Big data and ensemble learning for effective student orientation in Morocco

Morad Badrani[1], Adil Marouan[1], Nabil Kannouf[2], Abdelaziz Chetouani[1]
[1]Laboratory of LaMAO, ORAS Team, ENCG, Mohamed First University, Oujda, Morocco
[2]Laboratory of LSA, SOVAI Team, ENSA, Abdelmalek Essaadi University, Tetouan, Morocco

Article Info

Article history:

Received Jun 16, 2024
Revised Aug 20, 2024
Accepted Aug 31, 2024

Keywords:

Big data
Machine learning
Neural network
Random forest
Student's orientation

ABSTRACT

Guiding high school students toward suitable educational paths is a complex challenge, particularly influenced by academic performance. In Morocco, first-year high school students in the scientific branch face a crucial decision when selecting between science mathematics (SM), physics (SF), and Science of Life and Earth (SVT) paths. This decision is critical as it can significantly impact their future academic and professional success. To address the issue of suboptimal student orientation, this study proposes an automated, personalized approach leveraging big data technology. By employing ensemble learning techniques, including random forest and neural network models, we developed a classification system to predict students' optimal paths based on their academic performance. Our models achieved exceptional performance, with precision, accuracy, recall, and F-measure scores of approximately 98.59%, 98.60%, 98.60%, and 98.58%, respectively. This research demonstrates the potential of our approach to enhance educational support and decision-making, ultimately improving student outcomes and reducing dropout rates caused by wrong orientation.

Corresponding Author:

Morad Badrani
Laboratory of LaMAO, ORAS Team, ENCG, Mohamed First University
Oujda, Morocco
Email: m.badrani@ump.ac.ma

1. INTRODUCTION

The critical issue of ineffective student orientation in Morocco is a significant barrier to educational success. The alarming statistic of thousands of students dropping out is a critical issue underscores the gravity of this problem. Incorrect orientation not only contributes to this high dropout rate but also leads to a mismatch between educational qualifications and job market demands, hindering economic growth. Current practices rely heavily on traditional approaches, often influenced by parental expectations and neglecting individual student abilities and preferences. This misalignment between student potential and chosen paths results in academic struggles and a mismatch between education and career aspirations.

To address these challenges, this research proposes an automated orientation system that leverages big data and ensemble learning to provide predictive guidance for secondary track students. By analyzing academic performance and other relevant factors, the system aims to assist students in making informed decisions aligned with their abilities and interests. By combining two powerful machine learning techniques based on big data, this research aims to address the challenges of student orientation in Morocco and improve educational outcomes.

Limited research exists on predicting student orientation choices, with some studies focusing on traditional machine learning algorithms. For instance, Ouatik *et al.* [1] found Naïve-Bayes most suitable for student guidance using MapReduce with Hadoop. El Mrabet and Moussa [2] found decision tree (DT) outperformed other algorithms for student orientation. Another study by Badrani *et al.* [3] favored random forest for the same purpose. El Mrabet and Moussa [4] introduced the ETC guidance system utilizing internet of things (IoT) to reduce orientation failure rates. Kemper *et al*. [5] concluded DT facilitated student orientation best among k-nearest neighbor (KNN), support vector machine (SVM), and logistic regression. Ouatik *et al.* [6] achieved an 87.32% recognition rate predicting student success and failure using SVM. Mimis *et al.* [7] proposed an intelligent recommendation system for CPGE-Morocco students. Badrani *et al.* [8] conducted a comparison of DTs, Naïve-Bayes, and KNN algorithms, determining the DTs as most suitable for middle school student guidance. Zhou *et al.* [9] developed a chatbot-based vocational guidance system for Ben M'Sik Faculty of Sciences students. Nabil *et al.* [10] employed deep learning to forecast student success based on course data. Nabizadeh *et al.* [11] explored ML techniques to forecast final students' performances in gamified learning environments. Sánchez-Pozo *et al.* [12] compared supervised machine learning models to predict high school student academic success. Ujwal and Malik [13] determined critical student attributes for academic performance prediction to inform student orientation strategies. Rebai *et al.* [14] employed a visual machine learning approach to forecast secondary school performance, applicable to orientation systems. In another work, Hazra *et al.* [15] employed hybrid algorithms, including Naïve-Bayes, SVM to detect breast cancer.

