

ORIGINAL

## Utilizing Data Mining and Machine Learning for Enhancing Bachelor's Degree Outcomes and Predicting Students' Academic Success

### Utilizando la Minería de Datos y el Aprendizaje Automático para Mejorar los Resultados de los Programas de Licenciatura y Predecir el Éxito Académico de los Estudiantes

Mohamed Sabiri<sup>1</sup> , Yousef Farhaoui<sup>1</sup> , Agoujil Said<sup>2</sup>

<sup>1</sup>STI Laboratory, IDM, T-IDMS Faculty of Sciences and Techniques, Moulay Ismail University of Meknes. Morocco.

<sup>2</sup>Ecole Nationale de Commerce et de Gestion, Moulay Ismail University of Meknes. Morocco.

**Cite as:** Sabiri M, Farhaoui Y, Said A. Utilizando la Minería de Datos y el Aprendizaje Automático para Mejorar los Resultados de los Programas de Licenciatura y Predecir el Éxito Académico de los Estudiantes. Data and Metadata 2023;2:105. <https://doi.org/10.56294/dm2023105>.

Submitted: 22-08-2023

Revised: 07-11-2023

Accepted: 28-12-2023

Published: 29-12-2023

Editor: Prof. Dr. Javier González Argote 

**Note:** paper presented at the International Conference on Artificial Intelligence and Smart Environments (ICAISE'2023).

#### ABSTRACT

This paper aims to conceptualize, design, and implement a Data Mining (DM) system integrated with machine learning within the realm of school management. The primary objective is to support the educational community and decision-makers in addressing the issue of school dropout and enhancing success rates at the certificate levels in Morocco, specifically focusing on the bachelor's degree examination in the qualifying cycle. The proposed system categorizes students five months prior to the exam date, facilitating targeted academic interventions for those at risk of course repetition or discontinuation. The DM system, operational throughout the school year, enhances the precision and effectiveness of schools and provincial administrations by identifying areas requiring additional support to improve end-of-year success rates and student performance. Project development is rooted in the collection and analysis of existing data from various departmental information systems, utilizing classification and regression algorithms to predict learner performance, success rates, and overall outcomes at the conclusion of certificate levels.

**Keywords:** Data Mining; Machine learning; Prediction.

#### RESUMEN

Este artículo tiene como objetivo conceptualizar, diseñar e implementar un sistema de Minería de Datos (DM) integrado con Aprendizaje Automático en el ámbito de la gestión escolar. El objetivo principal es apoyar a la comunidad educativa y a los tomadores de decisiones en abordar el problema de la deserción escolar y mejorar las tasas de éxito en los niveles de certificación en Marruecos, centrándose específicamente en el examen de licenciatura en el ciclo clasificatorio. El sistema propuesto categoriza a los estudiantes cinco meses antes de la fecha del examen, facilitando intervenciones académicas dirigidas para aquellos en riesgo de repetición de cursos o abandono. El sistema de DM, operativo durante todo el año escolar, mejora la precisión y efectividad de las escuelas y las administraciones provinciales al identificar áreas que requieren apoyo adicional para mejorar las tasas de éxito al final del año y el rendimiento de los estudiantes. El desarrollo del proyecto se basa en la recopilación y análisis de datos existentes de diversos sistemas de información departamentales, utilizando algoritmos de clasificación y regresión para predecir el rendimiento de los estudiantes, las tasas de éxito y los resultados generales al final de los niveles de certificación.

**Palabras clave:** Minería de Datos; Aprendizaje Automático; Predicción.

## INTRODUCCIÓN

Technological advancements in hardware and software enable artificial intelligence, which allows all disciplines to predict future actions, outcomes and improve performance.<sup>(1)</sup> In fact, this could be the reason for changes in all spheres of life, including communication and the breakdown of barriers due to the Internet and the volume of data generated and stored by various information systems, whether they are specific to an establishment or are shared through social networks.<sup>(2)</sup>

One of the most pressing problems in secondary school is student dropout, which has significant negative social and financial effects. Lately, there has been a lot of interest in the early detection of students who are in danger of dropping out using machine learning methods.<sup>(3,4,5)</sup>

Data-driven educational research has a lot of potential thanks to the huge amount of data in school administrative systems. Artificial intelligence can play several roles in educational environments. For instance, the latest advancements in natural language processing allow for the automation of free text answer grading.<sup>(6,7)</sup>

