, based on patterns learned from labeled data. This classification process is employed in various domains, such as healthcare and sentiment analysis, to make data-driven decisions (Baradwaj & Pal, 2012). Goal of classification is to build a model based on input data, that explains said data. If we put new data in our model, the model outputs a solution based on input data it was built on. Usually, we have training data where are attributes and is a value of uknown function . Our goal is to find a function h that is the best approximation of function . Attributes (predictor variables) are independent (vectors), are target variables, function is called a model. Space of hypothesis expands very rapidly. If we have binary attributes, we would have different learning inputs and possible hypothesis. *To je bolj za v disertacijo kot za sam članek.*

Malo več o klasifikaciji ali ne? To je itak statistika, katere ne uvajamo v člankih, čeprav je tu bolj kompleksna?

### Connecting tandem learning and machine learning

Technologies can be considered in terms of whether they are mainly student teaching (with primarily instructionist approach), student supporting (primarily constructivist approach), or teacher supporting (which primarily help teachers do tasks they already do but faster or with less effort) (Holmes et al., 2019).

Beyond its broader applications, machine learning has been harnessed to predict student performance with remarkable precision (Ofori et al., 2020; Qazdar et al., 2019; Rastrollo-Guerrero et al., 2020; Yakubu & Abubakar, 2022). Leveraging the power of data analytics and advanced algorithms, machine learning models have been applied to forecast student success (Yakubu & Abubakar, 2022), identify at-risk learners (Adnan et al., 2021; Chui et al., 2020), and tailor educational interventions (Luan & Tsai, 2021; Stimpson & Cummings, 2014; Tsai et al., 2020; Yang, 2021). This transformative application of machine learning is exemplified by research conducted by Siemens & Gasevic (2012), which introduced the concept of "learning analytics" and demonstrated its potential in predicting student outcomes, which was demonstrated also by other studies (Abana, 2019; Bhusal, 2021; Cortez & Silva, 2008; Kotsiantis et al., 2004; Minaei-Bidgoli et al., 2003). Aside from forecasting success, machine learning can help us identify the most important variables that affect said forecast (Lu et al., 2020; Luan & Tsai, 2021). Multiple studies have dvelved into the analysis of crucial features in learning enviroment (Hodges, 2018; Humphrey et al., 2009; Moradi et al., 2018; Scribner & Donaldson, 2001), however only few have harnessed the power of modern algorithms, such as machine learning, which hold the potential for significantly enhanced sight. Therefore, it is of paramount importance to explore in great detail the feature selection problem.

## Empirical work

Research was conducted as he effectiveness of tandem learning in high school mathematics remains unclear due to the complexity of numerous variables influencing its success.

This study aims to leverage three state classification machine learning algorithms to analyze multifaceted variables to determine successfulness of tandem learning enviroment.

Given the scarcity of comprehensive studies employing machine learning algorithms to assess teaching method selections, this research acknowledges the potential limitations in drawing definitive conclusions. While aiming to discern patterns of predicting successfulness of tandem learning enviroment, this study will consider the need for cautious interpretation of results. Even in the absence of conclusive findings, insights gathered from the analysis will contribute to the ongoing discourse on the efficacy of tandem learning in mathematics education. Therefore we hypothesize that even when feeding machine learning models with diverse data sets encompassing student profiles and collaborative dynamics some algorithms will perform sub par, while others may perform fair.

## Methodology

In the present research, the causal descriptive method of pedagogical research is applied.

## Sample

In this retrospective study we used a dataset gathered at a high school in Slovenia, which is pulicly available at (Bregant, 2023). Sample comprised of 89 11th and 12th grade of a Slovenian Gymnasium (i.e., high school); 28 males and 61 females. All variables provided were used. Internal consistency of a dataset was already performed at (Bregant et al., 2024), with McDonal's Omega for continuous data and Gutman's Lambda for categorical variables all of them having fair results.

