Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 172 (2020) 235–239

Procedia Computer Science

www.elsevier.com/locate/procedia

9th World Engineering Education Forum 2019, WEEF 2019

Subjective Areas of Improvement: A Personalized Recommendation

Suganya G[a], Premalatha M[b], Piyush Dubey[c], Aryan Raj Drolia[c], Srihari S[c]

[a] *Associate Professor, Vellore Institute of Technology, Chennai, India*
[b] *Assistant Professor(SG), Vellore Institute of Technology, Chennai, India*
[c] *UG students, Vellore Institute of Technology, Chennai, India*

Abstract

Gen Z-ers, being independent, self-confident and autonomous as their key characteristics are proven to be technologically more advanced than their previous generations. To reach the milestone of understanding and satisfying the Gen-Z community, and to make use of the digital adherence of this generation, a personalized recommender system using machine learning is designed that can recommend the areas which the students should strengthen themselves to lead and be distinct among their peers. The proposed work uses Support Vector Machine to predict the area of improvement which the student need to focus on. The proposed system will also recommend the list of online courses and materials that the users can make use of to strengthen themselves. The model is built using Python flask and Jupyter notebook and is tested using one public dataset and a private dataset. The results are convincing and accurate enough in identifying the proper areas that require improvement.

Keywords: E-Learning; Support Vector Machine; Improvement strategies;

1. Introduction

The Common Sense Census, 2019 presents a clear picture of the evolution of media use among young people over time. Also, Gen-Z students don't want their weaknesses to be open to anyone including their parents and expects personalization in all their activities. Various researchers have contributed to personalization of job profiles, understanding level of learning, evaluating student performance etc., the age of innovations and applications. The amount of applications which has been developing and evolving has grown exponentially especially with the emergence of the smartphone generation. In this generation where applications have been a quintessential part there is a need and opportunity for this to expand and enhance in the education domain. This is because even though applications and ideas like smart classes etc., have been effective to an extent accessibility is still a question there.

10.1016/j.procs.2020.05.037

There have been a plethora of teaching methods which have been employed physically and virtually over the years to help out students to have their customized way for studying. Although these methods were employed these sometimes still doesn't prove that effective as an attempt to reach a broader appeal. So there is a need for an efficient, accurate and accountable application for this purpose.

Also personalization is a huge issue. These days personalization and customization works hugely thanks largely due to the impact of the data analytics and data science we have today. We can look deeper into to interest and needs of every individual with few amount of data extracted from them. Personalization is deeply rooted in the fact it is so successful among customers. The proposed work is an idea to solve the above problems and this proposes a personalized and customized learning application using machine learning methods. The data about performance of various students with various backgrounds at various levels will be analysed to give desired results. The given sections (II) background study discusses the various research works and motivation behind the work.(III) discusses the proposed architecture for the given system.(IV) discusses the various results for the given system.(V) discusses the conclusion for the system.

2. *Background*

Various papers were surveyed for the purpose of extracting information about how the working of student performance predictors and to understand about the research works on personalized student applications.

[1] C. L. Sa et. al. presented a predictive system that is able to predict the students' performance in course using classification systems. The performance in upcoming subject is predicted using previous semester performance through the rules generated via data mining technique. The method utilized in this venture is characterization, which groups the understudies dependent on understudies' evaluation. C.Kiu[2] used impact of student background, social activities and their coursework attainment for forseeing their academic performance. Different data mining algorithms like Naïve Bayes, Perceptron, Decision Tree and further Random Forest were used in predicting the performance in secondary school level. The prediction was performed on 2-level classification and 5-level classification and the results were recorded to be satisfactory.

Yaakub et. al. [3] investigated the factors that affect the performance of students in mathematics during their secondary education level to student's academic performance in electrical engineering study. Cumulative Grade Point Average (CGPA) is considered as the metric for engineering education and the findings showed that students from urban area achieved better compared to the students from rural area. Meanwhile, no strong relation between student's performances in engineering studies with school level is witnessed.

Preetha et. Al. [4] presented a REA framework for evaluating the comprehensive, problem solving, designing models, programming and application developing skills for all disciplines of engineering. This framework evolves in different modes like general questionnaire, questionnaires with tip-off mode, enhanced graphical mode and crossword mode to evaluate various aspects of engineering skills. The results show a considerable increase in performance of the students.

