IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Analyzing Blackboard Interactions For Forecasting Learning Outcomes.

S. Venkata Naresh[1] **Dr. P. Nirupuma**[2]

[1]PG student, Vemu Institute of Technology, P. kothakota.
[2]Professor, Vemu Institute of Technology, P. kothakota.

ABSTRACT:

This study delves into how student engagement with Blackboard, a prevalent Learning Management System in higher education, shapes their academic achievements. Employing a mixed-methods approach, it examines four deep learning models to predict student performance using Key Performance Indicators (KPIs) derived from Blackboard data. Analyzing data from seven courses, potential predictive KPIs are identified through documentary analysis. Correlational studies unveil significant associations between these factors and student performance metrics. The study demonstrates the efficacy of a combined CNN-LSTM predictive model in accurately forecasting outcomes. These findings advocate for the utilization of such models to optimize Blackboard's utility and bolster educational interventions in universities. Additionally, the integration of BI-LSTM extends the study's accuracy range to 95-100%, showcasing its potential in refining predictive capabilities further.

Keywords: Learning Outcomes, CNN, LSTM

INTRODUCTION:

In the realm of computer security, the detection of intrusions stands as a critical defense mechanism against unauthorized access to network systems, safeguarding the integrity, confidentiality, and availability of crucial data. Despite diligent efforts to fortify computer networks, the relentless surge of cyber threats persists, necessitating the evolution of more advanced detection tools. Two primary approaches emerge in the realm of intrusion detection: signature-based detection, reliant on historical attack data, and anomaly-based detection, which scrutinizes network behavior for irregularities. Leveraging Machine Learning (ML), particularly for anomaly detection, presents a promising avenue to bolster cybersecurity defenses. However, challenges such as misclassification and processing overheads prompt exploration into Deep Learning (DL) within the broader context of Artificial Intelligence (AI), aiming to refine intrusion detection systems for heightened accuracy and efficiency.

LITERATURE SURVEY

(N. Simanullang and J. Rajagukguk)et al

As author proposes a quasi-experimental study to assess the impact of Moodle, a leading Learning Management System (LMS), on student learning activities in online education. Moodle's versatile features, including video integration, discussion forums, chat, materials, and quizzes, make it ideal for facilitating diverse learning experiences. By focusing on student engagement within Moodle, the research aims to understand the effectiveness of LMS-based education. Through this investigation, the author seeks to gain insights into how students utilize Moodle's features for learning in an online environment, thereby contributing to a deeper understanding of effective online education strategies.

(R. M., N. F., and A. A)et al

The author suggests a framework for forecasting the academic success of first-year bachelor's students in computer science courses by leveraging data mining techniques on educational databases. Utilizing Decision Tree, Naïve Bayes, and Multi-Layer Perceptron classification methods via the WEKA Data Mining tool, the paper aims to establish an optimal prediction model for student performance. Through experimentation, the study identifies the most effective model and computes its accuracy. The insights gleaned from this predictive model will inform personalized student profiling, aiding educators in assessing students' potential success levels in their initial semester. This approach promises to enhance early intervention strategies and tailor educational support to individual student needs.

(M. J. Parker)et al

As an exploration of Biology 1310/1312 courses, pivotal for fulfilling core requirements at the University of Houston-Downtown, where over 14,000 undergraduates opt for these science electives. Recognizing them as crucial "gateway" courses, the study aims to enhance retention and graduation rates by fostering meaningful learning experiences for non-science majors. Leveraging BlackBoard Learn analytics, the research delves into quantifying engagement levels in online courses. By correlating engagement data with individual learner performance, the study seeks to ascertain the predictive value of online activity for success. This review underscores the significance of engagement/activity metrics in shaping both student and course learning outcomes in the online education landscape.

