International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 2s
DOI: https://doi.org/10.17762/ijritcc.v10i2s.5931
***Article Received**: 20 October 2022 **Revised**: 28 November 2022 **Accepted**: 18 December 2022*

Tukey Regressive Hoover Indexed Deep Shift-Invariant Neural Network for Student Behavior Prediction

R. Pushpavalli[1]*, C. Immaculate Mary[2]
[1]Assistant professor, Department of Computer Science
Sri Sarada College for Women(Autonomous),
Salem- 636016, Tamilnadu, India.
Email: pushpavallikumarapathy@gmail.com

[2]Head &Associate Professor, Department of Computer Science
Sri Sarada College for Women (Autonomous)
Salem- 636016, Tamilnadu, India
Email: cimmaculatemary@gmail.com

Abstract—Prediction of student performance in the academic field creates significant challenges in developing reliable and accurate diagnosis models. Through the use of online learning behavior data, this paper may assist teachers in identifying students with learning challenges in advance and providing timely assistance. A novel technique called Tukey Regressive Hoover indexed Deep Shift Invariant Structure Neural Network (TRHIDSISNN) Model is introduced for student behaviour analysis with lesser time consumption. Initially, the student data and features are collected and transmitted to the input layer. After that, the features of collected student data are analyzed in hidden layer 1 with help of the Tukey Regression. The correlation between one or more independent features is identified to find the dependent feature. The relevant features are sent to the hidden layer 2. In that layer, the Hoover index is applied for analyzing the training and testing features. Finally, the hidden layer result is sent to the output layer where the hyperbolic tangent activation function is used to classify the data that belongs to that particular class. Based on the classification, the student grade level is predicted as high, medium and low based on their behavior gets displayed. Experimental assessment is carried out using different parameters such as prediction accuracy, false-positive rate, prediction time, and space complexity with respect to the number of student data. The discussed results show that when compared to state-of-the-art approaches, the suggested TRHIDSISNN model achieves higher accuracy with shorter prediction times.

Keywords-Student Performance Prediction, Deep Shift Invariant Structure Neural Network, Tukey Regression, Hyperbolic Tangent Activation Function

I. INTRODUCTION

Predicting the student behavior in an earlier stage leads to evaluate the students' performances in the field of Educational Data Mining. Predicting student success is a key task in online learning systems that help to offer students access to active learning. Student performance prediction generally seeks to ascertain a student's capability for knowledge or learning. Several methods have been developed for predicting the students' learning performance.

A Deep Neural Network (DNN) framework was developed in [1] for early student performance prediction based on binary classification with two hidden layers. Though the framework enhances the accuracy, the complexity of performance prediction was not minimized. In [2], the Bidirectional Long Short Term Model (BLSTM) Model with Condition Random Field (CRF) techniques used for the early prediction of students' performance. Yet, it failed to improve the predictive ability with minimum time consumption.

Three common machine learning algorithms called decision tree, neural network, and support vector machine were developed in [3] to predict academic performance based on features. But it failed to include more data for predicting academic performance. An automated Machine Learning algorithm was designed in [4] for student performance prediction. However, it did not examine how various factors would affect the predictive model.

A novel approach was introduced in [5] to forecast the course achievement on Moodle based on their procrastinating performance. But it failed to consider more information on students' performance to possibly predict their performance using deep learning approaches. The effectiveness of machine learning algorithms was estimated in [6] to monitor students'

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 2s
DOI: https://doi.org/10.17762/ijritcc.v10i2s.5931
***Article Received**: 20 October 2022 **Revised**: 28 November 2022 **Accepted**: 18 December 2022*

academic progress. But the efficiency of the model was not improved.

A deep neural network (DNN) was introduced in [7] to enhance the predictive approaches for students' academic performance. But the performance of accuracy of prediction was not improved. The ICGAN-DSVM algorithm (improved conditional generative adversarial network-based deep support vector machine) was presented. introduced in [8] to forecast students' performance in supportive learning. However, it failed to generate significantly more data without taking into account the early students' performance prediction model.

A method known as Multiple Features Fusion Attention Mechanism Enhanced Deep-Knowledge-Tracing (MFA-DKT) was developed in [9] for student performance prediction based on both student behavior features and exercise. But it failed to consider the correlation between knowledge concepts of prediction. A deep artificial neural network was introduced in [10] based on a set of distinctive handcrafted features. But it failed to perform a more detailed analysis of the day-to-day activities of the student.

The main goal of the research described here.

- To predict the student's academic performance in an early stage through the behavior of the student with minimum time and space complexity, the TRHIDSISNN technique is introduced.
- To reduce time complexity by performing the feature selection process usingTukey Regressive.
- To recognize the student performance behavior, the Hoover index is utilized through feature matching.
- To strengthen the prediction accuracy by classifying grade level based on the behavior of the student using the Hyperbolic tangent activation function.

Numerous research works were introduced for student educational performance by using machine learning approaches. However, it failed to categorize the student behavior prediction performance. But, the prediction time was not reduced. The existing methods were unable to handle the early stage students' performance prediction and space complexity. Motivated by, a novel TRHIDSISNN model is introduced using machine learning for classifying an Educational dataset taken form UCI repository in an accurate manner.

The simulation analysis of proposed Tukey Regressive Hoover indexed Deep Shift Invariant Structure Neural Network (TRHIDSISNN) model is carried out by implementing using Java Language. For conducting simulation process, different number of student data is considered with Educational Process Mining (EPM): A Learning Analytics Data Set taken from the UCI machine learning repository [21].

The key contributions of this paper are as follows:

- We propose the TRHIDSISNN model to increase the student behavior prediction accuracy. Also, the architecture of our proposed model has strong scalability based on feature selection and classification.
- Our Tukey Regression is used to Deep Shift Invariant Structure Neural Network in the first hidden layer to pick the dependent features for data classification. This helps to reduce both the complexity of the space and the forecast time.
- Our Hoover index is employed in the second hidden layer to analyze the testing and training features. Therefore, the accuracy is said to be increased and the false positive rate is reduced.
- Our hyperbolic tangent activation function is applied in output layer to classify the data into a particular class. This activation function provides a similarity value in the range of -1 to +1. As a result, the student grade level is correctly predicted according to the similarity value.
- In order to estimate the performance of our TRHIDSISNN model method in comparison to the contemporary deep learning method based on various metrics, we experiment on a learning analytics dataset.