Although machine learning algorithms are used in the great majority of dropout prediction research to correctly classify prospective school's dropouts and graduates as feasible, the interpretation of the trained models and the reasons for their predictions are typically disregarded.<sup>(8)</sup> A strong decision support system, on the other hand, not only aids in classifying data efficiently but also in elucidating the reasoning behind a given prediction's output and recommending the best course of action for the affected person. The latter enables tutoring sessions, remedial classes, and individualized supervision to be given to the pupils.<sup>(9,10)</sup>

Post-hoc model-agnostic interpretable machine learning technologies have been used in a limited number of research in an educational context.<sup>(11,12)</sup> By locally approximating the predictions of the black-box model with an inherently interpretable model, such as linear regression, the local interpretable model-agnostic explanations method offers interpretable explanations of individual predictions.<sup>(13,14)</sup> As a case in point, this work will focus on the educational domain, where one of the main objectives is to ensure that learning processes enable us to understand students and their learning paths and also detect their performance. This is where educational data mining (EDM) brings fundamental value to educational institutions and all the entities that support different learning activity processes.<sup>(15)</sup> The aim of this work is to present the use case for data mining techniques in the analysis of learning assessments that occur in education. We also present the different DM models to be exploited in the educational domain to predict student success outcomes in order to deploy a more powerful model that helps educational decision-makers predict and make decision-making more effective.

This work enables school leaders and educational decision-makers to access student data, predict achievement levels and find students or groups of students who need special attention. For example, students who are weak in specific subjects, with the aim of preparing tutoring sessions at the right time during the school year, as well as combating school drop-out.<sup>(16)</sup>

This study is based on data from the Regional Academy for Education and Training (RAET) of Drira Tafilalt region (Morocco).

The Ministry of Education has set up integrated information systems (IS) for school management and monitoring.<sup>(17,18,19,20,21,22,23,24,25,26,27,28)</sup> The information systems on which data collection is based in this project are the MASSAR information system and the two information systems for managing certification examinations as shown in figure 1.



**Figure 1.** Information systems for managing certification examinations

The information systems listed above include data that is exploited and extracted in the formats of XML, CSV, and Access databases. The files need to be merged into an MS SQL Server relational database. The information systems that allow the categorization of learners collect information about each student's gender, the number of duplicates in the bachelor's degree year, notes from this year's major subjects, and information about each student's type of school (public or private) and environment (rural or urban) based on their outcome of earning a baccalaureate at the end of the academic year.

The modeling of each learner's is achieved by combining fifteen features, as illustrated in table 1, and the distribution of various subjects at different levels (Table 1).

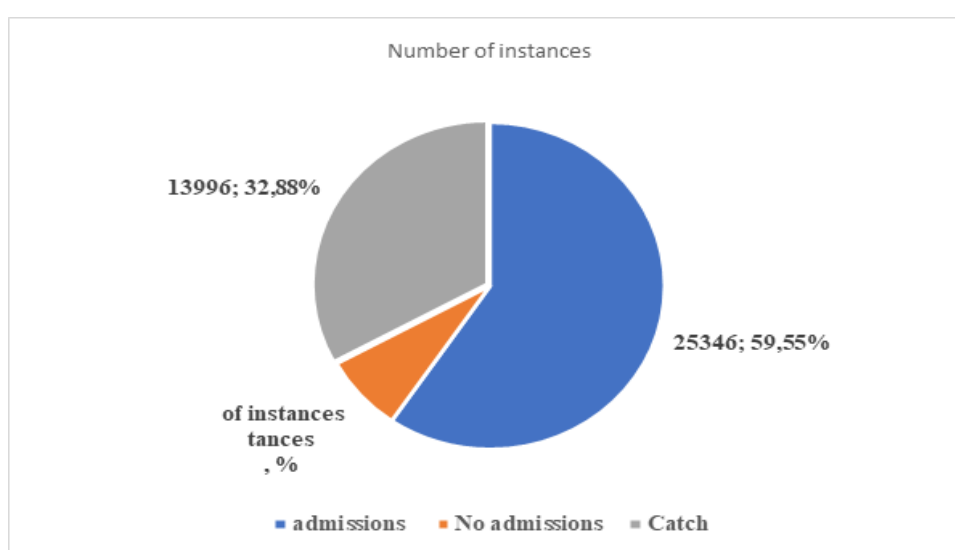
Table 1. Modeling of each learner’s is achieved by combining fifteen features				
Attributs		Type	Description	Value
Gender		Categorical	-	Male/Female
Level		Categorical	Discipline or BAC Type Name	See Table No. 02
Type Establishment		Categorical	Public/Private	Public/Private
natural environment		Categorical	Urban/rural	Urban/rural
Number of years Bac		Number	Number of years in 2nd Bac	Between 1 and 3
Main Subjects 1st year Bac	Subject No.01	Numeric	The grades obtained in the main subjects in continuous monitoring	Between 0 and 20
1st semester	Subject No.02	Numeric		
	Material No.03	Numeric		
	Subject No.04	Numeric		
Main Subjects 1st year Bac	Subject No.01	Numeric	The grades obtained in the main subjects in continuous monitoring	
2nd semester	Subject No.02	Numeric		
	Material No.03	Numeric		
	Subject No.04	Numeric		
Grade Exam Regional		Numeric	Grade Exam Regional	
Average 1st semester 2Bac		Numeric	Average 1st semester 2BAC	

