Webology (ISSN: 1735-188X)
Volume 18, Number 6, 2021

Time Division Ensemble Learning Classifier For Efficient Student Performance Analysis Using Convolution Neural Network

[1]Rasheed Mansoor Ali S , [2]Dr.S.Perumal

[1]Research Scholar , [2]Associate Professor

[1,2]Department of computer Science,

[1,2]Vels Institute of Science, Technology and Advanced Studies (VISTAS), Chennai, India.

Abstract:

The problem of performance analysis and ranking of students has been well studied. There are number of methods available and suffer to achieve higher performance in various factors. To improve the performance, an efficient Time Division Multi Feature Ensemble Learning Classifier (TDMFEL) model is presented. The method maintains the history of students' academic activity and performance achievements. The model monitors the academic activities and their performance in various corners. The history contains information related to each student which includes, number of seminars given, subject orient achievement, aptitude strength, extra curricular activities, knowledge on external world, leadership quality, presentation strength, number of admissions, number of placements and so on. Also, the model maintains the student records both on direct and online classes. Such traces are pre processed using Mandate Feature Analysis (MFA) to identify incomplete records. Further, the model extracts set of features like subject, score, learning activity, sports activity, seminar involvement, aptitude, achievements, external knowledge, number of presentations, number of appearances and so on. The features extracted are used in ensemble generation and trained with Convolution Neural Network. At the test phase, the neurons estimates support on Offline Learning Support (OLS), Online Learning Support (On LS), Offline Extra Curricular Support (OECS), Online Extra Curricular Support (On ECS), Sports Support (SS), Achievement Support (AS), Seminar and Presentation Support (SPS). According to the support measures computed, the models classify the student and predict the performance of the student. The proposed method improves the performance of prediction and ranking.

Index Terms: Performance Prediction, Student Performance Analysis, CNN, Ensemble Classifier, TDMFEL, MFA.

1. Introduction:

The technology development gives different shape for the educational sector in recent times. In earlier the students are educated with just using the blackboard teaching, where in recent times it has been given with E-Learning, Online learning and so on. Whatever the way, the students obtains knowledge through various models. But, there is a big difference in the performance of the students in academic, sports and others in terms of mode through which they get educated. The students have great impact from the society which affects their carrier as well as life. So, it is necessary to monitor the performance of the students in different scenario. By monitoring the student's performance, they can be predicted for different factors to support their development.

Data mining techniques are used in several problems and the same can be used in monitoring and analyzing the performance of students. By applying the datamining techniques, the hidden features can be identified to support the prediction process. It contains various approaches to support the growth of educational sector [1]. By analyzing the student performance, the performance of educational institutions can be ranked. Any country would have number of institutions but the performance of the student and the quality of skill studying in the institution differs. So, it is necessary to analyze the performance of students as well as institutions. The data mining techniques can be used in such process towards the development of educational sector. . The main functions of data mining are applying various methods and algorithms in order to discover and extract patterns of stored data [2]. Data mining and knowledge discovery applications have got a rich focus due to its significance in decision making and it has become an essential component in various organizations. There are multiple data classification techniques used for predicting the results each one having its own advantages and disadvantages. Various algorithms and techniques like Classification, Clustering, Regression, Artificial Intelligence, Neural Networks, Association Rules, Decision Trees, Genetic Algorithm, Nearest Neighbor method etc., are used for knowledge discovery from databases. These techniques and methods in data mining need brief mention to have better understanding [3].

The performance of students cannot be predicted just based on the marks obtained but it must be performed by considering various features. There exist several techniques which analyze the student performance with the score obtained but not efficient because, it misses various other features. Similarly, there are number of approaches exist in literature which uses time spent, seminars given, seminars attended and so on. The existing methods would miss several features in predicting the student performance, but it is necessary to consider entire features in analyzing the student performance effectively.

Machine learning algorithms are applied in several scientific problems. The performance analysis of students also can be approached with the same. In this way, the problem can be approached with pattern based approaches, fuzzy logic, linear regression, Support vector machine, artificial intelligence, neural network and so on. This article tends to approach the problem with Ensemble Learner based classification in predicting the student performance. Also analyze the effect of recent pandemic Covid-19 in the performance of student. The model focused on measuring the performance of any student in both the time line like at the physical class room performance, and online class performance. By monitoring the performance of student in both the time line, the performance of the student can be predicted in perfect way. The result of such analysis would help the educational institution in picking set of students for physical classes which is more essential. The detailed approach is presented in this section.

2. Related Works:

The problem of student performance analysis and prediction is approached with different techniques. Such methods are discussed in detail in this section.

In [4] they performed aspect extraction by using the dependency relation between opinion term and noun/noun phrase present within the sentence. They also developed concept ontology for extracting only domain relevant aspects and applied pruning of those aspects having a value less than the specified threshold value. Naive Bayes classifier was used for performing this task and for sentiment analysis they used an online sentiment analyzer. They achieved F-score 0.80 in aspect extraction and 0.72 in sentiment detection. Deep learning techniques are not yet explored in the academic domain. To the best of our knowledge, our study is a first attempt that is using LSTM for performing aspect based sentiment analysis on students' feedback for faculty teaching performance evaluation.

