

# A Predictive Model to identify possible affected Bipolar disorder students using Naïve Baye's, Random Forest and SVM machine learning techniques of data mining and Building a Sequential Deep Learning Model using Keras

S. Peerbasha<sup>1\*</sup>, M. Mohamed Surputheen<sup>2\*</sup>

<sup>1\*</sup>[bashapeer2003@gmail.com](mailto:bashapeer2003@gmail.com), <sup>2\*</sup>[msurfudeen@yahoo.com](mailto:msurfudeen@yahoo.com)

Research Scholar<sup>1\*</sup>, Associate Professor<sup>2\*</sup>,

P.G. & Research Department of Computer Science – Jamal Mohamed College (Autonomous),  
Affiliated to Bharathidasan University, Tiruchirappalli, Tamilnadu, India.

## Summary

Medical care practices include gathering a wide range of student data that are with manic episodes and depression which would assist the specialist with diagnosing a health condition of the students correctly. In this way, the instructors of the specific students will also identify those students and take care of them well. The data which we collected from the students could be straightforward indications seen by them. The artificial intelligence has been utilized with Naive Baye's classification, Random forest classification algorithm, SVM algorithm to characterize the datasets which we gathered to check whether the student is influenced by Bipolar illness or not. Performance analysis of the disease data for the algorithms used is calculated and compared. Also, a sequential deep learning model is build using Keras. The consequences of the simulations show the efficacy of the grouping techniques on a dataset, just as the nature and complexity of the dataset utilized.

**Keywords** *Artificial intelligence · Bipolar illness · Data mining techniques · Naive Baye's classification · Random forest classification – SVM classification*

## 1. Introduction

Because of the present day of life used by the students and psychological illness is increasing quickly. The food habits and the way of life of the students lead to creating an impact on their health causing bipolar confusion. Data mining procedure is perhaps the most difficult and driving research areas in clinical consideration because of the great significance of the valuable data [1]. The recent sprouting in the data mining approaches has given a strong stage to different applications in the medical care field. In medical care, data mining is playing a crucial part in different fields like interruption identification, design recognition, less expensive clinical therapies accessibility for the patients, infection diagnosing, discovering its procurement methods [2,3]. Artificial intelligence makes the model more delicate and activates the framework to think. In machine learning,

artificial intelligence acts as a sub-field to perform better forecasts [4].

It additionally accommodates the researchers in the field of medical services being developed of successful strategies and different systems to prevent bipolar disorder, early identification of illnesses can diminish the danger factor. The point of our work is to foresee bipolar disorder among the prepared dataset using classification algorithms. It has been trained the Naive Baye's, Random forest classifier model, SVM model with bipolar dataset and their performances are calculated. SVM model provides better forecast accuracy when compared with Naive Baye's and Random forest classifier model. Likewise, it has been applied few sample test data of the disease to those classified models to show whether the student data in the sample test are experiencing that bipolar sickness or not [6].

Artificial neural networks are the best effort classification algorithm for the forecast of clinical determination because of their best effectiveness boundary [7]. The feed-forward neural network with a support vector machine (SVM) is the best method for the prediction of cancer [8]. The ANN is utilized to group the named images dependent on the assurance of the true positive (TP) and false-positive (FP) detection rates. The recognition system is performed with the self-organized supervised learning algorithm. ANN approach gives a promising outcome for the identification of micro-calcifications features and biopsy detection [9]. The marked images are determined in a marginal space forming a hyperplane which reduces the generalization error [10]. A mixture (hybrid) classifier is anticipated here by hybridizing support vector machine and fake neural organization (ANN) [7]. A regular ANN consists of one input layer, one or more hidden layers, and one output layer as in Fig. 1. Each layer has few neurons,

and the neurons in one layer are associated with the neurons in the contiguous layer with its connection loads [11]. The accompanying figure shows the artificial neural network model with one input layer, one output layer, and one hidden layer. The neuron represents a node in the network. The input features are taken care of by the neurons in the input layer.