## Data preprocessing and data analysis

The gathered data was analysed using Python programming language, primarily using pandas (version 2.1.3) and scikit-learn (version 1.3.2) libraries. Machine learning code, as well as variable importance and other data analysis notebooks are openly accessible at (Bregant, 2023)

Dataset was already in the form of tidy data (Wickham, 2014). Categorical variables, which were a priori encoded using label encoding were transformed into dummy variables using one-hot as most machine learning algorithms require numerical input and do not handle well numerical data, that is not continuous (Tan & Pu, 2023; Wu et al., 2020).

We implemented 9 ML algorithms: Naive Bayes, K-Nearest Neighbors, Decision Tree, Gaussian Mixture, Linear Discriminant Analysis, Ada Boots (with 1000 estimators), Gradient Boosting (with 1000 estimators), Support Vector Machine (with RBF kernel) and Random Forest (with 1000 estimators). Said algorithms are capable of performing 3 state classification. Models were evaluated using 5x2-CV with stratified folds with the main metric of performance evaluation being F1-score through average fold repetition.

## Results

### Student sample and variables used

Target variable was a three state measure of successfulness on Likert scale, both regarding academic performance as well as general wellbeing during implementation of tandem learning. general sense (gender, class, professor, previous grade,) in psychological sense (MBTI variables: extroversion-introversion, sensing-intuition, thinking-feeling and judging-perceiving and other variables: mathematical anxiety and motivation) and in regards to tandem learning (qualitative interaction, quantitative interaction and whether student outperformed their partner). For a more detailed and precise dataset description, refer to (Bregant et al., 2024) for comprehensive insights.

For machine learning purposes that follows it is important to know that target variable was not balanced with represented classes being 6, 39 and 44 respectively on Likert scale.

Feature importance was already performed by (Bregant et al., 2024), indicating that variables in regards to tandem learning are most important, followed by variables in general sense and lastly variables regarding psychological profile of students were least important.

A group of graphs with different colored bars

Description automatically generated with medium confidence

Figure 2: Variables used

### Model performance

According to F1 metric, the three models with generally better performances in our experiments were Random forest, K-Nearest Neighbors, and Gradient boosting. Their performance are fair, as their accuracy was still better than number of samples in biggest class (49.4%) Torej bolje od ugibanja. Other models performed worse, their poor performance indicating, they might not be learning the underlying patterns effectively. Whole classification can be found in Table 1.

Table 1: Classification results

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Classifier | Accuracy | Precision | Recall | F1-score |
| Naive Bayes | 0.338 | 0.272 | 0.377 | 0.258 |
| K-Nearest Neighbors | 0.505 | 0.344 | 0.361 | 0.343 |
| Decision Tree | 0.471 | 0.344 | 0.337 | 0.312 |
| Gaussian Mixture | 0.023 | 0.017 | 0.050 | 0.024 |
| Linear Discriminant Analysis | 0.450 | 0.328 | 0.381 | 0.311 |
| AdaBoost | 0.484 | 0.307 | 0.350 | 0.316 |
| Gradient Boosting | 0.505 | 0.367 | 0.368 | 0.341 |
| Support Vector Machine | 0.450 | 0.159 | 0.304 | 0.208 |
| Random Forest | 0.550 | 0.427 | 0.397 | 0.351 |

In analyzing the model's performance, the plotted confusion matrices provide a visual representation of its classification accuracy, revealing the interplay between true positives, true negatives, false positives, and false negatives.

Figure 2: Confusion matrices of classification

A group of blue squares

Description automatically generated

To counter the imbalance in our dataset, we opted for a binary approach, merging classes 0 and 1 to create a more balanced representation. Therefore we got class 0 with 45 (50.6%) and class 1 with 44 (49.4%) representations. With this approach all algorithms, except AdaBoost performed fair. Support vector machine performed best, with moderate accuracy (0.59 averaged F1-score).