Veena Deshmukh, et al, 2018, implementing a student performance evaluation model using Mamdani Fuzzy Inference System (FIS) and Neuro Fuzzy system and comparing the results with classical averaging method for Network Analysis (NA) course studied by third semester Electronics and Communication Engineering students. This work explains the designing of scoring rubrics using Bloom's levels as the criteria of assessment for NA course. Also at initial stages of learning how students' strengths and weaknesses can be identified using rubrics and develop critical thinking skills. The five inputs identify, understand, apply, analyze and design/create are five levels of learning as per Bloom's Taxonomy. Fuzzy rules are applied and the evaluated results are expressed in both crisp and linguistic variables and compared with classical aggregate scores [5].

Abeer Badr El Din Ahmed et al, [6] currently the amount huge of data stored in educational database these database contain the useful information for predict of students performance. The most useful data mining techniques in educational database is classification. In this paper, the classification task is used to predict the final grade of students and as there are many approaches that are used for data classification, the decision tree (ID3) method is used here.

Students' feedback [7] is crucial for academic institutions in order to evaluate faculty performance. Handling the qualitative opinions of students efficiently while automatic report generation is a challenging task. Indeed, most of the organizations deal with quantitative feedback effectively, whereas qualitative feedback is either processed manually or ignored altogether. This research proposes a supervised aspect based opinion mining system based on two layered LSTM model. The first layer predicts the aspects described within the feedback and later specifies the orientation (positive, negative, neutral) of those predicted aspects. The model was tested on a manually tagged data set constructed from the last five years students' comments from Sukkur IBA University as well as on a standard SemEval-2014 data set. Unlike many other LSTM models proposed for other domains, the proposed model is quite simple in terms of architecture which results in less complexity. The system attains good accuracy using the domain embedding layer in both tasks: aspect extraction (91%) and sentiment polarity detection (93%). To the best of our knowledge, this study is a first attempt that uses deep learning approach for performing aspect based sentiment analysis on students' feedback for evaluating faculty teaching performance.

Isolate data among different campus information systems and not many effective information among the big data generated by these systems cause that it is a challenge for predicting achievement of students. This paper [8] designs a student achievement predicting framework, which includes data processing and student achievement predicting. In the data processing, data extraction, data cleaning, and feature extraction are designed. Using these data in data warehouse, we propose a layer-supervised multi-layer perception based method (LSMLP) to predict the achievement of students. Supervisions are fed to corresponding each hidden layer of MLP to improve the performance of student achievement prediction. Compared with SVM, Naive Bayes, Logistic Regression, and Multilayer Perceptron, our method get better performance.

Asif et.al used educational data to study the performance of undergraduate students. Their results indicated that few key courses impacted the students. In order to improve the performance of student, teachers should pay more attention on these key courses [9]. Jiang et.al used card consumption data to find the specific behavior of student. They discussed the behavior of consumption happened in canteen, supermarket and other place, and then summarized the significance of card analysis system [10].

Li et al [11] proposed a deep learning framework to predict student performance in the course. SPDN (Sequential Prediction based on Deep Network) uses the multi-source fusion CNN technique to predict students' online behavioral sequences and includes static information via bidirectional LSTM. Meanwhile, students' internet usage has a greater influence on their academic success. Though it predicts more student data we can use it for some course prediction only. . Khan et al. [12] published a paper on the estimation of student academic performance using a Bidirectional Long Short-Term Memory network (BiL STM). Mahareek et al [13] implement the simulated annealing algorithm to predict the student performance using SVM with Multilayer perceptron kernel (MLP kernel). Furthermore, several researchers have attempted to enhance student performance prediction using various deep learning approaches, but have yet to achieve a higher accuracy rate. To achieve a high rate of accuracy, we proposed the CNN-Multi-class LDA method for student performance prediction.

Al-Okaily et al [14], the University of Jordan students have experienced a variety of environmental, electronic, and mental challenges owing to COVID-19. Over 220,000 Jordanian university students took part in COVID-19 through an online survey using university websites and portals. Botao et al [15], proposed a paper to classify the accident narratives in construction using the CNN model. It provides vital knowledge to the manager for improving the safety on-site. Togascar et al [16], proposed a paper on detecting pneumonia by combining m RMR and machine learning models. The CNN deep features are applied to DT, kNN, LDA, LR, and SVM machine learning models. According to the findings of this study, deep features provided strong and consistent characteristics for pneumonia identification, and the m RMR technique improved classification efficiency.

In [17], an EKT: Exercise-Aware Knowledge Tracing for Student Performance Prediction model is presented, which explicitly track student's knowledge acquisition on multiple knowledge concepts, we extend EERNN to an explainable Exercise-aware Knowledge Tracing (EKT) framework by incorporating the knowledge concept information, where the student's integrated state vector is now extended to a knowledge state matrix.