The balance of this paper is structured into different sections as follows. Section 2 reviews the related works. The proposed TRHIDSISNN is described with a neat diagram in Section 3. Extensive experiments are presented in Section 4, followed by quantitative results analysis is presented in section 5. Finally, conclusions are presented in Section 6.

II. RELATED WORKDS -I

Students Performance prediction using Relational Association Rules (SPRAR) was developed in [11] for identifying the final result of a student. But the predictive performance of SPRAR was not improved. Ensemble meta-based tree model (EMT) classifier was developed in [12] for predicting student performance. But the designed classifier failed to based feature perform the dimensionality reduction for improving the prediction accuracy.

A student performance prediction model based on a supervised learning decision tree classifier was developed in [13]. But it failed to perform the data analysis on a large dataset with more attributes. A Black-Box Predictive Model was designed in [14] for Students' Behavior analysis. However, the model was not efficient to collect the time series data for improving the prediction. A Bidirectional Long Short-Term Memory (BiLSTM) network was introduced in [15] for predicting the performance grades based on historical data. But it failed to use fused learning models for achieving higher prediction performance.

An efficiency of transfer learning from deep neural networks was developed in [16] for the duties of students' performance prediction. But the performance of time complexity analysis of students' performance prediction was not executed. A new scalable algorithm called the random wheel was introduced in [17] for classification. However, the accuracy of classification was not improved.

A Dynamic Graph-based and Time-Edge Aware Model (DGTEAM) was designed in [18] to learn the student grade tendency from the dynamic graphs. But the model failed to perform the complexity analysis. Hybrid Neural Network Model based on High-Order Attention Mechanism, (HHA) was developed in [19] for academic abnormality prediction. However, it failed to select the expressive features for reducing the dimensionality of the dataset. An unsupervised ensemble clustering framework was introduced in [20] for predicting the student behavioral patterns. But it failed to extract more meaningful features for behavioral patterns prediction.

III. PROPOSAL METHODOLOGY -II

Automatic Student performance prediction plays a vital role due to the huge volume of data in the academic database. This process is being done by developing deep neural learning in educational data mining (EDM). In this section, a novel deep learning model called TRHIDSISNN is introduced for accurate prediction of Student academic performance with minimum time consumption. The TRHIDSISNN model has depended on two processes, one is feature selection, and the second one is classification module.

Figure 1 elucidates the architecture of the proposed TRHIDSISNN model includes two major processes namely feature selection and classification for improving student behavior prediction accuracy. Initially, the student raw data are collected from the dataset. Followed by, the feature selection process is performed using the Tukey Regressive Dimensionality Reduction resulting in minimizes the time complexity of behavior prediction. Finally, the classification is performed with given input data with higher accuracy. A Deep Shift Invariant Structure Neural Network is a class of deep neural networks used for feature selection and classification.

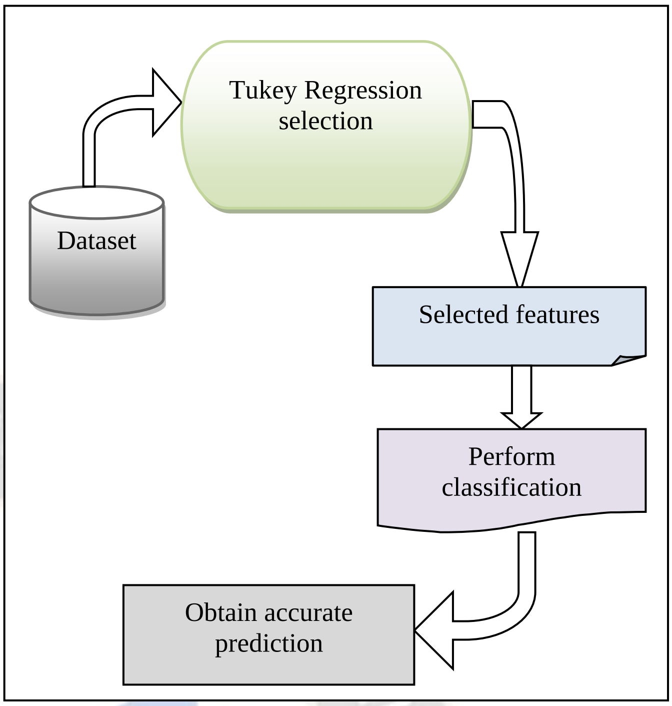

Figure 1 Architecture diagram TRHIDSISNN

Figure 2 shows the structural design of the Shift Invariant Deep neural learning to perform different processes with different layers. The design architecture consists of neurons like the nodes. The neuron in one layer is linked to another layer. The shift-invariant deep structure architecture includes the cascade of several processing layers such as an input, two hidden layers, and one output layer.

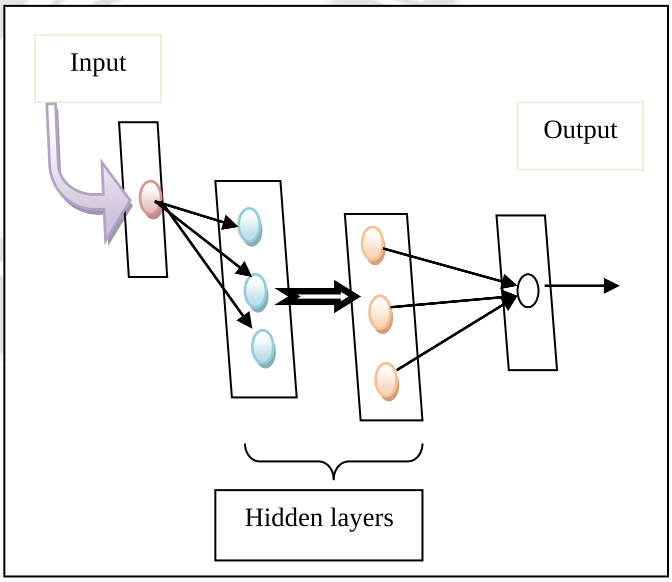

Figure 2 Structural design of the Shift Invariant Deep Neural Learning

3.1 Tukey Regression-Based Feature Selection

The proposed TRHIDSISNN performs the feature selection in the first hidden layer using Tukey Regression. Tukey

