International Journal For Recent Developments in Science & Technology
2581-4575
Crossref
A Peer Reviewed Research Journal

Student Performance Prediction in Online Courses Using Machine Learning Algorithms

[1]DR. E. GURUMOORTHI [2]M.VINAY KUMAR ,[3]D.VEDANTH REDDY & [4]B. UDAY KIRAN

[1]Associate Professor, Department of Information Technology, CMR College of Engineering & Technology

[2, 3, 4] B-Tech, Department of Information Technology, CMR College of Engineering & Technology

Abstract:

Advances in Information and Communications Technology (ICT) have increased the growth of Massive open online courses (MOOCs) applied in distance learning environments. Various tools have been utilized to deliver interactive content including pictures, figures, and videos that can motivate the learners to build new cognitive skills. High ranking universities have adopted MOOCs as an efficient dashboard platform where learners from around the world can participate in such courses. The student's learning progress is evaluated by using set computer marked assessments. In particular, the computer gives immediate feedback to the student once he or she completes the online assessments. The researchers claim that student success rate in an online course can be related to their performance at the previous session in addition to the level of engagement. Insufficient attention has been paid by literature to evaluate whether student performance and engagement in the prior assessments could affect student achievement in the next assessments.

INTRODUCTION

Problem Statement

With rapid advancements in technology, artificial intelligence has recently being applicable into all possible fields. The growth of Massive open online courses (MOOCs) applied in distance learning environments. The MOOCs present the course using digital tool materials in various forms such as visual, audio, video, and plain text. Most students prefer using video lectures to understand the contents of lessons over thoroughly reading plain text documents. The interactive video in the MOOCs could reduce students' stress, help them to feel relaxed and learn quickly.

OBJECTIVE

MOOCs can be classified into two distinct types mainly, connectivist Massive Open Online Courses (cMOOCs) and extended

Massive Open Online Courses (xMOOCs). The xMOOCs are learning paradigm based on the principles of cognitivist behaviorist theory. The structure of the courses is similar to the traditional course where the syllabus consists of a set of video lectures and a set of multiple-choice quizzes in addition to the final exam. The video lectures featuring the course instructor reviewing the content of the previous online lesson are released weekly. The participants can watch and pause the video at their own pace. Moreover, the students can socially interact with other participants and the instructor through posting in discussion forums. The instructors usually post questions, provide task solutions, and reply to student questions via these discussion forums as a consequence the discussion forums play a vital role in enhancing the course quality and make online sessions collaborative and engaging. Various tools have been utilized to deliver interactive content including pictures, figures, and videos that can motivate the learners to build new cognitive skills. The student's learning progress is evaluated by using set computer marked assessments. In particular, the computer gives immediate feedback to the student once he or she completes the online assessments. Two sets of experiments are conducted in this work. In the first set of experiments, regression analysis is implemented for estimation of students' assessment scores.

Project Scope and Limitations

Student Performance can be analyzed and assessed In the Online Courses With the regression analysis, the data does not provide the last date of students' activity prior to undertaken assessments. The findings' results have been recommended to consider the temporal features on predicting of subsequent assessments grades. So, Activity of The Students Cannot Be Tracked Accurately in This model.

IMPLEMENTATION

To achieve the project's aims, quantitative simulation research methods were conducted as suggested in the framework phases. In these phases the dataset will be prepared to be passed through visualization and clustering techniques, i.e. like heat map and hierarchical clustering, to extract the top correlated indicators. Then, the indicators will be used in different classification algorithms and the most accurate model will be the chosen for predicting student performance in dissertation projects and all courses grades. In between, and before the classification models' evaluation phase, the datasets will

International Journal For Recent Developments in Science & Technology
2581-4575
Crossref
A Peer Reviewed Research Journal

pass through a pre-processing (cleansing, missing data imputation, …) stage to make it ready for the analysis phase.

