Student Academic Performance Expectation Based on Multisource and Multi feature Communicative Data using machine learning

Dr.CHEN CHOU

[1.] **Professor**, Computer Science and Engineering Department, JNTUH CEJ, KRMR Dist.

[2.] M.Tech Student,Computer Science and Engineering Department, JNTUH College of Engineering Jagtial.

Abstract:

Digital data trails from disparate sources covering different aspects of student life are stored daily in most modern university campuses. However, it remains challenging to (i) combine these data to obtain a holistic view of a student, (ii) use these data to accurately predict academic performance, and (iii) use such predictions to promote positive student engagement with the university. To initially alleviate this problem, in this paper, a model named Augmented Education is proposed. In our study, first, an experiment is conducted based on a real-world campus dataset of college students that aggregates multisource behavioral data covering not only online and offline learning but also behaviors inside and outside of the classroom. Features representing dynamic changes in temporal lifestyle patterns are extracted by the means of long short-term memory (LSTM). Second, machine learning-based classification algorithms are developed to predict academic performance. Finally, visualized feedback enabling students to potentially optimize their interactions with the university and achieve a study-life balance is designed.

.

I.INTRODUCTION

As an important step to achieving personalized education, academic performance prediction is a key issue in the education data mining field. It has been extensively demonstrated that academic performance can be profoundly affected by the following factors:

Students' Personality

Personal Status (e.g., gender, age, height, weight, physical fitness, fitness, stress, intelligence, and executive function)

Lifestyle Behaviors (e.g., eating, physical activity, sleep patterns, social tie, and time management)

Learning Behaviors (e.g., class attendance).For example, investigated the incremental validity of the Big Five personality traits in predicting college GPA. Demonstrated that physical fitness in boys and obesity status in girls could be important factors related to academic achievement. Meanwhile, showed that a regular lifestyle could lead to good performance among college students. Showed that the degree of effort exerted while working could be strongly correlated with academic performance. Additionally, showed that compared with high- and medium-achieving students, low-achieving students were less emotionally engaged throughout the semester and tended to express more confusion during the final stage of the semester. By analyzing the effect of the factors influencing academic performance, many systems using data to predict academic performance have been developed in the literature. For instance, in, academic performance was predicted based on passive sensing data and self-reports from students' smart phones. In, a multitask predictive framework that captures intersemester and intermajor correlations and integrates student similarity was built to predict students' academic performance. In based on homework submission data, the academic performance of students enrolled in a blended learning course was predicted. According to their predicted academic performance, early feedbacks and interventions could be individually applied to at-risk students. For example, in, to help students with a low GPA, basic interventions are defined based on GPA predictions. However, the research on the feedback/intervention is still in the early stage, its achievements are relatively few. In recent years, compared with primary and secondary education, more and more attentions have been paid to the academic performance prediction for higher education. The reasons contributing to this phenomenon warrant further investigation and might include the following. First, for college students on a modern campus, life involves a combination of studying, eating, exercising, socializing, etc. All activities that students engage in (e.g., borrowing a book from the library) leave a digital trail in some database. Therefore, it is relatively easy to track college students' behaviors, e.g. online learning behaviors captured from massive open online courses (MOOC) and small private online courses (SPOC) platforms. Second, given the diverse range of activities listed above, it could be difficult for college students to maintain balanced, self-discipline, well-being university experiences, including excellent academic performance. Although many academic performance prediction systems have been developed for college students, the following challenges persist: (i) capturing a sufficiently rich profile of a student and integrating these data to obtain a holistic view; (ii) exploring the factors affecting students' academic performance and using this information to develop a robust prediction model with high accuracy; and (iii) taking advantage of the prediction model to deliver personalized services that potentially enable students to drive behavioral change and optimize their study-life balance. To address these challenges, four prediction systems (including one online system and three offline systems) are summarized in Table I. We first discuss the online prediction system,System A (proposed by Z. Liu). This system is relatively simple because its data is only captured from either SPOC or MOOC. Regarding the latter three offline prediction systems, i.e., Systems B ~ D (proposed by R. Wang, Y. Cao, and Z. Wang respectively), the number of data sources is reduced, while the corresponding scale size rapidly increases; Unfortunately, the number of different types of behaviors that could be considered is decreased. Ideally, multisource data at a medium/large scale could help lead to a better prediction system design. However, in practice, due to limitations, such as computing capability, either data diversity or the sample size is sacrificed during the system design process.

