Predicting and Improving Behavioural Factors that Boosts Learning Abilities in Post-Pandemic Times using AI Techniques

Dr. Jaya Gera¹, Dr. Ekta Bhambri Marwaha², Dr. Reema Thareja³, Dr. Aruna Jain⁴

Associate Professor, Department of Computer Science, Shyama Prasad Mukherji College, University of Delhi, India¹ Associate Professor, Department of Applied Psychology, Shyama Prasad Mukherji College, University of Delhi, India² Assistant Professor, Department of Computer Science, Shyama Prasad Mukherji College, University of Delhi, India³ Associate Professor, Department of Computer Science, Bharati College, University of Delhi, India⁴

Abstract-Quantifying student academic performance has always been challenging as it hinges on several factors including academic progress, personal characteristics and behaviours relating to learning activities. Several research studies are therefore being conducted to identify the factors so that appropriate measures can be conducted by academic institutions, family and the student to boost his/ her academic performance. The present study investigates personal characteristics, psychological factors, behavioural factors, social factors and learning capabilities, that directly or indirectly affect student's academic performance, which was tapped by administering a self-designed questionnaire. The data was collected from 214 undergraduate students studying in various streams of the University of Delhi and post that semi-structured interview was conducted to get in- depth information. The result proved the correlation between the aforementioned factors and the learning capabilities of the students. Using the results of analysis a machine learning model based on k-nn algorithm was formed to predict student performance. A chatbot is also proposed to provide guidance to students in strenuous situations, motivate them and interact with them without having personal bias.

Keywords—Academic performance; machine learning; chatbot; educational data mining; learning analytics

I. INTRODUCTION

In today's era of the information revolution, learning analytics, predictive analytics, educational data mining, and machine learning techniques has become a hot area of research [1, 2] as it is beneficial for Teachers, administrators, family and the student him/herself to provide a timely remedy. We found that students have frequent mood swings, internal conflicts and issues that keep them distracted. They are often not comfortable to share their problems with their family members and friends. Accumulation of these issues causes stress and anxiety. In such a situation students become pessimistic, spend more time on social media or get involved in substance use. It is therefore important to identify academic and non-academic parameters that affects a student's performance. Using these parameters, we need a ML model to predict a new student's performance so that teachers and mentors can address their concerns at the earliest and provide them every help to grow with a stable mind.

Supervised learning algorithms are being used to analyse student's learning behaviours in order to predict and classify the students' performance [13,17]. However, the challenging task is to find the optimal algorithm that gives best results. Machine learning algorithms like k-nn, Naïve Bayes, ANN, logistic regression, SVM, decision tree, random forest, etc are used for prediction but the accuracy of results obtained from each model depends on the size and quality of data [11,12].

In this paper, we have used data analytics to identify personal, behavioural and academic factors that affect student's learning capability and empirically proved that there exists a relationship between the identified factors and the learning capability. Then, an AI model is created which when fed with identified factors as input, predicts the how quick learner a student is using the k-nn algorithm. Moreover, to handle the mental health issues which in turn affects student's academic performance as well as their learning capability, an NLP based chatbot model has been proposed to interact and guide them especially in strenuous situations.

The paper is divided into six sections. The first section introduces the title of the paper. The second section presents a summarization of research already done in this area. In section, data science and data analytics have been used to mine information from the dataset and visually represent crucial numbers using tables and charts. The results of analysis are interpreted and explained in Section IV. The section also proposes an AI based solution to find a solution to the issues identified in Sections III and IV. Section V proposes an AI based machine learning model that can accurately predict a new student's future performance by analysing certain parameters and comparing them with data already stored in the database. The paper is finally concluded in Section VI.

II. LITERATURE REVIEW

The drop-out students' ratio from higher education institutions results in immense loss/resource wastage. It also affects the evaluation and assessment processes of these institutions. New technologies like data mining, machine learning should therefore be used to perform simple and effective analysis of student-performance data that could help to improve learning procedures and atmosphere [3,4]. In [5], researchers have explained the use of Bigdata methods for learning analysis that can be used for system performance prediction, visualization of data, student skills estimation, risk detection, course recommendation, grouping of students and collaboration with other students. Predictive analysis is done on student achievements, behaviour, and skill prediction.