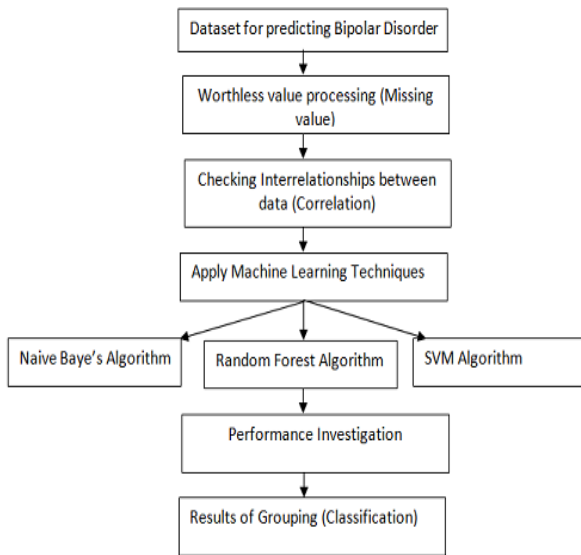


Fig. 1 Block diagram of proposed method

SVM is the managed (supervised) learning model, which can perform well even with the smaller data samples [12, 13]. SVM classifier has no curse of dimensionality since it can manage inadequate data in high-dimensional datasets [14]. Likewise, the SVM classifier has better generalization than the ANN and keeps away from the local minima problem. The rationale behind the SVM strategy is it makes the ideal (optimal) separation plane under linearly distinct conditions [12]. The hyperplane is upgraded by increasing the edge. The edge is the distance between the boundary and the nearest point of each class. These points nearest to the limit are called support vectors [15]. For medical care data examination, data mining techniques like grouping (classification), association rules, and clustering are normally utilized.

The primary commitment of this paper is summed up as follows: -

Initially, the student bipolar disorder dataset is taken as input for the framework to investigate so numerous data which are associated with the students (patients) health care and general data there in the dataset. These are life-threatening sicknesses for the students.

Data preprocessing is applied to the input datasets, it prompts to decrease the undesirable data for further investigation. Check the missing values and checking the interrelationship it assists to partition the training data into 80% unique data and 20% testing data, for efficient data investigation.

Data-mining algorithms like Naïve Baye's, Random Forest, and SVM are applied to assess the performance of the framework against the bipolar dataset. The grouping (classification) results are compared with existing outcomes, and it shows a better improvement.

## 2. LITERATURE SURVEY

Data mining is a developing field that changes a piece of data into valuable data. This strategy helps the approved person make informed options and make the right choices for their betterment [16]. It is used to comprehend, anticipate and control future behavior dependent based on the hidden patterns among large datasets. It prompts to offer tools for automated learning from the history of data and creating models to find the results of future situations. There are different tools for data mining machine learning algorithms to recognize and foresee the different illnesses as far as regression, decision tree, and Bayesian network [17]. Finding an illness required a distinctive test that brings about a variety of scenarios regarding the specific patient. By applying data mining, the idea for the concept for data analysis number of tests will be decreased. It assumes an important role in data investigation to improve performance and efficiency [18].

Variety of classification and clustering algorithms assumes a huge part in the forecast and conclusion of various kinds of sicknesses. Bayesian network classifiers and random forest classifiers are utilized to analyze the danger. [10, 19, 20]. The prediction exactness of the k-means algorithm is upgraded using both class and cluster methods and causing it to adjust to various datasets [21]. A group of classification algorithms excluding random forest algorithms is applied to bipolar data to analyze the danger. On comparing the presentation of each technique, the outcome shows that Random Forest was performed well in both exactness and ROC curve [8, 22, 23].

In ANN hybrid classifier, each neuron of the hidden and output layer gets signals from the past layer multiplied with loads of the interconnection. The neuron then delivers the output bypassing the added signal through the defined transfer function.

