International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-13 Issue-1, May 2024

A Framework to Optimize Student Performance using Machine Learning

Abhijeet Joshi, A. S. Kapse

Abstract*: For scholars, mining data and extracting information from huge databases has emerged as an intriguing field of study. Since a few decades ago, the concept of using data mining techniques to extract information has been around. The dataset was originally intended to be partitioned and the inherent features examined using classification and clustering algorithms. They base their predictions on these characteristics. These forecasts have been made in the area of educational data mining for a variety of reasons, including to predict student success based on personal characteristics and help students find the right professors and courses. These goals have been drawn from the attrition and retention of students. These objectives are the focus of our research on student attrition and retention. Additionally, we have found exciting variables that aid in predicting students' success, suggesting the most qualified instructors, and assisting them in course selection.

Keywords: Mining, Databases, Information, Dataset, Predictions, Performance

I. INTRODUCTION

Data mining and information extraction from massive databases have become fascinating academic fields of study. The idea of employing data mining techniques to extract information has existed for a few decades. The first plan was to divide up the dataset and use algorithms for classification and clustering to analyse the innate traits. They use these traits as the foundation for their predictions. These predictions have been made in the field of educational data mining for a number of purposes, such as predicting student success based on individual traits and assisting students in selecting the best instructors and courses [1]. These objectives were derived from student retention and attrition rates. Our study on student attrition and retention is centered on these goals. Additionally, we have discovered intriguing characteristics that support success prediction, instructor recommendation, and course recommendation for students [2][16][17][18][19][20].

Manuscript received on 09 April 2024 | Revised Manuscript received on 03 May 2024 | Manuscript Accepted on 15 May 2024 | Manuscript published on 30 May 2024.

*****Correspondence Author(s)*

Mr. Abhijeet Joshi*, Department of Computer Science & Engineering, Anuradha Engineering College Chikhli (Maharashtra), India. E-mail: abhijeet.joshi26@gmail.com, ORCID ID: 0009-0004-6150-404X

Dr. Avinash S. Kapse, Associate Professor & Head of Department Computer Science & Engineering, Anuradha Engineering College Chikhli (Maharashtra), India. E-mail: askapse@gmail.com

II. OBJECTIVES

- To focus on predicting the grades of low performance students by applying various machine learning techniques.
- To compare the techniques on the basis of mean absolute error.
- To make prediction made by machine learning algorithms will help the teachers and mentors to decide the course of education to be applied to specific students.
- To build such recommender system that helps students in selection and registration of the courses will be quite beneficial for the institutions.

III. LITERATURE REVIEW

Livieris et. al. states that Educational Data Mining (EDM) constitutes a new research field, which gained popularity in the modern educational era because of its potential to improve the quality of the educational institutions and system. During the last decade, this area of research field has grown exponentially, spurred by the fact that it enables all educational stakeholders to discover new, interesting and useful knowledge about students and potentially improve some aspects of the quality of education. The importance of EDM is founded on the fact that it allows educators and researchers to extract useful conclusions from sophisticated and complicated questions. More specifically, while traditional database queries can only answer questions such as “find the students with poor performance”, data mining can provide answers to more abstract questions like ”find the students who will exhibit poor performance”. Hence, the application of EDM is mainly concentrated on the development of accurate models that predict student characteristics and performances in order to improve learning experiences. The accurate prediction of students‟ academic performance is important for making admission decisions as well as providing better educational services [1]. Rudnichenko et. al. mentioned that Currently, there is a steady trend of regular growth in the volume of data collected in the various business organizations f production, operational and research activities processes. The sources of the such big data volumes (Big Data) appearance are often customers various behavioral factors, the frequency and size of payments for the services or goods, parameters and characteristics of installed technical equipment, medical indicators for diagnosing human health and others. Due to the statistical visibility and representativeness, the value of such data lies in the possibility of using it to search for hidden and unobvious relationships between individual factors (attributes) and target actions of clients to adjust and formulate business development strategies.

In fact, the implementation of such tasks becomes possible based on the use of data mining methods in order to extract new knowledge by forming and proving hypotheses about significant relationships between the individual attributes of data samples. Thus, the constant companies need to ensure a sufficient level of quality in the provision of goods and services, due to the high level of competition in the organization's business targets requires usage of the data science, data mining methods and technologies in key business processes. For this purpose, machine learning (ML) methods can be used to build various regression and predictive models for the business objectives [2].

