

# ICRAMCS 2026

THE EIGHTH EDITION OF THE INTERNATIONAL CONFERENCE ON  
RESEARCH IN APPLIED MATHEMATICS AND COMPUTER SCIENCE

April 23-24-25, 2026 | Marrakech, Morocco



## Predicting Low Achievement in PISA 2022 Morocco: A Machine-Learning Comparison of Random Forest, XGBoost, SVM, and MLP

### Communication Info

#### Authors:

Ibrahim ALLALI<sup>1</sup>  
Lahboub ZOUIRI<sup>1</sup>

<sup>1</sup> *Laboratory of Applied  
Economics, Faculty of Legal,  
Economic and Social Sciences –  
Agdal, Mohammed V University  
of Rabat, Morocco*

#### Keywords:

- (1) Machine learning
- (2) XGBoost
- (3) PISA 2022
- (4) Low achievement
- (5) Morocco

### Abstract

This study develops machine-learning early-warning models to identify Moroccan students at risk of low achievement in PISA 2022 [1]. Low achievement is defined as the bottom quartile of plausible-value-based performance distributions in mathematics and science. Using a merged student-school dataset ( $N = 6,867$ ) and 26 predictors capturing socioeconomic background, school climate, ICT resources, and school constraints, we compare Random Forest [2], XGBoost [3], radial-basis SVM [4], and a multilayer perceptron under an 80/20 stratified train-test split with training-only median/mode imputation. Performance is assessed using AUC and complementary threshold-based metrics used in educational data mining [5]. XGBoost performs best (AUC = 0.841 in mathematics; 0.796 in science), followed by Random Forest (0.824; 0.792). Risk stratification is strong: the top decile of predicted risk contains about two-thirds of observed low achievers ( $\approx 66.7\%$  in mathematics;  $\approx 68.1\%$  in science) versus  $\sim 20\%$  in the remaining students, implying enrichment ratios above 3. Importance and standardized mean-difference profiles consistently point to bullying exposure, student-teacher ratio, teacher qualification, ICT learning resources, school belonging, and early childhood education exposure as key correlates. These results support ML-based screening tools to guide targeted interventions addressing both learning support and school-climate/resource constraints.

© ICRAMCS 2026 Proceedings ISSN: 2605-7700

### References

- [1] Pisa, O. E. C. D., results (volume I): the state of learning and equity in education, PISA, 2023, Paris.
- [2] Breiman, L., Random forests, *Machine learning*, 45(1), 2001, 5-32.
- [3] Chen, T. and Guestrin, C., Xgboost: A scalable tree boosting system, *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, 2016, 785-794.
- [4] Cortes, C. and Vapnik, V., Support-vector networks, *Machine learning*, 20(3), 1995, 273-297.
- [5] Baker, R. S. and Yacef, K., The state of educational data mining in 2009: A review and future visions, *Journal of educational data mining*, 1(1), 2009, 3-17.