The impact of using programming platform by the students on course outcome is assessed by various researchers. By taking the snapshot of programming, the authors have predicted the skill level of the student. Vihavainen[5] does prediction of students performance using educational data mining. Experiments were done on batches/groups of engineering students to evaluate their comparative performance in terms of traditional mode of writing exams and the proposed framework. The authors have shown an improved performance of about 78% in engineering students.

Lee et. al.[6] predicted the job performance of a person using abilities of students in completing the task given by lecturer. The model predicts the future job performance of the students based on criteria like scores of AQ, Social, and Entrepreneur. Tan[7] suggest the data of students' examination score to verify whether there is a significant relationship between the students' performance in mathematical subjects and their choice of course. Yildiz et. al. [8] proposed the use of fuzzy decision support system and proposes a new approach called `refinement process' on student grades. The model was based on fuzzy multi criteria method and is developed for evaluating student performances in laboratory activities that combine theoretical knowledge with practical applications. The authors claim that fuzzy system can provide a better evaluation system than traditional systems. Azrilah [9] predicts the

student performance based on their mathematical scores. Rasch measurement model converts the results from the test into ration type data and sorts the ability according to Rasch-Guttman matrix.

Nyugen et. al. [10] proposes an approach that uses recommender system techniques for educational data mining to predict student performance. The results were compared with regression techniques and were presented satisfactory. Various researchers have presented methodologies to predict student performance based on parameters like, their place of living, performance in previous school etc., But, personalized recommendation based on the results is found to be lacking. Considering the requirement of this generation students, we propose a method to predict and hence recommend corresponding training material to students.

