Implementation of Latest Machine Learning Approaches for Students Grade Prediction

Mr. D.Sreenivasulu[1]
Assistant Professor
Department of CSE
(Data Science)
Institute of Aeronautical Engineering,
Hyderabad, India
dsnvas@gmail.com

Dr J Sirisha Devi[2]
Associate professor,
Department of Computer Science and Engineering,
Institute of Aeronautical Engineering, JNTU (H) India.
siri.cse21@gmail.com

Dr.P.Arulprakash[3]
Associate Professor&Head,
Department of CSE,
Rathinam Technical Campus,
Eachanari, Coimbatore-641021,
Tamil Nādu,India.
arulprakash247@gmail.com

Dr S. Venkataramana[4]
Professor,
Department of Information Technology,
SRKR Engineering College,
Chinamiram, Bhimavaram-534204,
Andhra Pradesh-India.
svramana@srkrec.ac.in

Dr Kutubuddin Sayyadliyakat Kazi[5]
Professor,
Department Of Electronics & Telecommunications,
Brahmdevdada Mane Institute of Technology,
Solapur, Maharashtra, India
drkkazi@gmail.com

Abstract:
Education is a necessary for a fruitful and happy life, and it enhances the quality of a person's existence in both worth and excellence. Self-confidence and the ability to participate in today's world are two reasons why education is so important. Education has had to deal with several difficulties over the years. In order to improve the quality of learning, several teaching and learning approaches are proposed. Computers and mobile devices are used in every aspect of daily life, and numerous materials may be found online at any time. When it comes to teaching and learning, Artificial Intelligence (AI) has had a remarkable rise in many areas including education. To improve learning and teaching, higher education institutions are using technologies into their traditional methods. Education systems are faced with difficult issue of predicting students' performance semester after semester. Education systems contain a ton of information that can be used by teachers to develop effective teaching practises. The primary goal of this research is to demonstrate that it is possible to train and model a dataset, and that a reliable prediction model can be built. Among other things, this research intends to discover the most significant factors that influence student achievement and to identify the most appropriate machine-learning method to forecast their performance. Data from student registries and web - based learning platforms are the most commonly used datasets in academic research. The importance of ML approaches in predicting at-risk pupils and dropout rates has been demonstrated, and this has led to an increase in student achievement.

Keywords: *machine learning, Student's Performance, Prediction, Learning Analytics, Visualization.*

1. Introduction:

As of today, there is an abundance of educational data coming from many sources. There is a great deal of important data held in traditional educational institutions' information systems repositories, for example. More and more educational data is being generated in a variety of formats due to the rapid growth of web-based learning platforms and electronic learning materials [1]. There is also a lot of data generated by newer learning settings such as virtual worlds and game-based learning as well as blended learning and studies play an important role. It is important for school institutions to have access to all education data in various forms and from numerous sources in order to make informed decisions that lead to optimal success [2].

The main challenge that educational institutions face is the analysis of vast amounts of data sources that develops exponentially. Human specialists and standard analytic approaches sometimes fail to uncover correlations, trends, and other discoveries that can be helpful to administrators as well as educators and students [3][21]. Because of this, it is necessary to employ automated analytic methods to extract the necessary relevant instructional knowledge from this type of data. Consequently, the subject of Educational Data Mining (EDM) has arisen to make the greatest use of educational data Finding patterns and connections in massive datasets derived from educational

DOI: 10.9756/INT-JECSE/V14I3.1141

settings is what we mean by EDM. Data mining (DM) and machine learning (ML) techniques such as categorization, grouping, clustering algorithms, regress, and sequential patterns are used to make sense of educational system data [4][20]. Data analysis has gotten more difficult as educational systems and learning platforms have evolved to meet the needs of students. As a result, EDM approaches based on ML have been put to good use. For example, classification algorithms serve primarily to address the challenge of predicting student performance (SPP) based on the provided datasets.