The paper is organized into three sections. The second part details the methodology, followed by results and discussion in the third section. The paper concludes with a final section.

2. METHOD

This research investigates the repercussions of incorrect student orientation on school dropout rates. We employ a hybrid methodology that leverages the strengths of both neural networks and random forest algorithms in our analysis. This combined strategy aims to offer a thorough and robust comprehension of the elements impacting student orientation and their connection to dropout rates within this specific setting.

2.1. Big data analytics

The concept of big data encompasses extensive datasets [16], including both structured and unstructured data, that can be analyzed to produce meaningful insights. Our study harnesses the power of big data to conduct a thorough examination of the complexities surrounding student orientation and its relationship with dropout rates. By integrating big data analytics into our hybrid approach, we aim to uncover nuanced patterns and insights may facilitate the creation of more effective interventions and strategies for lowering dropout rates.

2.1.1. Hadoop

Hadoop is a distributed computing platform for storing and processing large amounts of data in batches [17], [18], plays a crucial role in our research. This framework's core principle involves distributed processing across multiple nodes, significantly enhancing computing and storage capabilities to effectively manage extensive data volumes. This is particularly beneficial for our analysis considering student data across several academic years. Hadoop offers numerous advantages:

i) Rapid storage and processing: large datasets can be stored and processed efficiently, a crucial aspect for analyzing student data encompassing grades and chosen orientations over multiple years.
ii) Scalability: the system is inherently scalable; additional nodes can be seamlessly integrated as data needs grow, minimizing administrative effort. This flexibility is important as we anticipate potentially collecting even more data points in the future (e.g., extracurricular activities).
iii) Distributed computing and fault tolerance: distributed processing facilitates swift data analysis, with increased computation nodes translating to enhanced processing power. Additionally, Hadoop ensures data and application integrity by redirecting tasks to alternative nodes in the event of hardware failures, preventing disruption during distributed computation.
iv) Cost-effectiveness: as an open-source framework, Hadoop leverages standard machines for storing extensive data volumes, offering a cost-effective solution.

2.1.2. Hadoop distributed file system

Hadoop distributed file system (HDFS) serves as the underlying storage solution within our big data architecture [19]. Designed for commodity hardware, it provides robust fault tolerance capabilities. Each piece of data (student information for each year) is replicated across multiple locations within the Hadoop cluster, ensuring its availability under any circumstances and mitigating the risk of data corruption. Here's how HDFS specifically benefits our research:

- Optimized for large datasets: HDFS utilizes larger data block sizes compared to conventional file systems, allowing for efficient storage and retrieval of our student data encompassing grades and chosen orientations across multiple years [20].
- Data reliability and availability: the block replication feature ensures data reliability. Even if a node fails, student data can still be accessed from other nodes holding replicas of the same block. This is crucial for maintaining data integrity throughout our analysis.

Overall, HDFS provides a resilient and efficient storage solution for our large student dataset, facilitating reliable data access and retrieval for analysis using machine learning algorithms.

2.2. Classification algorithms

This research investigates the effectiveness of ensemble learning techniques for predicting optimal student orientation in Morocco. We propose a specific ensemble learning model and compare its performance with a selection of established classical classification algorithms. This comparative analysis allows us to identify the most suitable approach for student orientation prediction in this specific context.

2.2.1. Classical algorithms

Naïve-Bayes: is considered to be the most effective classification algorithm [21]. Its indication, by anticipating the likelihood that a record in the database is associated with a specific class, and this classification model is based on the statistical notion of Bayes' theorem.

$$P(Y|N) = \frac{P(N|Y) * P(Y)}{P(N)} \quad (1)$$

P(Y|N) is the probability of Y being true given that N is true.
$P(N|Y)$ is the opposite of the previous one. It is the probability of N being true given that Y is true.
P(Y) is the probability of Y being true.
P(N) is the probability of N being true.

The KNN: is a predictive method that involves comparing the features of a new record with those of existing records in the dataset [22]. By identifying the KNN based on similarity (typically using distance metrics), it predicts the unknown value of the new record. For classification tasks, it assigns the prevalent class within the k neighbors to the new record, while for regression tasks, it computes the average or weighted mean of the target values from the k neighbors as the predicted value.