The network is prepared for the given input iteratively. In every cycle, the mean square error (MSE) between the

target and the achieved output is calculated. The MSE for the  $j$ th iteration is characterized in Equation. (1) as follows,

$$F(X)=(t_i-a_i)^2$$

$t_j$  and  $a_j$  are the targeted and the achieved output, respectively. The network is prepared by changing the weights and the bias so that the MSE get limited. The MSE estimates the probability function for the classification issue. Despite the fact that ANNs are good classifiers, they require huge number of preparing sets to train for proper behavior. That is the reason it established fine if hybridized with the classifiers which could require a smaller number of training samples to classify appropriately. Neural network-based cancer classifiers are utilized with binary and multi-class problems to recognize the cancer samples [8].

Applied Naïve Baye's algorithm building an artificial intelligent framework, based on the comparison of specific boundaries used to oversee whether a person is having bipolar problem or not [2, 3]. The artificial intelligent-based methods are capable and fashionable one in recent years [24]. The diagnosis of diabetes and cancer forecasting the adaptive neuro-fuzzy inference framework shows better exactness [12]. Additionally shows the exactness data of Naive Baye's classifier and K-means algorithm. 80% exactness is obtained from this method [12]. Modified extreme learning machine and back-propagation neural network techniques are tended in forecasting of diabetes [13]. The data mining techniques such as K-means algorithms, decision tree algorithm and other classification algorithms provide consistent presentation in diagnosing the heart sickness [7, 16, 25, 26]. It helps a non-specialized specialist to make the correct choice about the disease risk level by generating original rules, pruned rules, classified rules and sorted rules [27].

### 3. PROPOSED METHOD

The Proposed method has been used with Anaconda tool (AEN 3.1 Version) for data analysis. Anaconda is a package management framework utilized for predictive analysis and data management [28]. It has been taken the students data such as Bipolar disorder disease data as input, the reason for choosing these datasets only because of life threatening characteristics and discover the effectiveness of the proposed technique in productive way and some relativity is there between the data. These data are loaded and verified if has any missing values. If any missing values are discovered, they are replaced to a null value. At that point it has been checked whether any segments in the information have any connection with another section in the information separately. If any correlation is found between two columns, one of those

columns is removed. In the event that any obvious and bogus worth is found in information, it is changed to 1 and 0 separately. It has split the original data into training data which has 80% of original data and test data which has 20% of the original data.

To check the quantity of valid and bogus cases in unique, preparing and test information of the three-class information, it has been trained our three different class data with the Naïve Baye's algorithm and calculate the results accuracy, this algorithm gives to the three classes separately using confusion matrix [29]. The block chart of the proposed technique is appeared in Figure-1. The exhibition report shows the presentation measurements of the precision estimation for each class information separately. Similarly, it has trained our three different class data with the Random forest algorithm and SVM algorithm. It has determined the output accuracy and this algorithm gives the three classes separately using confusion matrix. The presentation report shows the exhibition measurements of the exactness computation for every information independently.

It has taken not many example test information independently for each class information. Applying this example information on each prepared model of that infection shows us the outcomes if the information are related to that sickness. While comparing the results of both models for each class data, it can see that the model trained with SVM algorithm gives the accurate results of classification. To discover the productivity of the proposed strategy, the prepared information are contrasted independently and the proposed calculations and furthermore checked the presentation of test information. The proposed method is also applicable for testing the real-time disease data for classification and to identify whether the student is affected by bipolar disorder or not.

The data were collected under the following attributes such as Serial number, Name, Age, Short temper, Lack of concentration, Presence of mind, Sudden Laugh, Adamant, Depression, Change in mood, Presence of Energy, Sleeplessness, Extreme happiness, Hopelessness, Restlessness, Repeated Speech, Judgment, Panic, Slugish, Extraordinary confidence, Continuous talk, Misunderstanding, Subject-1, Subject-2, Subject-3, Subject-4, Scr1, Scr2, Lrcab, Time Stamp. Out of which 16 attributes were selected for the prediction.