3. Proposed System & Implementation

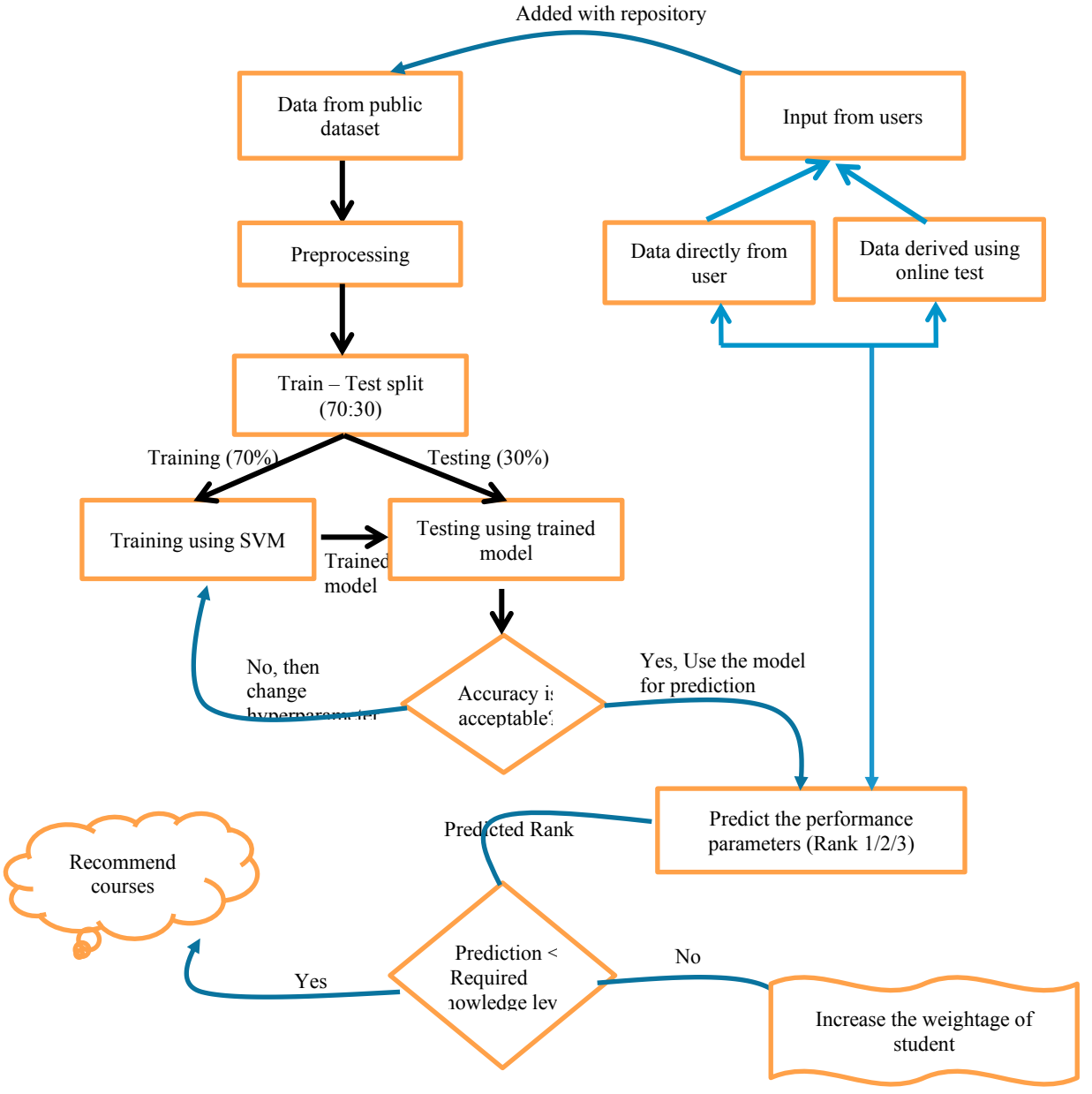

Fig. 1. Proposed Architecture

Personalized recommendations have been an essential part of today's generation and these recommendations are based on data, and more data in practice means improved recommendations. As the users of the system increases, the proposed methodology assumes the acquisition of data about the target group under study by storing the input details as part of the training dataset. The proposed methodology uses a two level mechanism to assess, predict and hence recommend different types of online materials to the users. The framework works in two stages: At the first stage, the methodology uses a public dataset collected from different students storing various types of marks. The dataset is then added with the inputs of all the new users registering with the system. In the second stage, appropriate recommendations are given to the user based on his/her performance given as input. Fig. 1 presents the overall architecture of the system.

Stage 1:

Initially, the dataset is formed by storing of all students' data including their personal information, marks of various subjects and their rank under study. Three levels of ranks are stored as output class label (1 – Good, 2 – Average, 3 – Poor). The dataset is then checked for Null values and if found, are filled with zeros. Normalization of columns is then performed to maintain uniformity among the range of values each attribute can take.

Fig. 2 Comparitive Performance Prediction

The normalized data is further split into training and testing data having 70% as the training proportion. Randomness is ensured while splitting the data to avoid bias towards any specific group. The machine learning model using Support Vector Machine (SVM) as the classifier is then trained using the training dataset. SVM is chosen as the training model due to its adherence to support non-linearity of data. The model is tested for accuracy using the test data got from the actual initial dataset. The model is fine tuned to ensure attainment of required accuracy level by tuning its hyper parameters. This training is performed to ensure a comparative performance tracking of new users with their common peer groups. Fig.2 presents the strategy of comparative performance calculation.

The users of the system are presented with the option to either enter the marks in various subjects directly of they can get themselves assessed using a Game solver as part of the system. The game solver takes the users serious of tests for their personal assessment. The results are then given as input to the system.

Stage 2:

It is really not worth to highlight someone as Good performer, average performer or as a poor performer; rather a mechanism to make the user improve to meet the talents of peers is necessary. As a follow-up, based on the prediction of the user in each subject, a web crawler is initiated to crawl the web. Appropriate materials corresponding to the level and subject of improvement will be recommended to the user. As part of continuous improvement process, users classified as "Good" are suggested with high level contents with more quizzes and assignments.

Implementation:

The entire framework was built using Python flask as the User Interface tool, jupyter notebook as the platform for implementing machine learning model. The gaming interface is developed using open source tools and the web service is written using Java.

4. Results and Discussion

Initially, the framework was tested with the dataset downloaded from the public repository and is cross-validated using 5-fold. To assess the real usage of the built framework, it was floated to a class of 60 students. The domains of subjects the students have in the current semester are chosen to be the attributes of the model. The dataset is formed using the performance of the seniors classifying them into three ranks (G/A/P). The system is then trained and tested using the current set of students.