DT: is a modeling technique depicted in a tree-like structure, as its name implies [23], [24]. Within this framework, each node represents a decision point based on specific features, and the branches signify possible outcomes or classifications. The terminal nodes, or leaves, correspond to the final predicted classes or values. DT algorithms are utilized to recursively split the dataset into subsets based on the most informative features, enabling effective classification and prediction tasks.

2.2.2. Ensemble learning

Our approach involves the implementation of ensemble learning, combining two powerful machine learning techniques. For each base classifier, such as random forest classifier and MLP classifier, we generate distinct training sets (D1 and D2) from the provided dataset D. The random forest classifier is trained using D1. In contrast, the MLP classifier is trained using D2. After training both base classifiers, their predictions are combined through a voting mechanism (refer to Figure 1). When a new data tuple is presented for classification, each base classifier offers its class prediction (orientation choice). Subsequently, the ensemble classifier aggregates these predictions via majority voting, ultimately yielding the final class prediction for the new data tuple.

- Random forest: is an ensemble learning technique that combines multiple DTs to make predictions [25]. It is a powerful method for handling both numerical and categorical features, enabling it to capture complex relationships within the data.
- Neural network: in our approach, we employ the MLPClassifier, a neural network model that utilizes deep learning techniques to comprehend intricate patterns within the data. Neural networks consist of interconnected nodes (neurons) arranged in layers and their parameters are optimized through backpropagation [26]. This allows the model to adeptly capture intricate relationships and patterns withinthe dataset.

2.3. Weka

In our work, we utilize Weka [27] for various tasks, including data preprocessing, feature selection, and implementing the base classifiers (random forest classifier and MLP classifier) before combining their predictions.Weka, also known as the Waikato environment for knowledge analysis, is a comprehensive set of Java-based tools for data manipulation and analysis. It offers a wide array of artificial intelligence algorithms, such as DTs and neural networks, for tasks like supervised or unsupervised classification. Weka provides Java classes for data loading and handling, various classification algorithms, attribute selection tools, statistical analysis features, and visualization capabilities.While Weka may face challenges with large volumes of data [28], particularly in the context of big data, it addresses this issue by implementing techniques and architectural solutions for effective data management. Additionally, Weka integrates with other tools and frameworks designed for handling large datasets, such as massive online analysis (MOA), an open-source framework containing a diverse set of learning algorithms and assessment tools. By leveraging Weka and complementary frameworks like MOA, researchers, and practitioners can navigate the complexities of big data analytics and derive meaningful insights from vast datasets.

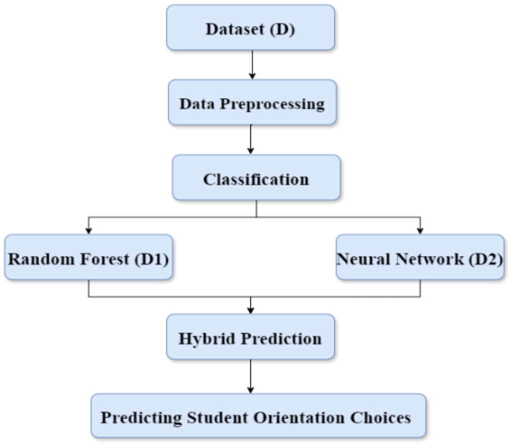

Figure 1. Workflow of combining classifiers for improved prediction

3. RESULTS AND DISSCUSSION

3.1. Data and processing

We collected student data from the Provincial Education Directorate in Nador, Morocco, using the MASSAR system. The dataset encompasses student data for several academic years, including the period from 2018 to 2022. It includes features such as student ID, grades in various subjects (Mathematics, English, Physics, History and Geography, Life and Earth Sciences, French, Physical Education, Islamic Education, Arabic), and their chosen academic orientations (Science Mathematics (SM), Physics (SF), or Science of Life and Earth (SVT)).

