STUDENT ACADEMIC PERFORMANCE PREDICTION

Mr. K. Kumara Swamy
Asst.Professor
Department of Information Technology
Malla Reddy Engineering College for Women
(UGC-Autonomous)
Maisammaguda, Hyd-500100, Telangana, India.

P. Sreekala
Student
Department of Information Technology
Malla Reddy Engineering College for Women
(UGC-Autonomous)
Maisammaguda, Hyd-500100, Telangana, India.

N. Manasvi
Student
Department of Information Technology
Malla Reddy Engineering College for Women
(UGC-Autonomous)
Maisammaguda, Hyd-500100, Telangana, India.

S. Aishwarya
Student
Department of Information Technology
Malla Reddy Engineering College for Women
(UGC-Autonomous)
Maisammaguda, Hyd-500100, Telangana, India.

Abstract— On most modern university campuses, digital data trails from many sources that cover various facets of student life are kept on a regular basis. However, it is still difficult to I combine these data to get a comprehensive picture of a student, (ii) utilise these data to predict academic achievement effectively, and (iii) use such predictions to encourage positive engagement of students with the university. In this project, a model called Augmented Education (AugmentED) is suggested as a first solution to this issue. In our work, (1) the first step is to run an experiment based on a real-world campus dataset of college students (N = 156) that compiles multisource behavioural data including both inside- and outside-the-classroom behaviour as well as online and offline learning. In particular, metrics measuring the linear and nonlinear behavioural changes (e.g., regularity and stability) of campus lifestyles are estimated, and features representing dynamic changes in temporal lifestyle patterns are extracted by the use of long short-term memory in order to gain in-depth insight into the features leading to excellent or poor performance (LSTM). (2) Next, classification algorithms based on machine learning are created to forecast academic success. (3) Visualized feedback is aimed to help students (particularly at-risk students) communicate with the university more effectively and strike a balance between their personal and academic lives. The results of the studies demonstrate how well the AugmentED model can forecast students' academic achievement.

INTRODUCTION

Academic performance prediction is a crucial problem in the field of education data mining since it is a crucial step toward obtaining personalised education. The following elements can have a significant impact on academic success, as has been well demonstrated:

Student personality traits like neuroticism, extraversion, and agreeableness;

Personal status traits like gender, age, height, and weight; physical fitness; cardiopulmonary fitness; aerobic fitness; stress; mental health; intellect; and executive function traits; and students' personality traits.

Learning Behaviors (such as class attendance, study time, library visits, and online learning); Lifestyle Behaviors (such as eating, physical exercise, sleep habits, social ties, and time management);

One study, for instance, looked into how well the Big Five personality qualities predicted college GPA. Others have shown that a boy's level of physical fitness and a girl's level of obesity may be significant predictors of academic success. Others, meantime, demonstrated that college students who have normal lives perform well or that the level of effort put in while working is directly associated to academic achievement. Another estimate revealed that low-achieving students were less emotionally invested throughout the semester and tended to exhibit more confusions in the last weeks of the semester compared to high- and medium-achieving students. Many systems employing data to predict academic achievement have been developed in the literature by examining the impact of the elements affecting academic performance. For instance, academic performance was predicted using passive sensing data and self-reports from students' smart phones. A multitask predictive framework that captures intersemester and intermajor correlations and integrates student similarity was also built to predict students' academic performance. Early feedback and interventions for at-risk students could be tailored to their expected academic achievement. Basic treatments are established based on GPA forecasts 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. The prediction of academic achievement for higher education has received increasingly more attention in recent years compared to primary and secondary school. Further research is necessary to determine the causes of this phenomena, which may include the following. First, life for college students on a contemporary

