Artificial intelligence predictive modeling for educational indicators using data profiling techniques

Soukaina Nai[1], Bahaa Eddine Elbaghazaoui[1,2], Amal Rifai[3], Abdelalim Sadiq[1]
[1]Laboratory of Computer Sciences, Faculty of Sciences, Ibn Tofail University, Kenitra, Morocco
[2]National School of Applied Sciences, Sultan Moulay Slimane University, Beni Mellal, Morocco
[3]Laboratory of Research in Engineering of Computing Environment for Human Learning,
Regional Center for the Professions of Education and Training, Rabat, Morocco

Article Info

Article history:

Received Oct 24, 2024
Revised Jun 16, 2025
Accepted Jul 10, 2025

Keywords:

Academic support
Artificial intelligence
Data profiling
Education
Machine learning
Prediction

ABSTRACT

In Morocco, the escalating challenges in the education sector underscore the necessity for precise predictions and informed decision-making. Effective management of the education system depends on robust statistical data, which is crucial for guiding decisions, refining policies, and improving both the quality and accessibility of education. Reliable indicators are vital for ensuring efficiency, equity, and accuracy in educational planning and decision- making. Without dependable data, implementing effective policies, addressing the needs appropriately, and achieving positive outcomes becomes difficult. This paper aims to identify the optimal machine learning model for analyzing educational indicators by comparing a range of advanced models across a comprehensive set of metrics. The objective is to determine the most effective model for profiling relevant information and addressing predictive challenges with high accuracy.

Corresponding Author:

Soukaina Nai
Laboratory of Computer Sciences, Faculty of Sciences, Ibn Tofail University
Kenitra 14000, Morocco
Email: soukaina.nai@gmail.com

1. INTRODUCTION

In Morocco, the education sector faces numerous challenges that require precise, data-driven intervention [1]. The rapid rise of issues such as educational access, quality disparities, and resource allocation has heightened the need for accurate forecasting and predictive analysis to inform policy decisions [2]. Effective management of education systems hinges on the availability of reliable statistical data [3], which not only informs policymakers but also plays a pivotal role in crafting strategies to enhance the overall quality and equity of education [4]. However, achieving these goals depends on the robustness and reliability of educational indicators [5].

Educational indicators, such as academic support [6], dropout [7], enrollment rates [8], literacy levels [9], government expenditure [10], and student performance metrics [11], are fundamental in assessing the efficiency and equity of educational systems [12]. The absence of dependable data can severely hinder the ability to implement effective policies and make informed decisions [13], leading to suboptimal outcomes and missed opportunities for improvement [14]. Thus, there is a pressing need for advanced analytical tools capable of profiling [15], and predicting key trends within the educational landscape [16].

This paper explores the application of machine learning techniques to educational data analysis [17], with the primary aim of identifying the optimal predictive model for analyzing and forecasting educational statistics in Morocco. By comparing a variety of advanced machine learning algorithms across a comprehensive set of educational indicators [18]–[20], this study seeks to determine which model offers the highest accuracy. It also aims to assess the reliability in predicting trends and guiding decision-making in the education sector.

2. RELATED WORKS

The use of predictive modeling and machine learning techniques in the education sector has garnered increasing attention in recent years [21]. Researchers have explored various methodologies to analyze educational data [22], develop predictive insights [23], and aid decision-makers in addressing critical challenges, such as resource allocation [24], student performance [25], and educational access [26]. This section reviews key studies and advances in the field, highlighting relevant machine learning models and their application to educational indicators [27].

One prominent area of research focuses on the predictive modeling of student performance [28]. Studies such as those by [29], [30] have explored machine learning algorithms, including decision trees, random forests, and support vector machines, to predict academic outcomes based on factors such as student demographics, prior academic history, and socio-economic status. These models have demonstrated considerable accuracy in forecasting student success and identifying at-risk students, providing valuable insights for educators and administrators.

Another significant line of research involves using machine learning to predict school dropout rates. For example, Alam *et al.* [31] utilized logistic regression and neural networks to identify students at risk of dropping out of school. The study found that machine learning models could effectively predict dropout rates by analyzing historical enrollment data, attendance records, and socio-economic factors. Similarly, Kuleto *et al.* [32] expanded this work by incorporating additional predictors such as student engagement metrics and teacher assessments, further improving the accuracy of dropout predictions.

Research has also investigated the optimization of resource allocation and expenditure in the education sector. The research in [33], [34] have employed machine learning techniques to analyze educational expenditures, staffing patterns, and school infrastructure needs. These works demonstrate that predictive modeling can support more efficient and equitable distribution of educational resources, ensuring that financial and human capital is allocated where it is most needed [35].

In the context of educational policy planning, machine learning models have been applied to forecast trends such as enrollment rates, literacy levels, and the demand for teachers. For instance, group of researcher developed a time-series forecasting model using recurrent neural networks (RNN) to predict future student enrollment across different educational levels [36]. Their model provided actionable insights into expected changes in enrollment trends, allowing policymakers to adjust strategies and resources accordingly. Similarly, Kangiwa *et al.* [37] explored how data-driven decision-making could enhance the effectiveness of educational interventions, suggesting that predictive models could serve as powerful tools for shaping national educational policies.

Despite the growing body of research in this field, there remain challenges in the accurate and reliable prediction of educational indicators [38], particularly when dealing with large and diverse datasets. Integrating data profiling with machine learning has the potential to improve the accuracy and usability of predictive models in the education sector [39]. As the availability of educational data continues to expand, researchers are increasingly focusing on refining these models to address complex issues such as inequality [40], dropout rates [41], and the efficient allocation of resources [42].

This paper builds on this body of work by employing a comparative approach to machine learning models [43], with the goal of identifying the most suitable techniques for profiling and predicting educational indicators in Morocco [44]. By analyzing a wide range of indicators, this study aims to contribute to the ongoing development of AI-driven solutions. These solutions can address the critical challenges faced by the education sector [45], particularly in the context of developing countries [46].