PROBLEM STATEMENT:

Studies predicting student success in offline education have typically collected measurements using validated questionnaires, interviews, and observational techniques, with relevant theoretical concepts in mind so that the measurement can be geared towards the concepts that the researcher thinks need to be measured. Many existing

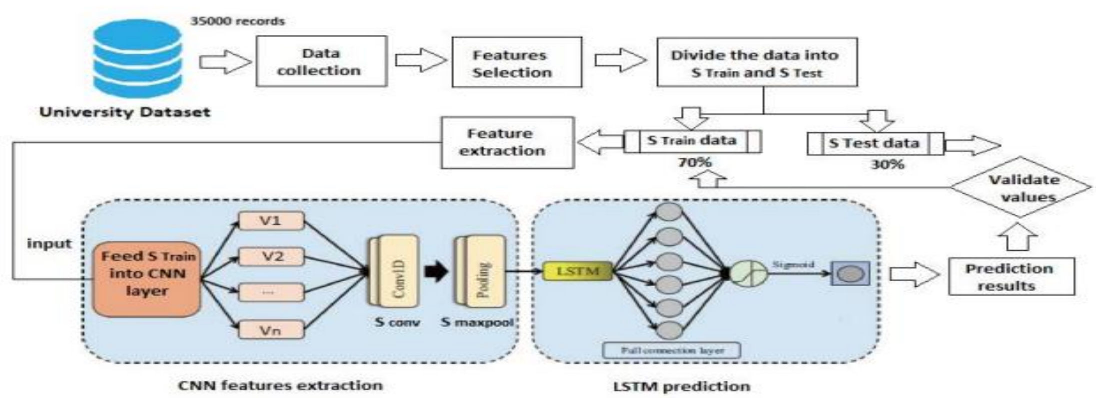
Machine and Deep learning algorithms are available but its prediction accuracy is not accurate.

PROPOSED METHOD:

In this paper author employing combination of CNN and LTSM to predict student Key Performance Indication (KPI) based on past performance such as interaction with teacher, number of hours studying etc. Predicting student performance help educational institution in improving their quality of education and spend more time on low performing students. So it helps both education institutions and students in improving their performance.

In propose paper author training dataset with CNN and then extracting trained features from CNN (as CNN use important features while training a model) and then retraining LSTM on extracted features. LSTM training in important features so its accuracy will be automatically high.

ARCHITECTURE



STUDENT PERFORMANCE DATASET:

```
gender,NationalITy,PlaceofBirth,StageID,GradeID,SectionID,Topic,Semester,Relation,raisedhands,VisITedResources,Announceme
M,KW,KuwaIT,lowerlevel,G-04,A,Online Learning,F,Father,15,16,2,20,Yes,Good,Under-7,M
M,KW,KuwaIT,lowerlevel,G-04,A,Online Learning,F,Father,20,20,3,25,Yes,Good,Under-7,M
M,KW,KuwaIT,lowerlevel,G-04,A,Online Learning,F,Father,10,7,0,30,No,Bad,Above-7,L
M,KW,KuwaIT,lowerlevel,G-04,A,Online Learning,F,Father,30,25,5,35,No,Bad,Above-7,L
M,KW,KuwaIT,lowerlevel,G-04,A,Online Learning,F,Father,40,50,12,50,No,Bad,Above-7,M
F,KW,KuwaIT,lowerlevel,G-04,A,Online Learning,F,Father,42,30,13,70,Yes,Bad,Above-7,M
M,KW,KuwaIT,MiddleSchool,G-07,A,Math,F,Father,35,12,0,17,No,Bad,Above-7,L
M,KW,KuwaIT,MiddleSchool,G-07,A,Math,F,Father,50,10,15,22,Yes,Good,Under-7,M
F,KW,KuwaIT,MiddleSchool,G-07,A,Math,F,Father,12,21,16,50,Yes,Good,Under-7,M
F,KW,KuwaIT,MiddleSchool,G-07,B,IT,F,Father,70,80,25,70,Yes,Good,Under-7,M
M,KW,KuwaIT,MiddleSchool,G-07,A,Math,F,Father,50,88,30,80,Yes,Good,Under-7,H
M,KW,KuwaIT,MiddleSchool,G-07,B,Math,F,Father,19,6,19,12,Yes,Good,Under-7,M
M,KW,KuwaIT,lowerlevel,G-04,A,IT,F,Father,5,1,0,11,No,Bad,Above-7,L
M,lebanon,lebanon,MiddleSchool,G-08,A,Math,F,Father,20,14,12,19,No,Bad,Above-7,L
F,KW,KuwaIT,MiddleSchool,G-08,A,Math,F,Mum,62,70,44,60,No,Bad,Above-7,H
F,KW,KuwaIT,MiddleSchool,G-06,A,IT,F,Father,30,40,22,66,Yes,Good,Under-7,M
M,KW,KuwaIT,MiddleSchool,G-07,B,IT,F,Father,36,30,20,80,No,Bad,Above-7,M
M,KW,KuwaIT,MiddleSchool,G-07,A,Math,F,Father,55,13,35,90,No,Bad,Above-7,M
F,KW,KuwaIT,MiddleSchool,G-07,A,IT,F,Mum,69,15,36,96,Yes,Good,Under-7,M
M,KW,KuwaIT,MiddleSchool,G-07,B,IT,F,Mum,70,50,40,99,Yes,Good,Under-7,H
F,KW,KuwaIT,MiddleSchool,G-07,A,IT,F,Father,60,60,33,90,No,Bad,Above-7,M
F,KW,KuwaIT,MiddleSchool,G-07,B,IT,F,Father,10,12,4,80,No,Bad,Under-7,M
M,KW,KuwaIT,MiddleSchool,G-07,A,IT,F,Father,15,21,2,90,No,Bad,Under-7,M
M,KW,KuwaIT,MiddleSchool,G-07,A,IT,F,Father,2,0,2,50,No,Bad,Above-7,L
```