The accuracy of the model is assessed by individual feedback of the students towards the system under study. Good rating(10/10) was given to proper recommendations whereas low rating(2/10) was given to moderate assessment with the ratings in-between. The accuracy is then calculated using mean analysis and is found to be satisfactory.

5. Conclusion and Future Works

Personalized recommendation without the necessity for peers interaction is the requirement of Gen-Z. The proposed framework presents a satisfactory model to predict the student performance through machine learning techniques. The model is useful to the users who needs self-assessment and further directions for improvement.

The model is a prototype and hence could be extended to include any type of data, model and type of recommendation (text, audio or Video material) to the user. In short, online personalized recommendation helps the student to improve their performance without the need for any kind of human mentor.

References

[1] C. L. Sa, D. H. b. Abang Ibrahim, E. Dahliana Hossain and M. bin Hossin, "Student performance analysis system (SPAS)," *The 5th International Conference on Information and Communication Technology for The Muslim World (ICT4M)*, Kuching, 2014, pp. 1-6. doi: 10.1109/ICT4M.2014.702066

[2] C. Kiu, "Data Mining Analysis on Student's Academic Performance through Exploration of Student's Background and Social Activities," *2018 Fourth International Conference on Advances in Computing, Communication & Automation (ICACCA)*, Subang Jaya, Malaysia, 2018, pp. 1-5. DOI: 10.1109/ICACCAF.2018.8776809

[3] *T. N. T. Yaakub, W. R. W. Ahmad, Y. Husaini and N. Burham, "*Influence Factors in Academic Performance among Electronics Engineering Student: Geographic Background, Mathematics Grade and Psycographic Characteristics*,"* 2018 IEEE 10th International Conference on Engineering Education (ICEED)*, Kuala Lumpur, Malaysia, 2018, pp. 30-33.* DOI: 10.1109/ICEED.2018.8626963

[4] P. K.G., S. S. and S. Asokan, "Enhanced Performance of Engineering Students through REA: A Comparative Analysis," *2012 International Conference on Advances in Computing and Communications*, Cochin, Kerala, 2012, pp. 178-181. DOI: 10.1109/ICACC.2012.41

[5] A. Vihavainen, "Predicting Students' Performance in an Introductory Programming Course Using Data from Students' Own Programming Process," *2013 IEEE 13th International Conference on Advanced Learning Technologies*, Beijing, 2013, pp. 498-499. doi: 10.1109/ICALT.2013.161

[6] M. F. Lee ; N.F. Mat Nawi ; C.S. Lai (2017), "Engineering Students' Job Performance Prediction Model based on Adversity Quotient & Career Interest", 2017 7th World Engineering Education Forum (WEEF), Vol. 1, Issue 1, Pp. 2-5. DOI 10.1109/WEEF.2017.8467175

[7] Mazwin Tan, Halina Hassan, Nurul Na'imy (2012), "Statistical analysis on students' performance and their preference of course in bachelor of Engineering Technology programmes in Malaysian Spanish Institute, Universiti Kuala Lumpur: Students performance versus programme preferences", Wan Journal - 2012 International Conference on Statistics in Science, Business and Engineering (ICSSBE), Vol. 1, Issue 1, Pp. 4. DOI 10.1109/ICSSBE.2012.6396625

[8] Z. Yıldız and A. F. Baba, "Evaluation of student performance in laboratory applications using fuzzy decision support system model," 2014 IEEE Global Engineering Education Conference (EDUCON), Istanbul, 2014, pp. 1023-1027. DOI: 10.1109/EDUCON.2014.6826230

[9] Azrilah Abdul Aziz, Azami Zaharim, Norain Farhana Ahmad Fuaad, Zulkifli Mohd Nopiah 2013, "Students' performance on engineering mathematics: Applying rasch measurement model", Information Technology Based Higher Education and Training (ITHET), Vol. 1, Issue 1, Pp. 4. DOI 10.1109/ITHET.2013.6671040

[10] NguyenThai-Nghe, LucasDrumond, Artus Krohn-Grimberghe, LarsSchmidt-Thieme 2010, "Recommender system for predicting student performance", 1st Workshop on Recommender Systems for Technology Enhanced Learning, Vol.1, Issue 2, pp. 2811 – 2819.