III. EXISTING SYSTEM:

- Random forest, Gradient Boosting Classifier, Support Vector Machine and the k-nearest neighbor algorithm, where 77% / 79.7% of accuracy is maximum for the above used algorithms. Datasets are available in. Kaggle.com.

IV. PROPOSED SYSTEM:

- I will analyze recent real-world campus dataset of college students that aggregates multisource behavioral data covering only inside and outside of the classroom. Features representing dynamic changes in temporal lifestyle patterns are extracted by the means of long short-term memory (LSTM).

V. SYSTEM ARCHITECTURE:

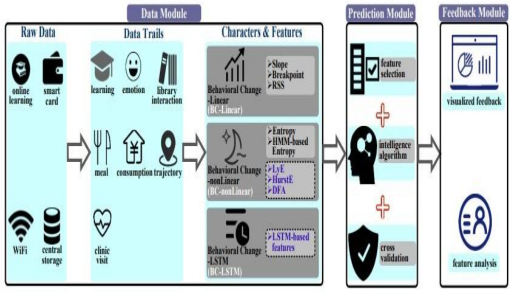


FIGURE 2. Overview of AugmentED. In the data module, the features blocked in dashed boxes (including LyE, HurstE, DFA, and LSTM-based features) are proposed in our study, to the best of our knowledge, which is used for the first time in student's behavioral analysis.

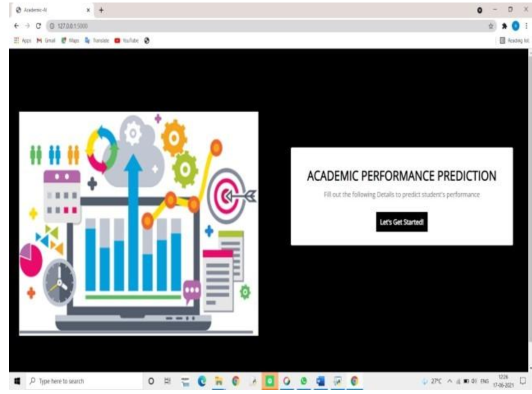
VI. EXPERIMENTS AND RESULT:

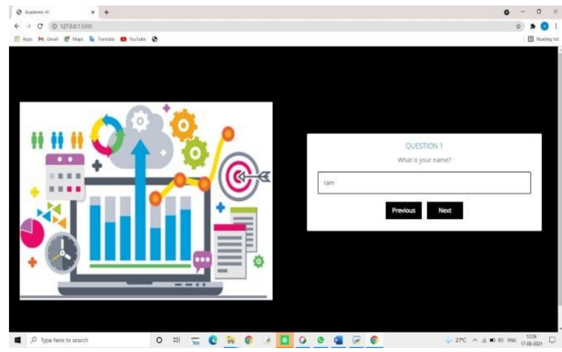
LSTM:

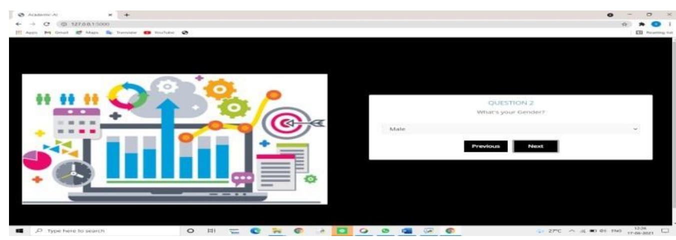
LSTM is a kind of deep neural network based on RNN. The core of LSTM is to add a special unit (memory module) to learn the current information and to extract the related information and rules between the data, so as to transfer the information. LSTM is more suitable for deep neural network calculation because of memory module to slow down information loss. Each memory module has three gates, including input gate (it), forget gate (ft), and output gate (ot). They are used to selectively memorize the correction parameters of the feedback error function as the gradient decreases..

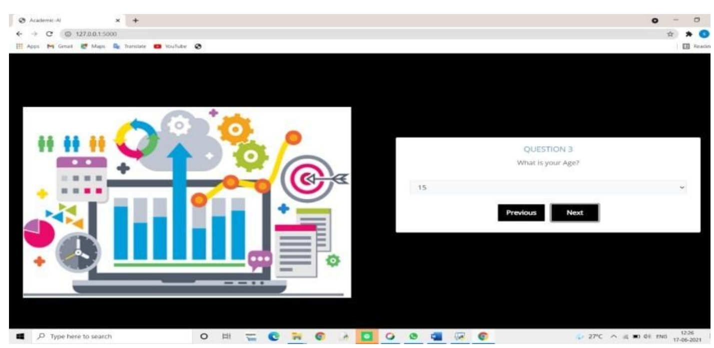
.