In [18], an Hybrid Regression and Multi-Label Classification model is presented which uses multi-label prediction of the influential factors is generated using an optimized self-organizing map. We empirically investigate and demonstrate the effectiveness of our entire approach on seven publicly available and varying datasets.

In [19], a Student Performance Prediction model is presented which uses Blended Learning. The model provides enlightenment for learning analysis and individualized teaching in blended learning. The combination of flipped classroom and SPOC is a good way to implement blended learning, but few studies have verified the predictability of learning performance in such a scenario to explore individualized teaching. Students' behavior in blended learning can be used to predict their learning outcomes, and the implementation method is reproducible.

In [20], Multiple Features Fusion Attention Mechanism Enhanced Deep Knowledge Tracing is presented towards Student Performance Prediction which making full use of both student behavior features and exercise features and combining the attention mechanism with the knowledge tracing model.

In [21], they propose a framework to automatically analyzing opinions of students expressed in reviews. Specifically, the framework relies on aspect-level sentiment analysis and aims to automatically identify sentiment or opinion polarity expressed towards a given aspect related to the MOOC. The proposed framework takes advantage of weakly supervised annotation of MOOC-related aspects and propagates the weak supervision signal to effectively identify the aspect categories discussed in the unlabeled students' reviews.

In [22], a Educational Game Using a Hidden Markov Model is presented towards performance prediction. Analyzing time series and the interaction between the students and the game data can result in valuable information that cannot be gained by only cross-sectional studies of the exams.

In [23], they propose a planned quantitative method for assessing students' gains in terms of programming performance and testing performance. Based on real data collected from students who engaged in our course, we use trend analysis to observe how students' performance has improved over the whole semester. By using correlation analysis, we obtain some interesting findings on how students' programming performance correlates with testing performance, which provides persuasive empirical evidence in integrating software testing practices into an Object-oriented programming curriculum.

In [24], a Nonparametric Analysis of the Effect of Knowledge Integration Activities on Third-Year Undergraduate Performance is conducted. In [25], an Unsupervised Ensemble Clustering Approach for the Analysis of Student Behavioral Patterns is presented which use student behavioral data to discover behavioral patterns. Because the behavioral data produced by students on campus are available in real time without intentional bias, clustering analysis can be relatively efficient and reliable. The proposed framework extracts behavior features from the two perspectives of statistics and entropy and then combines density-based spatial clustering of applications with noise (DBSCAN) and k-means algorithms to discover behavioral patterns.

The proposed model

The proposed model reads the student performance records and data set. The data set obtained has been pre processed with Mandate Feature Analysis (MFA). The pre processed data set has been used in feature extraction which extracts set of features like subject, score, learning activity, sports activity, seminar involvement, aptitude, achievements, external knowledge, number of presentations, number of appearances and so on. Extracted features are used in ensemble generation with Time line ensemble generation (TLEG) algorithm. Further, the model generates the neural network and train the network with the ensembles generated. At the testing phase, the test sample has been used in measuring support on Offline Learning Support (OLS), Online Learning Support (On LS), Offline Extra Curricular Support (OECS), Online Extra Curricular Support (On ECS), Sports Support (SS), Achievement Support (AS), Seminar and Presentation Support (SPS). According to the support measures computed, the models classify the student and predict the performance of the student. The detailed approach is presented in this section.