Table 2. Subjects by Level and Major								
Level	Level Technologie	Level Pro	Level SH	Level SVT	Level Lettre	Eco et gestion	Level S Math	Level PC
Economics and statistics	-	-	-	-	-	MP01	-	-
Economy and management of enterprises	-	-	-	-	-	MP02	-	-
Accounting and financial mathematics	-	-	-	-	-	MP03	-	-
Engennering sciences	MP01	-	-	-	-	-	-	-
History geograph	-	-	MP01	-	MP01	-	-	-
Mathematics	MP02	MP01	-	MP01	-	MP04	MP01	MP01
Philosophy	-	-	MP02	-	MP02	-	-	-
Physics and chemistry	MP03	MP02	-	MP02	-	-	MP02	MP02
English	MP04	MP03	MP03	MP03	MP03	-	MP03	MP03
Arabic	-	-	MP04	-	MP04	-	-	-
Professional materials	-	MP04	-	-	-	-	-	-
Life and Earth Sciences	-	-	-	MP04	-	-	MP04	MP04

The target variable (Target), with a total of 42 563 instances and fifteen attributes, constitutes the operating dataset (Table 3 and Figure 2). It is mentioned that the archive refers to the 2017-2018-2019 academic years.

**Table 3.** Summary of Admissions by Option and Total

	Catch		No admissions		Admissions		Total
Economics and Management	710	5,0729 %	177	5,4952 %	1 245	4,912 %	2 132
Letter Option	1 981	14,154 %	420	13,0394 %	4 296	16,9494 %	6 697
Option Physics Chemistry	1 900	13,5753 %	593	18,4104 %	6 093	24,0393 %	8 586
SVT option	3 679	26,2861 %	995	30,891 %	4 239	16,7245 %	8 913
Science Math Option	159	1,136 %	74	2,2974 %	1 512	5,9654 %	1 745
Option Humanities	5 349	38,2181 %	902	28,0037 %	7 311	28,8448 %	13 562
Prof Series	147	1,0503 %	36	1,1177 %	143	0,5642 %	326
Technology	71	0,5073 %	24	0,7451 %	507	2,0003 %	602
Total	13 996		3 221		25 346		42 563



**Figure 2.** Summary of Admissions by Option and Total

## EXPLORATION OF THE DATA

At this step, choosing the appropriate method for knowledge extraction (exploration) from the data is necessary. When techniques like clustering (segmentation), association rules, classification, and regression are used, it is understood that each one needs an initial look at the different instances and attributes of the data set, specifically the connections and correlations between each attribute.

Analyzing the relationship and influence between the various variables in the data set necessitates performing correlation analysis, which is an essential process.<sup>(16)</sup>

The Goodman and Kruskal gammas, two types of named ranking correlation values, are shown in this table. determining how comparable the data's order is when categorized according to each of the variables (Table 4). At the ordinal level, it measures the level of cross-data correlation between the two variables. It does not account for variations in the table. The values vary from -1 (perfect inversion or 100 % negative connection) to +1 (perfect agreement or 100 % positive association). Zero signals the absence of any relationship. The presence of a relationship (positive or negative association) between the variables is explained by correlation values that are either larger than or less than zero. For instance, there is a negative correlation between the variable number of years of bachelor's degree and the results of several topics, indicating that twice the candidate's level will have the opposite effect. Unlike other variables that have a positive link, the notes' values will impact the decision-making process and the correlations among the other subjects as they increase. The effect of the semester 1 CC scores (variable average half-year 1) with a correlation rate = 0,66 and the average of the regional examination with a value of correlation equal to 0,56 is the next situation we consider, where the correlation is considerably positive.