3. PROBLEMATIC

In Morocco, the education sector faces dynamic challenges that demand precise and forward-looking data to guide policy and improve outcomes [47]. Reliable forecasting of educational indicators is essential for effective decision-making [48], particularly in addressing issues related to access, quality, and resource allocation [49]. However, the selection of an appropriate machine learning model for predicting future educational values presents a complex problem [50]. The core issue lies in determining which machine learning model can deliver the most accurate and dependable forecasts based on historical educational data. This study aims to address this problem by analyzing a range of educational indicators literacy levels, and academic performance metrics, dropout, academic support and discussing their meanings and implications, and applying various advanced machine learning models to predict future trends. The research will involve comparing models on their predictive accuracy and suitability for different types of educational data. Additionally, the study aims to identify challenges related to data quality and model performance, offering insights into the practical application of machine learning in educational planning. By pinpointing the most effective machine learning model for this purpose, the study will contribute to more informed decision-making and strategic planning in Morocco's education sector, ultimately supporting efforts to enhance educational quality and equity.

4. METHODOLOGY

4.1. Analysis of educational data and indicators

This study aims to analyze historical educational data and apply various advanced machine learning models to predict future educational indicators in Morocco, including literacy levels, academic performance metrics, dropout rates, and participation in support programs. The objective is to evaluate and compare the predictive accuracy and reliability of different machine learning models in various educational contexts. Additionally, the research will explore challenges related to data quality and model performance, offering actionable insights for enhancing decision-making and strategic planning within Morocco's education sector. Ultimately, this study seeks to improve educational quality and equity by supporting the development of informed, data-driven policies.

4.2. Dataset

In recent years, the World Bank has made significant strides in enhancing transparency and accessibility of educational data globally. As part of this initiative, the World Bank has published a range of statistics specifically focused on education in Morocco [18]. These statistics, available through their data repository at World Bank education indicators for Morocco, providing a valuable resource for stakeholders interested in understanding the dynamics of Morocco's education system.

4.3. Indicators

The dataset includes detailed indicators on various aspects of education (253 indicators), such as enrollment rates, educational attainment, and the availability of educational resources. These indicators are crucial for evaluating the performance of the education system, identifying trends, and making informed decisions to enhance educational policies and practices. Among the 253 indicators, we focused on those related to academic support. To achieve this, we extracted the relevant data from the main file. Table 1 presents the results of this extraction.

Indicators for assessing educational outcomes include several key areas. Firstly, no education/dropout indicators measure the percentage of individuals who have either received no formal education or are not currently enrolled in school, reflecting the prevalence of educational disengagement. Secondly, teacher training indicators assess the qualifications and professional development of educators, which directly impact the quality of academic support provided to students. Education expenditure indicators quantify the financial resources allocated to education, influencing the availability of essential resources and support within schools. Additionally, benefiting from support programs tracks the proportion of students receiving supplementary academic assistance, which can enhance learning outcomes. Finally, international assessments evaluate student performance on global scales, providing insights into the effectiveness of academic support programs in improving educational achievements.

Table 1. List of the indicators related to school dropout and academic support

Indicator type	Indicator code (CSV file)	Description
No education (School dropout)	BAR.NOED.1519.FE.ZS, BAR.NOED.1519.ZS, BAR.NOED.15UP.FE.ZS.	Percentage of individuals with no education in various age groups, implying school dropout.
Out of school (Dropout)	SE.LPV.PRIM.OOS.FE, SE.LPV.PRIM.OOS.MA	Percentage of primary students who are out of school, reflecting potential dropouts.
Unenrollment (Dropout)	SE.PRM.UNER.ZS	Percentage of students unenrolled at the primary level, implying a relationship to school dropout.
Teacher training (Academic support)	SE.PRE.TCAQ.ZS, SE.PRM.TCAQ.ZS, SE.SEC.TCAQ.LO.ZS, SE.SEC.TCAQ.UP.ZS	Percentage of teachers who are trained/qualified at pre-primary, primary, and secondary levels.
Education expenditure (Support)	SE.XPD.CUR.TOTL.ZS, SE.XPD.PRIM.ZS, SE.XPD.SECO.ZS, SE.XPD.TERT.ZS	Public expenditure on education as a percentage of total government expenditure, by education level.
Benefiting from support programs	SE.LPV.PRIM.BMP.FE, SE.LPV.PRIM.BMP.MA	Percentage of primary students benefiting from academic support programs.
International assessments (Support)	LO.PIRLS.REA.INT, LO.TIMSS.MAT8.INT	Performance in international assessments (trends in international mathematics and science study (TIMSS), progress in international reading literacy study (PIRLS)), indirectly reflecting academic support.

4.4. Importance

Educational indicators are vital tools for measuring progress, identifying trends, and assessing the effectiveness of policies and interventions in the education sector. They provide essential data on key areas such as literacy rates, dropout rates, and academic performance, allowing policymakers and educators to understand the current state of the system. These indicators help track improvements or challenges over time, enabling targeted interventions where needed. For example, monitoring dropout rates can identify vulnerable regions or populations that require additional support, while literacy rates help evaluate the impact of national and local strategies.

Additionally, indicators enable data-driven decision-making, ensuring that resources are allocated efficiently and equitably. They also promote accountability and continuous improvement in educational outcomes. Ultimately, educational indicators are crucial for guiding strategic planning and achieving goals related to educational quality, equity, and access. Analyzing these statistics can offer critical insights into areas where improvements are needed, help monitor the effectiveness of current educational policies, and guide future planning and interventions. Leveraging machine learning techniques to analyze this data can further enhance the accuracy and depth of these insights, leading to more informed decision-making and targeted policy adjustments.