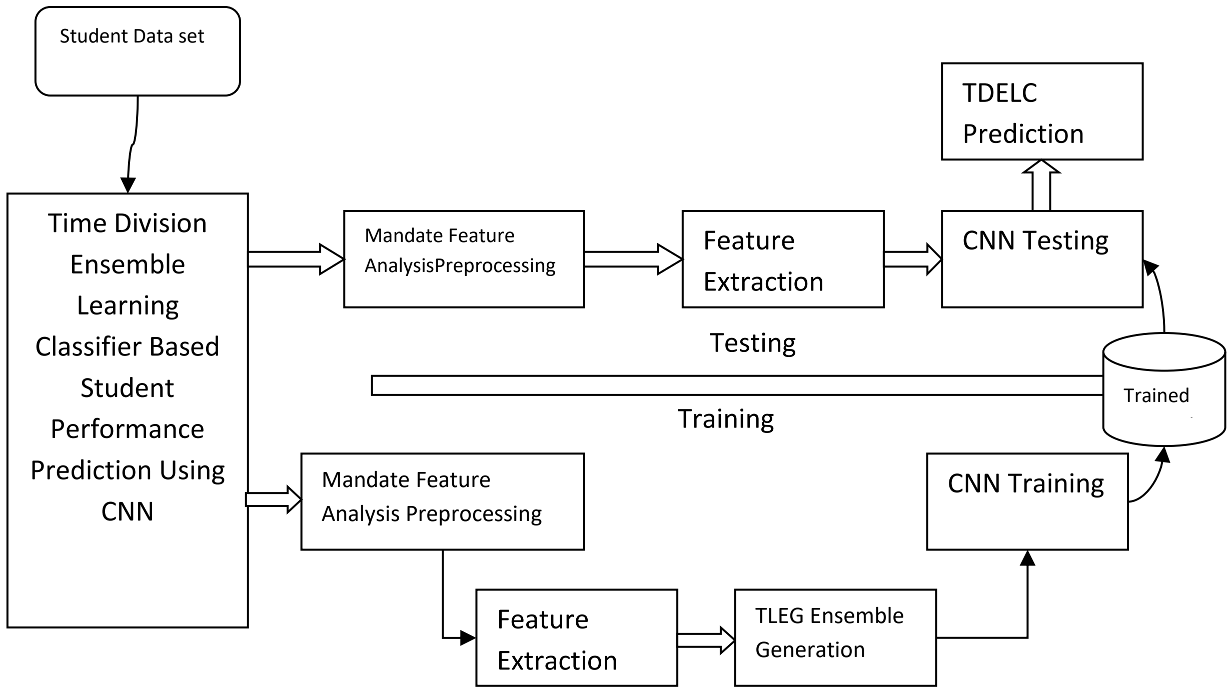

Figure 1: Architecture of Proposed TDELC-CNN Student Performance Prediction Model

The functional architecture of proposed TDELC-CNN model has been presented in Figure 1, where the functional models are detailed in this section.

Mandate Feature Pre processing:

The student data set considered for the performance prediction has been read. Initially, the method identifies set of features present in the overall data set. From the feature list identified, set of mandate features to be contained are identified with the use of data taxonomy. According to that, the method verifies the presence of set of all features to be mandate on each tuple. The tuple identified with missing features are removed from the data set. Second, for each feature identified and selected, the method computes the normalized feature value (NFV) according to the mean, max and minimum. The tuples with incomplete or missing values are set with the value of NFV computed at runtime based on the value of Mean, Min, and Max values. Pre processed data set has been used in student performance prediction in this model.

Algorithm:

Given: Student data set Sds, Feature Taxonomy FT

Obtain: Pre processed set Psds

Start

Read Sds.

Feature set Fes = $(\sum_{i=1}^{Size(Sds)} Features(Sds(i)) \ni Fes) \cup Fes$

for each tuple T

if T$\in \forall(Features(Fes))$ then

Psds = $(\sum Tuples \in Psds) \cup T$

End

End

For each tuple T

For each feature F

If T(F)==Null then

Compute Fe Mean = $\frac{\sum_{i=1}^{Size(Psds)} Psds(i)(F)}{size(Psds)}$

Compute Fe Min = $\underset{i=1}{\overset{Size(Psds)}{Min}}(Psds(i)(f))$

Compute Fe Max = $\underset{i=1}{\overset{Size(Psds)}{Min}}(Psds(i)(f))$

Generate a random R =Rand(1,10)

If R is odd then

Compute Normalized feature value NFV.

NFV = Random(Fe Mean, Fe Max)

Psds(T)(F) = NFV

Else

NFV = Random(Fe Mean, Fe Min)

Psds (T)(F) = NFV

End

End

End

End

Stop

The above discussed algorithm represents how the method performs pre processing of the data set given. The method eliminates the noisy data points and normalize the feature values according to the NFV value computed. The pre processed data set has been used towards student performance analysis.

Feature Extraction:

The pre processed data set has been used in extracting the features from the data set. The method extracts subject, score, learning activity, sports activity, seminar involvement, aptitude, achievements, external knowledge, number of presentations, and number of appearances. Extracted features are converted into to feature vector towards ensemble generation.

Time Line Ensemble Generation:

The data set would contain number of features and traces belong to various time line. The model split the data set into different time line according to the period of log generation. The logs would have generated at the covid pandemic as well as at normal time. So, the method first splits the pre processed logs into two different time line initially. Further, the features from each time line logs are extracted. Extracted features are converted in to ensemble. Generated ensemble set are used in training the neural network.

TLEG Algorithm:

Given: Pre processed data set Prds.

Obtain: Ensemble set Es.

Start

Read pre processed data set Prds.

Split data according to time line.

Normal Tme line set NTls = NTLS ∪ ($\sum_{i=1}^{size(Prds)} Prds(i).Time\ Line == Normal$)

Covid Time Line Set CTLS = CLTS ∪ ($\sum_{i=1}^{size(Prds)} Prds(i).Time\ Line == Covidl$)

For each trace T

Feature set Fs = Feature Extraction (NTLS(T))

Ensemble E = generate pattern with feature set Fs.

Add to ensemble set Es = ($\sum$ Ensembles $\in$ Es) $\cup$ E

End

For each trace T

Feature set Fs = Feature Extraction (CLTS(T))

Ensemble E = generate pattern with feature set Fs.

Add to ensemble set Es = ($\sum$ Ensembles $\in$ Es) $\cup$ E

End

Stop

The above discussed algorithm generates the ensembles according to different time line at which they are generated. Generated time line ensembles are used towards performance analysis.