While a Python environment on a single machine might be suitable for basic analysis of student grades over a limited timeframe, we opted for a Hadoop cluster with 3 datanodes and 1 namenode for the following reasons:

- Scalability: Hadoop provides a scalable architecture that can effortlessly accommodate future growth in the dataset. Suppose the Provincial Education Directorate expands or additional data points are collected in the future (e.g., extracurricular activities, data for even more years). In that case, our current system can handle the increased volume without significant modifications.
- Potential for complex analyses: including data for four academic years (2018-2022) allows for more intricate analyses of student progress and how their chosen orientation (SM, SF, or SVT) relates to their academic performance. We can explore trends, identify year-on-year variations in performance within chosen orientations, and potentially uncover factors influencing student success in their chosen paths. Additionally, we foresee the possibility of exploring more complex relationships within the data in the future, such as identifying factors influencing student success within their chosen paths. Hadoop’s parallel processing capabilities can be advantageous for computationally intensive tasks like machine learning algorithms or complex statistical modeling that might be employed in these future analyses.

3.2. Results

Our approach involves the implementation of ensemble learning, combining two powerful machine learning techniques: Neural network and random forest. The classification accuracy [29], precision [30], recall [31], and F-measure [32] were employed to contrast the effectiveness of our model with that of five classification algorithms, comprising Naïve-Bayes, KNN, DT, neural network, and random forest. Our obtained results are summarized in Table 1.

Table 1. Algorithms performance metrics

Variable	Accuracy (%)	Precision (%)	Recall (%)	F-measure (%)
Naïve-Bayes	70.67	68.96	51.67	58.30
DT	86.83	82.44	84.31	83.30
KNN	87.11	85.34	81.42	83.25
Random forest	92.73	91.31	89.71	90.40
Neural network	95.91	95.90	95.91	95.89
Our model	98.60	98.59	98.60	98.58

The findings of this research demonstrate the effectiveness of ensemble learning in predicting student orientation in the Moroccan educational context. The proposed model, which combines random forest and neural network algorithms, achieved outstanding performance metrics: precision (98.59%), accuracy (98.60%), recall (98.60%), and F1-score (98.58%). This exceptional performance surpasses the capabilities of individual algorithms, highlighting the ensemble approach’s strength in making accurate predictions for student orientation

This success can be attributed to the ensemble model’s ability to leverage the complementary strengths of its constituent algorithms. Random forest excels at handling high-dimensional data and identifying complex patterns, while the neural network offers enhanced learning capabilities. By combining these strengths, the ensemble model generates more robust and generalizable predictions compared to single algorithms.

Furthermore, the utilization of a big data framework like Hadoop and HDFS played a crucial role in managing and processing the large-scale educational dataset employed in this study. The volume and complexity of this data necessitated a robust infrastructure for efficient handling. By leveraging big data technologies, we were able to extract valuable insights from the dataset, allowing us to develop and train the ensemble learning model effectively. This successful integration of big data management with ensemble learning techniques paves the way for future research exploring the application of advanced machine learning on large educational datasets in Morocco and beyond.

3.3. Comparaison study

Table 2 summarizes the findings from previous studies that investigated student orientation prediction using various machine-learning techniques. As can be observed, our ensemble learning approach, which incorporates both random forest and neural network models, achieves the highest accuracy of 98.60% in 2024. This suggests that our approach is currently the most effective for students’ orientation. The accuracy of most approaches has improved over time. For example, SVM improved from 92.10% in 2021 to 92.73% in 2022. This superior performance underscores the effectiveness of our data-driven methodology in student orientation prediction compared to other studies in the field.

Table 2. Results of differents studies

Study	Year	Approach	Data type	Accuracy (%)	Big data
[1]	2021	Naïve Bayes	Synthetic	92.10	Yes
[8]	2022	Random forest	Real	92.73	No
Our study	2024	Ensemble learning	Real	98.60	Yes

4. CONCLUSION

This study aimed to address the challenge of guiding high school students towards suitable educational paths by developing a predictive model using big data and machine learning techniques. By comparing the performance of various classification algorithms, we found that an ensemble model combining random forest and neural networks significantly outperformed individual models in predicting student orientation choices based on academic indicators. This model achieved exceptional precision, accuracy, recall, and F-measure, demonstrating its potential to revolutionize the student guidance process.