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 2s
DOI: https://doi.org/10.17762/ijritcc.v10i2s.5931
***Article Received**: 20 October 2022 **Revised**: 28 November 2022 **Accepted**: 18 December 2022*

Regression is a machine learning technique that helps to analyze one or more independent features and dependent features. The dataset, in general, has a large number of features, but not all of which are always meaningful. Removing avoidable features for the trained model leads to a decrease in the whole accuracy and increases the complexity of the model. Therefore, the proposed TRHIDSISNN uses the Tukey Regression for analyzing the features and selecting the dependent features for further processing

Consider the quantity of features. $A_1, A_2, A_3, \ldots . A_m$ taken from the dataset. Then by applying the Tukey Regression, the correlation between the features are measured as given below,

$$Y = \sum_{i=1}^{n} \beta_i * Q[A_i - A_j] \quad (1)$$

From (1), D denotes an output of regression, β_i denotes a weight vector, Q indicates Tukey loss function, A_i, A_j denotes two features. The minimum deviations between the features are called dependent features and maximum deviation is said to be an independent feature. Then the dependent features are sent to the hidden layer 2 for predicting the user performance behavior.

Table 1 shows the Tukey correlation values for each attributes of given Input data. After calculating the correlation values based on Tukey correlation method, the features are selected for y>0.5 as dependent features.

Table 1Tukey correlation of input features

S.No	Input Features	Tukey Correlation
1	Session	0.839
2	Student-ID	0.759
3	Exercise	0.699
4	Activity	0.855
5	Start-time	0.185
6	End-time	0.213
7	Idle-time	0.460
8	Mouse-wheel	0.447
9	Mouse-wheel-click	0.098
10	Mouse-click-left	0.879
11	Mouse-click-right	0.303
12	Mouse-movement	0.779
13	Keystroke	0.817

They are Session (0.839), Student-ID(0.759), Exercise(0.699), Activity(0.855), Mouse-click-left(0.879), Mouse-movement(0.779), and Keystroke(0.817). And the Independent features are Start-time(0.185), End-time(0.213), Idle-time(0.460), Mouse-wheel(0.447), Mouse-wheel-click(0.0981), Mouse-click-right(0.303). The selected features are consideredfor Hoover index feature matching as a next phase.

3.2 *Hoover Indexive Feature Matching*

Following feature selection, the Hoover index is used for feature matching in the second hidden layer to determine the behaviour associated with student performance. The Hoover index is used to measure the relationships between features. The Hoover index is measured as follows,

$$\alpha = \frac{1}{2}\left[\frac{\sum_i |A_t - A_{tr}|}{\sum_i A_t}\right] (2)$$

From (2), 'α' symbolizes the Hoover index similarity coefficient, A_t denotes a testing feature, A_tr indicates a training feature. Therefore, the coefficient provides the outcomes in the ranges from 0 to +1.

It is computed by,

$$H_{(t)} = \sum_{i=1}^{n} D_i * w_1 + w_2 * h_{(t-1)} (3)$$

Where, H(t) denotes hidden layer, h(t-1) indicates the output over preceding hidden layer as well as 'w2 'indicates the weight of hidden layers, w_1 indicates weight among input as well as hidden layers, D_i indicates input data. Finally, similarity value was specified to output layer

Z_((t))=φ[w_3* H_((t))] (4)

From (4),Z_((t))represents the output layer φ denotes a Hyperbolic tangent function, w_3denotes weight among hidden as well as output layer, H_((t)) denotes hidden layer. These weights are said to be updated during training and is referred to as the learning rate. To be more specific, the learning rate is said to be a configurable hyper parameter utilized in the training of neural networks for feature matching possessing a small positive value ranging between 0.0 and 1.0. Higher learning rate results in the early convergence whereas the lower learning rate results in the process to get stuck. Hence, an optimal learning rate of 0.05 is selected in the proposed work. Followed by which, the hyperbolic tangent was employed for delivering better training performance for shift-invariant structure neural networks. The hyperbolic tangent activation was smoother as well as zero-centered varies among -1 to 1. The tangent activation function analyzes the similarity values as given below,

φ=(e^α-e^(-α))/(e^α+e^(-α)) (5)

Where ' φ 'denotes an activation function, α denotes the Hoover index similarity coefficient. Hyperbolic tangent activation function returns the value from -1 to +1. While similarity was better, then activation function returns value '+1'. Otherwise, it returns the value '-1'. The similarity is checked based on the threshold selection. Threshold selection in our work refers to the selection of the similarity value of the activation function. In our work, the selection of the threshold or the similarity value of activation function is selected as 0.5. The proposed neural network stops the training process when the error or the difference between desired output and expected output is below certain threshold level. In this way, student information was properly grouped. The accurate prediction was achieved with minimum time. Step by step procedure of TRHIDSISNN is described as a given below, Algorithm 1: A Tukey Regressive Hoover indexed Deep Shift Invariant Structure Neural Network

Input: Dataset 'D', features'$\{A_1, A_2, A_3, \ldots . A_m\}$', data $D_1, D_2, D_3, \ldots D_n$
Output: Improve prediction accuracy

Begin
1. Number of features'$\{A_1, A_2, A_3, \ldots . A_m\}$ are taken as input
2. For each features 'A_i' [Hidden layer 1]
3. Apply Tukey Regression
1. Find dependent and independent features
2. Select Dependent features and remove independent features
3. end for
4. for each data d_i' [Hidden layer 2]
5. Measure the similarity 'α' between testing and training features using (2)
6. end for
7. Apply Hyperbolic tangent activation function 'φ' -------- [output layer]
8. Analyze the similarity value 'α'
9. Obtain the classification results

End

Algorithm 1 provides student academic prediction by better accuracy with different procedures such as feature selection and classification for single epoch. An epoch refers to training of the entire process of neural network using all the comprehensive training data for single cycle. To be more specific, as far as epoch is concerned, the entire data is said to be utilized only once, including both forward pass and a backward pass. Initially, the numbers of features were composed over dataset as well as specified to input layer. Then, input features are transformed into first hidden layer. The feature selection is performed in that hidden layer with aid of Tukey Regression. The avoidable features for the trained model were eliminated with improved overall accuracy and lesser complexity. Dimensionality reduction was achieved by finding dependent and independent features. Next, the correlation values are measured through the Tukey correlation method to pick the dependent features and eliminate independent features. Next, dependent features were sent (list of feat…) within hidden layer. In the second hidden layer, Hoover Index was utilized to perform feature matching for detecting relationship between training as well as testing data. Finally, similarity was analyzed using hyperbolic tangent activation by performing classification to classify given information within various classes. Threshold selection is defined as the pick of the similarity value based on the activation function. In the proposed model, the threshold value is selected as 0.5 depending on the similarity value of the activation function. The proposed method is said to be stopped with the corresponding training process when the desired output and the expected output is said to be below the threshold level. A neural network is stopped training when the error, or the difference between the desired and expected output, is less than a certain threshold value, or when the number of iterations or epochs exceeds a certain threshold value. Depend on classification outcomes, accurate prediction was obtained by output layer. In this manner, the prediction was enhanced with diminished time.