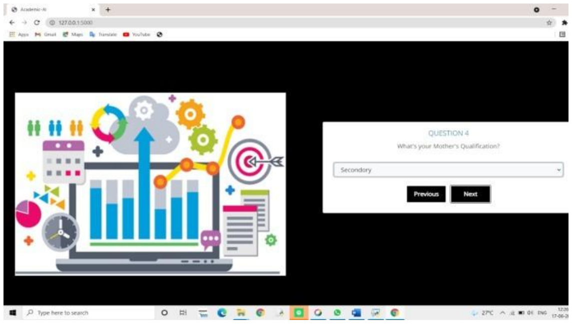
RESULT:

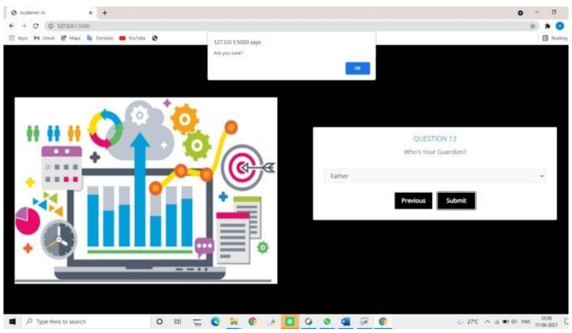


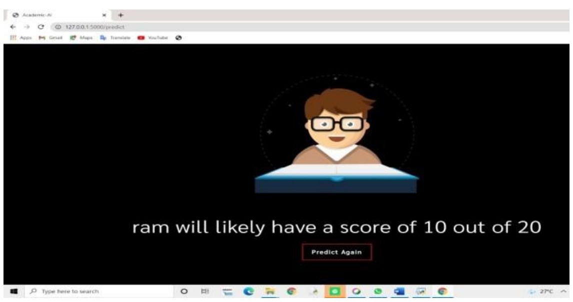












FUTUREWORK

In summary, the random forest-based architecture presented better adaptation at a cost of the training accuracy when handling the limited dataset. In order to overcome the over fitting problem and to improve the testing accuracy for multi-categories classification tasks, our future work will focus on collecting or generating more training data, integrating more information like airport traffic flow, airport visibility into our dataset, and designing more delicate network.

CONCLUSION:

As an important issue in the education data mining field, academic performance prediction has been studied by many researchers. However, due to lack of richness and diversity in both data sources and features, there still exist a lot of challenges in prediction accuracy and interpretability. This system can potentially lead to continual investigations on a larger scale. The knowledge obtained in this study can also potentially contribute to related research among students. Based architecture still needs to be solved.

.

REFERENCES:

[1] A. Furnham, and J. Monsen, “Personality traits and intelligence predict academic school grades," Learning and Individual Differences, vol. 19, no. 1, pp. 0-33, 2009.

[2] M. A. Conard, “Aptitude is not enough: How personality and behavior predict academic performance,” Journal of Research in Personality, vol. 40, no. 3, pp. 339-346, 2006.

[3] T. Chamorropremuzic, and A. Furnham, “Personality predicts academic

performance: Evidence from two longitudinal university samples," Journal of Research in Personality, vol. 37, no. 4, pp. 319- 338, 2003.

[4] R. Langford, C. P. Bonell, H. E. Jones, T. Pouliou, S. M. Murphy, and E. Waters, "The WHO health promoting s chool framework for improving the health and well-being of students and their academic achievement," Cochrane Database of Systematic Reviews, vol. 4, no.4, pp. CD008958, 2014.

[5] A. Jones, and K. Issroff, "Learning technologies: Affective and social issues in computer-supported collaborative learning," Computers & Education, vol. 44, no. 4, pp. 395-408, 2005.

[6] D. N. A. G. Van, E. Hartman, J. Smith, and C. Visscher, "Modeling relationships between physical fitness, executivefunctioning, and academic achievement in primary school children," Psychology of Sport & Exercise, vol. 15, no. 4, pp. 319-325, 2014.

[7] R. Wang, F. Chen, Z. Chen, T. Li, and A. T. Campbell, "StudentLife: Assessing mental health, academic performance andbehavioral trends of college students using smartphones," In Proc. of the ACM International Joint Conference on Pervasive & Ubiquitous Computing, Seattle, WA, USA, 2014