Table 3: Classification result of two classes

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Classifier | Accuracy | Precision | Recall | F1-score |
| Naive Bayes | 0.583 | 0.539 | 0.581 | 0.539 |
| K-Nearest Neighbors | 0.516 | 0.521 | 0.518 | 0.506 |
| Decision Tree | 0.518 | 0.511 | 0.517 | 0.504 |
| Gaussian Mixture | 0.540 | 0.581 | 0.532 | 0.447 |
| Linear Discriminant Analysis | 0.563 | 0.482 | 0.561 | 0.504 |
| AdaBoost | 0.473 | 0.421 | 0.469 | 0.420 |
| Gradient Boosting | 0.551 | 0.503 | 0.549 | 0.509 |
| Support Vector Machine | 0.607 | 0.638 | 0.607 | 0.586 |
| Random Forest | 0.561 | 0.513 | 0.560 | 0.520 |

We also performed ML algorithms with less predictor variables, as discussed. We dropped all six variables regarding psychological profile of students in sample and achieved low performance on all algorithms mentioned above.

Table 4: Classification results with selected features

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Classifier | Accuracy | Precision | Recall | F1-score |
| Naive Bayes | 0.339 | 0.260 | 0.376 | 0.250 |
| K-Nearest Neighbors | 0.494 | 0.343 | 0.359 | 0.340 |
| Decision Tree | 0.449 | 0.330 | 0.323 | 0.310 |
| Gaussian Mixture | 0.033 | 0.016 | 0.142 | 0.027 |
| Linear Discriminant Analysis | 0.394 | 0.239 | 0.340 | 0.251 |
| AdaBoost | 0.359 | 0.252 | 0.263 | 0.250 |
| Gradient Boosting | 0.405 | 0.255 | 0.294 | 0.265 |
| Support Vector Machine | 0.495 | 0.165 | 0.333 | 0.221 |
| Random Forest | 0.494 | 0.301 | 0.359 | 0.315 |

Given the suboptimal performance of our model, we employed t-SNE analysis to visually explore and comprehend the underlying patterns and relationships within the data, aiming to uncover potential complexities or overlaps that might have impacted the model's performance. It can be found in Appendix A, and shows that the patterns in data may be hard to distinguish between, as was speculated.

## Discussion

…

## Conclusions and limitations

This study demonstrates that machine learning algorithms can be used for personalized predictive estimation of whether students respond well to new teaching methods. As using AI models is a fairly new construct, especially in education research, there is still a lot of room for improvement. Key factors that influence results must be investigated further to achieve better model outcomes. The potential incorporation of the developed model in regards to teacher assistance needs to be investigated further.

Study does not include a prediction whether tandem learning is overall effective or not. It simply includes information about predicting certain student performance, which is a limitation. Some of the variables, likely relevant (glej clanek variable importance) were not used in our models because they were not obtained, which might have led to a slight bias in our study,

## Acknowledgements

We extend our heartfelt gratitude to the study participants and their dedicated professors for their invaluable contributions. Additionally, we deeply appreciate the insightful comments provided by our esteemed reviewers and editors.

## References

Abana, E. C. (2019). A Decision Tree Approach for Predicting Student Grades in Research Project using Weka. *International Journal of Advanced Computer Science and Applications*, *10*(7). https://doi.org/10.14569/IJACSA.2019.0100739

Adnan, M., Habib, A., Ashraf, J., Mussadiq, S., Raza, A. A., Abid, M., Bashir, M., & Khan, S. U. (2021). Predicting at-Risk Students at Different Percentages of Course Length for Early Intervention Using Machine Learning Models. *IEEE Access*, *9*, 7519–7539. https://doi.org/10.1109/ACCESS.2021.3049446

Baradwaj, B. K., & Pal, S. (2012). *Mining Educational Data to Analyze Students’ Performance* (arXiv:1201.3417). arXiv. https://doi.org/10.48550/arXiv.1201.3417

Bhusal, A. (2021). *Predicting Student’s Performance Through Data Mining*. https://doi.org/10.48550/ARXIV.2112.01247

Blažič, M., Ivanuš-Grmek, M., Kramar, M., & Strmčnik, F. (2003). *Didaktika: Visokošolski učbenik*. Visokošolsko središče, Inštitut za raziskovalno in razvojno delo.