campus consists of a variety of activities, such as studying, eating, exercising, and socialising. Every action that students take leaves a digital trace in a database, such as checking out a book from the library. As a result, tracking college students' actions is not too difficult. For instance, online learning behaviours can be collected through platforms for massive open online courses (MOOC) and small private online courses (SPOC). Second, it could be challenging for college students to maintain balanced, self-disciplined, and healthy university experiences, including outstanding academic performance, given the wide range of activities mentioned above. The following issues still exist despite the several academic achievement prediction systems that have been created for college students: I gathering comprehensive profiles of students; (ii) examining the variables influencing students' academic performance and using this data to create a reliable prediction model with high accuracy; and (iii) utilising the prediction model to deliver customised services that may help students change their behaviour and improve their study-life balance. First, we talk about System A, an online prediction system. This approach is rather straightforward because it simply collects data from MOOCs or SPOCs. The number of data sources is dropped while the matching scale size quickly grows for the last three offline prediction systems, or Systems B. Unfortunately, this also means that fewer diverse types of behaviours are now available for consideration. A more effective prediction system might theoretically be designed with the aid of multisource data on a medium to big scale. However, in reality, either data variety or the sample size is given up throughout the system design process due to constraints, such as processing power. We initially developed a paradigm called Augmented Education to address the issues outlined above (AugmentED). The following three modules make up the majority of this model. 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 visualised feedback. The Data Module evaluates characteristics and features that can represent students' behavioural change from three different perspectives. Finally, AugmentED is investigated using a dataset of college students from the actual world.

EXISTING SYSTEM

Nonlinear metrics have been utilised to identify nonlinear behavioural patterns in the students' behavioural time series. Entropy is used as an example because it has been suggested as a way to measure the regularity and orderliness of students' behaviour, and it has been shown that high regularity and excellent academic success are typically correlated with low entropy values. The two steps listed below are used to evaluate entropy: a behavioural time series' hidden states are extracted, and the entropy is then calculated. The following three new measures, which have not previously been used in student behavioural time series analysis, are also worthwhile to study in order to better understand students' activities and identify their nonlinear behavioural patterns. A time series' stability is gauged by the Lyapunov Exponent (LyE). The results show that a time series with a big LyE value is less stable than a series with a small LyE value, i.e., generally, a large LyE value suggests significant instability. For instance, LyE is used to evaluate the stability of a gait time series. LyE is therefore regarded as a stability risk indicator for falls in gait analyses that can separate healthy people from those who are at a greater risk of falling. A time series' predictability is measured by the Hurst Exponent (HurstE), which is also known as a time series' long-term memory in some research. The results show that a time series with a big HurstE value can be predicted more correctly than a series with a HurstE value close to 0.5, for instance, when HurstE is used to evaluate the predictability of a financial time series.

Disadvantages –

(i) The system currently in use cannot read Smart Card data.
(ii) Performs poorly because it lacks PREDICTION ALGORITHMS.

PROPOSED SYSTEM

In the proposed approach, a paradigm called Augmented Education (AugmentED) is suggested to initially ease the difficulties outlined above. This model primarily comprises of the following three modules of the proposed system: The characteristics/features that can represent students' behavioural change from three different perspectives are evaluated in the Data Module, Prediction Module, and Feedback Module. The Data Module aggregates and fuses multisource campus data covering a wide range of data trails, and the Prediction Module considers academic performance prediction to be a classification problem that is solved by machine learning (ML)-based algorithms. Finally, AugmentED is investigated using a dataset of 156 college students from the actual world.

Advantages – (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 personalize services that potentially enable students to drive behavioral change and optimize their study-life balance. An important outcome of preliminary investigation is the determination that the system request is feasible.

This is possible only if it is feasible within limited resource and time. The different feasibilities that have to be analyzed are

- Operational Feasibility
- Economic Feasibility
- Technical Feasibility

Operational Feasibility examines the potential of the system that will be created. Operationally, this approach relieves all of the administrator's stress and enables him to efficiently monitor the project's development. Automation of this nature will undoubtedly save time and energy that were previously expended on manual labour. The study's findings support the system's operational viability.

