International Journal of Engineering & Extended Technologies Research (IJEETR)

|ISSN: 2322-0163| www.ijeetr.com | A Bimonthly, Peer Reviewed, Scholarly Indexed Journal |

| Volume 8, Issue 2, March - April 2026 |

DOI:10.15662/IJEETR.2026.0802017

An Ensemble Approach for Predicting Student Performance using Learning Activities

Dr. S. Swetha [1], N. Ajay Kumar [2]

Assistant Professor, Dept. of CSE, Sree Rama College of Engineering, Tirupati, A.P, India[1]

PG Scholar, Dept. of CSE, Sree Rama College of Engineering, Tirupati, A.P, India[2]

Publication History: Received: 25.02.2026; Revised: 20.03.2026; Accepted: 25.03. 2026; Published: 28.03.2026.

ABSTRACT: One important field of study in current educational research is knowledge tracking. Knowledge tracking models that exhibit better prediction capabilities than conventional methods have been developed as a result of recent developments in deep learning. This paper presents a novel ensemble-based knowledge tracking model to address the issues of restricted interpretability and the complexity of lengthy sequence dependencies in current deep knowledge tracking frameworks. In order to better identify important characteristics at different temporal points during the learning process and improve predictive performance, this model uses the features of exercises in addition to students' learning capacities as input layer feature information.

KEYWORDS: Student Performance Prediction, Ensemble Learning, Machine Learning, Learning Analytics

I. INTRODUCTION

In order to improve learning outcomes and develop students' skills, personalized learning seeks to tailor the curriculum and learning route to each student's unique needs, traits, and learning progress. Many students' lifetime learning data is gathered and arranged by an artificial intelligence online learning platform. An important problem in educational data mining is how to evaluate and use these data to give users individualized instructional modalities. Conventional teaching assessments, which mostly rely on standardized tests or assignments, are unable to quickly give all the information required for personalized learning and cannot deliver feedback on students' current learning state. Knowledge tracking is a data-driven approach to learning evaluation that uses learners' answer history data to evaluate and forecast their level of knowledge mastery. Numerous methods have been developed by researchers to tackle the problem of knowledge tracking. Bayesian Knowledge Tracking (BKT) and Deep Knowledge Tracking (DKT) are the two most well-known models. When students come across newly unexplored concepts, these models frequently require assistance in making accurate predictions and capturing high-level semantic information, even though they successfully use the explicit links between exercises and concepts to forecast students' mastery of an idea. An developing learning analytics tool with a wide range of potential applications and advancements is deep knowledge tracking. In the future, DKT will realize more remarkable breakthroughs and applications and bring more innovations and changes to the education area thanks to the ongoing advancement of associated technologies and in-depth study.

Numerous methods have been developed by researchers to tackle the problem of knowledge tracking. Bayesian Knowledge Tracking (BKT) and Deep Knowledge Tracking (DKT) are the two most well-known models. When students come across newly unexplored concepts, these models frequently require assistance in making accurate predictions and capturing high-level semantic information, even though they successfully use the explicit links between exercises and concepts to forecast students' mastery of an idea. An developing learning analytics tool with a wide range of potential applications and advancements is deep knowledge tracking. In the future, DKT will realize more remarkable breakthroughs and applications and bring more innovations and changes to the education area thanks to the ongoing advancement of associated technologies and in-depth study.

II. LITERATURE SURVEY

A. Knowledge tracing: Modeling the acquisition of procedural knowledge

An attempt to simulate students' evolving knowledge states during skill learning is described in this work. In this study, students are using the ACT Programming Tutor (APT) to learn how to develop small programs. The ideal student model, a production rule cognitive model of programming knowledge, serves as the foundation for APT. With this

|ISSN: 2322-0163| www.ijeetr.com | A Bimonthly, Peer Reviewed, Scholarly Indexed Journal |

| Volume 8, Issue 2, March - April 2026 |

DOI:10.15662/IJEETR.2026.0802017

arrangement, the tutor can work through exercises with the student and offer help when needed. In a procedure known as knowledge tracing, the tutor keeps track of the likelihood that the student has mastered each of the ideal model's rules while the student works. Based on these probability estimates, the tutor gives the student a personalized series of exercises to do until the student has "mastered" every rule. The learning and performance assumptions, cognitive model, and programming tutor are explained. A number of research that look at the empirical validity of knowledge tracing and have changed the procedure are examined. As of right now, the model predicts test results fairly well. There is discussion of more modeling process changes that could raise performance levels.