This is an information data set of infection manifestation affiliations created by a computerized strategy dependent on data provided by the affected students of the various colleges of Tamilnadu especially from Jamal Mohamed College during 2018 – 2021. The

principal column shows the sickness, the second the quantity of affected students containing a positive and current mention of the disorder, and the flow notice of the disorder, and the related indication.

Kindly contact bashapeer2003@gmail.com for any inquiries with respect to the information data set.

### Steps required building a classifier:

Initialize: Model the classifier to be used

Train: Train the classifier utilizing a decent preparing information

Anticipate: Pass on to information X to the model that assesses the information to predict(X)

Evaluate: Evaluate the model

Using the various parameters of the bipolar disorder data set (independent variables), we have to classify a given dataset using Naïve Baye's, Random forest and SVM classification model. The following steps were involved:-

#### Step-1: Importing the libraries

We import the libraries using NumPy, Pandas and Matplotlib

```
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
```

#### Step-2: Importing the dataset

In this step, we import the bipolar.csv and save it to the variable dataset. After this, we assign the 4 independent variables to X and the dependent variable 'depression' to Y. The first five rows of the dataset used are displayed using the following code: -

```
dataset = pd.read_csv("path")
X = dataset.iloc[:,4].values
y = dataset['depression'].valuesdataset.head(5)
```

#### Step-3: Partitioning the dataset into the Training set and Test set

After obtaining the data set, we should split the data into the training set and the test set. In this data set, there are 2500 rows. As each class is given in a continuous order, we should randomly split the dataset. Here, we have the `test_size=0.2`, which means that **20%** of the dataset will be used for testing purpose as the **test set** and the remaining **80%** will be used as the **training set** for training the Naive Bayes, Random Forest and SVM classification model.

```
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, y,
test_size = 0.2)
```

#### Step-4: Feature scaling

It can be done using the following code:

```
from sklearn.preprocessing import StandardScaler
sc = StandardScaler()
X_train = sc.fit_transform(X_train)
X_test = sc.transform(X_test)
```

#### Step-5: Training the Naïve Baye's, Random Forest and SVM Classification model on the Training set

#### Step-6: Predicting the Test set results

After the model is trained, we should use the `classifier.predict()` to predict the values for the Test set and the values predicted are stored to the variable `y_pred`.

```
y_pred = classifier.predict(X_test)
y_pred
```

#### Step-7: Confusion matrix and accuracy

This step is used in almost all the classification techniques. Here, use the Accuracy of the trained model and plot the confusion matrix. The confusion matrix is used to show the number of correct and incorrect predictions on a classification problem when the real values of the Test Set are known. It is of the format

TP	FP
FN	TN

Where TP stands for True Positive, FP – False Positive, FN – False Negative and TN – True NegativeThe following code is used.

```
from sklearn.metrics import confusion_matrix
cm = confusion_matrix(y_test, y_pred)from sklearn.metrics
import accuracy_score
print ("Accuracy : ", accuracy_score(y_test, y_pred))
cm>>Accuracy : 0.92222222222222
```

**Step-8: Comparing the Real qualities with Predicted esteems**

Here, a Pandas DataFrame is made to compare the classified values of both the original Test set (*y\_test*) and the predicted results (*y\_pred*).

```
df = pd.DataFrame({'Real Values':y_test, 'Predicted
Values':y_pred})
df
```

**Building Sequential Deep Learning Model using Keras**

The following steps were carried out in building sequential deep learning model using Keras.

1. Necessary library functions should be imported
2. Preprocessing
3. Features Selection
4. Average score for each record to be found.
5. Forming of clusters based on average score and age.
6. Testing and Training data should be given.
7. Sequential Deep Learning Model is used with necessary number of input layers, hidden layers and output layers.

```
model = Sequential()

model.add(Dense(32, input_dim=16,
activation='relu'))

model.add(Dense(32, activation='relu'))

model.add(Dense(3, activation='softmax'))

model.summary()
```

8. Activation function should be used to introduce non-linearity into the output of a neuron for input layer, hidden layer and output layer.