IV. Experimental Scenario

1) In this section, simulation of TRHIDSISNN and DNN framework [1] BLSTM-CRF [2] were performed by using Java Language. For the experimental consideration, Educational Process Mining (EPM): The Learning Analytics bench mark Data Set is used wich is taken from UCI machine learning repository and it's source link is https://archive.ics.uci.edu/ml/datasets/Educational+Process+Mining+%28EPM%29%3A+A+Learning+Analytics+Data+Set. Dataset consists of 13 attributes as well as 230318 instances. This was considered on recordings of 115 students' over logging function using learning experiment. In addition, dataset comprises various students' time through six different periods. The associated tasks of the dataset are classification, regression, and clustering for detecting initial grade. Student-ID, Exercise, Activity, Start-time, End-time, and so on are the attributes. Simulation is conducted by using Java Language.

V. Qualitative Results And Discussions

a) Simulation of TFHIRCDBNN, as well as DNN framework [1] BLSTM-CRF [2], are analyzed by various parameters namely prediction accuracy, false-positive rate, prediction time, space complexity as well as precision.

5.1 Performance Analysis Of Prediction Accuracy

It was calculated by the number of student information that was correctly predicted via deep classification to the entire amount of student information. Therefore, accuracy was measured by,

PA= [ncp/n]*100 (6)

From (6), the prediction accuracy ' PA'is measured based on a number of data correctly predicted 'ncp' number of data, and 'n' indicates the number of data, it is expressed by percentage (%). The sample calculation of three methods are given below,

5.1.1 Sample Mathematical Calculation for Prediction Accuracy:

- **Existing DNN framework:**Total number of data =5000 and the number of data correctly predicted is 4500. Thus, the prediction accuracy is calculated as follows,
  PA= [4500/5000]*100 =90%
- **Existing BLSTM-CRF:**The number of data correctly predicted is 4400. Thus, the prediction accuracy is calculated as follows,
  PA= [4400/5000]*100=88%

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 2s
DOI: https://doi.org/10.17762/ijritcc.v10i2s.5931
***Article Received**: 20 October 2022 **Revised**: 28 November 2022 **Accepted**: 18 December 2022*

- **Proposed TRHIDSISNN:**The number of data correctly predicted is 4700. Thus, the prediction accuracy is calculated as follows,
  PA= [4700/5000]*100=94%

Table 2Performance Values of Prediction accuracy

Number of data	Prediction accuracy (%)		
	TRHIDSISNN	DNN framework	BLSTM-CRF
5000	94	90	88
10000	92	89	86
15000	94	88	84
20000	92	87	85
25000	94	89	86
30000	93	88	85
35000	94	87	83
40000	93	88	85
45000	94	89	88
50000	92	88	86

Table 1 reports prediction accuracy by the amount of student information ranges of 5000 to 50000. Accuracy was measured by TRHIDSISNN as well as DNN framework [1] BLSTM-CRF [2]. Among three methods, TRHIDSISNN provides superior performance than the existing methods. For a fair comparison, 10 different runs are noted and the corresponding accuracy results are observed. For example, 5000 data are considered for experimentation in the initial iteration. By using the TRHIDSISNN, 4700 data are correctly forecast and the accuracy is 94%, whereas 4500, and 4400 data are correctly classified and the existing methods' accuracy percentages [1] and [2] are 90% and 88%, respectively. In the second iteration, 5000 data are taken to conduct the experiments. By using TRHIDSISNN, 9200 data are properly discovered and the accuracy is 92%, whereas 8900, and 8600 data are properly determined and the existing methods [1] and [2] have accuracy percentages of 89% and 86%, respectively.Then, several performance results are noted for each method. Likewise, the remaining runs are performed and observed the accuracy results. Finally, the TRHIDSISNN was compared with the accuracy of conventional techniques. The prediction accuracy using TRHIDSISNN is considerably enhanced up to 6% and 9% compared with [1] and [2].

As illustrated in the above figure 3, the Convergence graph ofprediction accuracy by the amount of student information ranges of 5000-50000. Here, the x-axis represents the amount of student information, and the y-axis represents prediction accuracy. As shown in the graphical chart, there are three various colors of lines such as sky blue, yellow, and green that represent the prediction accuracy of all the three techniques namely TRHIDSISNN, DNN framework [1], and BLSTM-CRF [2] respectively. From the figure, it is shown that the TRHIDSISNN technique achieves higher prediction accuracy when compared with conventional techniques.

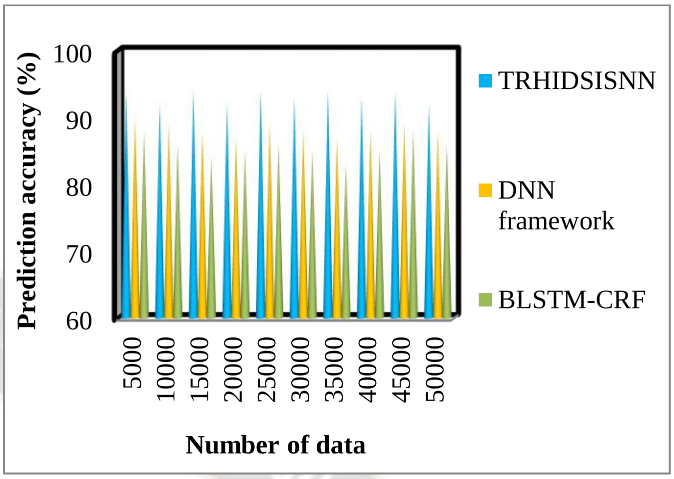

Figure 3 Convergence graph of prediction accuracy

This is due to the implementation of the TRHIDSISNN technique using the Tukey Regression, Hoover Index, and hyperbolic tangent activation function to predict student data. At first, Tukey Regression is employed to discover the dependent features and remove independent features. Followed by, Hoover Index is applied for obtaining relationship among education as well as difficult student information. Finally, similarity values are analyzed using the hyperbolic tangent activation function. The activation task detects relationship value as well as categorizes given information within various classes. The accurate prediction was obtained by the output layer with higher accuracy.