B. Deep knowledge tracing

A well-known issue in computer-supported education is knowledge tracing, in which a machine simulates a student's knowledge as they engage with coursework. Even while accurately simulating student knowledge would have a significant educational influence, there are numerous inherent difficulties with the task. In this work, we investigate the usefulness of modeling student learning with Recurrent Neural Networks (RNNs). The RNN family of models can capture more sophisticated representations of student knowledge and does not require the explicit encoding of human domain knowledge, which is a significant gain over earlier approaches. Prediction performance on a variety of knowledge tracing datasets is significantly improved by using neural networks. Additionally, the learnt model enables simple interpretation and the discovery of structure in student assignments, and it may be applied to intelligent curriculum design.

These findings point to an excellent application challenge for RNNs and a promising new direction for knowledge tracing research.

C. EKT: Exercise-aware knowledge tracing for Student performance prediction

Predicting student performance (e.g., scores) on future exercises is one of the essential tasks for providing proactive services (e.g., personalized exercise recommendation) to students in computer-supported intelligent education, where it is necessary to track the change in each student's knowledge acquisition during her exercising activities. Unfortunately, the problem of extracting rich information from the materials (e.g., knowledge concepts, exercise content) of exercises to achieve both more accurate prediction of student performance and more comprehensible analysis of knowledge acquisition remains underexplored. To the best of our knowledge, existing approaches can only exploit students' exercise records. In order to achieve this, we provide a comprehensive analysis of student performance prediction in this work. To immediately accomplish performance prediction's main objective. In EERNN, we only create a bidirectional LSTM to learn the encoding of each exercise from its content by summarizing each student's state into an integrated vector and tracking it with a recurrent neural network.

III. PROPOSED SYSTEM

The overview of our proposed system is shown in the below figure.

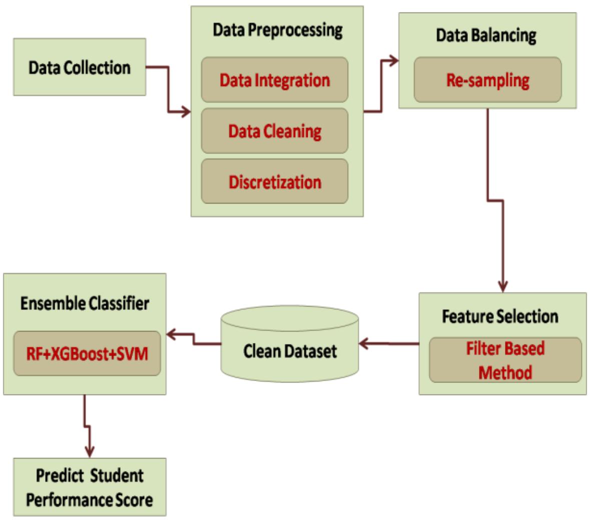

Fig. 1: System Overview

|ISSN: 2322-0163| www.ijeetr.com | A Bimonthly, Peer Reviewed, Scholarly Indexed Journal |

| Volume 8, Issue 2, March - April 2026 |

DOI:10.15662/IJEETR.2026.0802017

Implementation Modules

Data Collection

- Gather information about student learning activities from academic databases or openly accessible educational datasets.
- Kaggle.com provided the dataset.

Data Preprocessing

- Use mean, median, or mode imputation methods to deal with missing values.
- Get rid of records that are inconsistent and duplicate.
- Label encoding or one-hot encoding can be used to encode categorical variables.
- To guarantee a consistent scale, normalize or standardize numerical properties.

Feature Selection and Engineering

- Pick pertinent elements of learning activities that have a big impact on students' performance.
- Develop new features like performance trends or average study consistency.

Train and Test Model

- In this module, the service provider divided the dataset of learning activities into train and test data at a ratio of 70% and 30%, respectively. Thirty percent of the data is regarded as test data, which is used to test the models, and seventy percent is regarded as train data, which is used to train the model.

Prediction and Risk Identification

- Forecast the categories of student performance (e.g., High, Medium, Low).
- Early detection of students who are academically at danger.

Graphical Analysis

- Show the system's accuracy and predicted ratio graphs in this module. The graph analysis takes into account a number of parameters. Plot charts such as bar charts and others at this phase.