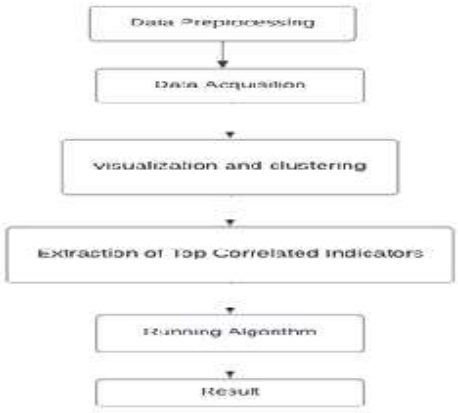


Figure1: Implementation

Flow of SVM Algorithm

PROPOSED SYSTEM

Objective of Proposed Model

Our Proposed System is to construct a machine learning model which is used for student performance prediction Using Gradient Boosting machine algorithm.The proposed technique has been successfully applied and also clearly demonstrates the performance of the Student Performance Prediction from the OULAD based data.

In this project the steps that are implemented are:

- Data Acquisition
- Feature Selection
- Matching
- Running Algorithm
- Result

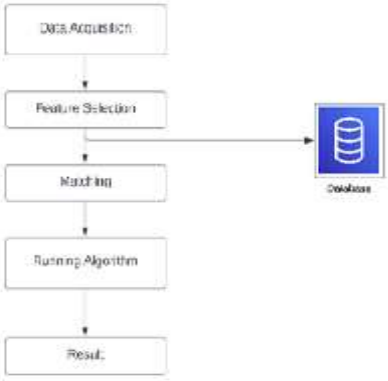


Figure.2: Flowchart of the running sequence of proposed solution

3.4. System Development Environment

Python:

Python is a general-purpose interpreted, interactive, object-oriented, and high-level programming language. It was created by Guido van Rossum during 1985- 1990. An interpreted language, Python has a design philosophy that emphasizes code readability (Notably using whitespace indentation to delimit code blocks rather than curly brackets or keywords), and a syntax that allows programmers to express concepts in fewer lines of code than might be used in languages such as C++or Java. It provides constructs that enable clear programming on both small and large scales. Python interpreters are available for many operating systems. Python, the

reference implementation of Python, is open-source software and has a community-based development model, as do nearly all of its variant implementations. Python features a dynamic type of system and automatic memory management. It supports multiple programming paradigms, including object-oriented, imperative, functional, and procedural, and has a large and comprehensive standard library.

History of Python:

Python was developed by Guido van Rossum in the late eighties and early nineties at the National Research Institute for Mathematics and Computer Science in the Netherlands. Python is derived from many other languages, including ABC, Modula-3, C, C++, Algol-68, Smalltalk, and Unix shell and other scripting languages. Python is copyrighted. Like Perl, Python source code is now available under the GNU General Public License (GPL). Python is now maintained by a core development team at the institute, although Guido van Rossum still holds a vital role in directing its progress.

Features of Python:

❖ Easy-to-learn − Python has few keywords, simple structure, and a clearly defined syntax. This allows the student to pick up the language quickly.

❖ Easy-to-read − Python code is more clearly defined and visible to the eyes.

❖ Easy-to-maintain − Python's source code is fairly easy-to-maintain.

❖ A broad standard library − Python's bulk of the library is very portable and cross-platform compatible on UNIX, Windows, and Macintosh.

❖ Interactive Mode − Python has support for an interactive mode which allows interactive testing and debugging of snippets of code.

❖ Portable − Python can run on a wide variety of hardware platforms and has the same interface on all platforms.

❖ Extendable − You can add low-level modules to the Python interpreter. These modules enable programmers to add to or customize their tools to be more efficient.

❖ GUI Programming − Python supports GUI applications that can be created and ported to many system calls, libraries and windows systems, such as Windows MFC, Macintosh, and the X Window system of Unix.

❖ Scalable − Python provides a better structure and support for large programs than shell scripting.

❖ Python is Interpreted − Python is processed at runtime by the interpreter. You do not need to compile your program before executing it. This is similar to PERL

and PHP.

❖ Python is Interactive − You can actually sit at a Python prompt and interact with the interpreter directly to write your programs.

❖ Python is Object - Oriented − Python supports Object-Oriented style or technique of programming that encapsulates code within objects.

❖ Python is a Beginner's Language − Python is a great language for the beginner-level programmers and supports the development of a wide range of applications from simple text processing to WWW browsers to games.