9. Optimizer should be used to reduce the losses of accuracy.
10. Represent the total number of epochs (iterations) to be used: 200 to find the accuracy and its loss.
11. Evaluation has been done and finally the prediction of bipolar disorder is made.

1.	Library functions imported	Pandas, Seaborn, Matplotlib, Numpy, Keras, Sklearn
2.	Preprocessing	Number of records before preprocessing: 2500 Number of records after preprocessing: 2406
3.	Original Number of Features (Sl.No, Name, Age, Stemp, Clack, Pmind, Slaugh, Admt, Dip, Cmoood, Penergy, Psleep, Ehapp, HopeFl, Rstls, RSpeech, Judmnt, Panic, Slugish, Exconf, Ctalk, Mis-han, Sub-1, Sub-2, Sub-3, Sub-4, Scr1, Scr2, Lrcab, Ts)	30
4.	Number of Features Selected (Stemp, Clack, Pmind, Slaugh, Admt, Dip, Cmoood, Penergy, Psleep, Ehapp, HopeFl, RStls, RSpeech, Panic, Exconf, Ctalk)	16
5.	Finding of average score for each record	df3["Total_Score"] = df3.mean(axis=1)
6.	Number of clusters formed based on average score	3 (Low, Moderate, High)
7.	Number of clusters formed based on age	3 (Minor, Major, Adult)
8.	Deep Learning Model used	Sequential model
9.	Number of Input Layers used  Number of Hidden Layers used  Number of Output Layers used	1 (16 Neurons used)  2 (Each with 32 Neurons)  3
10.	Activation functions used  Input Layer →First Hidden Layer  First Hidden Layer →Second Hidden Layer  Second Hidden Layer →Output Layer	Relu  Relu  Softmax
11.	Optimizer used	Adam
12.	Number of Epoch used	200
13.	Accuracy percentage  Loss percentage	99.71  0.29

**Table-2** Samples used in Sequential Deep Learning Model

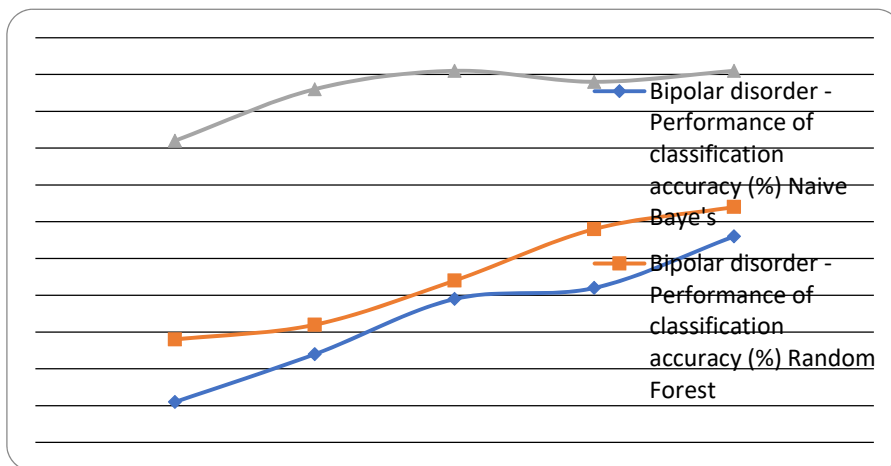
**4. RESULTS AND DISCUSSION**

According to the proposed methods, the collective bipolar disorder student dataset is processed with feature selection