5. IMPLEMENTATION

5.1. Environment

In this study, we will use Python as the primary programming environment due to its powerful libraries and tools for data analysis and machine learning. Specifically, we will leverage scikit-learn (sklearn), a versatile and user-friendly machine learning library within Python, which offers a wide range of algorithms for predictive modeling. This environment allows for seamless data preprocessing, model training, and evaluation, ensuring that we can efficiently compare various machine learning models in terms of their accuracy and performance. Python’s rich ecosystem, combined with sklearn’s robustness, provides the ideal platform for analyzing educational indicators and building reliable predictive models to address the challenges in Morocco’s education sector.

5.2. Data visualization

We aim to present the extracted data discussed in section 4.3. A particular focus is placed on academic support information in Morocco. This includes the analysis of data on support program participation as shown in Figure 1, and international assessments as shown in Figure 2.

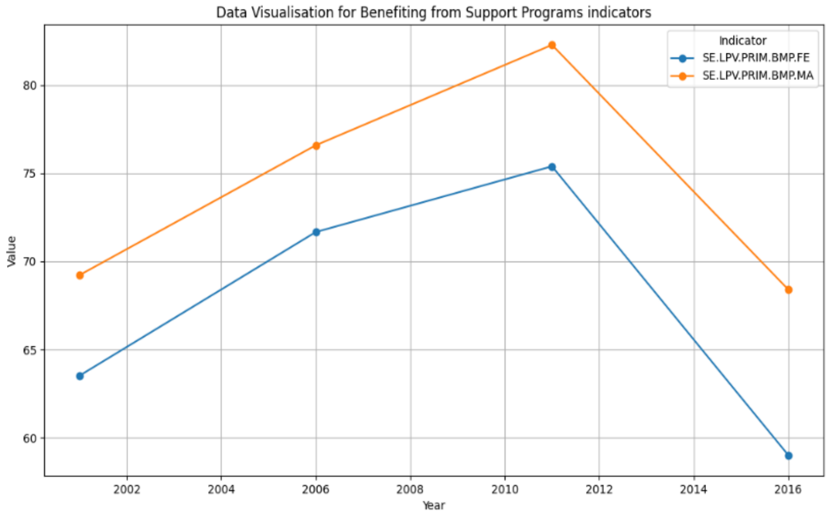

Figure 1. Data visualization for benefiting from support programs

The graph in Figure 1 illustrates the "Benefiting from support programs" indicators from 2002 to 2016 with two lines representing distinct metrics. The blue line (SE.LPV.PRIM.BMP.FE) begins at approximately 63 in 2002, rises to around 73 in 2010, and then declines sharply to about 58 by 2016. This indicates an initial improvement followed by a significant drop. Conversely, the orange line

(SE.LPV.PRIM.BMP.MA) starts at around 70 in 2002, steadily climbs to a peak of about 80 in 2012, and then falls to roughly 70 in 2016. This shows a continuous increase until 2012, followed by a decrease. The peak between 2010 and 2012 corresponds with the implementation of project E1.P5 under the 2009 to 2012 emergency program, which aimed to combat school repetition and dropout rates through personalized monitoring and support [6] The graph effectively contrasts the trends of both indicators, highlighting their rise and subsequent decline over time.

The graph in Figure 2 tracks two international assessment indicators over time: progress in international reading literacy study (LO.PIRLS.REA.INT) and trends in international mathematics and science study for 8th graders in mathematics (LO.TIMSS.MAT8.INT). The blue line for LO.PIRLS.REA.INT shows a significant decline from 2001, reaching its lowest around 2011, then rising sharply until 2016 before slightly declining by 2019. In contrast, the orange line for LO.TIMSS.MAT8.INT displays a steady increase from 2001 to 2006, a minor decline until 2011, and then a stabilization with a slight decrease by 2016. The graph reflects student performance trends in these assessments, revealing the lowest performance between 2010 and 2012. This drop is linked to the implementation of an alternative academic support program during that period.

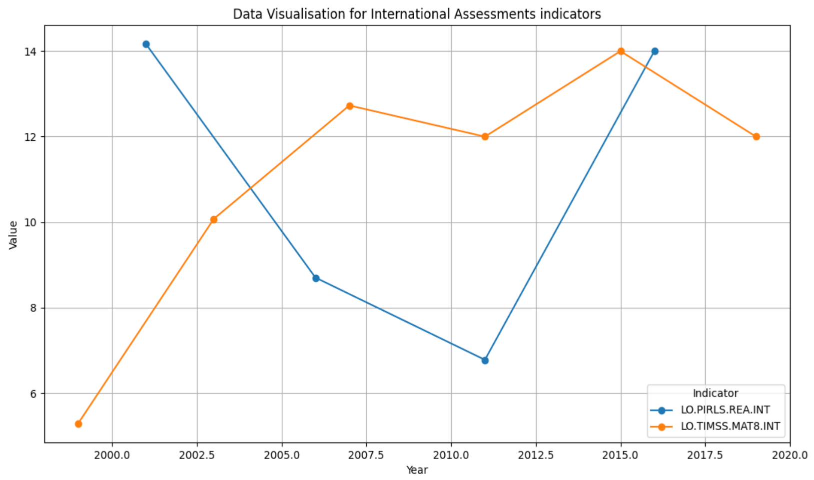

Figure 2. Data visualization for international assessments

5.3. Machine learning regression algorithms

To make prediction about the number of students who need benefiting from the support programs, machine learning algorithms are essential because they can process large and complex datasets to identify patterns and trends that traditional methods might miss. In fields like education, machine learning helps predict student performance, dropout rates, or the success of interventions, enabling more effective decision-making. Machine learning models continuously learn from data, adapting to changes over time and improving accuracy. Their ability to handle nonlinear relationships and diverse data makes them crucial for forecasting, allowing for better resource allocation and proactive strategies, ultimately driving improved outcomes and efficiency in decision-making processes. As consequence, we need to identify the optimal machine learning model for analyzing educational indicators. In this context we cite the list of algorithms that we will use in this paper, as presented in Table 2.