The findings of this research hold significant implications for educational institutions. By employing this big data-driven predictive model, schools can offer more personalized academic guidance, improve student satisfaction, and potentially reduce dropout rates. The accurate prediction of student orientation can optimize resource allocation and facilitate smoother transitions between educational levels. In our future research, we will explore the integration of additional factors such as student interests, personality traits, and socioeconomic background to enhance the predictive model’s accuracy. Developing a chatbot or interactive platform based on these findings could provide students with real-time personalized guidance, empowering them to make informed educational decisions.

REFERENCES

[1] F. O. Ouatik, M. E. Erritali, and M. J. Jourhmane, "Student orientation using machine learning under MapReduce with Hadoop," *Journal of Ubiquitous Systems & Pervasive Networks*, vol. 13, no. 1, pp. 21–26, Oct. 2020, doi: 10.5383/JUSPN.13.01.003.

[2] H. El Mrabet and A. A. Moussa, "A framework for predicting academic orientation using supervised machine learning," *Journal of Ambient Intelligence and Humanized Computing*, vol. 14, no. 12, pp. 16539–16549, Dec. 2023, doi: 10.1007/s12652-022-03909-7.

[3] M. Badrani, A. Marouan, N. Kannouf, and A. Chetouani, "Pathways to success: a study of classification algorithms for guiding middle school students in Morocco," in *Proceedings of the 6th International Conference on Networking, Intelligent Systems & Security*, May 2023, pp. 1–5, doi: 10.1145/3607720.3607748.

[4] H. El Mrabet and A. A. Moussa, "Smart school guidance and vocational guidance system through the internet of things," in *Proceedings of the 2nd International Conference on Networking, Information Systems & Security*, Mar. 2019, pp. 1–5, doi: 10.1145/3320326.3320404.

[5] L. Kemper, G. Vorhoff, and B. U. Wigger, "Predicting student dropout: a machine learning approach," *European Journal of Higher Education*, vol. 10, no. 1, pp. 28–47, Jan. 2020, doi: 10.1080/21568235.2020.1718520.

[6] F. Ouatik, M. Erritali, F. Ouatik, and M. Jourhmane, "Students' orientation using machine learning and big data," *International Journal of Online and Biomedical Engineering (iJOE)*, vol. 17, no. 01, pp. 111–119, Jan. 2021, doi: 10.3991/ijoe.v17i01.18037.

[7] M. Mimis, M. El Hajji, Y. Es-saady, A. O. Guejdi, H. Douzi, and D. Mammass, "A framework for smart academic guidance using educational data mining," *Education and Information Technologies*, vol. 24, no. 2, pp. 1379–1393, Mar. 2019, doi: 10.1007/s10639-018-9838-8.

[8] M. Badrani, A. Marouan, N. Kannouf, and A. Chetouani, "Smart techniques for moroccan students' orientation," in *International Conference on Emerging Technologies and Intelligent Systems*, vol. 573 LNNS, 2022, pp. 361–366.

[9] J. Zhou, A. H. Gandomi, F. Chen, and A. Holzinger, "Evaluating the quality of machine learning explanations: a survey on methods and metrics," *Electronics*, vol. 10, no. 5, p. 593, Mar. 2021, doi: 10.3390/electronics10050593.

[10] A. Nabil, M. Seyam, and A. Abou-Elfetouh, "Prediction of students' academic performance based on courses' grades using deep neural networks," *IEEE Access*, vol. 9, pp. 140731–140746, 2021, doi: 10.1109/ACCESS.2021.3119596.

[11] A. H. Nabizadeh, D. Goncalves, S. Gama, and J. Jorge, "Early prediction of students' final grades in a gamified course," *IEEE Transactions on Learning Technologies*, vol. 15, no. 3, pp. 311–325, Jun. 2022, doi: 10.1109/TLT.2022.3170494.

[12] N. N. Sanchez-Pozo, J. S. Mejia-Ordonez, D. C. Chamorro, D. Mayorca-Torres, and D. H. Peluffo-Ordonez, "Predicting high school students' academic performance: a comparative study of supervised machine learning techniques," in *2021 Machine Learning-Driven Digital Technologies for Educational Innovation Workshop*, Dec. 2021, pp. 1–6, doi: 10.1109/IEEECONF53024.2021.9733756.

[13] U. Ujwal and S. Malik, "A hybrid weight based feature selection algorithm for predicting students' academic advancement by employing data science approaches," *International Journal of Education and Management Engineering*, vol. 13, no. 5, pp. 1–22, Oct. 2023, doi: 10.5815/ijeme.2023.05.01.