Higher education students are sluggish when it comes to taking tests and maintaining a certain grade point average to graduate. It is a requirement of NEP2020 that all graduate students be educated at the level they desire. Students had no idea of their true intellectual standing until they obtained the unqualified grade [5-9]. When spoken instructions are delivered, they are quickly disregarded by the students. They can only be proficient in a 2nd or 3rd effort once they notice this and begin working on it, which has an effect on the success rate of the institution.

The length of the course is six months in the context of higher education. Each student must take all three of the course's continuous assessment tests, as well as three of the course's internal assignments and one of the course's external exams, in order to pass the course. Only after all of these examinations will the semester's grade be determined. Writing a supplement exam to make up missed work in the same topic the following semester is a necessary reaction. After the second continuous assessment, an algorithm is proposed that can predict the results of each student in each topic immediately. It aids mentors in providing the necessary assistance and training to students in order to obtain a credential.

With minimal settings, an alert can be issued in the learner learning system so that quick action can be made based on the prediction. Numerous studies have used a filter or wrapper method in spite of the enormous set of parameters and modeling techniques. The advantages of both systems are undeniable. By combining the working principles of the wrapper feature selection strategies into a single model, it is possible to reap the benefits of both methods while maintaining a minimal set of features. In order to train and evaluate several machine learning algorithms, the extremely relevant features were selected as a new feature set. Despite the fact that the list of parameters and the accuracy of the predictions determine the end strategy, supervised ml algorithms have a wide range. Using an ensemble model, you can get a lot of information out of a small quantity of data [10][22]. Prediction accuracy is influenced by the parameters of the algorithm used to run it [11][24]. With the use of its monitoring system, this study hopes to help students improve their academic ability throughout the course and ensure that all students pass on their first try.

2. Literature Survey:

Segmentation, categorization, time - series data mining, and association rules are just a few of the methods being used to construct automated educational systems in the EDM sector. An increasing number of organizations are embracing hybrid modeling techniques in an effort to create more accurate, robust, and efficient models. Earlier study on the subject is briefly discussed in this section. Predicting student performance has been the subject of previous research, which is the primary emphasis of this study.

The SPP problem has been addressed using a variety of classification algorithms, including Tree Structure (DT), ANN, NB, SVM, and ensembles (boosting and bagging). For example, Yang and Li [12][23] suggested a neural network-based categorization method (BP-NN). Students' future performance is predicted using their prior knowledge and the similarities they share with their peers. Quick and accurate predictions could be made using the model that was proposed. In a different method of forecasting graduation dates, Kesumawati and Utari [13] used NB and SVM. SVM was found to be superior to NB in a series of experiments. For the purpose of predicting student performance, Rana and Garg [14] used WEKA software and two machine learning algorithms, ANN and NB. Because of this, NB outperformed ANN in terms of performance. Research by tested the accuracy of ensemble approaches for forecasting student outcomes. To boost the effectiveness of ANN, DDT, and NB classifiers, ensemble approaches were used. Ensemble approaches proved to be more effective than traditional methods in a series of experiments. Stochastic training techniques were used with a feed-forward neural network (FNN) in the work of Thaher and Jayousi [15]. The tested datasets were re-balanced using the SMOTE approach. It was found that the proposed model was effective in solving the SPP problem.

Preprocessing steps such as Feature Selection(FS) are critical in reducing the dimensionality of data and

DOI: 10.9756/INT-JECSE/V14I3.1141

enhancing the performance of learning algorithms by removing redundant and irrelevant features. Wrapper and filter-based FS approaches have been used in certain SPP research. SPP models' performance has been shown to be enhanced by information gain filters, as demonstrated by the writers in [4] [15][25].

Students in MOOCs are being studied by He et al. [16]. Sequentially Smooth Regression Techniques (LR-SEQ) and Concurrently Smooth Logistic Regression (LR-SIM) were the two transfer learning techniques presented by the researchers. DisOpt 1 and DisOpt 2 datasets are used to test the proposed methods. AUC values for LR-SIM were higher in the first week than for LR-SEQ, when compared to the standard Logistic Regression (LR) algorithm's results. Early admittance results revealed a favorable projection.