Bregant, B. (2023). *Tandem learning: Student dataset* (1.0) [dataset]. GitHub. https://github.com/borbregant/ai\_tandem\_learning

Bregant, B., Doz, D., & Mešinović, S. (2024). *Influence of certain factors for tandem learning in mathematics* [Unpublished manuscript].

Candanedo, I. S., Nieves, E. H., González, S. R., Martín, M. T. S., & Briones, A. G. (2018). Machine Learning Predictive Model for Industry 4.0. In L. Uden, B. Hadzima, & I.-H. Ting (Eds.), *Knowledge Management in Organizations* (Vol. 877, pp. 501–510). Springer International Publishing. https://doi.org/10.1007/978-3-319-95204-8\_42

Chui, K. T., Fung, D. C. L., Lytras, M. D., & Lam, T. M. (2020). Predicting at-risk university students in a virtual learning environment via a machine learning algorithm. *Computers in Human Behavior*, *107*, 105584. https://doi.org/10.1016/j.chb.2018.06.032

Cortez, P., & Silva, A. (2008). *Using data mining to predict secondary school student performance*.

Hodges, L. C. (2018). Contemporary Issues in Group Learning in Undergraduate Science Classrooms: A Perspective from Student Engagement. *CBE—Life Sciences Education*, *17*(2), es3. https://doi.org/10.1187/cbe.17-11-0239

Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial Intelligence in Education. Promise and Implications for Teaching and Learning.*

Humphrey, N., Lendrum, A., Wigelsworth, M., & Kalambouka, A. (2009). Implementation of primary Social and Emotional Aspects of Learning small group work: A qualitative study. *Pastoral Care in Education*, *27*(3), 219–239. https://doi.org/10.1080/02643940903136808

Johnson, D. W., & Johnson, R. T. (2011). *Learning together and alone: Cooperative, competitive, and individualistic learning* (5. ed. [Repr.]). Allyn and Bacon.

Kotsiantis, S., Pierrakeas, C., & Pintelas, P. (2004). Predicting students’ performance in distance learning using machine learning techniques. *Applied Artificial Intelligence*, *18*(5), 411–426. https://doi.org/10.1080/08839510490442058

Lu, D.-N., Le, H.-Q., & Vu, T.-H. (2020). The Factors Affecting Acceptance of E-Learning: A Machine Learning Algorithm Approach. *Education Sciences*, *10*(10), 270. https://doi.org/10.3390/educsci10100270

Luan, H., & Tsai, C.-C. (2021). A Review of Using Machine Learning Approaches for Precision Education. *Educational Technology & Society*, *24*(1), 250–266.

Minaei-Bidgoli, B., Kashy, D. A., Kortemeyer, G., & Punch, W. F. (2003). Predicting student performance: An application of data mining methods with an educational web-based system. *33rd Annual Frontiers in Education, 2003. FIE 2003.*, *1*, T2A\_13-T2A\_18. https://doi.org/10.1109/FIE.2003.1263284

Moradi, S., Faghiharam, B., & Ghasempour, K. (2018). Relationship Between Group Learning and Interpersonal Skills With Emphasis on the Role of Mediating Emotional Intelligence Among High School Students. *SAGE Open*, *8*(2), 215824401878273. https://doi.org/10.1177/2158244018782734

Ofori, F., Maina, E., & Gitonga, R. (2020). Using Machine Learning Algorithms to Predict Students’ Performance and Improve Learning Outcome: A Literature Based Review. *Journal of Information and Technology*, *4*(1), Article 1.