5.2 Performance Analysis Of False-Positive Rate

It was calculated by the proportion of the amount of student information that was improperly detected through classification. It is computed by,

FPR = [nicp / n] * 100 (7)

From (7) false-positive rate ‘FPR was calculated by the amount of information that was improperly detected ‘nicp’ and the entire amount of the data 'n’. It was calculated by percentage (%). The sample calculation of three methods is given below,

5.2.1 Sample Mathematical Calculation for False-Positive Rate:

- **Existing DNN framework:**Total number of data =5000 and the number of data improperly identified is 500. Thus, the false-positive rate is measured as follows,
  FPR = [500/5000]*100 =10%
- **Existing BLSTM-CRF:**The number of data improperly identified is 600. Thus, the false-positive

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 2s
DOI: https://doi.org/10.17762/ijritcc.v10i2s.5931
***Article Received**: 20 October 2022 **Revised**: 28 November 2022 **Accepted**: 18 December 2022*

rate is measured as follows,
FPR = [600/5000]*100=12%

- **Proposed TRHIDSISNN:**The number of data improperly identified is 300. Thus, The rate of false-positives is calculated as follows,

FPR = [300/5000]*100=6%

Table 2 illustrates the performance of false-positive rate for 10 different numbers of runs with the data varied from 5000 to 50000. The observed result proves that the performance of false-positive rate by TRHIDSISNN was minimized by conventional techniques. Consider 5000 data for experimentation in the first iteration. From the input, 300 data are wrongly categorized and the false positive rate is 6% using the TRHIDSISNN technique. Similarly, using the DNN framework [1] BLSTM-CRF [2], 500 and 600 pieces of data are incorrectly classified, with false-positive rates of 10% and 12%.

Table 2 Comparison of False-positive rate

Number of data	False-positive rate (%)		
	TRHIDSISNN	DNN framework	BLSTM-CRF
5000	6	10	12
10000	8	11	14
15000	6	12	16
20000	8	13	15
25000	6	11	14
30000	8	12	15
35000	6	13	17
40000	7	12	15
45000	6	11	12
50000	8	12	14

In the second iteration, the 5000 data is considered to perform experimentation. From the input, 800 data are incorrectly classified and the false positive rate is 8% using the TRHIDSISNN technique. Similarly, 1100 and 1400 amount of data is wrongly predicted and the false-positive rates are 11% and 14% using the DNN framework [1] BLSTM-CRF [2]. From the observed result, it was noticed that the false positive rate is found to be considerably reduced using the proposed TRHIDSISNN technique than the conventional methods. For each method, ten false positive rate results are observed with different inputs. Similarly, the overall performance results indicate that the TRHIDSISNN of false positive rate reduces up to 41% and 52% when compared with existing [1] [2].

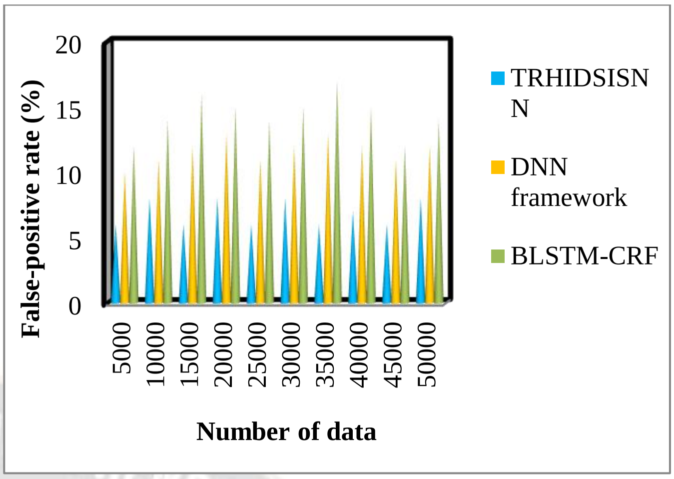

Figure 4 Convergence graph of false-positive rate

Figure 4 given above shows the convergence graph of the false-positive rate. As sown in figure 4, the 'x-axis represents the number of student data, and the ‘y’ axis represents the false positive rate. For conducting the experiments, the number of data is taken in the ranges from 5000 to 50000. From the figure, it is illustrative for different numbers of runs, the false positive rate is also said to be different. In the figure, the false positive rate is found to be minimized using the TRHIDSISNN when compared to the existing [1] and [2]. The cause for the lesser false positive rate is to apply the Tukey Regression, Hoover Index, and hyperbolic tangent activation function. Initially, the dependent features are selected and independent features are eradicated by the means of Tukey Regression. Hoover Index to accurately analyze the similarity value between the data for the training data and testing based on the Deep Shift Invariant Structure Neural Network. In addition, The hyperbolic tangent activation function improves classification results while reducing incorrect data classification. As a result, data is correctly classified into the appropriate classes, lowering the false positive rate.