CNN Training:

The ensembles generated at the TLEG ensemble generation algorithm has been used in training the network. From each set of ensembles, the method generates a neuron for each ensemble. The neurons generated are initialized with the ensemble features. The convolution neural network performs convolution at the feature level by choosing the required features. The features unwanted are eliminated by the neurons. This has been performed by applying the convolution operation at the selective features and the remaining features are given as it is. The neurons are feed by two different matrices where the first one forms the performance features at normal time line and second represents the performance in covid time line. The features are used at the Re LU layer to estimate the similarity measures. The network is framed with seven layers which includes input and output layer. The convolution layer applies convolution operation on the features and the pooling layer estimates weight measure on various features.

Performance Prediction:

The trained network has been used to perform testing towards disease prediction. Each layer of neurons measure the support measures on Offline Learning Support (OLS), Online Learning Support (On LS), Offline Extra Curricular Support (OECS), Online Extra Curricular Support (On ECS), Sports Support (SS), Achievement Support (AS), Seminar and Presentation Support (SPS). The value of offline learning support (OLS) is measured according to the number of times the student attend the class in offline and number of times the student score marks greater than specific threshold. Similarly, the online support score (On LS) is measured according to the number of times the class has been attended through online and number of times the score has been taken more than a threshold. The value of Online Extra curricular support (On ECS) is measured based on the frequency of extra curricular activities made on a session and scores more than threshold. Similarly, the value of online extra curricular support (On ECS) is measured on the same way. The sports support (SS) is measured based on the number of sports activities made on a session and scores obtained than a threshold. Finally, Seminar support (SPS) and Achievement Support (AS) are measured in the same way. According to the support measures, the method computes the value of weight measure to predict the performance of the student.

Performance Prediction Algorithm:

Given: Class Ensemble Set CES, Sample S, Performance Category Set PCS

Obtain: Performance Category C

Start

Read CES and S.

For each category C

Find Class Ensembles CE = $\sum_{i=1}^{Size(CES)} CES(i).Class == C$

Compute Offline Learning Support (OLS) = $\frac{\sum_{i=1}^{size(CE)} CE(i).Type==Offline \;\&\&\; CE(i).Score>TH}{Size(CE)}$

Compute Online Learning Support (On LS) = $\frac{\sum_{i=1}^{size(CE)} CE(i).Type==Online \;\&\&\; CE(i).Score>TH}{Size(CE)}$

Compute Offline Extra Curricular Support (OECS) = $\frac{\sum_{i=1}^{size(CE)} CE(i).Type==Offline \;\&\&\; CE(i).Extra==1 \;\&\&\; CE(i).Score>TH}{Size(CE)}$

Compute Online Extra Curricular Support (OECS) = $\frac{\sum_{i=1}^{size(CE)} CE(i).Type==Online \;\&\&\; CE(i).Extra==1 \;\&\&\; CE(i).Score>TH}{Size(CE)}$

Compute Sports Support SS= $\frac{\sum_{i=1}^{size(CE)} CE(i).Sport==1 \;\&\&\; CE(i).Score>TH}{Size(CE)}$

Compute Achievement Support AS= $\frac{\sum_{i=1}^{size(CE)} CE(i).Achievement==1 \;\&\&\; CE(i).Score>TH}{Size(CE)}$

Compute Seminar and Presentation support SPS= $\frac{\sum_{i=1}^{size(CE)} CE(i).seminar \,||\, CE(i).Presentation==1 \,\&\&\, CE(i).Score>TH}{Size(CE)}$

OLS = Dist (S.Offline, OLS)

On LS = Dist (S.Online, On LS)

OECS = Dist (S.ECA, OECS)

SS = Dist (S.sport, SS)

AS = Dist (S.Achievement, AS)

SPS = Dist (S.seminar, SPS)

Compute weight W = $\frac{OLS \times ONLS \times OECS \times SS \times AS \times SPS}{6}$

End

Class C =Choose the category with maximum weight.

Stop

The above discussed algorithm represents how the student's performance is predicted by computing various support measures. According to the support measures, the method computes the value of weight measures. Based on the values of weight measures the method performs performance prediction.

3. Results and Discussion:

The proposed Time Line Ensemble learner classification based student performance prediction approach has been implemented and evaluated for their performance in several constraints. The evaluation is carried by collecting the student traces from various educational sectors. The performance of the method is measured under various parameters and compared with the results of other methods. The details of data set are presented below:

Parameter	Value
Data Set	Kaggle Data Set
Number of Features	20
Number of Records	10000
Tool	Advanced Java

Table 1: Evaluation Details

The details of data set being used for performance evaluation are presented in Table 1. The evaluation is carried using Kaggle data set available as open source which contains more than 20 features which includes the features of Covid related student traces. Using the data set, the performance of the methods are measured under the following parameters.

Prediction Accuracy Vs No of Records			
Method	3000 Records	5000 Records	10000 Records
SPDN	68.15	71.32	73.51
Attention-based Bi LSTM	71.34	75.9	79.16
SVM(MLP kernel)	73.14	79.9	85.72
TDMFEL	86.5	90.94	96.5

Table 2: Analysis on Performance Prediction Accuracy

The accuracy of predicting the performance of students under various number of records are measured and presented in Table 2. The proposed TDMFEL algorithm has produced higher performance than other methods.