Table 2. Machine learning regression algorithms

Title	Description
Linear regression	A simple regression model that assumes a linear relationship between the input features and the target variable.
Polynomial regression	A variation of linear regression that models a nonlinear relationship by including polynomial features.
Random forest	An ensemble method that constructs multiple decision trees and merges them together to get more accurate and stable predictions.
Support vector regression	A regression algorithm that tries to fit the best hyperplane within a margin of tolerance while minimizing the error outside this margin.
Decision tree	A tree-based regression algorithm that splits the data into branches to predict the target variable by following simple decision rules.
Multi-layer perceptron (MLP) regressor	A type of neural network with multiple layers and nodes that approximates complex functions for regression tasks through backpropagation.

6. RESULTS AND DISCUSSION
6.1. Data visualization and prediction
6.1.1. Benefiting from support programs indicators

Using the indicators on percentage of primary students benefiting from academic support programs (Benefiting from support programs), we are going to apply the machine learning algorithms: linear regression as shown in Figure 3 and random forest as shown in Figure 4. This is done in order to make a comparison between them. The goal is to determine the best algorithm to make the prediction in this framework.

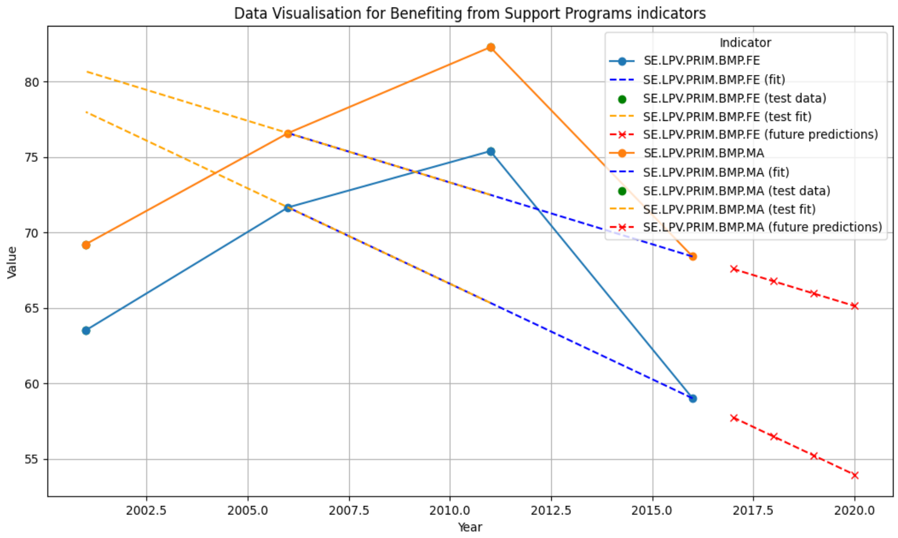

Figure 3. Prediction using linear regression

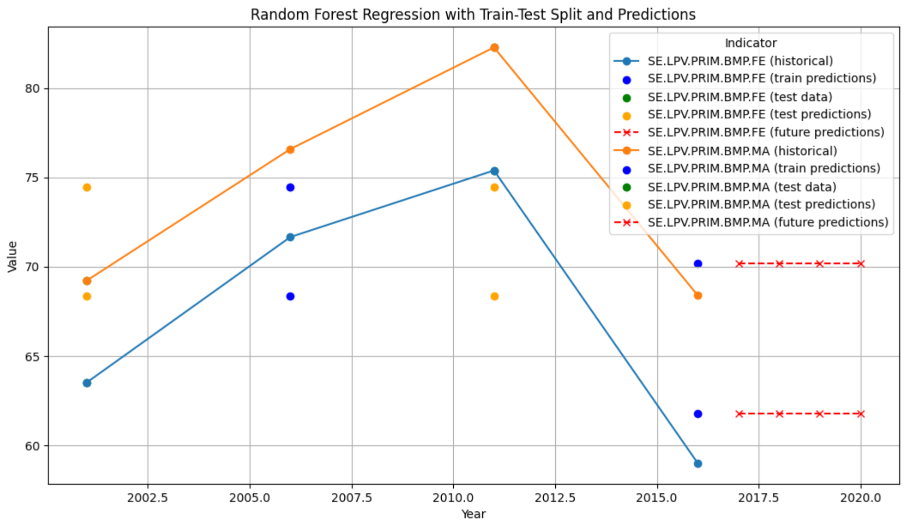

Figure 4. Prediction using random forest regression

The graph in Figure 3 is the result of a linear regression analysis, visualizes trends for two indicators: female primary school completion rate (SE.LPV.PRIM.BMP.FE) and male primary school completion rate (SE.LPV.PRIM.BMP.MA) related to support programs. Historical data (solid lines) is shown alongside model fits (dashed lines), and test data points are included. Predictions for future values extend from 2017 to 2020, represented by red crosses. The graph indicates declining trends for both male and female completion rates after 2015, as shown by the downward-sloping dashed and solid lines, suggesting a potential drop in educational outcomes without further intervention.

The graph in Figure 4 illustrates the results of a random forest regression model with a train-test split and predictions for two indicators: female primary school completion rate (SE.LPV.PRIM.BMP.FE) and male primary school completion rate (SE.LPV.PRIM.BMP.MA). Historical data is plotted from around 2002 to 2015, with train and test predictions overlayed. Future predictions for both indicators extend beyond 2017 to 2020. The orange and blue lines represent the historical and predicted trends for female and male completion rates, respectively. The red crosses represent future predictions, indicating a steady or decreasing trend after 2017 for both indicators.