		Number of years	Main Subjects 1st year Bac 1st semester				Main Subjects 1st year Bac 2nd semester				Grade Exam	Average 1st	Result
			MP01	NMP02	MP03	MP04	MP01	MP02	NMP03	NMP04			
Number of years Bac		1.00	-0.36	-0.33	-0.30	-0.37	-0.37	-0.34	-0.32	-0.38	-0.46	- 0.63	- 0.85
Main Subjects 1st year	MP01	- 0.36	1.00	0.47	0.29	0.31	0.54	0.38	0.27	0.29	0.33	0.38	0.42
	MP02	- 0.33	0.47	1.00	0.32	0.30	0.36	0.53	0.29	0.28	0.34	0.36	0.40
	MP03	- 0.30	0.29	0.32	1.00	0.53	0.26	0.29	0.60	0.40	0.42	0.35	0.40
	MP04	- 0.37	0.31	0.30	0.53	1.00	0.29	0.29	0.43	0.54	0.39	0.39	0.44
Main Subjects 1st year	MP01	- 0.37	0.54	0.36	0.26	0.29	1.00	0.46	0.27	0.30	0.32	0.39	0.42
	MP02	- 0.34	0.38	0.53	0.29	0.29	0.46	1.00	0.32	0.29	0.33	0.38	0.41
	MP03	- 0.32	0.27	0.29	0.60	0.43	0.27	0.32	1.00	0.48	0.41	0.36	0.40
	MP04	- 0.38	0.29	0.28	0.40	0.54	0.30	0.29	0.48	1.00	0.35	0.39	0.45
Grade Exam Regional		- 0.46	0.33	0.34	0.42	0.39	0.32	0.33	0.41	0.35	1.00	0.41	0.57
Average 1st semester 2Bac		- 0.63	0.38	0.36	0.35	0.39	0.39	0.38	0.36	0.39	0.41	1.00	0.66
Result		- 0.85	0.42	0.40	0.40	0.44	0.42	0.41	0.40	0.45	0.57	0.66	1.00

Figure 3. Academic Variable Correlation Matrix

### APPROACH OF DATA MINING: MACHINE LEARNING ALGORITHMS

Adaboost, SVM, KNN, Decision Tree, Logistics Regression, Naive Bayes, and Neural Network represent some of the classification methods investigated with Python and Orange Tools. An analysis was done on 42 563 educational results cases. One class, with more than fifteen attributes, has three values: admitted, not admitted, and caught. Performance results are shown in table 5 and figure 2. It is essential to remember that the cross-validation method—which should be used in the test-training process—allows for evaluating models with a restricted data supply.

Table 5. Algorithm Model					
Algorithm Model	Target Value	Correct classification	Incorrect classification	Target precision	Accuracy Model
AdaBoost	Admitted	21 598	3 877	85,78 %	74,66 %
	caught	8 878	5 245	63,34 %	
	Not Admitted	1 384	1 879	40,06 %	
SVM	Admitted	22 198	3 277	73,37 %	60,84 %
	caught	2 186	11 937	48,91 %	
	Not Admitted	1 425	1 838	22,61 %	

KNN	Admitted	23 122	2 353	83,06 %	77,59 %
	caught	9 474	4 649	67,38 %	
	Not Admitted	627	2 636	68,97 %	
Logistic Regression	Admitted	23 292	2 183	87,27 %	80,83 %
	caught	10 867	3 256	69,80 %	
	Not Admitted	467	2 796	81,42 %	
Decision Tree	Admitted	22 667	2 808	85,07 %	76,46 %
	caught	8 649	5 474	67,49 %	
	Not Admitted	1 448	1 815	42,32 %	
Naïve Bayes	Admitted	19 144	6 331	89,14 %	72,64 %
	caught	11 510	2 313	56,68 %	
	Not Admitted	529	2 734	46,56 %	
Neural Network	Admitted	23 360	2 115	88,46 %	82,12 %
	caught	10 252	3 871	72,92 %	
	Not Admitted	1 406	1 857	66,11 %	

