mis4TEL
International Conference

An advanced explainable and interpretable ML-based framework for educational data mining

13th International Conference on Methodologies and Intelligent Systems for Technology Enhanced Learning

Ioannis LIVIERIS[*], **Nikos KARACAPILIDIS**[**], **Georgios DOMALIS**[*], **Dimitris TSAKALIDIS**[*] ([*] Novelcore, [**] University of Patras)

July 10, 2023

Table of Contents

Introduction

1 Motivation

Educational data mining

- Educational Data Mining (EDM) is a field of research that builds on Machine Learning (ML) and other data analysis techniques to analyze educational data for gaining insights into the learning process

Primary goals

- Identify patterns, relationships and factors that affect the outcomes of the learning process, aiming to predict the learners' future performance and accordingly use this information to improve educational practices and policies

Explainable EDM

Motivation

- In recent years, there is an increasing trend in EDM to provide meaningful explanations to educators, administrators and students about how the underlying ML models arrive at their predictions or decisions

Importance

- **Improve decision making:** Provide educational stakeholders insights about how complex EDM prediction works and how confident they are in their predictions
- **Explainability is important for ethical reasons:** ensure that ML model are fair and unbiased, and that they do not perpetuate or amplify existing biases in the associated data

Table of Contents

Analysis

2 Related work

Hur P., Lee H., Bhat S. and Bosch N.
Using machine learning explainability methods to personalize interventions for students
International Educational Data Mining Society (2022).

Ramaswami, G., Susnjak, T., and Mathrani, A.
On developing generic models for predicting student outcomes in educational data mining
Big Data and Cognitive Computing 6(1), 6 (2022).

Ribeiro, M.T., Singh, S., and Guestrin, C.
"Why should i trust you?" explaining the predictions of any classifier
Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining.(2016).

Tampakas, V., Livieris, I.E., Pintelas, E., Karacapilidis, N., and Pintelas, P.
Prediction of students' graduation time using a two-level classification algorithm
Technology and Innovation in Learning, Teaching and Education. pp. 553–565. Springer (2019).

Limitations

- These works provide none/limited explainability abilities.
- No recommendations is presented for assisting the education process

Table of Contents

Our approach

3 Contribution

Contribution

- We propose an advanced explainable framework for predicting students' performance, which provides accurate, reliable and interpretable predictions

Main components

- **Prediction model** based on NGBoost algorithm
- **Local and Global explainability**, through the use of LIME and SHAP importance methods, respectively

Advantage of the proposed approach

- Predict students' performance with high accuracy and reliability
- It is able to provide local and global explainability, presenting a complete picture as well as with human-interpretable insights about why a particular decision was made by the model

Proposed framework

3 Contribution

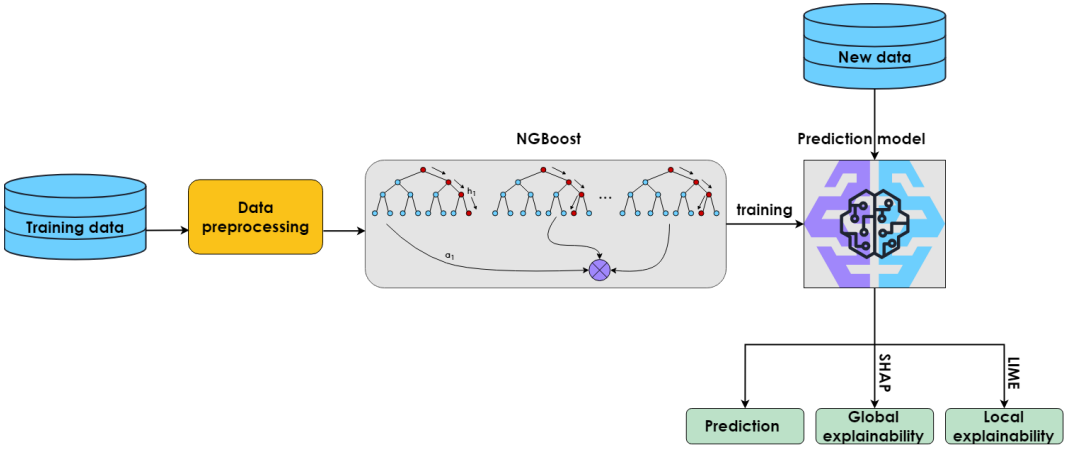


Table of Contents

Numerical experiments

4 Evaluation

Numerical experiments

- Performance evaluation of the NGBoost algorithm against SoA ML algorithms
- Performance metrics
  - Accuracy, Area under Curve (AUC), Geometric Mean (GM), Precision and Recall
- Educational datasets
  - DATASET1: prediction of University students' final performance on an academic course
    - *(Pass, Fail)*
  - DATASET2: prediction of secondary school students' performance from the first two out of three semesters
    - *(Fail, Good, Very Good, Excellent)*

Use case scenarios

- Demonstrate the applicability of our framework and give insights about the recommendations that can be provided to educators and students

Algorithm	Accuracy	AUC	GM	Precision	Recall
LGBM	60.4%	0.633	8.034	0.581	0.633
Random-Forest	70.8%	0.641	8.094	0.605	0.641
XGBoost	66.6%	0.640	8.156	0.592	0.640
NGBoost	**71.2%**	**0.649**	**8.221**	**0.608**	**0.648**