5.3 *Performance Analysis Of Prediction Time*

It is defined by number of time taken with an algorithm for detecting student performance through classification. Prediction time was measured by,

PT=[n]*time (predict one student data) (8)

From (8), ‘PT’ indicates prediction time, ‘n’ denotes number of data, time indicates time to predict student behaviour performance. It was calculated by milliseconds (ms). The sample calculation of three methods is given below,

5.3.1 Sample Mathematical Calculation for Processing Time:

- **Existing DNN framework**: Total number of data =5000 and the time involved in detecting student performance is 0.006ms. Then, the prediction time is computed as follows,

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 2s
DOI: https://doi.org/10.17762/ijritcc.v10i2s.5931
***Article Received**: 20 October 2022 **Revised**: 28 November 2022 **Accepted**: 18 December 2022*

PT = 5000*0.006ms = 30 ms

- **Existing BLSTM-CRF**: The time involved in detecting student performance is 0.0066ms. Then, the prediction time is computed as follows,
  PT= 5000*0.0066ms = 33 ms
- **Proposed TRHIDSISNN:** The time involved in detecting student performance is 0.005ms. Then, the prediction time is computed as follows,
  PT= 5000*0.005ms = 25 ms

Table 3 Comparison of Prediction time

Number of data	Prediction time (ms)		
	TRHIDSISNN	DNN framework	BLSTM-CRF
5000	25	30	33
10000	30	35	37
15000	33	36	39
20000	40	44	46
25000	43	45	48
30000	45	48	51
35000	49	53	56
40000	52	56	60
45000	54	59	63
50000	58	60	65

Above table 3 illustrates the performance outcomes of prediction time using TRHIDSISNN and the existing DNN framework [1] BLSTM-CRF [2]. Prediction time is considerably reduced by conventional techniques. Let us consider the ‘5000' number of information, the time is taken to discover the student performance is 25ms by using the TRHIDSISNN technique. Likewise, prediction times for student performance using [1], [2] are 30ms and 33ms by applying the same counts of the input data. In the second iteration, with consideration of 5000 data, the predicting time of existing [1] [2] is obtained as 35ms and 37 ms respectively. Besides, the proposed TRHIDSISNN technique obtains the processing time as 30ms. The above-mentioned statistical result validates that TRHIDSISNN reduces prediction time. As shown in the table, the prediction time for all three classification methods gradually increases while increasing the number of data as the number of data counts increases for each run. Overall comparison of ten results confirms that the time consumption for predicting the student performance level is considerably decreased using TRHIDSISNN by 11% when compared to [1] and 18% when compared to [2] respectively.

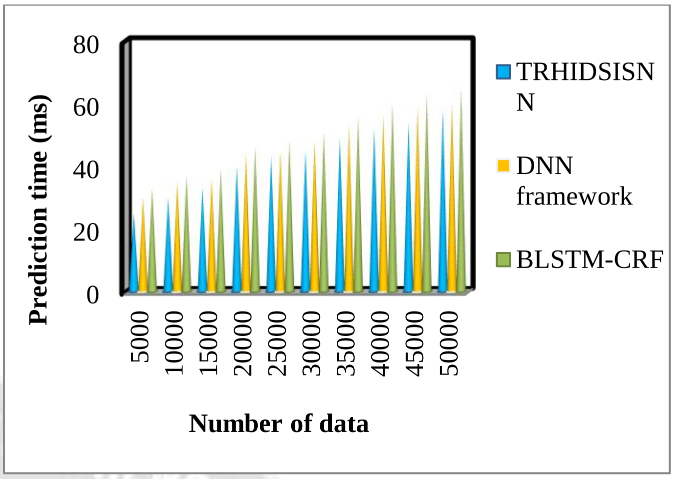

Figure 5 Convergence graph of prediction time

Figure 4 exhibits prediction time for a different amount of user data. For experimental purposes, the amount of data varies of 5000 to 50000 are considered. The x-axis refers to the number of data and the y-axis refers to the prediction time of three methods. As shown in the graph, three different colours of cones, such as blue, yellow, and green, indicate the prediction time of three methods such as the TRHIDSISNN, existing [1] and [2] respectively. In the above figure, the prediction time of the proposed TRHIDSISNN is lesser than other conventional techniques. In other words, as the number of data increases, the prediction time of each method increases. The reason behind the TRHIDSISNN uses Tukey Regressive Dimensionality Reduction. By using this function, the numbers of features are analyzed and redundant features are eliminated with Tukey Regression. In addition, the dependent and independent features are identified. Based on regression analysis, the dependent features are selected for classifying the student information.

5.4 Performance analysis of space complexity

It was defined by number of memory space taken with the algorithm for detecting student performance on data classification. It was measured by,

SC=[n]*Mem (predict one student data) (9)

Where SC indicates space complexity, ‘n’ indicates number of data, ‘Mem' indicates memory space taken to detect data. It is calculated by Megabytes (MB). The sample calculation of three methods is given below,

5.4.1 Sample Mathematical Calculation for Space Complexity:

- **Existing DNN framework**: Total number of data =5000 and the space consumed to discover data is 0.0056 MB. Hence, the space complexity is evaluated as follows,
  SC= 5000*0.0056 MB = 28MB
- **Existing BLSTM-CRF**: The space consumed to

ISSN: 2321-8169 Volume: 10 Issue: 2s
DOI: https://doi.org/10.17762/ijritcc.v10i2s.5931
***Article Received**: 20 October 2022 **Revised**: 28 November 2022 **Accepted**: 18 December 2022*

discover data is 0.006 MB. Hence, the space complexity is evaluated as follows,
SC = 5000*0.006 MB = 30 MB

- **Proposed TRHIDSISNN:** The space consumed to discover data is 0.005 MB. Hence, the space complexity is evaluated as follows,
SC = 5000*0.005 MB = 25 MB

Table 4 Performance Values of Space complexity

Number of data	Space complexity (MB)		
	TRHIDSISNN	DNN framework	BLSTM-CRF
5000	25	28	30
10000	27	30	35
15000	33	36	39
20000	34	38	40
25000	38	40	43
30000	42	45	48
35000	44	46	51
40000	48	52	54
45000	55	57	59
50000	59	63	65

Table 4 portrays performance results of space complexity for detecting student performance behavior. The amount of space complexity was calculated by megabytes (MB). From the observed results, the graphical plot indicates that the amount of space complexity gets increases while increasing the tasks. But comparatively, the TRHIDSISNN Model reduces the storage space for processing multiple student data. Let us consider the 5000 student data for calculating the space complexity. Firstly, the TRHIDSISNN Model consumes 25MB of memory for predicting the data. Next, the space complexity of the DNN framework [1] BLSTM-CRF [2] is observed by 28MB and 30MB respectively. Though the space complexity is decreasing with the increase in the data, with simulations conducted using 5000 data, space complexity using the three methods are 0.0054MB, 0.006MB and 0.007MB. Therefore, the space complexity was observed to be 27ms using TRHIDSISNN Model, 30 MB using [1], and 35 MB using [2]. As a result, the overall performance comparison results incurred using the TRHIDSISNN Model are found to be considerably minimized up to 7% and 12% compared with the two existing techniques.