Implementation Algorithms

Support Vector Machine

- Support-vector machines (SVMs, also called support-vector networks) are supervised learning models in machine learning that analyze data for classification and regression using related learning techniques. An SVM training method builds a model that assigns fresh samples to either category, making it a non-probabilistic binary linear classifier.

Logistic Regression

- Under the heading of supervised learning, logistic regression is one of the most used machine learning algorithms. The categorical dependent variable is predicted using a specific set of independent factors.
- Logistic regression predicts the output of a categorical dependent variable. Consequently, the outcome must be a discrete or category value. It is possible to use yes or no, 0 or 1, true or false, etc., however probabilistic values between 0 and 1 are given instead of the exact values of 0 and 1.

XGBoost

- XGBoost is a distributed gradient boosting library that has been developed for effective and scalable machine learning model training.
- It is an ensemble learning technique that generates a stronger prediction by combining the predictions of several weak models.
- XGBoost, which stands for "Extreme Gradient Boosting," is one of the most well-liked and extensively applied machine learning algorithms because of its capacity to manage sizable datasets and attain cutting-edge results in a variety of machine learning tasks, including regression and classification.

International Journal of Engineering & Extended Technologies Research (IJEETR)

|ISSN: 2322-0163| www.ijeetr.com | A Bimonthly, Peer Reviewed, Scholarly Indexed Journal |

| Volume 8, Issue 2, March - April 2026 |

DOI:10.15662/IJEETR.2026.0802017

IV. RESULTS

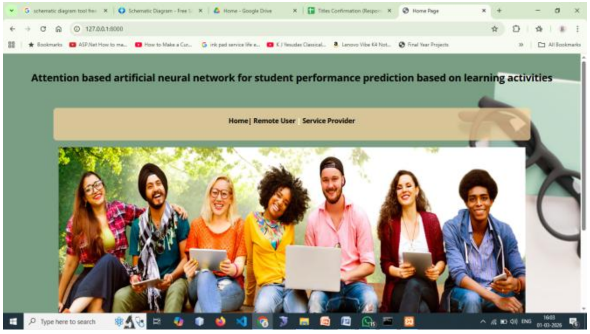

Fig. 2: Home Page

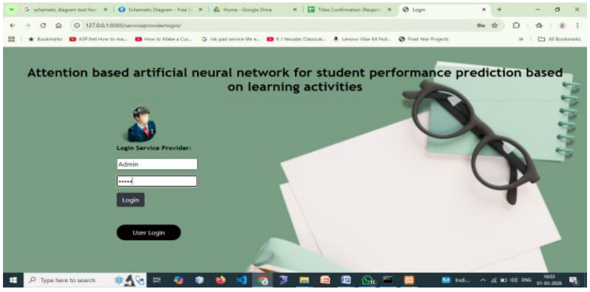

Fig. 3:Service Provider Login

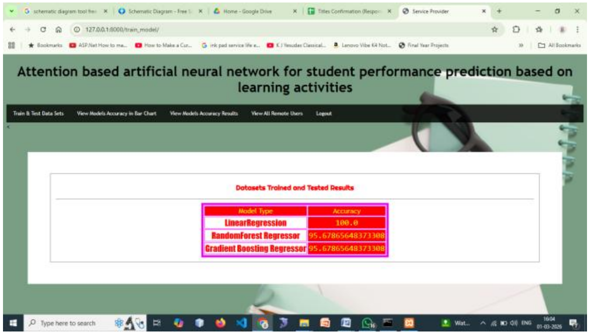

Fig. 4: Model Accuracy

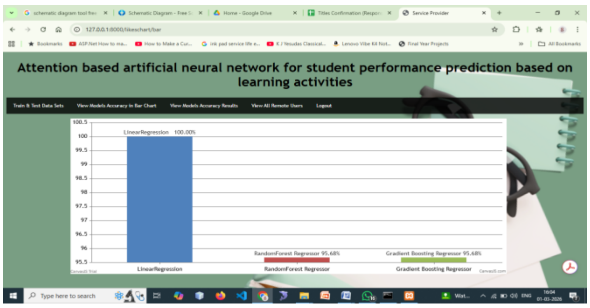

Fig. 5: Model Accuracy Results

|ISSN: 2322-0163| www.ijeetr.com | A Bimonthly, Peer Reviewed, Scholarly Indexed Journal |