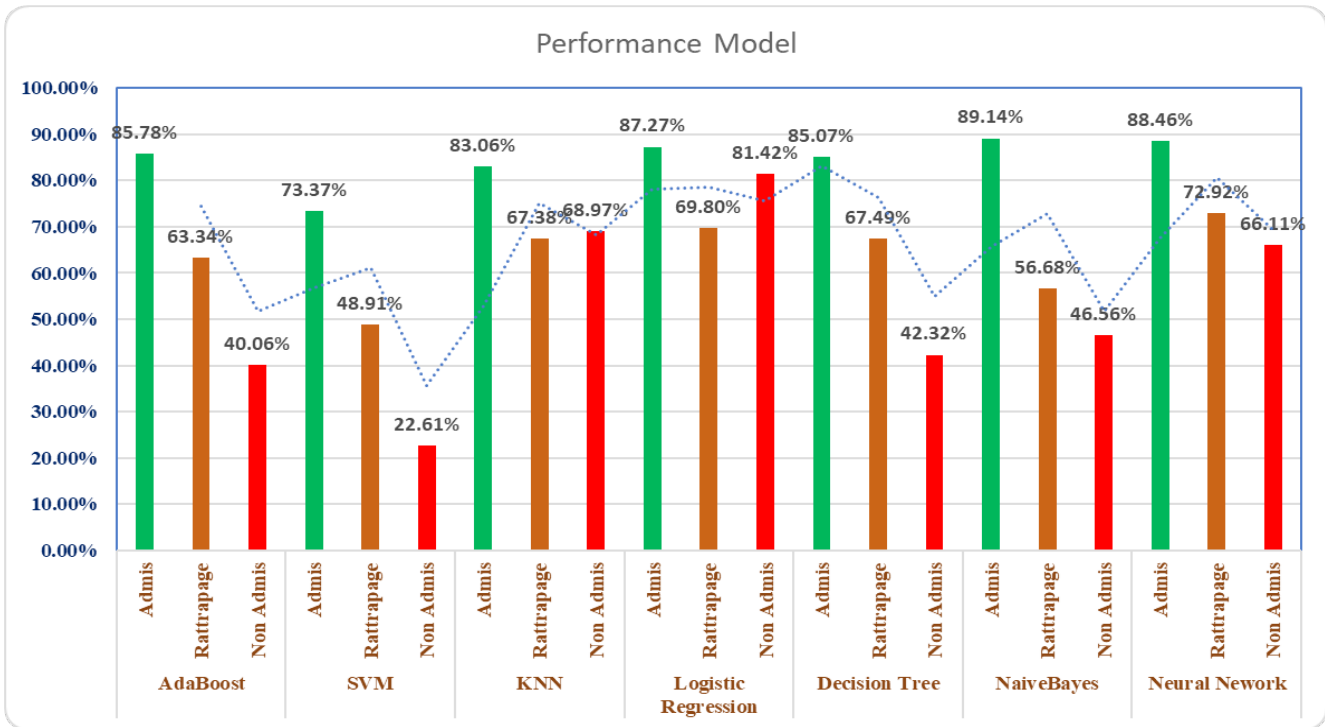


Figure 4. Performance Model

Table 6. Evaluation of each model a) Evaluation of target: admitted					
Model	AUC	CA	F1	Precision	Recall
Adaboost	0,82154346	0,82678892	0,85356424	0,85782817	0,84934249
kNN	0,9155134	0,83474487	0,86708326	0,83062487	0,90688911
Logistic Regression	0,94330544	0,86955507	0,8927962	0,87266662	0,91387635
Naive Bayes	0,90649511	0,79811484	0,81575641	0,89139134	0,75195289
Neural Network	0,95204102	0,87772101	0,89878918	0,88455983	0,91348381
SVM	0,81747503	0,70707636	0,76363499	0,73370234	0,79611384
Tree	0,80851368	0,84295747	0,87105611	0,85066976	0,89244357