6.1.2. International assessments (support) indicators

Using the performance indicators from international assessments such as TIMSS and PIRLS, which indirectly reflect academic support, we applied machine learning algorithms. Specifically, we used linear regression as shown in Figure 5 and random forest as shown in Figure 6. This was done to compare the effectiveness of both algorithms in predicting outcomes within this framework and to identify the most suitable model.

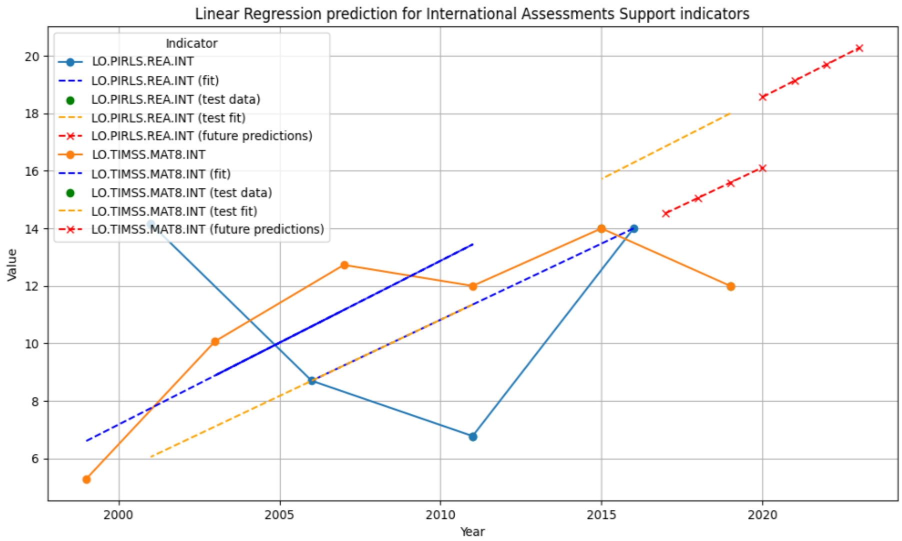

Figure 5. Prediction using linear regression

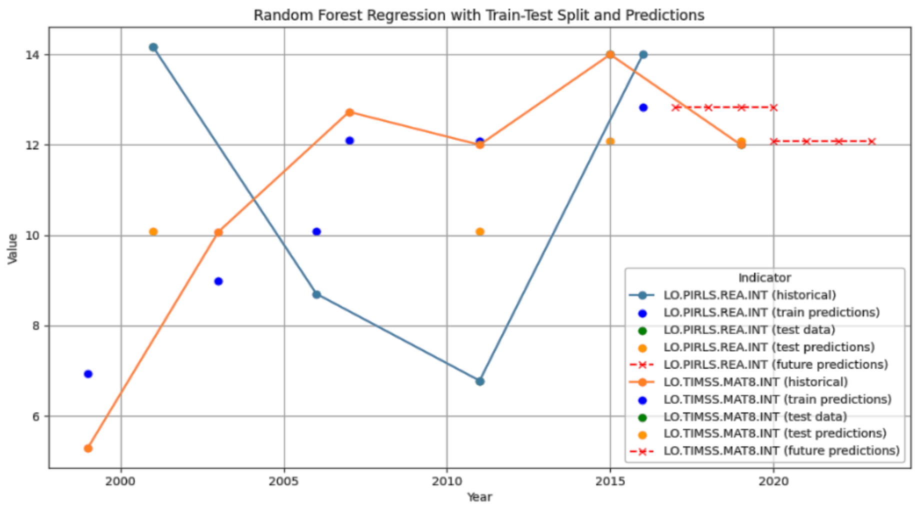

Figure 6. Prediction using random forest regression

The graph in Figure 5 derived from a linear regression analysis, forecasts trends for two international assessment support indicators: LO.PIRLS.REA.INT and LO.TIMSS.MAT8.INT. The historical data (solid lines) and test data points reveal the performance trends over time, while the dashed lines represent the model fits for these indicators. The graph spans from around 2000 to 2015 for observed data, with future predictions extending to 2020, represented by red crosses. The PIRLS scores exhibit early fluctuations, but the general trend for both indicators is upward, suggesting an overall improvement in reading and mathematics performance. The future predictions also indicate a steady rise in both assessment scores, showing potential progress in these international evaluations.

The graph in Figure 6 illustrates the results of a random forest regression model with a train-test split and predictions for two indicators: LO.PIRLS.REA.INT and LO.TIMSS.MAT8.INT. Historical data is depicted from approximately 2002 to 2015, with train and test predictions overlaid. Future predictions for both indicators extend from 2017 to 2020. The orange and blue lines correspond to the historical and predicted trends for female and male completion rates, respectively. The red crosses denote future predictions, suggesting either a stable or declining trend beyond 2017 for both indicators.

Result: through this analysis, we conclude that it is insufficient to determine the best machine learning model based on just two models. In consequence, to identify the most effective model for predicting educational indicators, we will apply various machine learning algorithms to the entire dataset obtained from the World Bank website. Then, we will compare these models based on accuracy measures to determine which one performs the best.

6.2. Machine learning comparison

To identify the best machine learning model for predicting educational indicators, we will apply multiple algorithms to the full dataset obtained. This approach will allow us to evaluate each model's performance and determine which is most effective for making accurate predictions based on the data. Our program, written in Python language and using the same machine, produces the Table 3. Analysis:

– Mean absolute error (MAE): measures the average magnitude of errors in predictions without considering their direction. Lower values indicate better performance.
  Best: random forest (3.14)
  Worst: MLP regressor (21.10)
– Mean squared error (MSE): measures the average of the squares of the errors, giving more weight to larger errors. Lower values indicate better performance.
  Best: random forest (87.62)
  Worst: support vector regression (9159.70)
– Root mean squared error (RMSE): square root of MSE, providing an error metric in the same unit as the target variable. Lower values indicate better performance.
  Best: random forest (3.91)
  Worst: MLP regressor (23.94)
– R^2 score: represents the proportion of variance in the dependent variable that is predictable from the independent variable. 0.92 indicates that the model random forest has converged successfully.