| Volume 8, Issue 2, March - April 2026 |

DOI:10.15662/IJEETR.2026.0802017

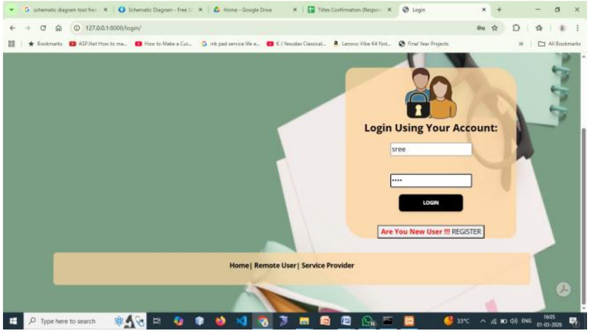

Fig. 6: User Login

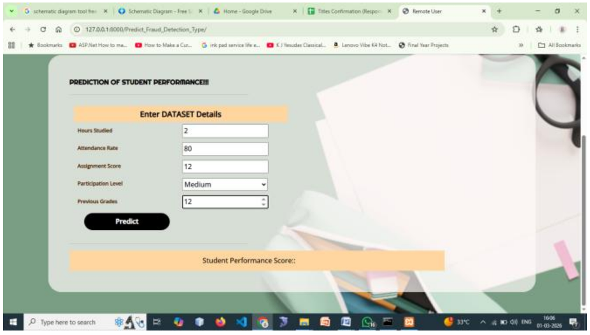

Fig. 6: Predict Score

V. CONCLUSION

Researchers are investigating a variety of EL methods in order to create a trustworthy model that can accurately forecast students' academic performance. Additional hybrid EL methods for feature selection, like as voting and stacking, are also being investigated for potential implementation with CAE. The results of this study show that algorithms that use hybridized EL and the correlation attribute evaluator produce better classification results. This is in contrast to algorithms that do not utilize feature selection or EL. Accuracy in classification or prediction can be greatly increased by using ensemble learning in combination with feature selection. More precisely, the method's conceptual framework is divided into four different layers, each of which addresses a different problem. Following their analysis and application of that, the writers have reached the following conclusion: Stacking and voting are two instances of heterogeneous ensemble techniques that routinely perform better than conventional classification algorithms. Furthermore, methods for feature selection, like the correlation attribute evaluator, are employed to obtain the most accurate prediction result.

REFERENCES

[1] K. Muldner, M. Wixon, D. Rai, W. Burleson, B. P. Woolf, and I. Arroyo, ‘‘Exploring the impact of a learning dashboard on student affect,’’ in Proc. Int. Conf. Artif. Intell. Educ., 2015, pp. 1–28.
[2] Y. A. Saputra, M. L. Singgih, E. Latiffianti, and E. Suryani, ‘‘A study on course management system implementation in Indonesian higher education institutions,’’ Proc. IOP Conf. Ser., Mater. Sci. Eng., vol. 337, no. 1, Apr. 2018, Art. no. 012059.
[3] X. Bai, F. Zhang, J. Li, T. Guo, A. Aziz, A. Jin, and F. Xia, ‘‘Educational big data: Predictions, applications and challenges,’’ Big Data Res., vol. 26, Nov. 2021, Art. no. 100270.
[4] L. Calvet Liñán and Á. A. Juan Pérez, ‘‘Educational data mining and learning analytics: Differences, similarities, and time evolution,’’ RUSC. Universities Knowl. Soc. J., vol. 12, no. 3, p. 98, Jul. 2015.
[5] M.Hlosta, Z. Zdrahal, and J. Zendulka, Ouroboros: Early Identification of At-Risk Students Without Models Based on Legacy Data. New York, NY, USA: Association for Computing Machinery, 2017, pp. 6–15.
[6] F. Marbouti, H. A. Diefes-Dux, and K. Madhavan, ‘‘Models for early prediction of at-risk students in a course using standards-based grading,’’ Comput. Educ., vol. 103, pp. 1–15, Dec. 2016.

[7] Y. Zhang, Y. Yun, R. An, J. Cui, H. Dai, and X. Shang, ‘‘Educational data mining techniques for student performance prediction: Method review and comparison analysis,’’ Frontiers Psychol., vol. 12, pp. 1–20, Dec. 2021.