RESULTS:

Comparison of Existing Solutions

First Solution Which Is Prediction Using SVM Algorithm Has Major Flaws Which Include Computation Time of the result and The Second one is Size of Dataset. SVM Algorithm can Handle only Small Datasets Compared to other efficient algorithms.In FAM model in the existing work, the system is poor performance in which the assessments are not Tutor Marked Assessment (TMA), Computer Marked Assessment (CMA). This system is less performance due to Lack of Massive open online courses (MOOCs). Decision Trees, comes with a few serious disadvantages, including overfitting, error due to bias and error due to variance. Overfitting happens for many reasons, including presence of noise and lack of representative instances. It's possible for overfitting with one large (deep) tree. Bias error happens when you place too many restrictions on target functions. For example, restricting your result with a restricting function (e.g., a linear equation) or by a simple binary algorithm (like the true/false choices in the above tree) will often result in bias. Variance error refers to how much a result will change based on changes to the training set. Decision trees have high variance, which means that tiny changes in the training data have the potential to cause large changes in the final results.

View All Student Assessments Grades

Figure 3: View All Student Assessments Grades

International Journal For Recent Developments in Science & Technology
2581-4575
Crossref
A Peer Reviewed Research Journal

View All Remote Users

Figure 4: View All Remote Users

View All Student Performance by Line Chart

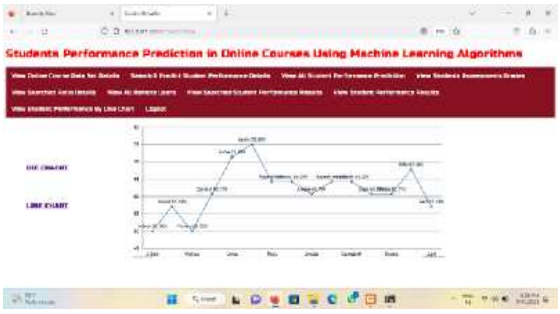
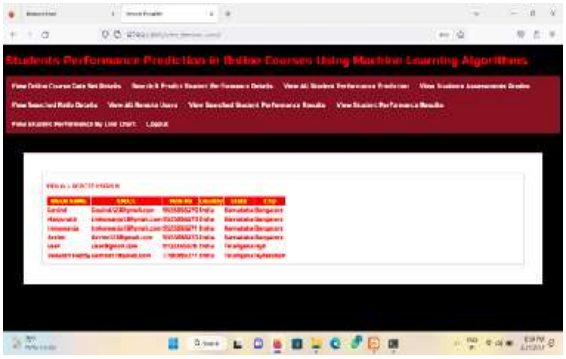


Figure 5: View All Student Performance by Line Chart

CONCLUSION

Two sets of exterminates have been carried out in this study using regression and classification analysis. The results of predicting students' assessments grades model show that the students' performance in a particular assignment relies on students' mark in the previous assignment within single Courses. The researchers conclude that students' prior grade point average (GPA) with a low mark is considered as a significant factor of withdrawal from the next course in the traditional classroom setting, both conventional classroom setting, and virtual class share similar characteristic in term of the effective of pervious performance into student learning achievement in the future. The final student performance predictive



model revealed that student engagement with digital material has a significant impact on their success in the entire course. The findings' results also demonstrate that long-term students' performance achieves better accuracy than students' assessments grades prediction model, due to the exclusion of temporal features in regression analysis. The date of student deregistration from the course is a valuable predictor that is significantly correlated with student performance. With the regression analysis, the data does not provide the last date of students' activity prior to undertaken assessments. The findings' results have been recommended to consider the temporal features on predicting of subsequent assessments grades.

FUTURE ENHANCEMENT

Future research direction involves the use of temporal features for predicting students' assessments grades model. With temporal feature time series analysis will

be untaken, might be more advanced machine leering will be utilized.

REFERENCES

[1] P. and K. Al-Shabandar, R., Hussain, A.J., Liatsis, "Detecting At-Risk Students with Early Interventions Using Machine Learning Techniques," IEEE Access, vol. 7, pp. 149464–149478, 2019.