Result: random forest is the best-performing model across all provided metrics, with the lowest MAE, MSE, and RMSE. This suggests that it has the most accurate and consistent predictions among the models tested.

Table 3. Comparison of machine learning for predicting

Algorithm	MAE	MSE	RMSE	R^2
Linear regression	5.14	539.95	5.88	0.75
Polynomial regression	4.43	140.25	5.04	0.85
Random forest	3.14	87.62	3.91	0.92
Support vector regression	14.12	9159.70	17.45	0.10
Decision tree	3.43	136.60	4.12	0.84
MLP regressor	21.10	8479.91	23.94	0.15

7. CONCLUSION

In summary, our paper explores how machine learning models can improve prediction accuracy in Morocco's education sector. By comparing various algorithms across different educational indicators, we aim to identify the best model for guiding data-driven decisions. This research will help us develop better, fairer, and more accurate educational policies to address key challenges in Morocco's education system. In our future work, we will focus on incorporating larger, real-time datasets to enhance the precision of our models. Additionally, we plan to develop personalized education system suggestions to tailor recommendations based on individual needs and learning styles. This approach aims to further improve the effectiveness of educational interventions and policy decisions.

FUNDING INFORMATION

Authors state no funding involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Soukaina Nai	✓	✓	✓	✓	✓	✓		✓	✓	✓			✓	
Bahaa Eddine Elbaghazaoui	✓	✓	✓	✓		✓	✓	✓		✓	✓	✓		
Rifai Amal	✓		✓	✓			✓			✓	✓	✓	✓	
Abdelalim Sadiq	✓			✓						✓		✓	✓	

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

The data that support the findings of this study are openly available in [The World Bank] at http://doi.org/10.26599/BDMA.2019.9020007, reference number [18].

REFERENCES

[1] R. Ejjami, "Revolutionizing Moroccan education with AI: a path to customized learning," *International Journal For Multidisciplinary Research*, vol. 6, no. 3, pp. 1–32, 2024, doi: 10.36948/ijfmr.2024.v06i03.19462.

[2] A. Zine and A. Kaaouachi, "Key determinants of learning analytics adoption in Moroccan universities," *Journal of Ecohumanism*, vol. 3, no. 5, pp. 1–12, 2024, doi: 10.62754/joe.v3i5.3634.

[3] A. Qazdar, O. Hasidi, S. Qassimi, and E. H. Abdelwahed, "Newly proposed student performance indicators based on learning analytics for continuous monitoring in learning management systems," *International Journal of Online and Biomedical Engineering*, vol. 19, no. 11, pp. 19–30, 2023, doi: 10.3991/ijoe.v19i11.39471.

[4] M. Bougroum and A. Ibourk, "Access and equity in financing higher education: the case of Morocco," *Prospects*, vol. 41, no. 1, pp. 115–134, 2011, doi: 10.1007/s11125-011-9184-8.

[5] J. Brasse, M. Förster, P. Hühn, J. Klier, M. Klier, and L. Moestue, "Preparing for the future of work: a novel data-driven approach for the identification of future skills," *Journal of Business Economics*, vol. 94, no. 3, pp. 467–500, 2024, doi: 10.1007/s11573-023-01169-1.

[6] N. A. Dahri *et al.*, "Investigating AI-based academic support acceptance and its impact on students' performance in Malaysian and Pakistani higher education institutions," *Education and Information Technologies*, vol. 29, no. 14, pp. 18695–18744, 2024, doi: 10.1007/s10639-024-12599-x.

[7] A. Ibourk and S. Raoui, "Inclusive education and school dropout of special needs students in Morocco: a spatial analysis," *Review of Education*, vol. 12, no. 1, 2024, doi: 10.1002/rev3.3453.

[8] K. Saoudi, R. Chroqui, and C. Okar, "Student achievement in Moroccan educational reforms: a significant gap between aspired outcomes and current practices," *Interchange*, vol. 51, no. 2, pp. 117–136, 2020, doi: 10.1007/s10780-019-09380-2.

[9] N. Morchid, "Investigating quality education in Moroccan educational reforms from 1999 to 2019," *IOSR Journal of Research & Method in Education (IOSR-JRME)*, vol. 10, no. 1, pp. 54–61, 2020, doi: 10.9790/7388-1001015461.

[10] O. Elkhalfi, R. Chaabita, K. Zahraoui, and H. El Alaoui, "Public spending on human capital and economic growth in Morocco," *International Journal of Economics and Financial Issues*, vol. 13, no. 4, pp. 102–110, 2023, doi: 10.32479/ijefi.14374.

[11] I. Tammouch, A. Elouafi, S. Eddarouich, and R. Touahni, "Identifying low-performing regions in Moroccan education: a deep learning approach using the PISA dataset," *International Journal of Advanced and Applied Sciences*, vol. 10, no. 7, pp. 138–144, 2023, doi: 10.21833/ijaas.2023.07.015.

[12] H. Boudine, H. Sajid, M. Bentaleb, M. Tayebi, and D. El Karfa, "M-learning and autonomous education: the impact of the Moroccan digital classroom project on science subject's learning," *International Journal of Civilizations Studies & Tolerance Sciences*, vol. 1, no. 1, pp. 44–54, 2024, doi: 10.54878/1qcw9564.

[13] R. Boukbech and M. Liouaeddine, "Economic analysis of educational efficiency in Morocco: an application of the DEA method," *International Journal of Strategic Management and Economic Studies (IJSMES)*, vol. 2, no. 3, pp. 1534–1547, 2023, doi: 10.5281/zenodo.8337959.

[14] H. Berbar, S. Lotfi, and M. Talbi, "Validation and development of a Moroccan school quality evaluation system," *Education Research International*, vol. 2021, no. 1, 2021, doi: 10.1155/2021/1829259.