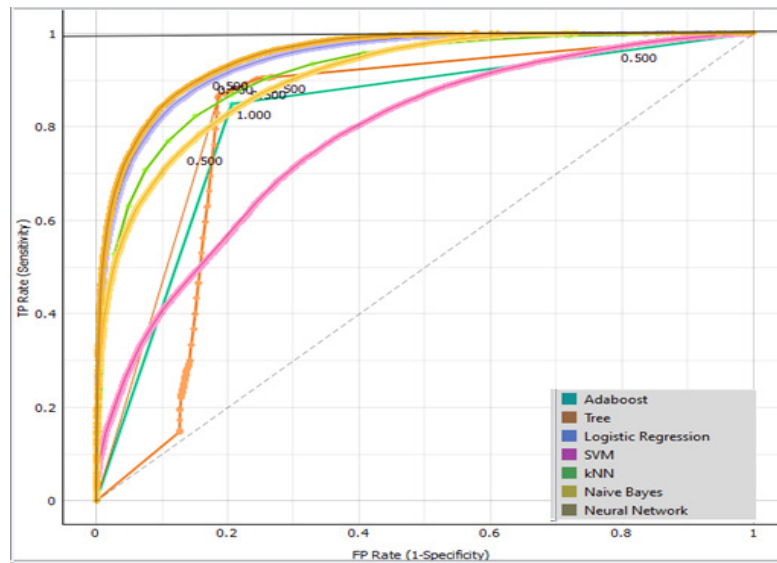


Figure 5. Evaluation of each model a) Evaluation of target: admitted

Model	AUC	CA	F1	Precision	Recall
Adaboost	0,68700322	0,9077483	0,41317898	0,40057554	0,42660129
kNN	0,85101302	0,93182614	0,29793369	0,68965517	0,19000919
Logistic Regression	0,89767917	0,9325494	0,25006485	0,81418919	0,14771683
Naive Bayes	0,86602198	0,92177037	0,26678329	0,46564885	0,18694453
Neural Network	0,92522256	0,93968876	0,5183529	0,66112167	0,42629482
SVM	0,80276907	0,84377406	0,29737671	0,22610499	0,43426295
Tree	0,65520242	0,91171461	0,43131951	0,423179	0,43977934

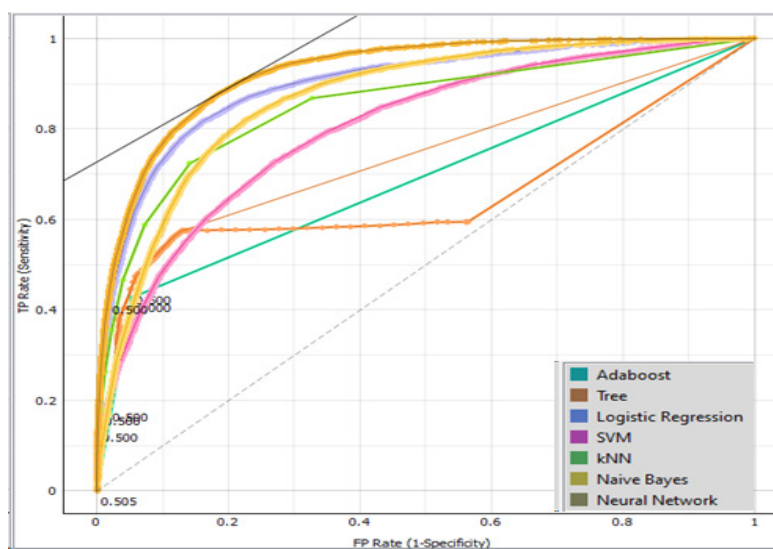


Figure 6. Evaluation of each model b) Evaluation of target: not admitted

Model	AUC	CA	F1	Precision	Recall
Adaboost	0,72727006	0,75866172	0,63430672	0,633411	0,63520498
kNN	0,85031281	0,785236	0,67440133	0,67380548	0,67499823
Logistic Regression	0,89005838	0,81456336	0,73251666	0,69803092	0,77058699
Naive Bayes	0,79847797	0,73299736	0,6652235	0,56677135	0,80506974
Neural Network	0,89986758	0,82501575	0,73750525	0,72918541	0,74601714
SVM	0,72710913	0,66592007	0,37945829	0,48905273	0,3099908
Tree	0,67900923	0,77450363	0,64026501	0,67490584	0,60900659