IV. USE CASE DIAGRAM

A use case diagram in the Unified Modeling Language (UML) is a type of behavioral diagram defined by and created from a Use-case analysis. Its purpose is to present a graphical overview of the functionality provided by a system in terms of actors, their goals (represented as use cases), and any dependencies between those use cases. The main purpose of a use case diagram is to show what system functions are performed for which actor. Roles of the actors in the system can be depicted [3].

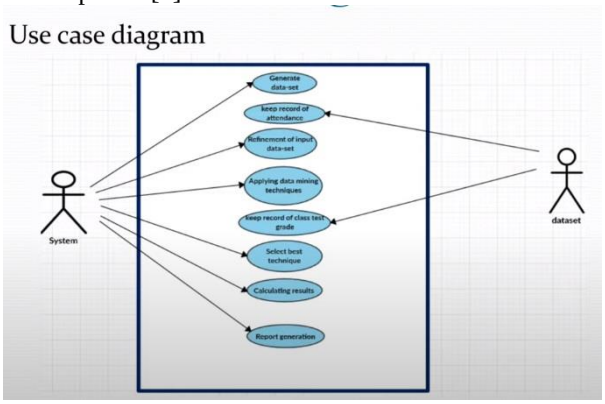

Fig. 1 Use case diagram

V. ALGORITHM AND TECHNIQUE

Nave Based and J48 classification algorithms have been compared in order to assess our research. These classifiers were chosen for a variety of factors, including the fact that they support categorical and nominal data. The programme is open source and accessible online; it is typically used for data training and testing, classification, and machine learning algorithms. The data can then be shown graphically as well. The created attribute combinations have been assessed in the current study to address our research questions. In order to enter the data into the dataset for this purpose, it was converted to csv format, and the output from various filters and classifiers was compiled [4].

A. Naïve Bayes

Naïve bayes algorithm is comparatively fast algorithm in terms of classification. It works faster on huge datasets by using Bayes algorithm of probability. Bayes algorithm generally used to predict the class of unknown dataset4. Naïve based algorithm works on assumptions to label an item whose features are known but name is unknown. For example; a fruit is labeled as an apple if it is round and red in color and its size is 3 inches in diameter. These features of apple will raise the probability of this fruit that it is an apple.

B. J48

J48 decision tree is used to predict the target variable of new dataset. If dataset contains predictors or independent variables and set of target or dependent variables, then this algorithm is applied to extract the target variable of new dataset5 [5].

C. Linear Regression

Generally linear regression classification algorithm is an approach to identify the relationship between dependent and independent variable. It is generally used for predictive analysis and has two main points. One is to check whether predictor variable does a good job in predicting the expected outcome variable. Second main thing that linear regression does is the identification of variable that are significant predictors of dependent variables. At the end, the regression equation is used which helps to determine the set of predictors which are used to predict the outcome. In this research, algorithms are being used to compare the trend and pattern of the factors with other approaches like non-linear regression and SMO [6].

VI. RESULT AND DISCUSSION

Entering various field into the system which helps to predict the current scenario of the grade of the students. It helps the students to judge the grade and study accordingly with due to respect to increase or decrease in the parameters schedule

A. Study Time from Terminal

```
Mean squared error is: 5.524221095174995
Weights: [-0.00482364 -0.24366995  0.10044543  0.0527645   0.11366621  0.99885866]
intersect: 2.2046770814240126
Enter Study Time:-
```

Fig. 2 Study Time from Terminal

B. Age From Terminal

```
Mean squared error is: 5.524221095174995
Weights: [-0.00482364 -0.24366995  0.10044543  0.0527645   0.11366621  0.99885866]
intersect: 2.2046770814240126
Enter Study Time:- 5
Enter student's age:-
```

Fig. 3 Age from Terminal

C. Travel Time to and Fro from Terminal

```
Mean squared error is: 5.524221095174995
Weights: [-0.00482364 -0.24366995  0.10044543  0.0527645   0.11366621  0.99885866]
intersect: 2.2046770814240126
Enter Study Time:- 5
Enter student's age:- 22
Enter travel time:-
```