Kovacic, Z. [17] used machine learning approaches to study the initial assessment of student progress in their study. As part of the review, socio-demographic characteristics, such as educational attainment and employment position as well as course aspects, such as course programmer and course block were examined for their capacity to predict success. The Open University Zealand provided these features for the dataset. Selecting the most important features determining a student's achievement can be done using machine learning methods. Study results showed that ethnicity, curriculum, and block of courses were the three most important factors determining students' achievement[26].

During an introductory programming course, students' activity logs were examined by Watson et al. [18] to determine their predicted performance. An automated predictor, rather than a direct basis, was recommended in this review to track pupils' success over time. WATWIN, their suggested grading algorithm, assigns a numerical value to each of the various programming activities that students complete. The grading algorithm takes into account the student's ability to cope with programming errors and the time it takes to address such issues, amongst other variables. The dataset for this study was compiled from the programming activity logs of 45 students across 14 sessions[27]. Watwin scores were awarded to each student's activities, and these were employed in linear regression. Linear regression utilizing the WATWIN score has a 76% success rate. The data must be balanced in order to make accurate predictions. Each of the forecast categories has an equal amount of attributes in a balanced dataset.

3. Methodology:

The research utilizes the acquired data in the training phase for the suggested technique in order to predict student success for the courses. Preprocessing will actually occur following the collection of real data from a diverse university. The data will then be separated into two sets, one for training and the other for testing, based on when the data was collected. During the testing phase, the suggested prediction model's accuracy is evaluated.

3.1. Dataset Preprocessing and Transformation.

Pre-processing is necessary to address undesirable issues[19] including such redundant attributes and noise in the collected database because of the wealth of information it contains:

First step:get rid of clichés like "course," "lecturer," and "student."

Second step:clear records that are superfluous or unneeded, such as courses that students have registered for but never finished, such as exams or exemptions, and so on.
Third step: When the number of registered pupils falls below 15, several universities do not offer the course. In our instance, the courses that have been disregarded are viewed as a nuisance and will not be taken.

Fourth step:convert text or string values into numbers.

Table 1 illustrates examples of data used in the literature, which were utilized to choose the training model input properties. These characteristics were picked after careful consideration of the results of experiments and previous models for predicting student success.
The suggested prediction model will employ quarter conversion (QTF) with min-max scalar (MMS) to create and convert all values in the range of values where deep learning methods have converged because there is a varied distribution for the various attributes.
One of the best preprocessing strategies is a nonlinear transformation QTF, which reduces outlier effects

DOI: 10.9756/INT-JECSE/V14I3.1141

significantly. It is possible for new data, such as validation/test data, to have values greater or lower than those of the fitted boundaries. The prevalence and breadth of each attribute differ significantly before data modification. As a result of the QTF data transformation, each feature scalar's distribution will be more normal, and the range of each feature's data transformation will be between 0 and 1.

Machine Learning and regression analysis were used to create a model that predicts student performance. 1D CNN receives a 1D set of data points of 21 features; next, it passes through a stack of 64 node each with 3 kernels in a convolution layer. Once the Rec plan also provides Linear Unit (ReLU) operational amplifier has been established for each convolution, 64 Tanh units and a one-time step are included in LSTM. 1D CNN and LSTM can be trained using the Sigmoid illustrated by equation (1), which results in an output range of 0 to 1. Mathematical formula (1) will yield an output multiplied by four in order to replicate grades from 0.0 to 4.0.

$$Sigmoid = \frac{1}{1+e^{-x}} \rightarrow (1)$$

Table 1:Literature Data Set

Data Set	Training (%)	Size
Education	68	289,178
Environment	72	356,789
Economics	87	789.125
Foreign Languages	76	189,458
Law	86	689,785
Political Sciences	48	756,145
Agriculture	56	478,987