During the forecast of student performance [6] datamining techniques were used to build a predictor framework for the final-marks based on students' achievements and features including features, non-courses variables, out of class student conduct, video watching and post-school tutoring.

Researchers in [7] have evaluated students at the beginning of an academic session to forecast their achievements using collaborative filtering technique that is based on their academic history. Going further, in [8], the author has used historical academic data of students to evaluate a students' performance. The study relied on a factorization of low-rangematrixes and dispersed linear model.

In [9], a student classification system was developed using naive Bayes and decision tree algorithms that used features like the occupation of parents to enhance the correctness of the grade-prediction framework. The Decision Tree classifier performed better than the naive Bayes Classification in terms of accuracy [10].

III. DATA ANALYSIS

In the first stage, the raw datasets are collected from 214 students studying different courses in the University of Delhi. The data was then pre-processed to transform categorical data into numerical. Data analytics techniques were then applied using programming languages R and Python. The observations are reported in Table I.

Query	Response	Query	Response		
When you listen to a spiritual leader / motivational speaker, what are your topics of interest?	 36.4% for Handling Relationships 66.3% for Enhancing Concentration 57.3% for overcoming anxiety 72.4% for staying motivated 12% do not listen to any spiritual leader 	How many hours do you spend on social media per day?	13.1% spend less than 1 hour daily31.8% spend 1 to 2 hours daily35.5% spend 2 to 3 hours daily19.6% spend more than 3 hours daily		
How many close friends do you have?	11.7% have no close friend 65.4% have 1 to 3 close friends 15.9% have 4 to 5 close friends 7% have more than 5 close friends	How frequently do you meet your friends in a week?	8.4% meet once in a week5.1% meet twice in a week6.1% meet thrice in a week26.2% meet more than three times54.2% meet them occasionally		
Do you discuss your personal issues with your close friends?	50.5% says certain issues not all 35% feel free to discuss 14.5% do not share	How frequently do you have mood swings?	36.4% experience mood swings multiple times in a day15.4% once in a day13.6% once in a week28.5% once every fortnightRest occasionally		
What is the state of your mind in general?	43.3% varies frequently 24.3% calm 16.8% happy 14.9% highly frustrated and disappointed	How many hours do you devote every day for self-study?	 16.8% spend less than 1 hour daily 36.9% spend 1 to 2 hours daily 29% spend 2 to 3 hours daily 12.6% spend 3 to 4 hours daily 4.7% study for more than 4 hours daily 		
Which method of study do you prefer the most?	59.3% Reading Books or any other sort of Reading Material29.9% Watching Videos7.5% Group Studies with Friends2.3% Listening Audio PodcastsRest Educational Programs on TV	Who is your role model?	 43% Family Member 42.5% have no role model 4.7% sees their teacher as their role model 6.5% sees Sportsperson / army / any eminent person from history as their role model 3.3% sees Actor / Actress as their role model 		
How do you find your classes?	15.9% find them Boring and unproductive 50% Not boring but putting extra burden 34.1 %Exciting and triggering my thought process	Has your learning capabilities been affected due to the pandemic?	41.6% Badly affected51.4% Slightly affected3.3% Not at all affected3.7% Have become better than before		
Were you comfortable with online teaching?	29.9% Yes 41.6% No Rest was unsure	Which mode of teaching do you prefer now?	10.7% Online 47.7% Offline 41.6% Hybrid		
Did your learning capabilities improve with digital technology?	53.7% Yes 20.6% No 25.7% are unsure	Did any of the cases given below affect your learning capabilities in the last 2 years?	 71.4% were affected due to Stress and Anxiety 21.9% were affected due to Death in the Family 44.8% were affected due to Strained Financial Conditions 63.08% were affected due to Problems in Relationships 72.4% were affected due to Increased exposure to Electronic Gadgets 		