Fig. 4 Travel Time to and Fro from Terminal

D. Absentee From Terminal

```
Mean squared error is: 5.524221095174995
Weights: [-0.00482364 -0.24366995  0.10044543  0.0527645   0.11366621  0.99885866]
intersect: 2.2046770814240126
Enter Study Time:- 5
Enter student's age:- 22
Enter travel time:- 20
Enter school absences:-
```

Fig. 5 Absentee from Terminal

Retrieval Number: 100.1/ijrte.A805213010524
DOI: 10.35940/ijrte.A8052.13010524
Journal Website: www.ijrte.org

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

E. Predicted Grade According to System

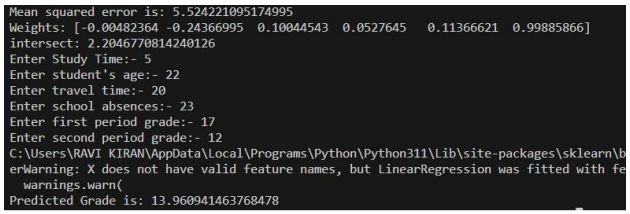

Fig. 9 Predicted Grade According to System

After finding the grade it will be easy to understand the flow of the student of performance and issue related to its performance and can be optimised introducing elements that can be used by educational institutions to uncover hidden patterns in academic databases. Every time the result was evaluated by the internal factors but the proposed system allows to evaluate and optimised the performance of the student by considering external factors which also effects the study patterns and resulted in unaware outcomes. So, the proposed system considered both internal and external factors for the purpose and improving the performance of student results [7][8].

VII. CONCLUSION

In order to improve the accuracy of the student performance system, we employ an improved platform that can access information intelligently to supply and carry out a number of tasks using the online educational system. The primary goals are to enhance the real-world experience for both students and teachers and to develop understanding and optimize the performance of the students. Making students' grasping abilities stronger will help them improvise more broadly and attain desirable results without stress. This study puts someone on the list of candidates for those who will devote enough time and effort to finish the course successfully.

ACKNOWLEDGMENT

We are thankful for the institutions for allowing us to submit a paper and interested to contribute and to evaluate the result of students which helps for the performance f the students in many ways [9][10].

DECLARATION STATEMENT

We hereby declared that we cited the suthors work and try to improvised the work done by many researcher in the field of education for its betterment.

Funding	No, we did not receive.
Conflicts of Interest	No conflicts of interest to the best of our knowledge.
Ethical Approval and Consent to Participate	No, the article does not require ethical approval and consent to participate with evidence.
Availability of Data and Material	Not relevant.
Authors Contributions	All authors have equal participation in this article.