Updated hidden Layers

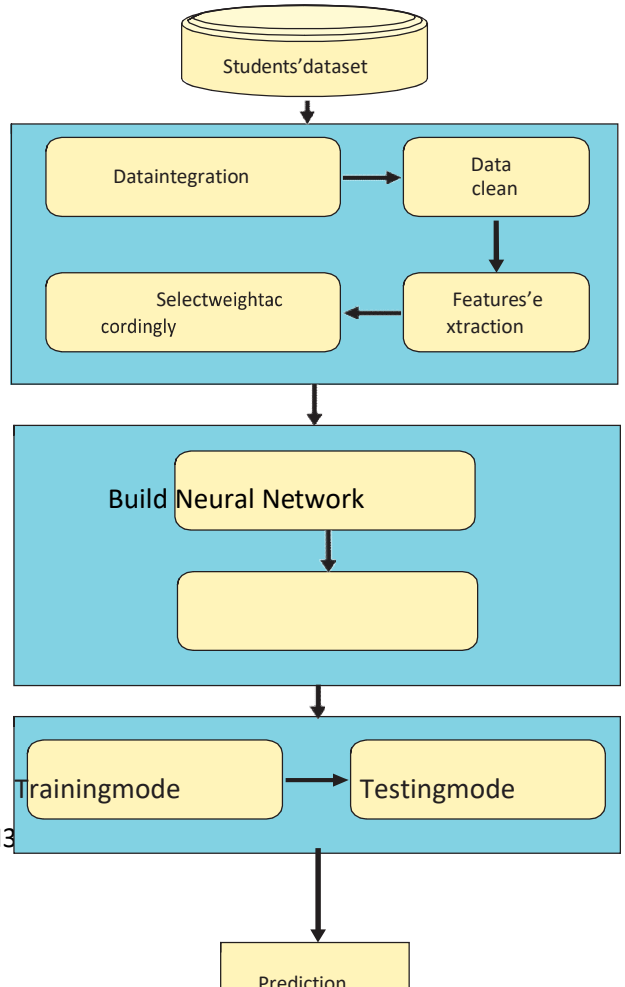

Figure 1:Overall System Architecture

DOI: 10.9756/INT-JECSE/V14I3.1141

4. Performance Evaluation

After processing the existing information set on a Single - layered Perception using the WEKA tool, a confusion matrix was obtained. True Positive (TP, a number of correctly identified positive cases), False Positive (FP, a number of wrongly classified positive instances), False Negative (FN, a number of mistakenly categorized negative instances) and True Negative (TP, a number of correctly classified negative instances).The 10-fold cross validation approach is used to filter the data for CNN training and testing. One portion of the data is used for testing, while the other nine portions are used for training. The neural network is trained using the CNN technique. Table 2 explains these numbers in detail.

Table 2. Confusion Matrix

		Observed	
		True	False
Predicted	***True***	True Positive (TP)	False Positive (FP)
	False	False Negative (FN)	True Negative (TN)

From the values of Table2 sensitivity, specificity and accuracy are calculated.
Sensitivity = TP / (TP + FN)→(2)
Specificity = TP / (TP + FP)→(3)
Sensitivity and specificity can be determined using the equations 2 and 3 provided above for successfully categorized and wrongly classified cases, respectively.Eq. 4 calculates accuracy, which is defined as the classifier's overall success rate.
Accuracy = (TP + TN) / (TP + FP + FN + TN)→(4)