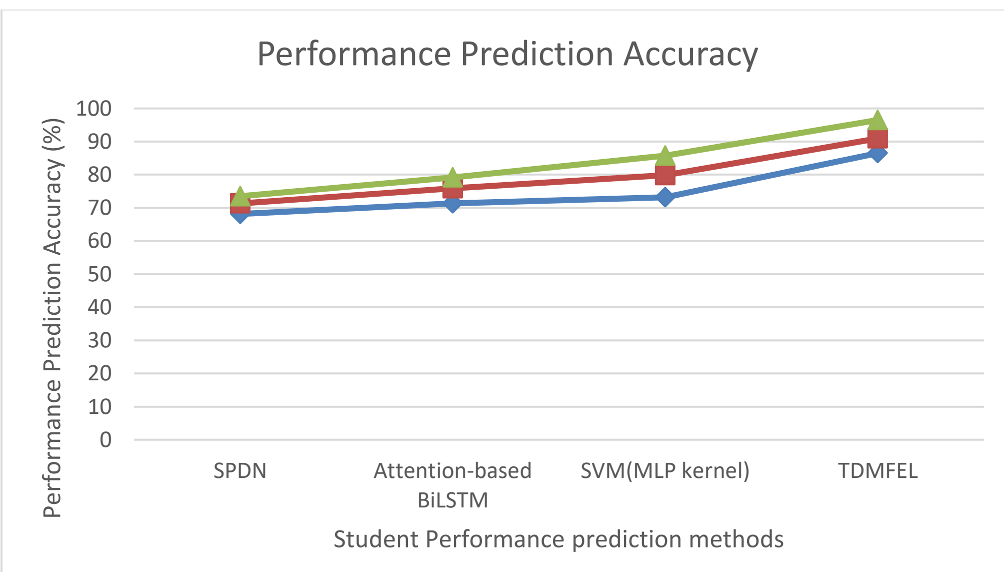

Figure 2: Accuracy of performance Prediction Analysis

The analysis on student performance prediction accuracy is measured and compared in Figure 23. The proposed TDMFEL algorithm has produced higher performance than other methods.

Prediction Accuracy Vs No of Records			
Method	3000 Records	5000 Records	10000 Records
SPDN	31.85	28.68	26.49
Attention-based BiLSTM	28.66	24.1	20.84
SVM(MLP kernel)	26.86	20.1	14.28
TDMFEL	13.5	9.6	3.5

Table 3: Analysis on False Ratio

The ratio of false classification made by various approaches is measured at the presence of different number of records in the data set. The proposed TDMFEL algorithm has produced less false ratio and classification than other approaches.

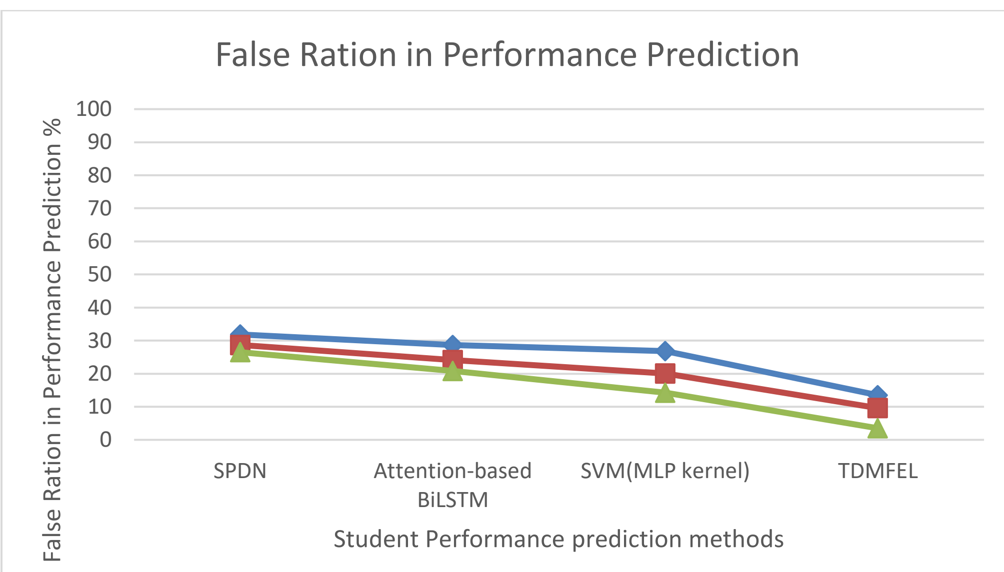

Figure 3: Analysis on False Ratio in Performance Prediction

The ratio of false prediction generated by various methods are measured and presented in Figure 3. The proposed TDMFEL algorithm has produced less false ratio than other approaches.