TABLE I. QUERIES AND CORRESPONDING RESPONSES

			18.69% were affected due to All of the above			
During pandemic, your studies were affected the most due to which of the following reasons?	 46.77% Lack of personal space 20.56% Inadequate ventilation 25.7% Power Cut 54.7% Noise from surroundings 62.14% Technical Issues 77.17% Disturbed routine 42.05% Unhealthy eating habit 	Which activity do you practice to keep your mind stable?	27.57% practice Yoga 39.71% practice Meditation 58.4% practice Singing 54.2% practice Dancing 31.7% get involved in Sports Activities			
Do you want any type of counselling to handle your stress?	57.5% Yes Rest No	What do you do to burst your stress?	 64.4% prefer Talking to someone who is close to me 65.8% prefer Keeping Quiet and Sitting Alone 35.9% prefer Reading Shopping 33.17% prefer Eating too much 68.2% prefer Watching TV or Web Series 54.2% prefer Walking or any Physical Activity 			
As a student, when was your level of motivation and attention higher?	49.5% Pre-pandemic times 7.5% During pandemic times 14.5% post-pandemic times Rest opted for Never 24.3% Always the same	Do you feel that your memorizing capabilities have been adversely affected due to COVID-19?	52.8% Yes 24.3% No Rest believes Maybe			
During online classes, was your learning affected due to any of the following reasons?	 28.03% Classes were not conducted on online platform regularly 61.68% Recording of lecturers were not available 50% Sessions were not interactive 70.09% Could not interact with other students in the class 51.8% Could not attend due to poor net connection 42.98% Did not have enough devices to attend classes 35.98% Study material was not available 80.8% Difficult to concentrate online beyond a certain limit 	Are you facing self-regulation challenges?	 40.65% Often miss the deadlines given by teachers 52.33% Did not get appropriate help during online classes 48.5% Lack the ability to control my own thoughts, emotions, and actions during classes 63.08% Have limited preparation before a class. 60.28% have poor time management skills during online classes. 55.14% fail to properly use online peer learning strategies (i.e., learning from one another to better 			
Do you face Technological literacy and competency challenges?	 27.10% Lack competence and proficiency in using various interfaces or systems 21.49% resist learning technology. 30.37% are distracted by an overly complex technology. 19.6%I have difficulties in learning a new technology. 22.42% lack the ability to effectively use technology to facilitate learning. 23.83% lack knowledge and training in the use of technology. 20.09% are intimidated by the technologies used for learning. 28.5% resist and/or am confused when getting appropriate help during online classes. 30.84% have poor understanding of directions and expectations during online learning. 29.43% perceive technology as a barrier to getting help from others during online classes 					

IV. DATA INTERPRETATION

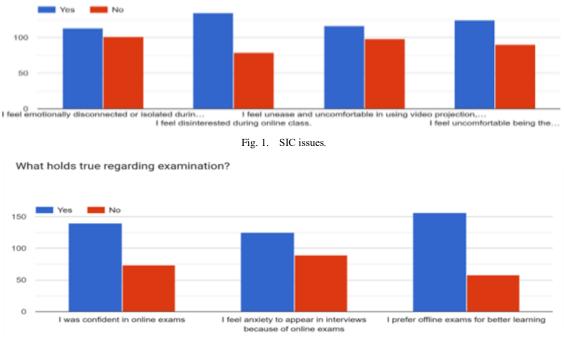
When we analysed the data collected, we observed an alarming situation. 90% of the students admitted that their learning capabilities have been affected due to the pandemic. Out of these, 71.4% of the students were affected due to Stress and Anxiety and 41.6% of these students were never comfortable with online teaching.

49.5% students believe that their level of motivation and attention was higher in pre-pandemic times and 52.8% feel that their memorizing capabilities have been adversely affected due to COVID-19.

During online classes, learning abilities of the students were affected as their classes were not conducted on online platform regularly (28.03%), recorded lecturers were not available (61.68%), teaching sessions were not interactive (50%), no interactions with peer group (70.09%), poor net

connection (51.8%) and lack of concentration (80.08%). In fact, every student experienced Student isolation challenges (SIC). Due to this, students feel emotionally disconnected or isolated and uncomfortable during online classes. They preferred face-to-face interaction with teachers and traditional classroom methods for socialization [11]. This is evident from the graph shown in Fig. 1.

During pandemic, studies were affected majorly due to lack of personal space (46.77%), noise from surroundings (54.7%), technical Issues (62.14%), disturbed routine (77.17%) and unhealthy eating habits (42.05%). In fact, the students admitted that during exams, they experienced anxiety issues when appearing for interviews or viva. So, they preferred to have offline teaching for better outcomes (refer Fig. 2). Did you experience Student isolation challenges (SIC)?