An evaluation of the financial basis for a computer-based project is known as economic feasibility or cost-benefit analysis. Since hardware was deployed from the start and for many purposes, the project's hardware cost is modest. Any number of employees connected to the LAN within that organisation can utilise this tool at any time because the system is network-based. The organization's current resources will be used to create the virtual private network. The project is therefore financially viable. According to Roger S. Pressman, technical feasibility is the evaluation of an organization's technological resources. The company requires computers that are IBM compatible and have an Intranet and Internet connection as well as a graphical web browser. The system was made for a platform-neutral environment. The system was created using Java Server Pages, JavaScript, HTML, SQL Server, and WebLogic Server. A technical feasibility study has been completed. Technically speaking, the system can be constructed and used with the facility that already exists.

SYSTEM ENVIRONMENT

PYTHON:

This programming language is interpretable, high-level, and all-purpose. Python's design philosophy places a strong emphasis on code readability and makes considerable use of whitespace. Its language constructs and object-oriented methodology are designed to aid programmers in creating clean, comprehensible code for both little and big projects. There are Python interpreters for many different operating systems. The nonprofit Python Software Foundation is in charge of running Python. Python has an autonomous memory management mechanism and a static type system. It includes a sizable and thorough standard library and supports a variety of programming paradigms, including procedural, functional, and object-oriented programming. Python is a programming language that is appealing for application development since it is simple to learn but strong and flexible.

PYTHON IDLE: IDLE stands for Integrated Development and Learning Environment.

The origin of the moniker IDLE is comparable to that of Python. Python was given that name by Guido Van Rossum in honour of the British comic troupe Monty Python, while IDLE was chosen as a tribute to Eric Idle, one of the group's original members. Since the 01.5.2b1 release, IDLE is included with the Python language's standard implementation. Many Linux, Windows, and Mac distributions include it in their Python packaging as an optional component.

DJANGO:

A high-level Python web framework called Django promotes quick iteration and logical, elegant design. It was created by seasoned programmers and handles a lot of the hassle associated with Web development, freeing you up to concentrate on building your app without having to invent the wheel. It is open source and free. Django's main objective is to make it simpler to create intricate, database-driven websites. Django places a strong emphasis on quick development, reuse and "pluggability" of components, and the don't repeat yourself maxim. Everywhere, including in the configuration files and data models, Python is used.

IMPLEMENTATION

SERVICE PROVIDER :

The administrator must log in to this module using a valid user name and password. After successfully logging in, he can carry out certain actions including viewing all of a student's academic information, to predict student behaviour, View All LSTM Behavioral Change, View All Users, View Academic Performance Prediction Results, and Extract Prediction of Poor Academic Performance.

VIEWING AND AUTHORIZING USERS:

In this module, the provider views all users details and authorize them for login permission. User Details such as User Name, Address, Email Id and Mobile Number.

USER:

There are n numbers of users present in this module. Users should sign up before carrying out any actions.

Once a user registers, the database will record their information. After successfully registering, he must log in using an authorised user name and password.

After successfully logging in, the user can carry out several actions including SEARCH STUDENT DATA SET, VIEW ALL STUDENT RECORD DETAILS, and BROWSE STUDENT DATA SETS.

VIEWING PROFILE DETAILS:

In this module, the user can see their own profile details, such as their address, email, mobile number, profile Image.

CONCLUSION

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. The first step in solving this issue is to build a solid model for predicting academic achievement in order to obtain a thorough understanding of student behavioural trends and possibly assist students in maximising their interactions with the university. College students' academic success is predicted

using a model called Augment ED. Our contributions in this regard come from three different places. To the best of our knowledge, this effort is the first to collect, evaluate, and apply multisource data for academic performance prediction that includes campus-life activities both inside and outside of the classroom in addition to online and offline learning. One can create a rich profile of a pupil using this multisource data. Second, in terms of the feature evaluation, linear, nonlinear, and deep learning (LSTM) methodologies are used to evaluate behavioural change, providing a systematic picture of students' behavioural patterns. In particular, it is the first time that the behavioural 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 individualised feedback for students who are at danger of failing or who lack self-discipline. Our study does have certain shortcomings, though. The generalisation of AugmentED may be adversely affected by this restriction. Additionally, we primarily concentrated on behavioural transformation. Other qualities/features that merit attention were not assessed, including peer effect and sleep. In conclusion, the foundation of our research is a comprehensive passive daily data collection system found in the majority of contemporary colleges. This system might result in more extensive, ongoing investigations.