In above student performance dataset first row contains column names and remaining rows contains dataset values as student MARKS and other details and in last column we have student performance label as H (high), L (low) and M (medium). So by using above dataset we will trained a model and then we will apply test data on trained model to predict student performance.

METHODOLOGY:

Data Collection and Preparation

Data collection begins with accessing the Student Performance dataset sourced from Kaggle. This dataset encompasses various attributes such as student marks, interactions, and performance labels categorized as High, Low, or Medium.

Data Preprocessing

In preparation for modeling, non-numeric data is encoded using Label Encoding to convert categorical variables into numerical format. Additionally, data normalization is implemented to standardize numeric features, ensuring uniformity across the dataset.

Model Training

The project adopts three distinct models: Convolutional Neural Network (CNN), CNN-LSTM, and Bidirectional LSTM (BI-LSTM). Initially, the CNN model is trained on the dataset to extract significant features. These extracted features are then utilized to retrain the LSTM model for improved performance. The BI-LSTM, an advanced variant of LSTM, is employed as an extension to further enhance accuracy.

Model Evaluation

Evaluation of the trained models involves several metrics, including Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), training loss, and accuracy. The accuracy metric serves as a primary indicator of model effectiveness, with lower RMSE, MAPE, and loss values signifying superior performance.

Results Analysis

Upon evaluation, the CNN model achieves an accuracy range of 90% to 93%, while the CNN-LSTM model demonstrates improved accuracy, ranging between 95% to 98%. The introduction of the BI-LSTM extension yields even higher accuracy, ranging from 95% to 100%, accompanied by RMSE and MAPE values reaching 0%. Comparative analysis of accuracy and loss across models offers insights into their respective strengths and weaknesses.

Visualization and Interpretation

To facilitate interpretation, accuracy and loss graphs are plotted for each model, allowing for visual assessment of their performance. These visualizations aid in understanding the comparative efficacy of CNN, CNN-LSTM, and BI-LSTM in predicting student performance based on the provided dataset.

EVALUATION:

Accuracy

To calculate accuracy, use the accuracy_score function from sklearn.metrics. This function compares the predicted labels with the true labels and returns the accuracy.