Number of records	SPDN [2]	Attention-based Bi LSTM [14]	SVM(MLP kernel) [15]	TDMFEL
3000	5.3	5.1	5	4.3
5000	6.5	6.2	5.6	4.7
10000	7.5	7.3	6.8	5.8

Table 4: Time Complexity Analysis

The time complexity produced by various approaches are measured and presented in Table 4. The proposed TDMFEL algorithm has produced less time complexity in all the test cases considered.

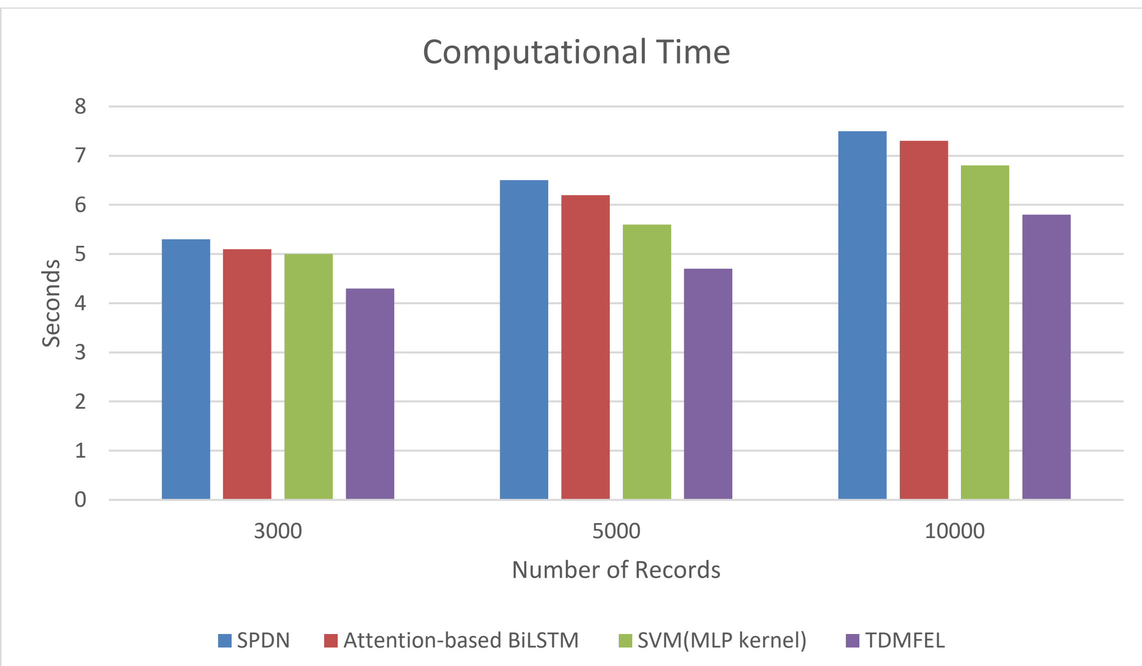

Figure 4: Computational Time of TDMFEL method compared with the Existing methods (SPDN, Attention-based Bi LSTM, and SVM)

We evaluated the time complexity of the proposed TDMFEL method for different records in Table 2 and Figure 6. It founded that the proposed method's computational time is lesser than the existing methods.

4. Conclusion:

This paper introduced a novel Time Division Multi Feature Ensemble learner Classifier based student performance analysis model using Convolution neural network. The method reads the student data set and performs pre-processing with mandate feature analysis algorithm. Further, the features are extracted and perform time line ensemble generation to generate the ensembles. Using the ensembles the method generates the convolution neural network and trains the network. At the test phase, the method use the input sample and with the features, the neurons estimates various support measures on different online and offline features, activity, sports, extracurricular and so on. Using all of them, the method computes the value of weight to decide the class or rank of the student. The method introduces higher prediction performance compare to various other approaches.

References:

1. Bhavesh Patel et al, International Journal of Computer Science and Mobile Computing, Vol.6 Issue.5,pg. 64-71, May- 2017.
2. Brijesh Kumar Baradwaj, Saurabh Pal, “Mining Educational Data to Analyze Students‟ Performance”, International Journal of Advanced Computer Science and Applications, Vol. 2, No. 6, 2011.
3. Astha Soni, Vivek Kumar, Rajwant Kaur, D. Hemavathi, “predicting student performance using data mining techniques”, International Journal of Pure and Applied Mathematics, Vol. 119, No. 12, pp. 221-227, 2018.
4. G. S. Chauhan, P. Agrawal, and Y. K. Meena, "Aspect-Based Sentiment Analysis of Students’ Feedback to Improve Teaching Learning Process,” Information and Communication Technology for Intelligent Systems Smart Innovation, Systems and Technologies, pp. 259-266, 2018.
5. Veena Deshmukh, Srinivas Mangalwede, Dandina Hulikunta Rao, “ Student Performance Evaluation Using Data Mining Techniques for Engineering Education”, Advances in Science, Technology and Engineering Systems Journal Vol. 3, No. 6, 259-264 (2018).
6. Abeer Badr El Din Ahmed, Ibrahim Sayed Elaraby, “Data Mining: A prediction for Student's Performance Using Classification Method”, World Journal of Computer Application and Technology 2(2): 43-47, 2014.
7. Irum Sindhu, Sher Muhammad Daudpota, Kamal Badar, Maheen Bakhtyar, Junaid Baber, and Mohammad Nurunnabi, “Aspect Based Opinion Mining On Student’s Feedback For Faculty Teaching Performance Evaluation”, IEEE Access, 2017.