[2] R. Alshabandar, A. Hussain, R. Keight, A. Laws, and T. Baker, "The Application of Gaussian Mixture Models for the Identification of At-Risk Learners in Massive Open Online Courses," in 2018 IEEE Congress on Evolutionary Computation, CEC 2018 - Proceedings, 2018.

[3] W. Xing and D. Du, "Dropout Prediction in MOOCs:Using Deep Learning for Personalized Intervention,"*J. Educ. Comput. Res. p.0735633118757015.*, 2018.

[4] H. B. Shapiro, C. H. Lee, N. E. Wyman Roth, K. Li, M. Çetinkaya-Rundel, and D. A. Canelas, "Understanding the massive open online course (MOOC) student experience: An examination of attitudes, motivations, and barriers," Comput. Educ., vol. 110, pp. 35–50, 2017.

[5] J.-L. Hung, M. C. Wang, S. Wang, M. Abdelrasoul, Y. Li, and W. He, "Identifying At-Risk Students for Early Interventions—A Time-Series Clustering Approach," IEEE Trans. Emerg. Top. Comput., vol. 5, no. 1, pp. 45–55, 2017.

[6] Z. Wang, T. Anderson, L. Chen, and E. Barbera, "Interaction pattern analysis in cMOOCs based on the connectivist interaction and engagement framework,"Br. J. Educ. Technol., vol. 48, no. 2, pp. 683–699,2017.

[7] J. Renz, F. Schwerer, and C. Meinel, "openSAP: Evaluating xMOOC Usage and Challenges for Scalable and Open Enterprise Education.," Int. J. Adv.Corp. Learn., vol. 9, no. 2, pp. 34–39, 2016.

[8] S. Li, Q. Tang, and Y. Zhang, "A Case Study on Learning Difficulties and Corresponding Supports for Learning in cMOOCs| Une étude de cas sur les difficultés d'apprentissage et le soutien correspondant pour l'apprentissage dansles cMOOC," Can. J. Learn.Technol. Rev. Can. l'apprentissage la Technol., vol.42, no. 2, 2016.

[9] Vinay, R. & Soujanya, K.L.S. & Singh, P.. (2019). Disease prediction by using deep learning based on patient treatment history. International Journal of Recent Technology and Engineering. 7. 745-754.

[10] C. M. Latha, S. Bhuvaneswari and K. L. S. Soujanya, "Stock Price Prediction using HFTSF Algorithm," *2022 Sixth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-*

SMAC), Dharan, Nepal, 2022, pp. 1053-1059, doi: 10.1109/I-SMAC55078.2022.9987378.

[11] Shareef, S.K., Sridevi, R., Raju, V.R., Rao, K.S.S., 2022, An Intelligent Secure Monitoring Phase in Blockchain Framework for Large Transaction, International Journal of Electrical and Electronics Research, 10.37391/IJEER.100322

[12] Patel, P., Sivaiah, B., Patel, R., 2022, Approaches for finding Optimal Number of Clusters using K-Means and Agglomerative Hierarchical Clustering Techniques, 2022 International Conference on Intelligent Controller and Computing for Smart Power, ICICCSP 2022, 10.1109/ICICCSP53532.2022.9862439

[13] Rao, G.S., Kalyan, C.N.S., Kumar, C.V.V., Goud, B.S., Kumar, M.K., Reddy, C.R., 2022, Automatic Voltage Regulator Using Global Optimization Algorithms Based on Traditional Controller, 2022 International Conference on Intelligent Controller and Computing for Smart Power, ICICCSP 2022, 10.1109/ICICCSP53532.2022.9862470

[14] Sai Kalyan, C.H.N., Srinivasa Rao, G., Rambabu, K., Kumar, M.K., Goud, B.S., Reddy, C.R., 2022, Exhibiting the Effect of AVR Coupling on the Performance of LFC in Multi Area Hybrid Power System, 2022 3rd International Conference for Emerging Technology, INCET 2022, 10.1109/INCET54531.2022.9824930

[15] Yuan, S., Chen, Y., Ye, C., Ansari, M.D., 2022, Edge detection using nonlinear structure tensor, Nonlinear Engineering, 10.1515/nleng-2022-0038