Figure 6 portrays the performance results of the space complexity regarding the number of data. The data is taken in the horizontal direction and the space complexity is observed at the vertical axis. To explore the significance of the TRHIDSISNN Model, we show the average processing time of TRHIDSISNN, existing [1] and [2] with varying data.As illustrated by the graphs above, the proposed TRHIDSISNN Model performs better than the other two classification algorithms in terms of minimising space complexity. By applying the Deep Shift Invariant Structure Neural Network, the student information is accurately predicted with minimum space complexity. Therefore, the comparison results clearly show that the space complexity of the proposed TRHIDSISNN Model is lesser.

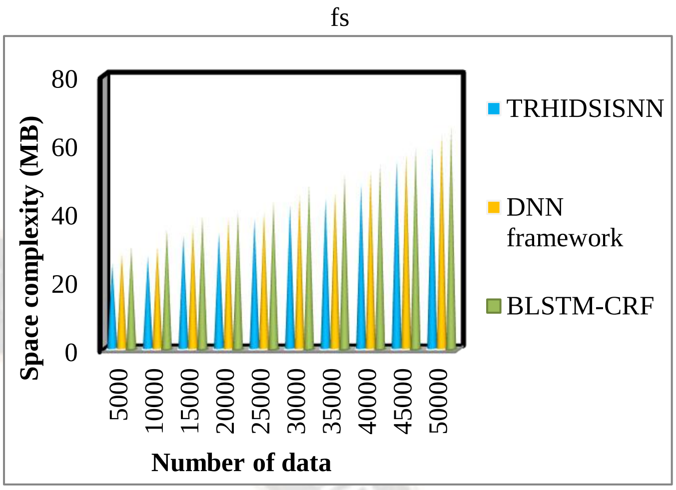

Figure 6 Convergence graph of space complexity

5.5 Performance analysis of precision

It was calculated as the percentage of relevant student data that was correctly and incorrectly classified out of the total number of student data collected for the experimental evaluation. The precision is calculated mathematically as shown below.

P= [TP/(TP+FP)]*100 (10)

Where, P denotes precision, TP indicates a true positive (i.e.correctly classified), and FP represents a false positive (incorrectly identified). It was measured by percentage (%).The sample calculation of three methods are given below,

5.5.1 Sample Mathematical Calculation for Precision:

- **Existing DNN framework:**Total number of data =5000 and the number of data correctly classified is 4300 and the incorrectly classified data is 700. Thus, the precision is estimated as,
P= [4300/(4300+700)]*100=86%
- **Existing BLSTM-CRF:**The number of data correctly classified is 4100 and the incorrectly classified data is 900. Thus, the precision is estimated as,
P= [4100/(4100+900)]*100=82%
- **Proposed TRHIDSISNN:**The number of data correctly classified is 4500 and the incorrectly classified data is 500. Thus, the precision is estimated as,
P= [4500/(4500+500)]*100=90%

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 2s
DOI: https://doi.org/10.17762/ijritcc.v10i2s.5931
***Article Received**: 20 October 2022 **Revised**: 28 November 2022 **Accepted**: 18 December 2022*

Table 5Performance Values of Precision

Number of data	Precision (%)		
	TRHIDSISNN	DNN framework	BLSTM-CRF
5000	90	86	82
10000	88	85	80
15000	90	83	81
20000	89	81	80
25000	91	84	82
30000	90	82	81
35000	92	85	84
40000	91	84	82
45000	93	86	84
50000	90	85	81

Table 5 explains the performance of three alternative methods of precision, namely TRHIDSISNN as well as DNN framework [1] BLSTM-CRF [2]. The numbers of data are considered in the ranges from 5000 to 50000. Each run's various precision results are collected and displayed in the table. The precision is computed using true positives and false positives. Let us consider the 5000 data from the dataset, the precision of TRHIDSISNN is 90 whereas, the precision value of the two conventional methods is 86 and 82 respectively. Similar to this, but with regard to a different count of data, the nine different conclusions are achieved. The outcomes of the proposed methodology are then contrasted with those of existing approaches.The average of the comparison data shows that, in comparison to the DNN framework [1] and the BLSTM-CRF [2], the TRHIDSISNN approach boosts recall by 8% and 11%, respectively.

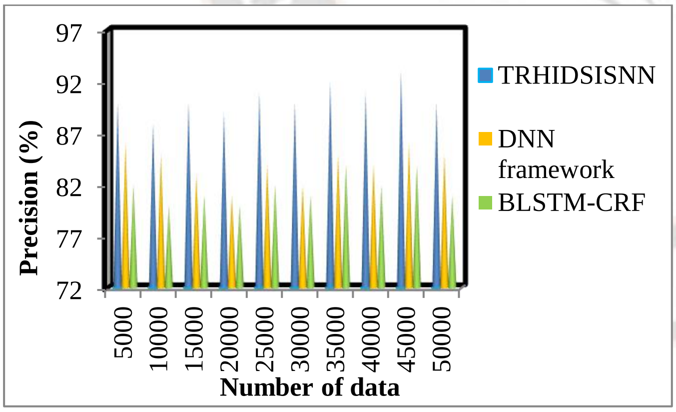

Figure 7Convergence graph of precision

Figure 7 given above illustrates the comparative results of precision based on the number of data. The input is taken in a horizontal manner, and the precision results are shown in the vertical axis of the upper column chart. The TRHIDSISNN methodology offers higher recall values than the other two classification methods. By using the similarity function to compare the similarities between training and testing student data, a considerable improvement is made. Next, the hyperbolic tangent activation function displays the classified student data according to the similarity value. In this way, the precision is enhanced.

V. CONCLUSION

Data mining was a significant systematic instrument to resolve difficulty for detecting vast number of information accessed within a database to decision-making procedure and detecting students' performance by early period. The novel TRHIDSISNN Model is introduced for predicting the students' academic performance by lesser time and complexity. TRHIDSISNN Model uses the Deep Shift Invariant Structure Neural Network for student behavior analysis with lesser time consumption. After that, the Tukey Regression is applied for selecting the dependent features in hidden layer 1. Followed by, the correlation between one or more features is identified using the Hoover index. At last, hyperbolic tangent activation utility was used for classifying data and identifies student grade level with higher accuracy. Experimental assessment was performed with different parameters by the amount of data. TRHIDSISNN Model enhances the accuracy as well as reduces false-positive rate, prediction time and space complexity when compared with state-of-the-art techniques. In the future, we plan to expand on the current work by introducing virtual assessments with varying levels of difficulty. This might be useful in developing a complete algorithm to identify students with poor learning abilities.