40.65% of the students feel disappointed as they often miss the deadlines given by teachers and 48.5% find it difficult to control their thoughts, emotions, and actions during classes. However, after the colleges reopened post COVID, the situation has become little better. More than 76% of the students agreed that a personal eye contact with the teacher is required for effective learning [12] (refer Fig. 3).

In fact, after colleges started in offline, only 8.9% of the students stated that small intra-class as well as inter-class competitions will not help to augment effective learning.

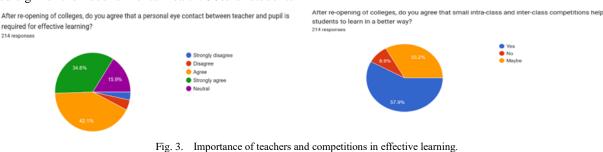
A. Stress and Anxiety Hampers Success

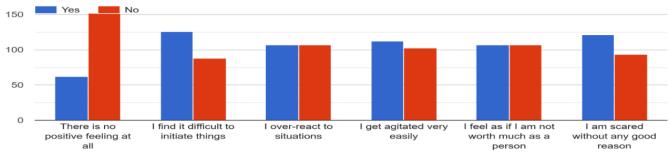
The students accepted that due to stress and anxiety they are unable to perform to their maximum potential. 51.8% of students experience mood swings at least once in a day [13]. 43.3% believes that their state of mind is not calm and happy. Rather, it varies frequently. The situation is worsened as 14.9% of students accept that they are highly frustrated and disappointed.

42.5% of students do not have a role model. This may not be a good sign for their sound mental health. 50% of students

feel that their classes add an extra burden to their daily lives. 57.8% students admitted that they need counselling to handle their issues. When asked additional questions related to their mental health, students' responses were recorded as shown in Fig. 4.

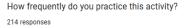
To handle such a strenuous situation, students are taking measures to overcome factors affecting their studies, which are in line with other research studies. They are listening to and motivational speakers for spiritual enhancing concentration and overcoming stress as well as anxiety. As evident from the pie chart shown in Fig. 5, more than 75% of the students have expressed that they listen to a spiritual/ motivational speaker at least once in a week. 58.4% practice Singing and 54.2% of the students practice Dancing to keep stress at bay. In extreme strenuous conditions, 64.4% prefer talking to a close friend/relation, 65.8% prefer keeping quiet and sitting alone, 68.2% watch TV or Web Series and 54.2% prefer walking or any other physical activity, which helped improve the mood and reduce stress among students [14].





Does any of the following hold true for you?

Fig. 4. Analysing mental health of students in post-pandemic times.



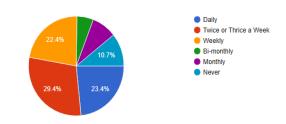


Fig. 5. Frequency of listening to spiritual / motivational speakers.

B. Solution Proposed to Boost Learning Capabilities

In the study, it was observed that 65.4% of the students have 1 to 3 close friends but 54.2% of them meet them occasionally. Moreover, 65% of the students admitted that they feel uncomfortable to discuss issues that disturb them with their close friends.

Moreover, in the post-pandemic times, more than 56% of the students are facing issues due to fatigue. And only 37% of the students are finding their studies interesting. Rest feels overwhelmed due to stress and thus feel fatigued after studying for few minutes.

To help students handle their stress and anxiety issues, we have proposed a chatbot based on the model depicted in Fig. 6. The role of this chatbot is to provide guidance to students in strenuous situations, motivate them and interact with them without having personal bias. Our study reveals that students are not comfortable discussing their personal issues even with their close friends. In such a scenario, they can at least share their feelings with an AI assisted chatbot that can communicate like a human and provide counselling or guidance.

C. Data Exploration to Build the Model

Before designing the machine learning model to make predictions, we first need to explore data to understand the relationships that exists between different parameters [15]. Therefore, we calculated co-relation between some key variables. The co-relation values between two variables are given in Fig. 7. The same values when plotted using a heatmap can be visualized. From the heat map plotted, it is very clearly evident that there is a strong relationship between Hours_of_Self_Study and Optimistic, Consistent_Learner and Quick_Learner, Stable_Mind and Optimistic, Stable_Mind and Hours_of_Self_Study. To study their underlying relationship in depth, we have plotted bar graphs. Identifying relationships between the variables and their impact on being quick and consistent learner is important, as a successful student has to be quick to learn and learn consistently. The conclusion that a consistent learner is quick to learn and vice versa has been drawn from the heatmap.