Qazdar, A., Er-Raha, B., Cherkaoui, C., & Mammass, D. (2019). A machine learning algorithm framework for predicting students performance: A case study of baccalaureate students in Morocco. *Education and Information Technologies*, *24*(6), 3577–3589. https://doi.org/10.1007/s10639-019-09946-8

Rastrollo-Guerrero, J. L., Gómez-Pulido, J. A., & Durán-Domínguez, A. (2020). Analyzing and Predicting Students’ Performance by Means of Machine Learning: A Review. *Applied Sciences*, *10*(3), 1042. https://doi.org/10.3390/app10031042

Roschelle, J., Rafanan, K., Bhanot, R., Estrella, G., Penuel, B., Nussbaum, M., & Claro, S. (2010). Scaffolding group explanation and feedback with handheld technology: Impact on students’ mathematics learning. *Educational Technology Research and Development*, *58*(4), 399–419. https://doi.org/10.1007/s11423-009-9142-9

Scribner, J. P., & Donaldson, J. F. (2001). The Dynamics of Group Learning in a Cohort: From Nonlearning to Transformative Learning. *Educational Administration Quarterly*, *37*(5), 605–636. https://doi.org/10.1177/00131610121969442

Siemens, G., & Gasevic, D. (2012). Guest Editorial—Learning and Knowledge Analytics. *Educational Technology and Society*, *15*(1–2).

Singhal, S., & Jena, M. (2013). A study on WEKA tool for data preprocessing, classification and clustering. *International Journal of Innovative Technology and Exploring Engineering*, *2*(6), 250–253.

Slavin, R. E. (1996). Research on Cooperative Learning and Achievement: What We Know, What We Need to Know. *Contemporary Educational Psychology*, *21*(1), 43–69. https://doi.org/10.1006/ceps.1996.0004

Slavin, R. E., Hurley, E. A., & Chamberlain, A. (2003). Cooperative Learning and Achievement: Theory and Research. In I. B. Weiner (Ed.), *Handbook of Psychology* (1st ed., pp. 177–198). Wiley. https://doi.org/10.1002/0471264385.wei0709

Stimpson, A. J., & Cummings, M. L. (2014). Assessing Intervention Timing in Computer-Based Education Using Machine Learning Algorithms. *IEEE Access*, *2*, 78–87. https://doi.org/10.1109/ACCESS.2014.2303071

Tan, S., & Pu, Y. (2023). Frac-Vector: Better Category Representation. *Fractal and Fractional*, *7*(2), 132. https://doi.org/10.3390/fractalfract7020132

Tomić, A. (2002). *Spremljanje pouka* (1. natis). Zavod Republike Slovenije za šolstvo.

Tsai, S.-C., Chen, C.-H., Shiao, Y.-T., Ciou, J.-S., & Wu, T.-N. (2020). Precision education with statistical learning and deep learning: A case study in Taiwan. *International Journal of Educational Technology in Higher Education*, *17*(1), 12. https://doi.org/10.1186/s41239-020-00186-2

Webb, N. M. (1991). Task-Related Verbal Interaction and Mathematics Learning in Small Groups. *Journal for Research in Mathematics Education*, *22*(5), 366. https://doi.org/10.2307/749186

Wickham, H. (2014). Tidy Data. *Journal of Statistical Software*, *59*, 1–23. https://doi.org/10.18637/jss.v059.i10

Wu, H., Liu, Y., & Wang, J. (2020). Review of Text Classification Methods on Deep Learning. *Computers, Materials & Continua*, *63*(3), 1309–1321. https://doi.org/10.32604/cmc.2020.010172

Yakubu, M. N., & Abubakar, A. M. (2022). Applying machine learning approach to predict students’ performance in higher educational institutions. *Kybernetes*, *51*(2), 916–934. https://doi.org/10.1108/K-12-2020-0865

Yang, S. (2021). Guest Editorial: Precision Education - A New Challenge for AI in Education. *Educational Technology and Society*, *24*(1), 105–108.

## Appendix A: t-SNE

A diagram of a graph

Description automatically generated with medium confidence

A diagram of a graph

Description automatically generated