REFERENCES

1. Ioannis E. Livieris, Tassos A. Mikropoulos, Panagiotis Pintelas , “A decision support system for predicting students’ performance “, Themes in Science & Technology Education, 9(1), 43-57, 2016
2. Nicolay Rudnichenko, Vladimir Vychuzhanin, Igor Petrov, Denis Shibaev, “Decision Support System for the Machine Learning Methods Selection in Big Data Mining”, CEUR-WS, Vol-2608, 2020 https://doi.org/10.32782/cmis/2608-65
3. R. Sasi Regha, Dr R. Uma Rani, “Optimization Feature Selection for classifying student in Educational Data Mining”, International Journal of Innovations in Engineering and Technology (IJIET) Volume 7, Volume 7 Issue 4 December 2016 ISSN: 2319 – 1058
4. Rasha Radi Shalabi, “The Importance and Applications of Decision Support Systems (Dss) in Higher Education”, DOI: 10.6084/m9.figshare.12465599.v1, 2020
5. Khalid A. Fakeeh, “Decision Support Systems (DSS) in Higher Education System”, International Journal of Applied Information Systems (IJAIS) – ISSN : 2249-0868 Foundation of Computer Science FCS, New York, USA Volume 9 – No.2, June 2015
6. G Abaya, S. A., & Gerardo, B. D. (2013, September). An education data mining tool for marketing based on C4. 5 classification technique. In e- Learning and e-Technologies in Education (ICEEE), 2013 Second International Conference on (pp. 289-293). IEEE. https://doi.org/10.1109/ICeLeTE.2013.6644390
7. Aher, S. B., & Lobo, L. M. R. J. (2012). A comparative study of association rule algorithms for course recommender system in e-learning. International Journal of Computer Applications, 39(1), 48-52. https://doi.org/10.5120/4788-7021
8. Al-Barrak, M. A., & Al-Razgan, M. (2016). Predicting student’s final GPA using decision trees: a case study. International Journal of Information and Education Technology, 6(7), 528. https://doi.org/10.7763/IJIET.2016.V6.745
9. Anand, S. S., Bell, D. A., & Hughes, J. G. (1996). EDM: A general framework for data mining based on evidence theory. Data & Knowledge Engineering, 18(3), 189-223. https://doi.org/10.1016/0169-023X(95)00038-T
10. Azarcon Jr, D. E., Gallardo, C. D., Anacin, C. G., & Velasco, E. (2014).
11. Attrition and retention in higher education institution: A conjoint analysis of consumer behavior in higher education. Asia Pacific Journal of Education, Arts and SCIENCE, 1(5), 107-118.
12. Baker, R. S., & Inventado, P. S. (2014). Educational data mining and learning analytics. In learning analytics (pp. 61-75). Springer New York. https://doi.org/10.1007/978-1-4614-3305-7_4
13. Baradwaj, B. K., & Pal, S. (2012). Mining educational data to analyse students' performance. arXiv preprint arXiv: 1201.3417.
14. Bydžovská, H. (2013). Course Enrolment Recommender System (Doctoral dissertation, Masarykova univerzita, Fakulta informatiky).
15. Cabrera, A. F., Nora, A., & Castaneda, M. B. (1993). College persistence: Structural equations modeling test of an integrated model of student retention. The journal of higher education, 64(2), 123-139. https://doi.org/10.2307/2960026
16. Gupta, R., Gowalker, N., Patil, D. S., & Joshi, Dr. S. D. (2019). Predicting Risk in Sentiment Analysis using Machine Learning. In International Journal of Engineering and Advanced Technology (Vol. 9, Issue 1, pp. 455–460). https://doi.org/10.35940/ijeat.a9540.109119
17. Radhamani, V., & Dalin, G. (2019). Significance of Artificial Intelligence and Machine Learning Techniques in Smart Cloud Computing: A Review. In International Journal of Soft Computing and Engineering (Vol. 9, Issue 3, pp. 1–7). https://doi.org/10.35940/ijsce.c3265.099319
18. Nikam, S. S., & Dalvi, Prof. R. (2020). Fake News Detection on Social Media using Machine Learning Techniques. In International Journal of Innovative Technology and Exploring Engineering (Vol. 9, Issue 7, pp. 940–943). https://doi.org/10.35940/ijitee.g5428.059720
19. Sharma, P., & Site, S. (2022). A Comprehensive Study on Different Machine Learning Techniques to Predict Heart Disease. In Indian Journal of Artificial Intelligence and Neural Networking (Vol. 2, Issue 3, pp. 1–7). https://doi.org/10.54105/ijainn.c1046.042322
20. Behera*, D. K., Das, M., & Swetanisha, S. (2019). A Research on Collaborative Filtering Based Movie Recommendations: From Neighborhood to Deep Learning Based System. In International Journal of Recent Technology and Engineering (IJRTE) (Vol. 8, Issue 4, pp. 10809–10814). https://doi.org/10.35940/ijrte.d4362.118419

AUTHORS PROFILE

Mr. Abhijeet Joshi, a final year post graduate student from Anuradha Engineering College, Chikhli affiliated to Sant Gadge Baba Amravati University, interested in the field of education, and want to provide better solution to the development of students by optimizing student performance by considering various factors. Focusing mainly on student performance in various categories also studied gradation system which helps student to improve result, generated result from multiple inputs before going through the practical assumption.

Dr. Avinash S. Kapse is an Associate Professor & Head of Department Computer Science & Engineering, Anuradha Engineering College Chikhli MS India and Research Supervisor of Sant Gadge Baba Amravati University Amravati in Computer Science & Engineering and Information Technology. He is a second term Board of Studies Member (BoS Member) in Information Technology of Sant Gadge Baba Amravati University Amravati. He has 18+years' experience in teaching and research. He published 53+ nos. research papers in national and international journals and conferences. He published 6 nos. patents & 7 Copyrights.