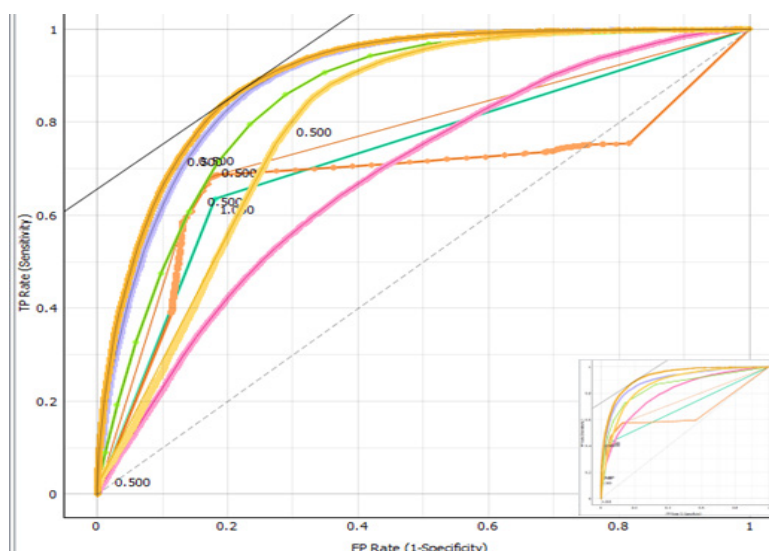


Figure 7. Evaluation of each model c) Evaluation of target: caught

## DISCUSSION

In Morocco, the Baccalaureate Examination is a significant national event enabling second-year students to obtain the graduate degrees they need for acceptance to universities. This section of the performance evaluation of the various models or classifiers notes, first, that the target values (admitted, not admitted, and caught) vary because each student in the regular session is subjected to one of three decisions. Additionally, based on the presentation and comparison of performance results tables, there is a significant variance in performance between the models, with the neural network algorithm being the most performing model in comparison to the other six models with an accuracy of 82,12 %, an F1 rating of 0,816, and an AUC of 0,926, indicating very good performance. As the primary goal of this study suggests, pupils who are predicted to be neither admitted or excluded at the end of the school year will most likely never get it. The same classifier model performs exceptionally well when it comes to predicting the list of learners in the class: no admitted or caught. This allows us to examine the grades of the subjects in which we must assist in terms of school support and to provide extra preparation before the baccalaureate exams. It also helps to increase success rates and reduce school dropouts in the qualifying secondary cycle.

## CONCLUSION

Exploration of educational data is a relatively new field that offers enormous potential for anyone involved in the learning process. Techniques enabling data exploration have been developed to automatically identify patterns in the data and discover submerged knowledge. Classifying and predicting student performance, dropout rates, pupils who are at risk of abandoning school, and teacher performance can all be done through the exploration of educational data. This can assist teachers in monitoring student progress to improve instruction, increase graduation rates, and reduce crowding in the classroom. It can also assist students in making better decisions about their courses, school assistance programs, and educational administration.

To increase school profitability, educational data exploration can be used to draw in and educate students. Finding, identifying, and comprehending the most effective teaching approaches requires analyzing student data. We demonstrated in this work how a data mining system based on information from the “Draa Tafilalet”



Region was implemented for Moroccan certification levels.

In order to identify students who are at risk of dropping out of school by double-failing their classes, the primary goal of this paper is to identify and analyze student performance based on the core subjects of each level and orientation series. This will allow educational decision-makers to implement strategies that can assist the affected students at the most advantageous times during the school year.

## ACKNOWLEDGMENTS

We deeply appreciate the director of the AREF DRAA TAFILALET-MOROCCO for his support and permission to use the different data sets.

## REFERENCES

1. Susnjak T. Beyond Predictive Learning Analytics Modelling and onto Explainable Artificial Intelligence with Prescriptive Analytics and ChatGPT. *Int J Artif Intell Educ* 2023. <https://doi.org/10,1007/s40593-023-00336-3>.
2. Sunyaev A. *Internet Computing: Principles of Distributed Systems and Emerging Internet-Based Technologies*. Cham: Springer International Publishing; 2020. <https://doi.org/10,1007/978-3-030-34957-8>.
3. Zapata-Rivera D. Open Student Modeling Research and its Connections to Educational Assessment. *Int J Artif Intell Educ* 2021;31:380-96. <https://doi.org/10,1007/s40593-020-00206-2>.
4. Nagy M, Molontay R. Interpretable Dropout Prediction: Towards XAI-Based Personalized Intervention. *Int J Artif Intell Educ* 2023. <https://doi.org/10,1007/s40593-023-00331-8>.
5. Gray CC, Perkins D. Utilizing early engagement and machine learning to predict student outcomes. *Comput Educ* 2019;131:22-32.
6. Howard E, Meehan M, Parnell A. Contrasting prediction methods for early warning systems at undergraduate level. *Internet High Educ* 2018;37:66-75.
7. Rienties B, Køhler Simonsen H, Herodotou C. Defining the boundaries between artificial intelligence in education, computer-supported collaborative learning, educational data mining, and learning analytics: A need for coherence. *Front. Educ.*, vol. 5, Frontiers Media SA; 2020, p. 128.
8. Eglington LG, Pavlik PI. How to Optimize Student Learning Using Student Models That Adapt Rapidly to Individual Differences. *Int J Artif Intell Educ* 2023;33:497-518. <https://doi.org/10,1007/s40593-022-00296-0>.
9. Houari R, Bounceur A, Kechadi T. Nouvelle approche de prétraitement pour les fouilles de données numériques. *Proc. 2ieme Édition Conférence Natl. L'informatique Destin. Aux Étud. Grad. Postgraduation JEESI'12*, 2012.
10. Miguéis VL, Freitas A, Garcia PJ, Silva A. Early segmentation of students according to their academic performance: A predictive modelling approach. *Decis Support Syst* 2018;115:36-51.
11. Vieira CP, Digiampietri LA. Machine Learning post-hoc interpretability: a systematic mapping study. *XVIII Braz. Symp. Inf. Syst., Curitiba Brazil: ACM*; 2022, p. 1-8. <https://doi.org/10,1145/3535511,3535512>.
12. Mitros J, Mac Namee B. *A Categorisation of Post-hoc Explanations for Predictive Models* 2019.
13. Zafar MR, Khan N. Deterministic local interpretable model-agnostic explanations for stable explainability. *Mach Learn Knowl Extr* 2021;3:525-41.
14. Messalas A, Kanellopoulos Y, Makris C. Model-agnostic interpretability with shapley values. *2019 10th Int. Conf. Inf. Intell. Syst. Appl. IISA, IEEE*; 2019, p. 1-7.
15. Romero C, Ventura S. Educational data mining and learning analytics: An updated survey. *WIREs Data Min Knowl Discov* 2020;10:e1355. <https://doi.org/10,1002/widm.1355>.
16. Möller J, Zitzmann S, Helm F, Machts N, Wolff F. A Meta-Analysis of Relations Between Achievement and Self-Concept. *Rev Educ Res* 2020;90:376-419. <https://doi.org/10,3102/0034654320919354>.