We observed that students who believe in group studies with friends are not quick and consistent learners. Quick and consistent learners prefer to study from books or any sort of reading material. Few of them also prefer to learn by watching educational videos on the World Wide Web. A quick and consistent learner spends at least 1 to 3 hours daily (refer Fig. 8) and spends majority of the time reading books/ study material followed by watching educational videos (refer Fig. 9). Even in the digital age, students prefer to read hard bound books or printed text material. They watch videos only for certain topics that they find too hard to understand. Moreover, in a country like India students find difficult to understand terminologies of hard-core English language, so they watch videos in local languages to understand as well as validate their understanding. One more interesting thing that we concluded from our studies is that negligible number of students prefers watching educational programs on Television and listen to podcast for effective learning.

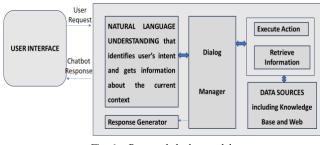
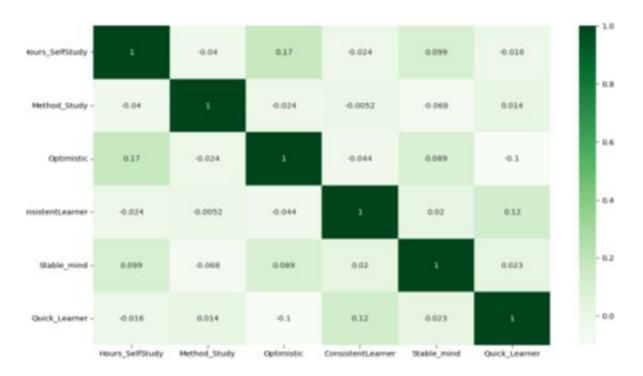
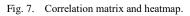


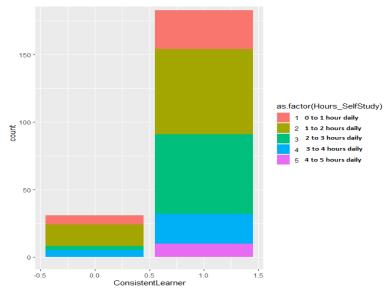
Fig. 6. Proposed chatbot model.

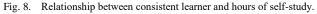
(IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 14, No. 11, 2023

	Hours_SelfStudy	Method_Study	Optimistic	ConsistentLearner	Stable_mind	Quick Learner
Hours_SelfStudy	1.000000	-0.039559	0.174594	-0.023668	0.098585	-0.015722
Method_Study	-0.039559	1.000000	-0.023559	-0.005205	-0.068424	0.014068
Optimistic	0.174594	-0.023559	1.000000	-0.043759	0.089492	-0.101285
ConsistentLearner	-0.023668	-0.005205	-0.043759	1.000000	0.019614	0.124775
Stable_mind	0.098585	-0.068424	0.089492	0.019614	1.000000	0.022564
Quick_Learner	-0.015722	0.014068	-0.101285	0.124775	0.022564	1.000000









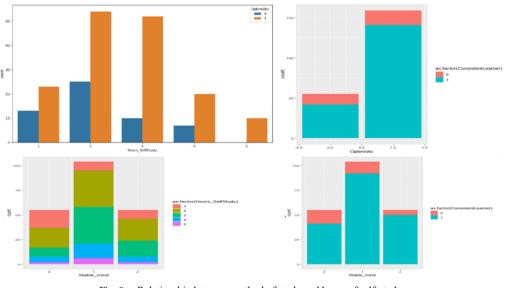
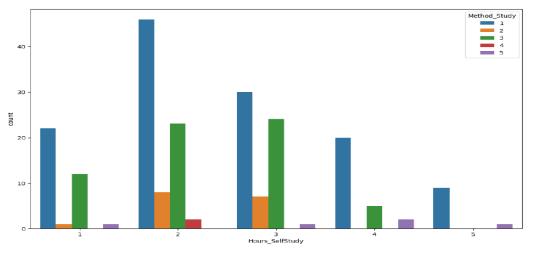
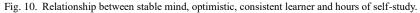


Fig. 9. Relationship between method of study and hours of self-study.