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 school 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, executive functioning, 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 and behavioral trends of college students using smartphones,” In Proc. of the ACM International Joint Conference on Pervasive & Ubiquitous Computing, Seattle, WA, USA, 2014.

[8] R. Wang, G. Harari, P. Hao, X. Zhou, and A. T. Campbell, “SmartGPA: How smartphones can assess and predict academic performance of college students,” In Proc. of the ACM International Joint Conference on Pervasive & Ubiquitous Computing, Osaka, Japan, 2015.

[9] M. T. Trockel, M. D. Barnes, and D. L. Egget, “Health-related variables and academic performance among first-year college students: Implications for sleep and other behaviors,” Journal of American College Health, vol. 49, no. 3, pp. 125-131, 2000.

[10] D. M. Hansen, S. D. Herrmann, K. Lambourne, J. Lee, and J. E. Donnelly, “Linear/nonlinear relations of activity and fitness with children’s academic achievement,” Med Sci Sports Exerc. vol. 46, no. 12, pp. 2279-2285, 2014.

[11] A. K. Porter, K. J. Matthews, D. Salvo, and H. W. Kohl, “Associations of physical activity, sedentary time, and screen time with cardiovascular fitness in United States adolescents: Results from the NHANES national youth fitness survey (NNYFS),” Journal of Physical Activity and Health, pp. 1-21, 2017.

[12] K. N. Aadland, O. Yngvar, A. Eivind, K. S. Bronnick, L. Arne, and G. K. Resaland, “Executive functions do not mediate prospective relations between indices of physical activity and academic performance: the active smarter kids (ask) study,” Frontiers in Psychology, vol. 8, pp. 1088, 2017.

[13] M. Credé, S. G. Roch, and U. M. Kieszczynska, "Class attendance in college: A meta-analytic review of the relationship of class attendance with grades and student characteristics," Review of Educational Research, vol. 80, no. 2, pp. 272-295, 2010.

[14] S. P. Gilbert, and C. C. Weaver, “Sleep quality and academic performance in university students: A wake-up call for college psychologists,” Journal of College Student Psychotherapy, vol. 24, no. 4, pp. 295-306, 2010.

[15] A. Wald, P. A. Muennig, K. A. O"Connell, and C. E. Garber, “Associations between healthy lifestyle behaviors and academic performance in U.S. undergraduates: A secondary analysis of the American college health association\"s National College Health Assessment II,”

American Journal of Health Promotion, vol. 28, no. 5, pp. 298-305, 2014.

[16] P. Scanlon, and A. Smeaton, "Identifying the impact of friends on their peers academic performance," In Proc. of the IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), San Francisco, CA, USA, 2016.

[17] E. L. Faught, J. P. Ekwaru, D. Gleddie, K. E. Storey, M. Asbridge, and P. J. Veugelers, "The combined impact of diet, physical activity, sleep and screen time on academic achievement: A prospective study of elementary school students in Nova Scotia, Canada," International Journal of Behavioral Nutrition and Physical Activity, vol. STUDENT ACADEMIC PERFORMANCE PREDICTION 2022 DEPARTMENT OF INFORMATION TECHNOLOGY, MRECW Page 33 14, no. 1, pp. 29, 2017.

[18] E. L. Faught, D. Gleddie, K. E. Storey, C.M. Davison, and P. J. Veugelers, "Healthy lifestyle behaviours are positively and independently associated with academic achievement: an analysis of self-reported data from a nationally representative sample of Canadian early adolescents," Plos One, vol. 12, no. 7, pp. e0181938, 2017.

[19] V. Kassarnig, E. Mones, A. Bjerre-Nielsen, P. Sapiezynski, D. D. Lassen, and S. Lehmann, "Academic performance and behavioral patterns," Epj Data Science, vol. 7, no. 1, pp. 10, 2017