Calculate accuracy for CNN

cnn_accuracy = accuracy_score(predicted_label_cnn, true_label)

print("CNN Accuracy = ", cnn_accuracy)

Calculate accuracy for Propose CNN-LSTM

propose_cnn_lstm_accuracy = accuracy_score(predicted_label_propose_cnn_lstm, true_label)

print("Propose CNN-LSTM Accuracy = ", propose_cnn_lstm_accuracy)

Calculate accuracy for Extension CNN-Bidirectional-LSTM

extension_bilstm_accuracy = accuracy_score(predicted_label_extension_bilstm, true_label)

print("Extension CNN-Bidirectional-LSTM Accuracy = ", extension_bilstm_accuracy)

Mean Squared Error (MSE):

MSE is calculated by taking the average of the squared differences between the predicted values and the actual values.

$$\mathrm{MSE} = \frac{1}{n} \sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2$$

 # Calculate Mean Squared Error (MSE)

defmean_squared_error(actual, predicted):

mse = np.mean((actual - predicted)**2)

returnmse

Root Mean Squared Error (RMSE):

RMSE is the square root of the MSE and provides a measure of the average magnitude of the errors in the predictions.

$$\mathrm{RMSE} = \sqrt{\mathrm{MSE}}$$

 # Calculate Root Mean Squared Error (RMSE)

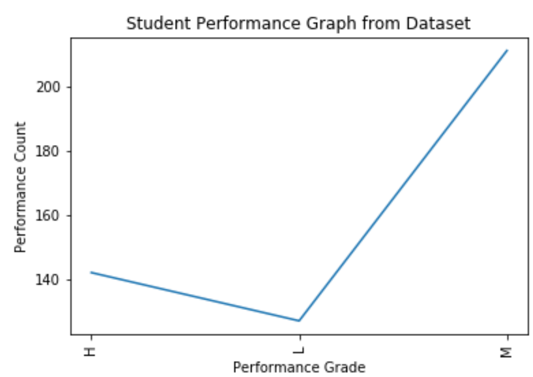
defroot_mean_squared_error(actual, predicted):

mse = mean_squared_error(actual, predicted)

rmse = np.sqrt(mse)

returnrmse

RESULTS:



In above graph x-axis represents labels as H, L and M and y-axis represents number of students with high, low and medium performance available in dataset which we are plotting as line graph

```
CNN Accuracy = 0.9375
CNN RMSE     = 0.46770717334674267
CNN MAPE     = 10.416666666666668
CNN Loss Value = 0.3278464161946128
```

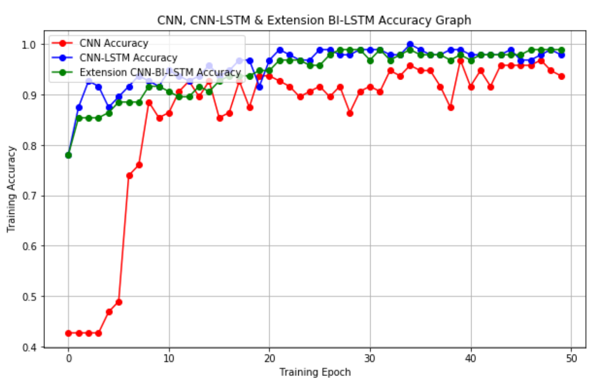
CNN accuracy as 93% and we can see other metrics also

```
Propose CNN-LSTM Accuracy = 0.9895833333333334
Propose CNN-LSTM RMSE     = 0.10206207261596575
Propose CNN-LSTM MAPE     = 0.5208333333333333
Propose CNN-LSTM Loss Value = 0.10764585249550389
```

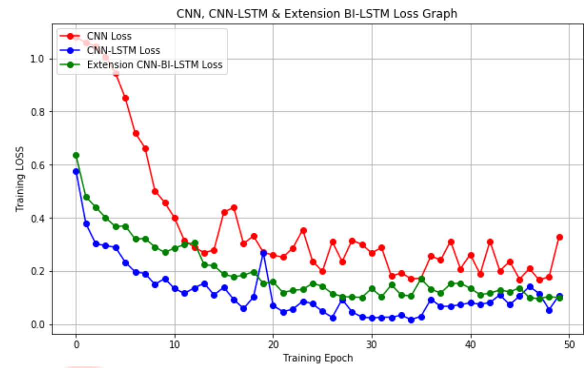
CNN-LSTM we got 98% accuracy and we can see other metric output also

```
Extension CNN-Bidirectional-LSTM Accuracy = 1.0
Extension CNN-Bidirectional-LSTM RMSE     = 0.0
Extension CNN-Bidirectional-LSTM MAPE     = 0.0
Extension CNN-Bidirectional-LSTM Loss Value = 0.09878107971235295
```