[8]C.Nagarajan and M.Madheswaran - ‘Stability Analysis of Series Parallel Resonant Converter with Fuzzy Logic Controller Using State Space Techniques’- Taylor &Francis, Electric Power Components and Systems, Vol.39 (8), pp.780-793, May 2011. DOI: 10.1080/15325008.2010.541746

[9]C.Nagarajan and M.Madheswaran - ‘Experimental verification and stability state space analysis of CLL-T Series Parallel Resonant Converter’ - Journal of Electrical Engineering, Vol.63 (6), pp.365-372, Dec.2012. DOI: 10.2478/v10187-012-0054-2

[10]C.Nagarajan and M.Madheswaran - ‘Performance Analysis of LCL-T Resonant Converter with Fuzzy/PID Using State Space Analysis’- Springer, Electrical Engineering, Vol.93 (3), pp.167-178, September 2011. DOI 10.1007/s00202-011-0203-9

[11]S.Tamilselvi, R.Prakash, C.Nagarajan,“Solar System Integrated Smart Grid Utilizing Hybrid Coot-Genetic Algorithm Optimized ANN Controller” Iranian Journal Of Science And Technology-Transactions Of Electrical Engineering, DOI10.1007/s40998-025-00917-z,2025

[12]S.Tamilselvi, R.Prakash, C.Nagarajan,“ Adaptive sliding mode control of multilevel grid-connected inverters using reinforcement learning for enhanced LVRT performance” Electric Power Systems Research 253 (2026) 112428, doi.org/10.1016/j.epsr.2025.112428

[13]S.Thirunavukkarasu, C. Nagarajan, 2024, “Performance Investigation on OCF and SCF study in BLDC machine using FTANN Controller," Journal of Electrical Engineering And Technology, Volume 20, pages 2675–2688, (2025), doi.org/10.1007/s42835-024-02126-w

[14]C. Nagarajan, M.Madheswaran and D.Ramasubramanian- ‘Development of DSP based Robust Control Method for General Resonant Converter Topologies using Transfer Function Model’- Acta Electrotechnica et Informatica Journal , Vol.13 (2), pp.18-31,April-June.2013, DOI: 10.2478/aeei-2013-0025.

[15]C.Nagarajan and M.Madheswaran - ‘DSP Based Fuzzy Controller for Series Parallel Resonant converter’- Springer, Frontiers of Electrical and Electronic Engineering, Vol. 7(4), pp. 438-446, Dec.12. DOI 10.1007/s11460-012-0212-0.

[16]C.Nagarajan and M.Madheswaran - ‘Experimental Study and steady state stability analysis of CLL-T Series Parallel Resonant Converter with Fuzzy controller using State Space Analysis’- Iranian Journal of Electrical & Electronic Engineering, Vol.8 (3), pp.259-267, September 2012.

[17]C.Nagarajan and M.Madheswaran, “Analysis and Simulation of LCL Series Resonant Full Bridge Converter Using PWM Technique with Load Independent Operation” has been presented in ICTES’08, a IEEE / IET International Conference organized by M.G.R.University, Chennai.Vol.no.1, pp.190-195, Dec.2007

[18]Suganthi Mullainathan, Ramesh Natarajan, “An SPSS and CNN modelling based quality assessment using ceramic materials and membrane filtration techniques”, Revista Materia (Rio J.) Vol. 30, 2025, DOI: https://doi.org/10.1590/1517-7076-RMAT-2024-0721

[19]M Suganthi, N Ramesh, “Treatment of water using natural zeolite as membrane filter”, Journal of Environmental Protection and Ecology, Volume 23, Issue 2, pp: 520-530,2022

[20] M. Riestra-González, M. D. P. Paule-Ruíz, and F. Ortin, ‘‘Massive LMS log data analysis for the early prediction of course-agnostic student performance,’’ Comput. Educ., vol. 163, Apr. 2021, Art. no. 104108.

[21] A. Rivas, A. González-Briones, G. Hernández, J. Prieto, and P. Chamoso, ‘‘Artificial neural network analysis of the academic performance of students in virtual learning environments,’’ Neurocomputing, vol. 423, pp. 713–720, Jan. 2021.

[22] S. Leelaluk, T. Minematsu, Y. Taniguchi, F. Okubo, and A. Shimada, ‘‘Predicting student performance based on lecture materials data using neural network models,’’ in Proc. 4th Workshop Predicting Perform. Based Anal. Reading Behav., 2022, pp. 1–18.