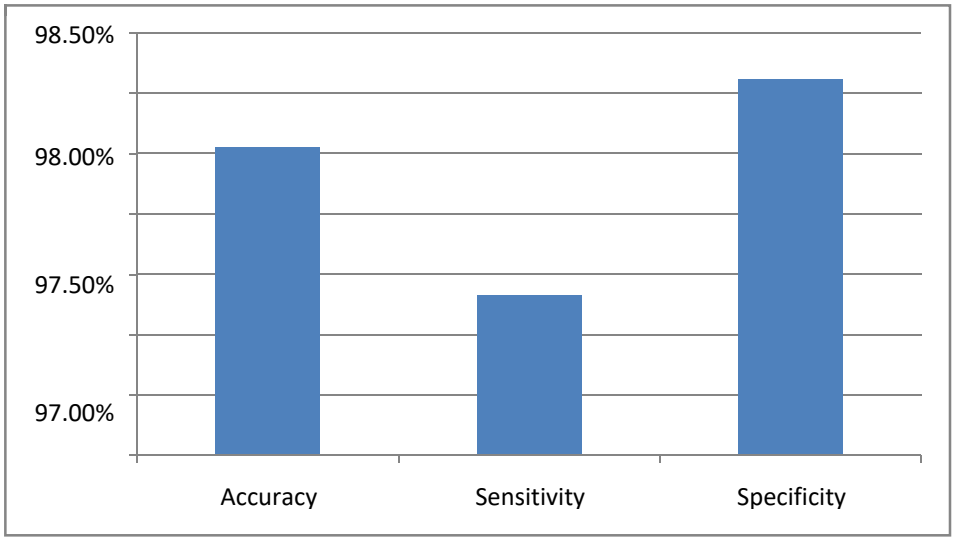

Figure 2: Assessment of the Proposed System

5. Conclusion:

E-learning tools and other web-based educational platforms create enormous amounts of data that can be used to improve education. As a result, predicting student performance is a difficult challenge for educational institutions. The quality of the educational process can be improved by giving decision-makers, teachers, and students with useful prediction models. As a means of increasing the model's accuracy, the Convolution Neural Network Method was employed in this study. The efficiency, generality, and simplicity of CNN modeling are only a few of the advantages it offers. To better represent complicated systems, CNN's properties make it an appealing option, and

DOI: 10.9756/INT-JECSE/V14I3.1141

this method attained an accuracy rate of 98.35 percent. An improved method of machine learning in the future could yield even more accurate results.

References:

1. C. Romero and S. Ventura, "Educational data science in massive open online courses," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. WIREs Data Mining Knowl Discov, 01 2017.
2. P. Cortez and A. Silva, "Using data mining to predict secondary school student performance," in In A. Brito and J. Teixeira Eds., Proceedings of 5th FUture BUsiness TEChnology Conference (FUBUTEC 2008), Porto, Portugal, 2008, pp. 5–12.
3. R. Baker, "Challenges for the future of educational data mining: The baker learning analytics prizes," Journal of Educational Data Mining, vol. 11, pp. 1–17, 08 2019.
4. E. Amrieh, T. Hamtini, and I. Aljarah, "Mining educational data to predict student's academic performance using ensemble methods," International Journal of Database Theory and Application, vol. 9, pp. 119–136, 09 2016.
5. Badugu S, Rachakatla B. Student's performance prediction using machine learning approach. In Data engineering and communication technology. Singapore: Springer; 2020. p. 333–340.
6. Sunday K, Ocheja P, Hussain S, et al. Analyzing student performance in programming education using classi- fication techniques. International Journal of Emerging Technologies in Learning (iJET). 2020;15(2):127–144.
7. Hooshyar D, Pedaste M, Yang Y. Mining educational data to predict students' performance through procras- tination behavior. Entropy . 2020;22(1):12.
8. Zulfiker MS, Kabir N, Biswas AA, et al. Predicting students' performance of the private universities of Bangladesh using machine learning approaches. Inter- national Journal of Advanced Computer Science and Applications. 2020;11(3):672–679.
9. Tatar AE, Düştegör D. Prediction of academic per- formance at undergraduate graduation: course grades or grade point average? Applied Sciences. 2020;10(14): 4967.
10. Gajwani J, Chakraborty P. Students' performance pre- diction using feature selection and supervised machine learning algorithms. In International Conference on Innovative Computing and Communications (pp. 347- 354). Springer, Singapore; 2021.
11. Ajibade SSM, Ahmad NB, Shamsuddin SM. A heuris- tic feature selection algorithm to evaluate the academic performance of students. In 2019 IEEE 10th Control and System Graduate Research Colloquium (ICSGRC) (pp. 110-114). IEEE; 2019, August.
12. F. Yang and F. W. Li, "Study on student performance estimation, student progress analysis, and student potential prediction based on data mining," Computers & Education, vol. 123, pp. 97 – 108, 2018.
13. A. Kesumawati and D. T. Utari, "Predicting patterns of student gradua- tion rates using na¨ıve bayes classifier and support vector machine," AIP Conference Proceedings, vol. 2021, no. 1, p. 060005, 2018.
14. S. Rana and R. Garg, "Student's performance evaluation of an institute using various classification algorithms," in Information and Communi- cation Technology for Sustainable Development, D. K. Mishra, M. K. Nayak, and A. Joshi, Eds. Singapore: Springer Singapore, 2018, pp. 229–238.
15. T. Thaher and R. Jayousi, "Prediction of student's academic performance using feedforward neural network augmented with stochastic trainers," in 2020 IEEE 14th International Conference on Application of Information and Communication Technologies (AICT), 2020, pp. 1–7.
16. He, J.; Bailey, J.; Rubinstein, B.I.; Zhang, R. Identifying at-risk students in massive open online courses. In Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence, Austin, TX, USA, 25–30 January 2015.
17. Kovacic, Z. Early Prediction of Student Success: Mining Students' Enrolment Data. In Proceedings of the Informing Science and Information Technology Education Joint Conference, Cassino, Italy, 19–24 June 2010.
18. Watson, C.; Li, F.W.; Godwin, J.L. Predicting performance in an introductory programming course by logging and analyzing student programming behavior. In Proceedings of the IEEE 13th International Conference on Advanced Learning Technologies, Beijing, China, 15–18 July 2013; pp. 319–323.
19. Shuping Li, Taotang Liu, "Performance Prediction for Higher Education Students Using Deep Learning", *Complexity*, vol. 2021, ArticleID 9958203, 10 pages, 2021.