Extension BI-LSTM we got 100% accuracy and RMSE and MAPE as 0 on test data. So extension can get accuracy between 95 to 100%



CNN, CNN-LSTM and CNN-BILSTM training accuracy graph where x-axis represents training EPOCH and y-axis represents accuracy and red line is for CNN accuracy and blue line is for CNN-LSTM and green line is for extension CNN-BILSTM and we can see extension green line is little higher than other 2 algorithms



Plotting LOSS graph and we can see extension green line is closer to propose CNN-LSTM so for any algorithm LOSS must be less and accuracy must be high

	Algorithm Name	Accuracy	RMSE	MAPE	LOSS
0	CNN	0.937500	0.467707	10.416667	0.327846
1	Propose CNN-LSTM	0.989583	0.102062	0.520833	0.107646
2	Extension CNN-BI-LSTM	1.000000	0.000000	0.000000	0.098781

CNN, CNN-LSTM and Extension CNN-BILSTM accuracy and other metric and we can see extension got high accuracy and less RSME and other values

PREDCTION:

```
Student Test Data ['M' 'KW' 'KuwaIT' 'lowerlevel' 'G-05' 'A' 'English' 'F' 'Father' 7 10 1
 30 'No' 'Bad' 'Above-7'] Predicted Performance ====> L

Student Test Data ['F' 'KW' 'KuwaIT' 'HighSchool' 'G-12' 'A' 'English' 'F' 'Mum' 70 4 39 90
 'Yes' 'Good' 'Under-7'] Predicted Performance ====> H

Student Test Data ['F' 'KW' 'KuwaIT' 'HighSchool' 'G-12' 'A' 'English' 'F' 'Mum' 13 80 40 88
 'Yes' 'Good' 'Under-7'] Predicted Performance ====> H

Student Test Data ['M' 'KW' 'KuwaIT' 'HighSchool' 'G-09' 'A' 'IT' 'F' 'Father' 20 80 33 33
 'Yes' 'Good' 'Under-7'] Predicted Performance ====> M

Student Test Data ['M' 'KW' 'KuwaIT' 'HighSchool' 'G-11' 'A' 'Quran' 'F' 'Father' 13 3 11 9
 'No' 'Bad' 'Above-7'] Predicted Performance ====> L

Student Test Data ['F' 'lebanon' 'lebanon' 'MiddleSchool' 'G-07' 'B' 'Math' 'F' 'Mum' 80 90
 49 55 'Yes' 'Bad' 'Under-7'] Predicted Performance ====> H
```

In above screen in square bracket we can see student TEST data and after arrow symbol =➔ we can see predicted performance as 'H or L or M'. H means high and L means Low and M means Medium

CONCLUSION

Integration of CNN and LSTM in predicting student Key Performance Indicators (KPI) offers significant potential for educational institutions and students alike. By leveraging past performance metrics, such as teacher interactions and study hours, the model enhances educational quality and directs attention to underperforming students. While existing machine learning algorithms have limitations in accuracy, the combined CNN-LSTM approach presented in this paper shows promise. Evaluation metrics including RMSE, MAPE, and training loss underscore the effectiveness of the model, with CNN-LSTM achieving an accuracy of 95-98%. Furthermore, extending the model to BI-LSTM enhances accuracy to 95-100%, offering a robust framework for performance prediction in education.

REFERENCES:

[1] N. Simanullang and J. Rajagukguk, ''Learning management system (LMS) based on Moodle to improve students' learning activity,'' J. Phys., Conf., vol. 1462, no. 1, pp. 120–144, 2020.

[2] A. Aldiab, H. Chowdhury, A. Kootsookos, F. Alam, and H. Allhibi, ''Utilization of learning management systems (LMSs) in higher education system: A case review for Saudi Arabia,'' Energy Proc., vol. 160, pp. 731–737, Feb. 2019.