[14] S. Rebai, F. B.Yahia, and H. Essid, "A graphically based machine learning approach to predict secondary schools performance in Tunisia," *Socio-Economic Planning Sciences*, vol. 70, p. 100724, Jun. 2020, doi: 10.1016/j.seps.2019.06.009.

[15] A. Hazra, S. Kumar, and A. Gupta, "Study and analysis of breast cancer cell detection using Naïve Bayes, SVM and ensemble algorithms," *International Journal of Computer Applications*, vol. 145, no. 2, pp. 39–45, Jul. 2016, doi: 10.5120/ijca2016910595.

[16] N. Seman and N. A. Razmi, "Machine learning-based technique for big data sentiments extraction," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 9, no. 3, pp. 473–479, Sep. 2020, doi: 10.11591/ijai.v9.i3.pp473-479.

[17] I. Hamdaoui, M. El Fissaoui, K. El Makkaoui, and Z. El Allali, "Hadoop-based big data distributions: a comparative study," in *International Conference on Networking, Intelligent Systems and Security*, 2022, pp. 242–252, doi: 10.1007/978-3-031-15191-0_24.

[18] S. Gao, L. Li, W. Li, K. Janowicz, and Y. Zhang, "Constructing gazetteers from volunteered big geo-data based on Hadoop," *Computers, Environment and Urban Systems*, vol. 61, pp. 172–186, Jan. 2017, doi: 10.1016/j.compenvurbsys.2014.02.004.

[19] A. Jain, V. Sharma, and V. Sharma, "Big data mining using supervised machine learning approaches for Hadoop with Weka distribution," *International Journal of Computational Intelligence Research*, vol. 13, no. 8, pp. 2095–2111, 2017.

[20] A. Kala Karun and K. Chitharanjan, "A review on hadoop - HDFS infrastructure extensions," in *2013 IEEE Conference on Information and Communication TechnologiES*, Apr. 2013, pp. 132–137, doi: 10.1109/CICT.2013.6558077.

[21] B. Mahesh, "Machine learning algorithms - a review," *International Journal of Science and Research (IJSR)*, vol. 9, no. 1, pp. 381–386, Jan. 2020, doi: 10.21275/ART20203995.

[22] O. Mahmoudi, M. F. Bouami, and M. Badri, "Arabic language modeling based on supervised machine learning," *Revue d'Intelligence Artificielle*, vol. 36, no. 3, pp. 467–473, Jun. 2022, doi: 10.18280/ria.360315.

[23] T. Vafeiadis, K. I. Diamantaras, G. Sarigiannidis, and K. C. Chatzisavvas, "A comparison of machine learning techniques for customer churn prediction," *Simulation Modelling Practice and Theory*, vol. 55, pp. 1–9, Jun. 2015, doi: 10.1016/j.simpat.2015.03.003.

[24] A. Marouan, M. Badrani, N. Kannouf, A. Zannou, and A. Chetouani, "Blockchain-based e-voting system in a university," *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*, vol. 34, no. 3, pp. 1915–1923, Jun. 2024, doi: 10.11591/ijeecs.v34.i3.pp1915-1923.

[25] S. Sabahi and M. M. Parast, "The impact of entrepreneurship orientation on project performance: a machine learning approach," *International Journal of Production Economics*, vol. 226, p. 107621, Aug. 2020, doi: 10.1016/j.ijpe.2020.107621.

[26] M. Kurucan, M. Özbaltan, Z. Yetgin, and A. Alkaya, "Applications of artificial neural network based battery management systems: a literature review," *Renewable and Sustainable Energy Reviews*, vol. 192, p. 114262, Mar. 2024, doi: 10.1016/j.rser.2023.114262.

[27] S. Hussain, N. A. Dahan, F. M. Ba-Alwi, and N. Ribata, "Educational data mining and analysis of students' academic performance using WEKA," *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*, vol. 9, no. 2, pp. 447–459, Feb. 2018, doi: 10.11591/ijeecs.v9.i2.pp447-459.