For effective learning, a quick and consistent learner devotes 1 to 3 hours daily. Studying beyond 3 hours does not add to their learning productivity. Research has indicated a favourable correlation between study duration and scholastic achievement [16]. A specific amount of study time will be reached before the grade no longer demonstrates a discernible shift. A quick and consistent learner experience less stress and anxiety and therefore has a stable mindset and is usually optimistic. Optimistic student prefers self-study using books or study material and is not dependent on peer group for learning. These numbers are evident from graphs given in Fig. 10.

V. PROPOSED ML MODEL TO PREDICT LEARNING CAPABILITIES

Using these conclusions, we can create a machine learning model (refer Fig. 11). The model divides the entire data set into two groups- training dataset and testing dataset. While training dataset has 70% of the data records chosen randomly from the original dataset, testing data set on the other hand has rest of the 30% records. The training dataset is utilized in the learning phase of the machine learning model. Correspondingly, the testing dataset is used in the evaluation phase to predict values for the 30% rows in the dataset. The ML model classifies a student either as a Quick Learner or a Slow Learner. Values predicted are then compared with actual values. The number of correctly predicted values are then compared with those predicted wrongly to determine the accuracy of the model. Higher the accuracy, higher is the probability that a new student will be accurately classified being a quick and consistent learner.

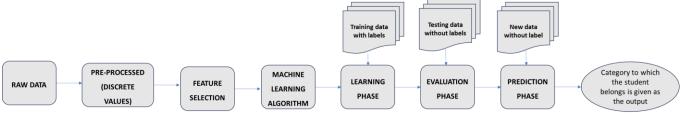


Fig. 11. Proposed ML model.

An accurate machine learning model will help the teachers to focus on inconsistent and slow learners and at the same time plan certain activities or assignments to further polish and accelerate the learning of quick and consistent learners.

The proposed machine learning model uses K Nearest Neighbour Classification algorithm. We also designed ML model using Naïve Bayes algorithm but accuracy of predictions was higher in case of knn algorithm [17]. For better and efficient results, we had done hyper parameter tuning. The confusion matrix of the k-nn algorithm is given in Fig. 12.

Accuracy (all correct / all) = (TP + TN) / (TP + TN + FP + FN).

Accuracy = (26 + 23) / (26 + 23 + 11 + 3) = 49 / 63 = 0.7777.

 $\label{eq:misclassification (all incorrect / all) = (FP + FN) / (TP + TN + FP + FN).$

Misclassification = 14 / 63 = 0.2222

Precision (true positives / predicted positives) = TP / (TP + FP)Precision = 26 / 37 = 0.7027.

Sensitivity aka Recall (true positives / all actual positives) = TP / (TP + FN).

Recall = 26 / 29 = 0.8965

Specificity (true negatives / all actual negatives) =TN / (TN + FP).

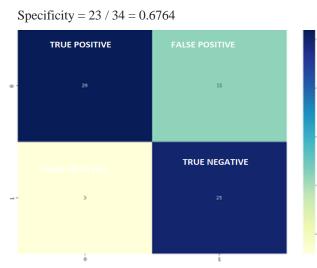


Fig. 12. Confusion matrix.

VI. CONCLUSION AND FUTURE SCOPE

The paper presents a study on 214 students in Higher Education. Several questions were asked to get an insight into their learning capabilities, behaviour, mental health, academic profile and their study habits. It was observed that students are facing issues including loneliness, stress, anxiety, resistivity to learn new things, lack of confidence, competence and proficiency. The students are aware about these challenges and are trying several techniques to deal with them and stay motivated as these issues are impacting their learning capabilities and success.

In the paper, we have proposed an AI Chatbot based on Natural Language Processing to interact with the students, guide them and motivate then whenever they feel low. The paper also proposes a machine learning model using knn algorithm that could classify a new student as a quick learner or a slow learner based on several factors related to mental health, hours of self-study, and preferred way to study. The model could predict data with an accuracy of 78%. In our subsequent study, we shall try to improve the accuracy by collecting more data about these factors affecting a student's academic performance. Since data is the fuel of AI applications, more the data better are the results.

