IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Predicting Performance Of Students In Online Classes Using ML

Ms. V. Sravani (M.C.A). Rajeev Gandhi Memorial college Of Engineering and Technology, Nandyal

*Mrs. B.Swetha M.Tech, (Ph.D). Rajeev Gandhi Memorial college Of Engineering and Technology, Nandyal

Abstract

This study suggests a concept called Augmented Education (Augmented) to initially address the issues raised above. This model primarily comprises of the following three modules, as indicated in Fig. 2: Three modules are included: (1) a Data Module that combines data from multiple sources on campus, (2) a Prediction Module that uses machine learning (ML)-based algorithms to predict academic performance, and (3) a Feedback Module that delivers visualized feedback. The Data Module evaluates characteristics and features that can represent students' behavioral change from three different perspectives. In the end, AugmentED is evaluated using data from 156 actual college students. In this work, a model for the proposed system called Augmented Education (AugmentED) is put forth in an effort to initially address the issues raised above. This model primarily comprises of the following three modules of the proposed system.

1. INTRODUCTION

1.1 Introduction

Academic performance prediction is a crucial problem in the field of education data mining since it is a crucial step toward obtaining personalized education. The following elements can have a significant impact on academic success, as has been well demonstrated: Personality traits of students (such as agreeableness, extraversion, and neuroticism). Gender, age, height, weight, physical fitness, cardiopulmonary fitness, aerobic fitness, stress, mood, mental health, IQ, and executive functions are a few examples of personal status. Lifestyle behaviors (such as eating, exercising, sleeping, socializing, and time management) and learning behaviors (such as class attendance, study time, library visits, and online learning) are both examples of lifestyle behaviors. examined the Big Five personality traits' incremental validity, for instance, in predicting college GPA. demonstrated that obesity status in girls and boys could be significant determinants of academic success. Meanwhile, research indicates that college students can perform well when they lead a normal lifestyle. be closely tied to academic achievement. Additionally, the results indicated that low-achieving students were less emotionally invested throughout the semester than high- and medium-achieving students and tended to report more perplexity in the last weeks of the semester. Many systems employing data to predict academic success have been developed in the literature by examining the impact of the elements affecting academic performance. For instance, academic achievement was predicted using passive sensing data and student smartphone self-reports. To forecast students' academic success, a multifunctional prediction framework that incorporates student similarity and intertester and InterMoor correlations was created. The academic achievement of students enrolled in a blended learning course was predicted using data on homework submission. Early feedback and interventions for at-risk students could be tailored to their expected academic achievement. For instance, basic interventions are designed based on GPA forecasts in order to assist students with low GPAs. The research on feedback and intervention is still in its early stages, and there have been very few successes. demonstrated the level of effort put in while working.

2. Literature Survey

• **[1] predicting student performance:**things the working group's specialists were already aware of. We tried several search phrases using this corpus of well-known publications by looking at the following three indexes: (1) Scopus, (2) IEEE, and (3) ACM. In order to get started, we used the search phrases from the prior surveys (see Section 2). Results in a nutshell [1]. The numerous data mining strategies used in various student performance prediction models were closely examined by Shahiria A.M. et al. They examined the effectiveness of using these forecasting algorithms to locate the most crucial data points in the database of pupils [2]. In this analysis, they noted that the internal assessment and cumulative grade point average (CGPA) were the two criteria that the majority of researchers considered when creating their models. An analysis of the predictive modeling method for A. Mat et al.'s monitoring of students' academic performance places particular emphasis on the numerous learning activities used in both the development and deployment of the prediction models.

[2] A Case Study on Learning Difficulties and Corresponding Supports for Learning in cMOOCs:cMOOCs, which are founded on connectivist learning theory, present learners with difficulties as well as chances for introspection. However, these studies do not offer any in-depth, empirical explorations of student issues or assistance techniques. Previous research has suggested that learners in cMOOCs may experience learning challenges. raising the standard of academic instruction; helping students while they study; and giving tutors more options when instructing their students. There have been numerous works published recently that deal with this topic. The case study on student challenges and support needs at the start of a cMOOC is presented in this paper. Messages made by students and professors in four major online learning environments, including Moodle, blogs, Facebook, and Twitter, were examined for their content. conducted. In this work, three queries are investigated: Results: (1) What sorts of challenges do students face when they first enroll in a cMOOC? (2) Which of these challenges apply to most students? (3) How are these challenges addressed and supported in the cMOOC environment? In the final section of the paper, we offer some observations on learning assistance for cMOOCs based on the research findings of this study, as well as a discussion of the research itself.

3. OVERVIEW OF THESYSTEM

3.1 Existing System

Nonlinear metrics have been utilized to identify nonlinear behavioral patterns in the students' behavioral time series. We use entropy as an illustration. It was shown that a low entropy score typically translates into high regularity and strong academic performance. Entropy is recommended to quantify the regularity/orderliness of students' behaviors. Another illustration is entropy determined using a Hidden Markov Model (HMM) analysis, which is referred to in our paper as HMM-based entropy for simplicity.