Table 1. Performance evaluation of classification algorithms on DATASET$_1$

Algorithm	Accuracy	AUC	GM	Precision	Recall
LGBM	86.9%	0.911	44.01	0.862	0.867
Random-Forest	89.4%	0.923	44.99	**0.906**	0.884
XGBoost	89.7%	0.924	45	0.905	0.885
NGBoost	**90.0%**	**0.927**	**45.24**	0.897	**0.927**

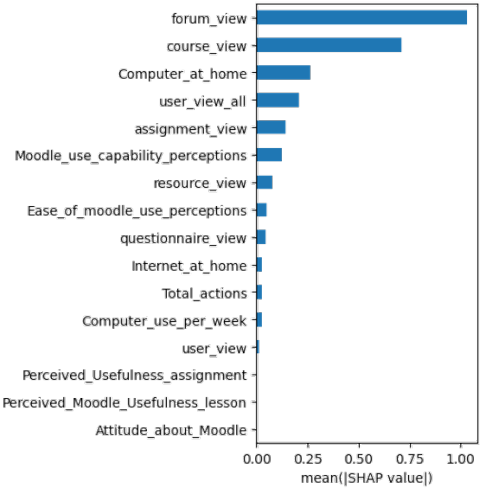
Table 2. Performance evaluation of classification algorithms on DATASET$_2$

Use case scenarios

- NGBoost demonstrates the best overall performance based on Accuracy, AUC and GM metrics
- NGBoost presents the best *Precision-Recall balance* among all evaluated algorithms

Use case 1: Global explainability

4 Evaluation



Conclusions

- The global explainability component shows that the three features that influence most the model are *forum_view*, *course_view* and *computer_at_home*, while the three least important are *Perceived_Moodle_Usefulness*, *Perceived_Usefulness_assignment* and *Attitude_about_Moodle*

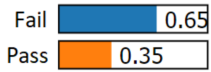
Possible recommendation

Educator should provide special attention to

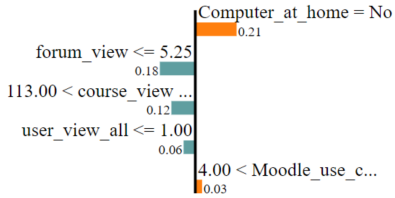
- the number of times each student accesses the description of the forum,
- the number of accesses to the basic material of each week's laboratory session,
- the number of accesses to the forum section,
- and if he/she owes a computer

Use case 1: Local explainability

4 Evaluation



Feature	Value
Computer_at_home	No
forum_view	1.00
course_view	139.00
user_view_all	0.00
Moodle_use_capability_perceptions	5.00



Conclusions

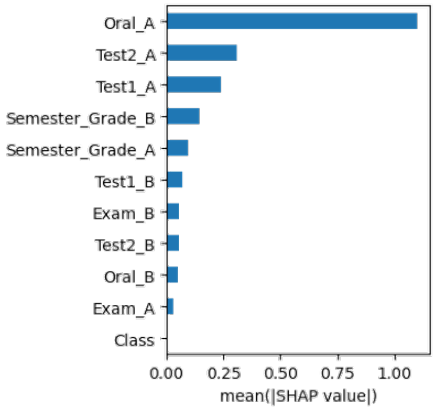
- The model predicts that the student will fail to the pass the class with probability 65%
- The *student has no computer at home* has the highest impact on the model's decision

Possible recommendation

- The student is at high risk of failing
- The student should pay more attention to the basic material of each week's laboratory

Use case 2: Global explainability

4 Evaluation



Conclusions

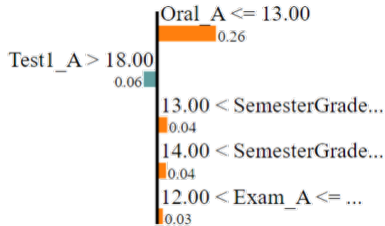
- The global explainability component shows that the three features that influence most the model are *Oral_A*, *Test2_A* and *Test1_A*, while the three least important ones are *Oral_B*, *Exam_A* and *Class*

Possible recommendation

- The students' performance on the first semesters, especially their performance on oral examinations and on 15-minutes tests seem to considerably affect the students' performance on the final examination

Use case 2: Local explainability

4 Evaluation



Conclusions

- The model predicts that the student will have “Good” performance on the final examinations
- The student’s 1st oral grade has the highest impact on the model’s decision

Possible recommendation

- Improve student’s oral argumentation skills

Table of Contents

Summary

5 Conclusion

Contribution

- We proposed an advanced explainable framework for predicting students' performance, which provides accurate, reliable and interpretable predictions

Main components

- A **prediction model** based on NGBoost algorithm
- **Local and Global explainability**, through the use of LIME and SHAP importance methods, respectively

Advantages

- Accurate & reliable predictions since NGBoost is able to outperform traditional state-of-the-art ML algorithms
- The adoption of SHAP ensures that the importance scores are fair and unbiased, while the adoption of LIME provides a flexible, fast and reliable technique for interpreting a single predictions

Next steps

5 Conclusion

Limitations

- Proposed framework was evaluated only on two real-world datasets
- Datasets contained a limited number of attributes

Future work

- Elaborate additional educational datasets, and accordingly evaluate the proposed framework across diverse challenging issues in the educational domain
- Automatic generation of recommendations in natural language (*highly human-interpretable*) based on the feedback provided by the proposed framework

Thank you

6 Acknowledgements

This work received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101061509, for the project:
augMENTOR: Augmented Intelligence for Pedagogically Systained Training and Education