[15] K. Ghanouane and T. Benkaraache, “The role of analytical skills in big data-driven decision-making in african firms: evidence from Morocco,” *Journal of Economics and Management*, vol. 20, pp. 75–97, 2024.
[16] M. Liouaeddine, M. Elatrachi, and E. M. Karam, “The analysis of the efficiency of primary schools in Morocco: modelling using TIMSS database (2011),” *Journal of North African Studies*, vol. 23, no. 4, pp. 624–647, 2018, doi: 10.1080/13629387.2017.1422978.
[17] Y. Mourdi, M. Sadgal, W. B. Fathi, and H. El Kabtane, “A machine learning based approach to enhance MOOC users’ classification,” *Turkish Online Journal of Distance Education-TOJDE*, vol. 21, no. 2, pp. 47–68, 2020.
[18] The World Bank, “Morocco - education,” *OCHA Services*. 2024, doi: 10.26599/BDMA.2019.9020007.
[19] H. Barakat, “Public-sector education in Morocco: the perspective of the higher council for education, training and scientific research (CSEFRS),” *Revue Marocaine d’évaluation et de la Recherche Educative (RMERE)*, pp. 77–96, 2021, doi: 10.48423/IMIST.PRSM/rmere-v0i0.32908.
[20] Z. Boussouf, H. Amrani, M. Z. Khal, and F. Daidai, “Artificial intelligence in education: a systematic literature review,” *Data and Metadata*, vol. 3, 2024, doi: 10.56294/dm2024288.
[21] A. Ibourk, K. Hnini, and I. Ouaadi, “Analysis of the pedagogical effectiveness of teacher qualification cycle in Morocco: a machine learning model approach,” in *International Conference on Advanced Intelligent Systems for Sustainable Development*, 2023, pp. 344–353, doi: 10.1007/978-3-031-26384-2_30.
[22] A. Ouajdouni, K. Chafik, and O. Boubker, “Measuring e-learning systems success: data from students of higher education institutions in Morocco,” *Data in Brief*, vol. 35, 2021, doi: 10.1016/j.dib.2021.106807.
[23] A. Sadqui, M. Ertel, H. Sadiki, and S. Amali, “Evaluating machine learning models for predicting graduation timelines in Moroccan universities,” *International Journal of Advanced Computer Science and Applications*, vol. 14, no. 7, pp. 304–310, 2023, doi: 10.14569/IJACSA.2023.0140734.
[24] K. Hamdani and S. Koubaa, “The strategic planning of university transformation: the case of Moroccan public universities,” *Projectics / Proyéctica / Projectique*, vol. 28, no. 1, pp. 51–68, 2021, doi: 10.3917/proj.028.0051.
[25] A. Chatri, O. Chahbi, and M. Snihji, “The multilevel analysis of students’ achievement: evidence from Morocco,” *African Development Review*, vol. 33, no. 1, pp. 117–129, 2021, doi: 10.1111/1467-8268.12497.
[26] R. El Hatimi, C. F. Choukhan, and M. Esghir, “An improved k-means clustering algorithm towards an efficient educational and economical data modeling,” *International Journal of Advanced Computer Science and Applications*, vol. 15, no. 1, pp. 1104–1114, 2024, doi: 10.14569/IJACSA.2024.01501109.
[27] S. Mittal, S. Mahendra, V. Sanap, and P. Churi, “How can machine learning be used in stress management: a systematic literature review of applications in workplaces and education,” *International Journal of Information Management Data Insights*, vol. 2, no. 2, 2022, doi: 10.1016/j.jjimei.2022.100110.
[28] Z. Khoudi, M. Nachaoui, and S. Lyaqini, “Identifying the contextual factors related to the reading performance of Moroccan fourth-grade students from a machine learning-based approach,” *Education and Information Technologies*, vol. 29, no. 3, pp. 3047–3073, 2024, doi: 10.1007/s10639-023-11881-8.
[29] G. W. Dekker, M. Pechenizkiy, and J. M. Vleeshouwers, “Predicting students drop out: a case study,” in *2nd International Conference on Educational Data Mining*, 2009, pp. 41–50.
[30] A. J. Bowers, R. Sprott, and S. A. Taff, “Do we know who will drop out ?: a review of the predictors of dropping out of high school: precision, sensitivity, and specificity,” *The High School Journal*, vol. 96, no. 2, pp. 77–100, 2013, doi: 10.1353/hsj.2013.0000.
[31] T. M. Alam, M. Mushtaq, K. Shaukat, I. A. Hameed, M. U. Sarwar, and S. Luo, “A novel method for performance measurement of public educational institutions using machine learning models,” *Applied Sciences*, vol. 11, no. 19, 2021, doi: 10.3390/app11199296.
[32] V. Kuleto *et al.*, “Exploring opportunities and challenges of artificial intelligence and machine learning in higher education institutions,” *Sustainability*, vol. 13, no. 18, 2021, doi: 10.3390/su131810424.
[33] M. P. Ilić, D. Păun, N. P. Šević, A. Hadžić, and A. Jianu, “Needs and performance analysis for changes in higher education and implementation of artificial intelligence, machine learning, and extended reality,” *Education Sciences*, vol. 11, no. 10, 2021, doi: 10.3390/educsci11100568.
[34] K. L. M. Ang, F. L. Ge, and K. P. Seng, “Big educational data & analytics: Survey, architecture and challenges,” *IEEE Access*, vol. 8, pp. 116392–116414, 2020, doi: 10.1109/ACCESS.2020.2994561.
[35] O. B. Joseph and N. C. Uzondu, “Integrating AI and machine learning in STEM education: challenges and opportunities,” *Computer Science & IT Research Journal*, vol. 5, no. 8, pp. 1732–1750, 2024, doi: 10.51594/csitrj.v5i8.1379.
[36] A. J. Ganimian and R. J. Murnane, “Improving education in developing countries: lessons from rigorous impact evaluations,” *Review of Educational Research*, vol. 86, no. 3, pp. 719–755, 2016, doi: 10.3102/0034654315627499.
[37] B. I. Kangiwa, O. E. Oludare, H. S. Nassarawa, N. S. Abubakar, E. L. Efeoma, and H. A. Enefola, “Leveraging artificial intelligence for enhancing entrepreneurship and creativity in STEM education,” *Journal of Educational Research and Practice*, vol. 4, no. 8, pp. 149–162, 2024.
[38] M. Skittou, M. Merrouchi, and T. Gadi, “A model of an integrated educational management information system to support educational planning and decision making: a Moroccan case,” in *WITS 2020: Proceedings of the 6th International Conference on Wireless Technologies, Embedded, and Intelligent Systems*, 2022, pp. 167–177, doi: 10.1007/978-981-33-6893-4_16.
[39] H. Almaghrabi, B. Soh, and A. Li, “Using ML to predict user satisfaction with ICT technology for educational institution administration. information,” *Information*, vol. 15, no. 4, 2024, doi: 10.3390/info15040218.
[40] A. Yassine and F. Bakass, “Youth’s poverty and inequality of opportunities: empirical evidence from Morocco,” *Social Sciences*, vol. 12, no. 1, 2023, doi: 10.3390/socsci12010028.
[41] M. Berriane, H. De Haas, and K. Natter, “Social transformations and migrations in Morocco,” *International Migration Institute Working Paper Series*, vol. 171, pp. 1–47, 2021.
[42] W. A. Owings and L. S. Kaplan, *Equity audits and school resource allocation: applying critical resource theory to increase equal opportunity in schools*. New York, United States: Routledge, 2024, doi: 10.4324/9781003493907.
[43] Y. Chen and L. Zhai, “A comparative study on student performance prediction using machine learning,” *Education and Information Technologies*, vol. 28, no. 9, pp. 12039–12057, 2023, doi: 10.1007/s10639-023-11672-1.
[44] M. Akdim *et al.*, “Scale’s impact in education system’s performance: Cases in Draa-Tafilalet, Morocco,” *International Journal of Evaluation and Research in Education*, vol. 12, no. 4, pp. 2374–2386, 2023, doi: 10.11591/ijere.v12i4.25100.
[45] UNESCO, UNICEF, and World Bank, *The state of the global education crisis: a path to recovery*. Washington, United States: The World Bank, UNESCO and UNICEF, 2021, doi: 10.54675/JLUG7649.