REFERENCES

- S. Slater, S. Joksimovic, V. Kovanovic, R.s Baker, and D. Gasevic (2016), "Tools for Educational Data Mining: A Review", Journal of Educational and Behavioral Statistics, Vol. 42, No. 1, 2016, pp. 88-106.
- [2] Pooja Thakar, Anil Mehta, and Manisha (2015) "Performance Analysis and Prediction in Educational Data Mining: A Research Travelogue", International Journal of Computer Application, Vol. 100, No.12, January 2015, pp. 60-68.
- [3] Buenano-fernandez D (2019) The use of tools of data mining to decision making in engineering education: a systematic mapping study. Comput Appl Eng Edu. https:// doi. org/ 10. 1002/ cae. 22100.
- [4] Ahmed M, Najmul Islam AKM (2019) Deep learning: hope or hype. Ann Data Sci 7(3):427–432. https:// doi. org/ 10. 1007/ s40745- 019-00237-0.
- [5] Gil D, Fernandez-aleman JL, Trujillo J, Garcia-mateos G, Lujan-mora S, Toval A (2018) The effect of green software: a study of impact factors on the correctness of software. Sustainability. https:// doi. org/ 10. 3390/ su101 03471.
- [6] Lu OHT, Huang AYQ, Huang JCH, Lin AJQ, Yang SJH (2018) Applying learning analytics for the early prediction of students ' academic performance in blended learning. Edu Technol Soc 21(2):220– 232.
- [7] Bydžovska H (2015) Are collaborative filtering methods suitable for student performance prediction?pp 425–430. doi: https:// doi. org/ 10. 1007/978-3-319-23485-4.
- [8] Polyzou A, Karypis G (2016) Grade prediction with models specific to students and courses. Int J Data Sci Anal 2(3):159–171. https:// doi. org/ 10. 1007/ s41060- 016- 0024-z.

- [9] Khan B, Sikandar Hayat Khiyal M, Daud Khattak M (2015) Final grade prediction of secondary school student using decision tree. Int J Comput Appl 115(21):32–36. https:// doi. org/ 10. 5120/ 20278- 2712.
- [10] Shatrma N., Appukutti S., Garg U., Mukherjee J., Mishra S., (2023), Analysis of Student's Academic Performance based on their Time Spent on Extra-Curricular Activities using Machine Learning Techniques, I.J. Modern Education and Computer Science, Vol. 1, pp 46-57.
- [11] Al-Azzam N, Elsalem L, Gombedza F. (2020) A cross-sectional study to determine factors affecting dental and medical students' preference for virtual learning during the COVID-19 outbreak. Heliyon. 2020;6(12): e05704.
- [12] Jum'ah AA, Elsalem L, Loch C, Schwass D, Brunton PA. (2020). Perception of health and educational risks amongst dental students and educators in the era of COVID-19. Eur J Dent Educ.

- [13] Matud, M. P., Díaz, A., Bethencourt, J. M., and Ibáñez, I. (2020). Stress and psychological distress in emerging adulthood: a gender analysis. J. Clin. Med. 9, 2859. doi: 10.3390/jcm9092859.
- [14] Ekta Bhambri Marwaha. Subjective well-being among university students & physical activity patterns(2019). Int J Appl Res 2019;5(5):319-325. DOI: 10.22271/allresearch.2019.v5.i5d.11004.
- [15] Yunusovich A. V., Ahmedov V., Norboyev K., Zakirov F., (2022), Analysis of Experimental Research Results Focused on Improving Student Psychological Health, I.J. Modern Education and Computer Science, Vol 2, pp 14-30.
- [16] Mukun Liu (2022) The Relationship between Students' Study Time and Academic Performance and its Practical Significance.BCP Education & Psychology, 7:412-415. DOI:10.54691/bcpep.v7i.2696.
- [17] Vidushi Singla, Rashi Thareja, Reema Thareja,(2022) "Psychometric Analysis Using Computational Intelligence for Smart Choices", International Journal of Modern Education and Computer Science(IJMECS), Vol.14, No.2, pp. 65-75, 2022.DOI: 10.5815/ijmecs.2022.02.06.