DOI: 10.9756/INT-JECSE/V14I3.1141

20. P. Yogendra Prasad ,"Implementation of Machine Learning Based Google Teachable Machine in Early Childhood Education",International Journal of Early Childhood Special Education,Vol 14, Issue 03,PP 4132-4138,2022.
21. Sushama, C., P. Arulprakash, M. Sunil Kumar, D. Ganesh, and Kamepalli Sujatha. "The Future of Education: Artificial Intelligence based Remote Learning." International Journal of Early Childhood 14, no. 03: 2022.
22. Ganesh, D., Kumar, M.S., Reddy, M.P.V., Kavitha, S. and Murthy, D.S., Implementation of AI Pop Bots and its allied Applications for Designing Efficient Curriculum in Early Childhood Education. International Journal of Early Childhood, 14(03), p.2022.
23. Prakash, Mr S. Shiva, Manasa Bandlamudi, and Ragitha Radhakrishnan. "Educating and communicating with deaf learner's using CNN based Sign Language Prediction System."Vol 14, Issue 02,PP: 2624-2629, 2022.
24. Kumar, M.S. and Harika, A., 2020. Extraction and classification of Non-Functional Requirements from Text Files: A Supervised Learning Approach. PSYCHOLOGY AND EDUCATION, 57(9), pp.4120-4123.
25. Natarajan, M. K. V. A., & Macha Babitha, D. G. Machine Learning Based Identification of Covid-19 From Lung Segmented CT Images Using Radiomics Features. Biosc. Biotech. Res. Comm. Special Issue, 14(07), 350-355.
26. Ganesh, D., Kumar, T. P., & Kumar, M. S. (2021). Optimised Levenshtein centroid cross-layer defence for multi-hop cognitive radio networks. IET Communications, 15(2), 245-256.
27. Gampala V, Kumar MS, Sushama C, Raj EF. Deep learning based image processing approaches for image deblurring. Materials Today: Proceedings. 2020 Dec 26.

DOI: 10.9756/INT-JECSE/V14I3.1141