[46] M. Ennebch, A. Bassibas, Y. H. Amar, M. Ragbi, and M. A. Berdaa, "Spatial analysis of youth tvet education, insertion, migration and regional economic growth in Morocco using spatial econometrics," in *The Migration Conference 2023 Selected Papers*, 2023, pp. 13–63.

[47] J. Amaghouss and M. Zouine, "A critical analysis of the governance of the moroccan education system in the era of online education," in *Socioeconomic Inclusion During an Era of Online Education*, New York, United States: IGI Global, 2022, pp. 156–176, doi: 10.4018/978-1-6684-4364-4.ch008.

[48] Y. Safsouf, K. Mansouri, and F. Poirier, "Smart learning environment, measure online student satisfaction: a case study in the context of higher education in Morocco," in *2020 International Conference on Electrical and Information Technologies (ICEIT)*, 2020, pp. 1–5, doi: 10.1109/ICEIT48248.2020.9113189.

[49] A. Fahim, Q. Tan, B. Naz, Q. U. Ain, and S. U. Bazai, "Sustainable higher education reform quality assessment using SWOT analysis with integration of AHP and entropy models: a case study of Morocco," *Sustainability*, vol. 13, no. 8, 2021, doi: 10.3390/su13084312.

[50] Y. T. Badal and R. K. Sungkur, "Predictive modelling and analytics of students' grades using machine learning algorithms," *Education and Information Technologies*, vol. 28, no. 3, pp. 3027–3057, 2023, doi: 10.1007/s10639-022-11299-8.

BIOGRAPHIES OF AUTHORS

Soukaina Nai is an IT Engineer and a research student at Department of Computer Science, Faculty of Sciences, Ibn Tofail University (Kenitra Morocco). Her research interests include machine learning, integrated management systems, decision-making systems, evolutionary algorithms and education systems. She has published numerous articles in several journals. She can be contacted at email: soukaina.nai@gmail.com.

Bahaa Eddine Elbaghazaoui is a Research Professor at ENSA Beni Mellal and a passionate full-stack developer with a specialization in AI, DevSecOps, and software architecture. His work blends academia with industry, focusing on solving complex challenges through innovative technologies. With over 6 years of experience, he has worked on various projects involving AI, DevSecOps, software architecture, and digital transformation. He continues to contribute to academia and the industry, teaching, and mentoring future engineers while actively participating in research. He can be contacted at email: elbaghazaoui.bahaa@gmail.com and b.elbaghazaoui@usms.ac.ma.

Amal Rifai is a Professor of Computer Science at the Centre régional des métiers de l'éducation et de la formation (RCPET), Rabat, Morocco. She is also head of the research team in Computer Environment Engineering for Human Learning. Her research interests include artificial intelligence, machine learning, digital systems and technology, evolutionary algorithms, and educational systems. She has published extensively in numerous journals. She can be contacted at email: rifaiamal10@gmail.com.

Abdelalim Sadiq is a Professor of Higher Education of Computer Science, at Ibn Tofail University in Kinetra, and head of the Computer Science Research Laboratory, his research interests are artificial intelligence, machine learning, digital systems and technology, evolutionary algorithms. He has published numerous articles in a wide range of journals. He can be contacted at email: a.sadiq@uit.ac.ma.