[3] N. Kerimbayev, J. Kultan, S. Abdykarimova, and A. Akramova, ''LMS Moodle: Distance international education in cooperation of higher education institutions of different countries,'' Educ. Inf. Technol., vol. 22, no. 5, pp. 2125–2139, Sep. 2017.

[4] R. M., N. F., and A. A., ''Predicting and analysis of students' academic performance using data

mining techniques,'' Int. J. Comput. Appl., vol. 182, no. 32, pp. 1–6, Dec. 2018.

[5] M. J. Parker, ''Can BlackBoard engagement (Activity) suggest connections to student success in an online course and offer predictive value?'' in Proc. E-Learn, World Conf. E-Learn. Corporate, Government, Healthcare, Higher Educ., 2018, pp. 1284–1289.

[6] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning. Cambridge, MA, USA: MIT Press, 2016.

[7] Y. LeCun, Y. Bengio, and G. Hinton, ''Deep learning,'' Nature, vol. 521, no. 7553, pp. 436–444, Sep. 2015.

[8] R. Conijn, C. Snijders, A. Kleingeld, and U. Matzat, ''Predicting student performance from LMS data: A comparison of 17 blended courses using Moodle LMS,'' IEEE Trans. Learn. Technol., vol. 10, no. 1, pp. 17–29, Jan./Mar. 2017.

[9] Z. Unal and A. Unal, ''Investigating and comparing user experiences of course management systems: BlackBoard vs. Moodle,'' J. Interact. Learn. Res., vol. 25, no. 1, pp. 101–123, 2014.

[10] A. Ali and K. A. Sheirah, ''Utilizing the blackboard learning management system at the university of Ha'il from the perspective of faculty members using Internet of Things (IoT),'' Int. J. Educ. Inf. Technol., vol. 14, pp. 214–222, Feb. 2021.

[11] J. Uziak, M. T. Oladiran, E. Lorencowicz, and K. Becker, ''Students' and instructor's perspective on the use of blackboard platform for delivering an engineering course,'' Electron. J. e-Learn., vol. 16, no. 1, pp. pp1–15, Feb. 2018.

[12] J. A. Alokluk, ''The effectiveness of blackboard system, uses and limitations in information management,'' Intell. Inf. Manage., vol. 10, no. 6, p. 133, 2018.

[13] I. Osman, M. Nasir, and R. Alzoubi, ''Blackboard usage: An investigative study among CCSE female faculty staff and students at University of Hail,'' J. Econ. Manage. Perspect., vol. 11, no. 2, pp. 508–515, 2017.

[14] D. Carmichael and C. MacEachen, ''Heuristic evaluation of the use of Blackboard & Facebook groups in computing higher education,'' Int. J. Modern Educ. Comput. Sci., vol. 9, no. 6, pp. 1–8, Jun. 2017.

[15] M. J. Verdu, J.-P. De Castro, L. M. Regueras, and A. Corell, ''MSocial: Practical integration of social learning analytics into Moodle,'' IEEE Access, vol. 9, pp. 23705–23716, 2021.

[16] C. Greenhow and E. Askari, ''Learning and teaching with social network sites: A decade of research in K-12 related education,'' Educ. Inf. Technol., vol. 22, no. 2, pp. 623–645, Mar. 2017.

[17] M. D. Lytras, N. R. Aljohani, A. Visvizi, P. Ordonez De Pablos, and D. Gasevic, ''Advanced decision-making in higher education: Learning analytics research and key performance indicators,'' Behav. Inf. Technol., vol. 37, nos. 10–11, pp. 937–940, Nov. 2018.

[18] X. Zhang, Y. Meng, P. O. de Pablos, and Y. Sun, ''Learning analytics in collaborative learning supported by slack: From the perspective of engagement,'' Comput. Hum. Behav., vol. 92, pp. 625–633, Mar. 2019.

[19] A. Van Barneveld, K. E. Arnold, and J. P. Campbell, ''Analytics in higher education: Establishing a common language,'' Educause Learn. Initiative, vol. 1, no. 1, pp. 1–2, 2012.

[20] R. Ferguson, ''Learning analytics: Drivers, developments and challenges,'' Int. J. Technol. Enhanced Learn., vol. 4, nos. 5–6, pp. 304–317, 2012.