17. Gonzalez-Argote D, Gonzalez-Argote J. Generation of graphs from scientific journal metadata with the OAI-PMH system. *Seminars in Medical Writing and Education* 2023;2:43-43. <https://doi.org/10,56294/mw202343>.
18. Uman JMM, Arias LVC, Romero-Carazas R. Factores que dificultan la graduación: El caso de la carrera profesional de contabilidad en las universidades peruanas. *Revista Científica Empresarial Debe-Haber* 2023;1:58-74.
19. Inastrilla CRA. Big Data in Health Information Systems. *Seminars in Medical Writing and Education* 2022;1:6-6. <https://doi.org/10,56294/mw20226>.
20. Contreras JG, Rodríguez AU, Gaviño AS. Comportamiento Organizacional para el Balance Integral Humano desde la NOM-035 en escenario post-pandemia COVID-19. *Revista Científica Empresarial Debe-Haber* 2023;1:41-57.
21. Rodríguez FAR, Flores LG, Vitón-Castillo AA. Artificial intelligence and machine learning: present and future applications in health sciences. *Seminars in Medical Writing and Education* 2022;1:9-9. <https://doi.org/10,56294/mw20229>.
22. Alaoui, S.S., and all. "Hate Speech Detection Using Text Mining and Machine Learning", *International Journal of Decision Support System Technology*, 2022, 14(1), 80. DOI: 10,4018/IJDSST.286680
23. Alaoui, S.S., and all, "Data openness for efficient e-governance in the age of big data", *International Journal of Cloud Computing*, 2021, 10(5-6), pp. 522-532, <https://doi.org/10,1504/IJCC.2021,120391>
24. El Mouatasim, A., and all. "Nesterov Step Reduced Gradient Algorithm for Convex Programming Problems", *Lecture Notes in Networks and Systems*, 2020, 81, pp. 140-148. [https://doi.org/10,1007/978-3-030-23672-4\\_11](https://doi.org/10,1007/978-3-030-23672-4_11)
25. Tarik, A., and all. "Recommender System for Orientation Student" *Lecture Notes in Networks and Systems*, 2020, 81, pp. 367-370. [https://doi.org/10,1007/978-3-030-23672-4\\_27](https://doi.org/10,1007/978-3-030-23672-4_27)
26. Sossi Alaoui, S., and all. "A comparative study of the four well-known classification algorithms in data mining", *Lecture Notes in Networks and Systems*, 2018, 25, pp. 362-373. [https://doi.org/10,1007/978-3-319-69137-4\\_32](https://doi.org/10,1007/978-3-319-69137-4_32)
27. Canova-Barrios C, Machuca-Contreras F. Interoperability standards in Health Information Systems: systematic review. *Seminars in Medical Writing and Education* 2022;1:7-7. <https://doi.org/10,56294/mw20227>.
28. Farhaoui, Y., "Securing a Local Area Network by IDPS Open Source", *Procedia Computer Science*, 2017, 110, pp. 416-421. <https://doi.org/10,1016/j.procs.2017,06.106>

## FINANCING

No financing.

## CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

## AUTHORSHIP CONTRIBUTION

*Conceptualization:* Mohamed Sabiri, Yousef Farhaoui, Agoujil Said.

*Research:* Mohamed Sabiri, Yousef Farhaoui, Agoujil Said.

*Drafting - original draft:* Mohamed Sabiri, Yousef Farhaoui, Agoujil Said.

*Writing - proofreading and editing:* Mohamed Sabiri, Yousef Farhaoui, Agoujil Said.