REFERENCES

[1] F. Giannakas, C. Troussas, I. Voyiatzis, C. Sgouropoulou, "A deep learning classification framework for early prediction of team-based academic performance", *Applied Soft Computing, Elsevier*, Vol. 106, pp.1-17, 2021

[2] DiaaUliyan, AbdulazizSalamahAljaloud, Adel Alkhalil, Hanan Salem Al Amer, MagdyAbdElrhmanAbdallah Mohamed and AzizahFhad Mohammed Alogali, "Deep Learning Model to Predict Students Retention Using BLSTM and CRF", *IEEE Access*, Vol. 9, pp.135550 – 135558, 2021

[3] Xing Xua, JianzhongWanga, HaoPengd, RuilinWua, "Prediction of academic performance associated with internet usage behaviors using machine learning algorithms", *Computers in Human Behavior, Elsevier*, Vol. 98, pp. 166-173, 2019

[4] Hassan Zeineddine, UdoBraendle, Assaad Farah, "Enhancing prediction of student success: Automated machine learning approach", *Computers & Electrical Engineering*, Vol.89, pp.1-22, 2021

[5] Yeongwook Yang, DanialHooshyar, MargusPedaste, Minhong Wang, Yueh-Min Huang and Heuiseok Lim "Predicting course achievement of university students based on their procrastination behaviour on Moodle", *Soft*

Computing, Springer, pp.1-17, 2020

[6] Ijaz Khan, Abdul Rahim Ahmad, NafaaJabeur and Mohammed Najah Mahdi, "An artificial intelligence approach to monitor student performance and devise preventive measures", *Smart Learning Environments, Springer*, Vol. 8, pp.1-18, 2021

[7] Aya Nabil, Mohammed Seyam, Ahmed Abou-Elfetouh, "Prediction of Students' Academic Performance Based on Courses' Grades Using Deep Neural Networks", *IEEE Access*, Vol.9, pp.140731 – 140746, 2021

[8] Kwok Tai Chui, Ryan Wen Liu, Mingbo Zhao, Patricia Ordóñez De Pablos, "Predicting Students' Performance With School and Family Tutoring Using Generative Adversarial Network-Based Deep Support Vector Machine", *IEEE Access*, Vol.8, pp. 86745 – 86752, 2020

[9] Dong Liu, Yunping Zhang, Jun Zhang, Qinpeng Li, Congpin Zhang, Yu Yin, "Multiple Features Fusion Attention Mechanism Enhanced Deep Knowledge Tracing for Student Performance Prediction", *IEEE Access* , Vol. 8, pp.194894 – 194903, 2020

[10] HajraWaheed, Saeed-Ul Hassan, NaifRadiAljohani, Julie Hardman, Raheel Nawaz, "Predicting Academic Performance of Students from VLE Big Data using Deep Learning Models", *Computers in Human Behavior, Elsevier*, Vol. 104, pp.1-34, 2020

[11] Gabriela Czibulaa, Andrei Mihaia, Liana Maria Criveia, "S PRAR: A novel relational association rule mining classification model applied for academic performance prediction", *Procedia Computer Science, Elsevier*, Vol.159, pp.20-29, 2019

[12] AmmarAlmasri , ErbugCelebi, and Rami S. Alkhawaldeh, "EMT: Ensemble Meta-Based Tree Model for Predicting Student Performance", *Scientific Programming, Hindawi*, Vol.2019, pp.1-13, 2019

[13] Muhammad Imran, ShahzadLatif, Danish Mehmood, Muhammad Saqlain Shah, "Student Academic Performance Prediction using Supervised Learning Techniques", *International Journal of Emerging Technologies in Learning (iJET)*, Vol. 14, No. 14, pp.92-103, 2019

[14] LipeDwan Pereira, Samuel C. Fonseca, Elaine H. T. Oliveira , Alexandra I. Cristea, HenrikBellhäuser ,Luiz Rodrigues, David B. F. Oliveira, Seiji Isotani, and Leandro S. G. Carvalho, "Explaining Individual and Collective Programming Students' Behavior by Interpreting a Black-Box Predictive Model" *IEEE Access* ,Vol. 9, pp.117097 – 117119, 2021

[15] Bashir Khan Yousafzai , SherAfzal Khan, TajRahman , Inayat Khan, InamUllah , Ateeq Ur Rehman,MohammedBaz , HabibHamam and Omar Cheikhrouhou, "Student-Performulator: Student Academic Performance Using Hybrid Deep Neural Network", *Sustainability*, Vol. 13, pp.1-21, 2021

[16] Maria Tsiakmaki, GeorgiosKostopoulos , Sotiris Kotsiantis and OmirosRagos, "Transfer Learning from Deep Neural Networks for Predicting Student Performance", *Applied Science*, Vol. 10, pp.1-12, 2020

[17] Anupam Khan, Soumya K. Ghosh, DurgadasGhosh, ShubhamChattopadhyay, "Random wheel: An algorithm for early classification of student performance with confidence", *Engineering Applications of Artificial Intelligence, Elsevier*, Vol. 102, pp.1-15, 2021

[18] Xuansheng Lu, Yanmin Zhu, YananXu, Jiadi Yu, "Learning from multiple dynamic graphs of student and course interactions for student grade predictions", *Neurocomputing , Elsevier*,Volum 431 , 2021, Pages 23-33

[19] YawenZeng, Yong Ouyang, RongGao, Ye Qiu, Yonghong Yu, Chunzhi Wang , "HHA: An Attentive Prediction Model for Academic Abnormality", *IEEE Access*, Vol.8, pp.124755 – 124766, 2020

[20] Xiaoyong Li, Yong Zhang, Huimin Cheng, Feifei Zhou, Baocai Yin, "An Unsupervised Ensemble Clustering Approach for the Analysis of Student Behavioral Patterns", *IEEE Access*, Vol. 9, 7076 – 7091, 2021.

[21] https://archive.ics.uci.edu/ml/datasets/Educational+Process+Mining+%28EPM%29%3A+A+Learning+Analytics+Data+Set.