[28] N. Kannouf, M. Labbi, Y. Chahid, M. Benabdellah, and A. Azizi, "A key establishment attempt based on genetic algorithms applied to RFID technologies," *International Journal of Information Security and Privacy*, vol. 15, no. 3, pp. 33–47, Jul. 2021, doi: 10.4018/IJISP.2021070103.

[29] A. Gunawardana and G. Shani, "A survey of accuracy evaluation metrics of recommendation tasks," *Journal of Machine Learning Research*, vol. 10, no. 12, pp. 2935–2962, 2009.

[30] O. Rainio, J. Teuho, and R. Klén, "Evaluation metrics and statistical tests for machine learning," *Scientific Reports*, vol. 14, no. 1, p. 6086, Mar. 2024, doi: 10.1038/s41598-024-56706-x.

[31] T. Winarti, H. Indriyawati, V. Vydia, and F. W. Christanto, "Performance comparison between naive bayes and k-nearest neighbor algorithm for the classification of Indonesian language articles," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 10, no. 2, pp. 452–457, Jun. 2021, doi: 10.11591/ijai.v10.i2.pp452-457.

[32] A. Merhej Radhi, "Risk assessment optimization for decision support using intelligent model based on fuzzy inference renewable rules," *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*, vol. 19, no. 2, pp. 1028–1035, Aug. 2020, doi: 10.11591/ijeecs.v19.i2.pp1028-1035.

BIOGRAPHIES OF AUTHORS

Morad Badrani is currently pursuing a Ph.D. in Computer Science within the ORAS team at LaMAO Laboratory, ENCG, Mohamed First University in Oujda, Morocco. His research primarily revolves around leveraging machine learning and deep learning algorithms to address various challenges in the education sector. Holding a master's degree in information engineering from Abdelmalik Essadi Tetouan University since 2013, He possesses a strong academic background in this domain. His academic interests encompass data science, machine learning, deep learning, and blockchain technologies. He is deeply committed to harnessing the potential of these technologies to foster innovation and bring about positive changes in the realm of education. He can be contacted at email: m.badrani@ump.ac.ma.

Adil Marouan is a Ph.D. student currently affiliated with the Applied Mathematics Laboratory of the Oriental (LaMAO) at ENCG Oujda. His research focus centers around Blockchain Technology for Education. Marouan's academic journey is marked by a Master's Specialization in Computer Engineering (M.2.I) from the Faculty of Sciences in Oujda (F.S.O) at Mohammed Premier University (U.M.P). Prior to this, he earned a fundamental degree in Mathematical Sciences and Computer Science (S.M.I). His diverse educational background reflects his commitment to advancing his knowledge and skills in the realm of information technology and engineering. He can be contacted at email: adil.marouan@ump.ac.ma.

Prof. Dr. Nabil Kannouf is a distinguished professor of Computer Science, holding a Ph.D. in Computer Sciences. He earned his Master's degree in Computer Engineering from the Faculty of Sciences at the University of Mohamed Premier in Oujda, Morocco, in 2010. Since 2020, he has been serving as a professor of computer science at Abdelmalek Essaidi University in Morocco. With a prolific academic career, he has authored and published over 10 research papers. For further communication, he can be contacted at email: nabil.kannouf@gmail.com.

Prof. Dr. Abdelaziz Chetouani born on May 10, 1973, in Oujda, Morocco, is a highly accomplished academic and researcher. He began his academic journey by obtaining a Bachelor's degree in Applied Mathematics from the Department of Mathematics at the Faculty of Sciences, Mohamed I University, Oujda, in 1995. Continuing his academic pursuits, he earned a Master's degree in Applied Mathematics from the same institution in 1997. In 2003, Abdelaziz successfully completed his Doctorate in Numerical Methods of Nonlinear Parabolic Problems with Nonlocal Boundary Conditions and Applications at the Faculty of Sciences, Mohamed I University, Oujda. Currently serving as a Professor in the Department of CTAM at ENCGO, Mohamed I University, Oujda, since 2007, her research interests encompass numerical analysis, artificial intelligence, and biomathematics. His impressive academic track record includes the authorship of 19 papers and 2 theses. Additionally, her actively contributes to scientific organizations, holding memberships in CIMPA (2019, 2022), CIMAM (2022), and ROSA (2022). For further inquiries, he can be contacted at email: a.chetouani@ump.ac.ma.