8. Shaojie.Qu, Kan Li, Shuhui. Zhang, Yongchao. Wang, “Predicting Achievement of Students in Smart Campus”, IEEE Access 2018.

9. Asif, Raheela, et al, "Analyzing undergraduate student’s performance using educational data mining", Computers & Education, 2017.

10. Jiang, Tongjun, et al. "Analysis and Data Mining of Students' Consumption Behavior Based on a Campus Card System." International Conference on Smart City and Systems Engineering IEEE Computer Society, pp. 58-60, 2017.

11. Li X., Zhu X., Zhu X., Ji Y., Tang X. (2020) Student Academic Performance Prediction Using Deep Multi-source Behavior Sequential Network. In: Lauw H., Wong RW., Ntoulas A., Lim EP., Ng SK., Pan S. (eds) Advances in Knowledge Discovery and Data Mining. PAKDD 2020. Lecture Notes in Computer Science, vol 12084. Springer, Cham. https://doi.org/10.1007/978-3-030-47426-3_44.

12. Khan, Bashir &Afzal, Sher & Rahman, Taj& Khan, Inayat&Ullah, Inam & Rehman, Ateeq & Baz, Mohammed &Hamam, Habib & Cheikhrouhou, Omar. (2021). Student-Performulator: Student Academic Performance Using Hybrid Deep Neural Network. Sustainability. 13. 9775. 10.3390/su13179775.

13. Mahareek, E. A.; Desuky, A. S.; El-Zhni, H. A., "Simulated annealing for svm parameters optimization in student’s performance prediction", Bulletin of Electrical Engineering and Informatics ; 10(3):1211-1219, 2021.

14. Al-Okaily, M. Alqudah, H. Matar, A. Lutfi, A. Taamneh, A. Dataset on the Acceptance of e-learning System among Universities Students under the COVID-19 Pandemic Conditions. Data Brief 2020, 32, 106176.

15. Botao Zhong, Xing Pan, Peter E.D. Love, Lieyun Ding, Weili Fang, "Deep learning and network analysis: Classifying and visualizing accident narratives in construction", Automation in Construction, Volume 113, 2020,103089, ISSN 0926-5805, https://doi.org/10.1016/j.autcon.2020.103089.

16. Toğaçar, Mesut, “A Deep Feature Learning Model for Pneumonia Detection Applying a Combination of m RMR Feature Selection and Machine Learning Models.” Irbm 41 (2020): 212-222.

17. Q. Liu et al., "EKT: Exercise-Aware Knowledge Tracing for Student Performance Prediction," IEEE (TKDE), Volume 33, Number 1, pp. 100-115, 2021.

18. A. Alshanqiti and A. Namoun, "Predicting Student Performance and Its Influential Factors Using Hybrid Regression and Multi-Label Classification," IEEE, Volume. 8, pp. 203827-203844, 2020.

19. Z. Xu, H. Yuan and Q. Liu, "Student Performance Prediction Based on Blended Learning," IEEE (TE), Volume 64, Number 1, pp. 66-73, 2021.

20. D. Liu, Y. Zhang, J. Zhang, Q. Li, C. Zhang and Y. Yin, "Multiple Features Fusion Attention Mechanism Enhanced Deep Knowledge Tracing for Student Performance Prediction," IEEE, Volume. 8, pp. 194894-194903, 2020.
21. Z. Kastrati, A. S. Imran and A. Kurti, "Weakly Supervised Framework for Aspect-Based Sentiment Analysis on Students' Reviews of MOOCs," IEEE, Volume. 8, pp. 106799-106810, 2020.
22. M. Tadayon and G. J. Pottie, "Predicting Student Performance in an Educational Game Using a Hidden Markov Model," IEEE (TE), Volume. 63, Number. 4, pp. 299-304, 2020.
23. Q. Sun, J. Wu and K. Liu, "Toward Understanding Students' Learning Performance in an Object-Oriented Programming Course: The Perspective of Program Quality," IEEE, Volume. 8, pp. 37505-37517, 2020.
24. C. Robbiano, A. A. Maciejewski and E. K. P. Chong, "Nonparametric Analysis of the Effect of Knowledge Integration Activities on Third-Year Undergraduate Performance," IEEE (TE), Volume. 63, Number. 4, pp. 305-313, 2020.
25. X. Li, Y. Zhang, H. Cheng, F. Zhou and B. Yin, "An Unsupervised Ensemble Clustering Approach for the Analysis of Student Behavioral Patterns," IEEE, Volume. 9, pp. 7076-7091, 2021.