3.1.1 Disadvantages of Existing System

1.In the existing work, the system does not read Smart Card data.

2.This system is less performance due to lack of

Prediction Algorithms.

3.2 Proposed System

In this work, a model for the suggested system called Augmented Education (Augmented) is put forth in order to initially address the issues raised above. This model primarily comprises of the following three modules of the proposed system: Three modules are available: Data, Prediction, and Feedback. Data gathers and fuses campus-wide multisource data from a wide range of data trails, while Prediction and Feedback consider academic performance prediction to be a classification problem that can be solved by machine learning (ML)-based algorithms. Feedback Module delivers visualized feedback to each student individually. Finally, a real-world dataset of 156 college students is used to examine Augmented.

3.3 Methodology

In this project work, I used five modules and each module has own functions, such as:

1 PREDICTION MODULE

The main task of this module is to select features and use these features to train the prediction algorithm.

2. FEATURE SELECTION

In our study, 708 different types of features are extracted, including 510 linear features, 119 nonlinear features, 50 LSTM-based features, and 29 basic features (including e.g. frequency and duration, gender, age, and grade). For instance, because multiple behaviors are involved in our study, there are 20 DFA related features in total to quantify long-range correlation for each behavior individually (e.g. library entry). The distributions of the evaluated features and GPA are spread in different value scopes. Therefore, to eliminate a potential effect on the correlation analysis, both the features and GPA are normalized by min-max normalization. Additionally, to improve the performance of the prediction algorithms, the top 130 features with the most significant effect on academic performance are selected by the SelectKBest function in a python library named scikit-learn.

3. PREDICTION ALGORITHM

Subsequently, the selected features are used to train the MLbased classification algorithm for the academic performance prediction. Specifically, in our study, five ML algorithms are

applied, including RF (random forest), GBRT (gradient boost regression tree), KNN (k-nearest neighbor), SVM, and XGBoost (extreme gradient boosting). The hyper parameters of the ML and LSTM algorithms are optimized by GridSearchCV in scikit-learn.

4. CROSS VALIDATION

Our dataset is divided into a training set and a test set at the ratio of 7:3. The classification algorithm is first trained and then applied to the test set to predict academic performance.

Finally, the robustness of the algorithm is tested by 10-fold cross validation.

5. VISUALIZATION MODULE

The main task of this module is to provide personalized feedback, including GPA prediction and a visualized summary of the students' behavioral patterns.

.

4 Architecture

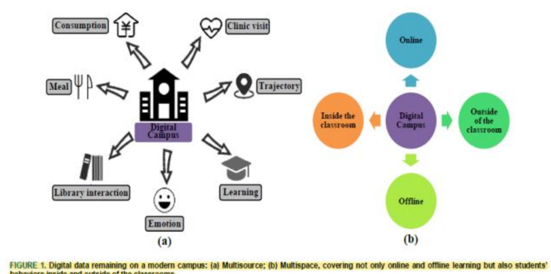
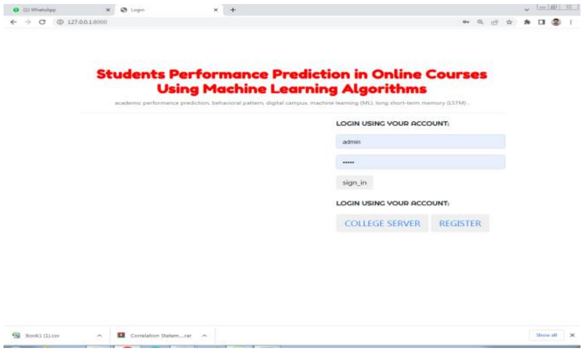


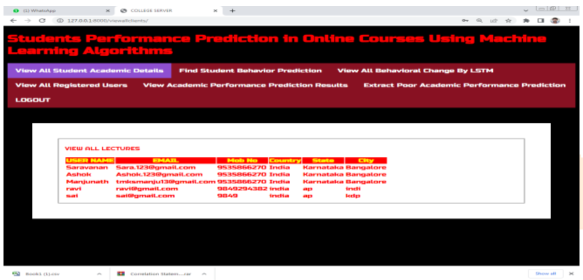
Fig 1: Frame work of proposed method

5 RESULTSSCREEN SHOTS

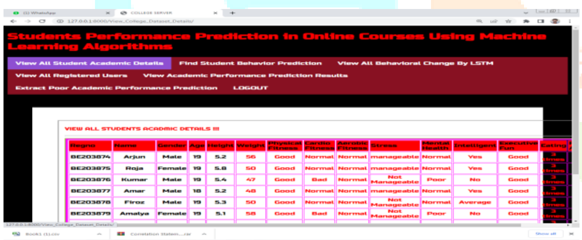
Home Page:



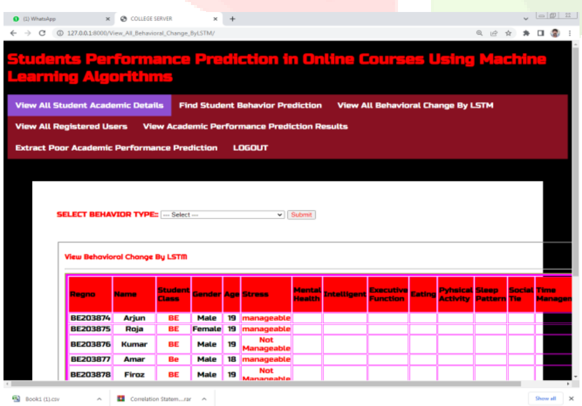
Upload Data:



View Students details:



Predict Result:



7. CONCLUSION

✓ Academic performance prediction has drawn a lot of attention from scholars since it is a crucial topic in the field of education data mining. However, there are still many difficulties with prediction accuracy and interpretability because to the lack of richness and diversity in both data sources and features. Our study aims to get a thorough understanding of student behavioral patterns in order to enable students better understand how to interact with the university. This will help to initially relieve the problem.In our work, a model called AugmentED is put out to forecast college students' academic achievement. Our contributions to this work come from three different places. To our knowledge, our effort is the first to acquire, analyze, and utilise many sources of data for data fusion. Data on campus life behaviors both inside and outside of the classroom as well as online and offline learning are covered for the purpose of predicting academic performance. One can create a rich profile of a pupil using these multisource facts. Second, when it comes to the feature evaluation, the approaches used to assess behavioral change—linear, nonlinear, and deep learning (LSTM)—provide a systematic look at students' behavioral patterns. In particular, it is the first time that the behavioral time series analysis of students has used three innovative nonlinear metrics (LyE, HurstE, and DFA) and LSTM. Third, the findings of our experiments show that AugmentED can predict academic achievement with a high degree of accuracy, which may be used to create individualized feedback for students who are at danger of failing or who lack self-discipline.

Future Enhancement

✓ However, there are certain restrictions with our study as well. By limiting the dataset to only include student-generated data from a single course, we were able to obtain a multisource dataset. This restriction might have a detrimental effect on AugmentED's ability to be used more widely. Additionally, behavioral modification is the main emphasis of this study. This study did not analyze other meritorious attributes or features, such as peer effect or sleep. Finally, our study is based on a comprehensive passive daily data collection system that is present at the majority of contemporary universities. This system might result in ongoing, more extensive inquiries. The information obtained from this study may also help with similar studies involving K–12 pupils.

8. References

[1] WHO Pneumonia. World Health Organization. (2019), https://www.who.int/news-room/fact-sheets/ detail/pneumonia

[2] Neuman M., Lee E., Bixby S., Diperna S., Hellinger J., Markowitz R., et al. Variability in the interpretation of chest radiographs for the diagnosis of pneumonia in children. Journal of Hospital Medicine. 7, 294–298 (2012) https://doi.org/10.1002/jhm.955 PMID: 22009855

[3] Williams G., Macaskill P., Kerr M., Fitzgerald D., Isaacs D., Codarini M., et al. Variability and accuracy in interpretation of consolidation on chest radiography for diagnosing pneumonia in children under 5 years of age. Pediatric Pulmonology. 48, 1195–1200 (2013) https://doi.org/10.1002/ppul.22806 PMID: 23997040

[4] Kermany D., Zhang K. & Goldbaum M. Labeled Optical Coherence Tomography (OCT) and Chest X-ray Images for Classification. (Mendeley, 2018)

[5] Lal S., Rehman S., Shah J., Meraj T., Rauf H., Damasˇevičius R., et al. Adversarial Attack and Defence through Adversarial Training and Feature Fusion for Diabetic Retinopathy Recognition. w

[6] Rauf H., Lali M., Khan M., Kadry S., Alolaiyan H., Razaq A., et al. Time series forecasting of COVID-19 transmission in Asia Pacific countries using deep neural networks. Personal and Ubiquitous Computing. pp. 1–18 (2021) https://doi.org/10.1007/s00779-020-01494-0 PMID: 33456433

[7] Deng J., Dong W., Socher R., Li L., Li K. & Fei-Fei, L. Imagenet: A large-scale hierarchical image database. 2009 IEEE Conference on Computer Vision and Pattern Recognition. pp. 248-255 (2009)

[8] Dalhoumi S., Dray G., Montmain J., Derosière, G. & Perrey S. An adaptive accuracy-weighted ensemble for inter-subjects classification in brain-computer interfacing. 2015 7th International IEEE/EMBS Conference on Neural Engineering (NER). pp. 126-129 (2015)

[9] Albahli S., Rauf H., Algosaibi A. & Balas V. AI-driven deep CNN approach for multi-label pathology classification using chest X-Rays. PeerJ Computer Science. 7 pp. e495 (2021) https://doi.org/10.7717/ peerj-cs.495 PMID: 33977135.