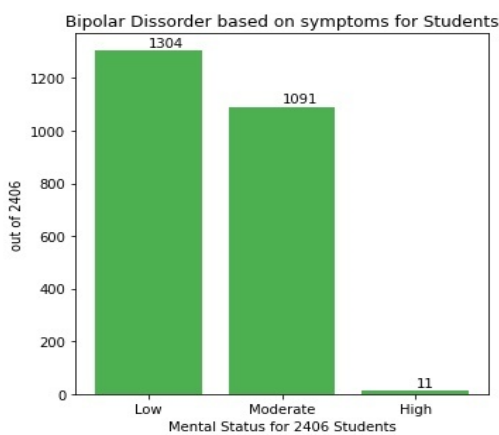
and classification. The implementation algorithm is test with the confusion matrix and comparison is made between Naïve Baye’s, Random Forest and SVM algorithms. The performance evaluation comparisons of the proposed system and the conventional system are tabulated below.

**Table-1** Performance of classification accuracy

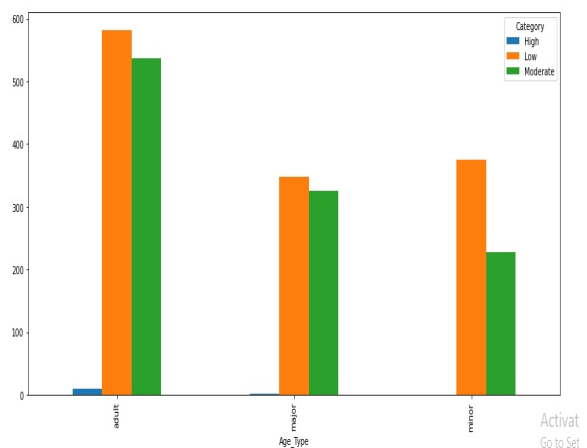
The number of records used	Impact of Classification Accuracy (%)		
	Naïve Baye’s	Random Forest	SVM
500	83.1	84.8	90.2
1000	84.4	85.2	91.6
1500	85.9	86.4	92.1
2000	86.2	87.8	91.8
2500	87.6	88.4	92.1



**Figure-2** Performance of Classification Accuracy



**Figure-3** Mental Status of the Students



**Figure-4** Age based mental status



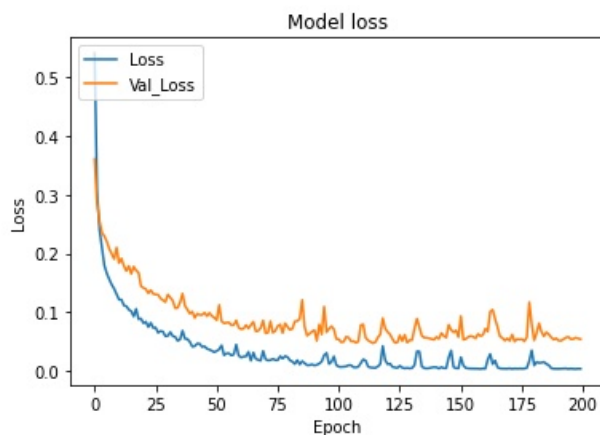


Figure-5 Model Loss

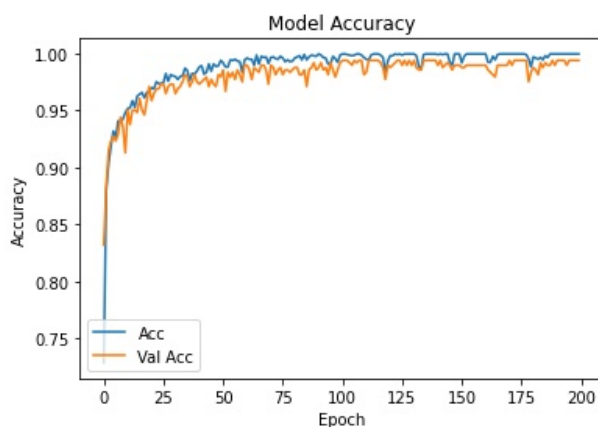


Figure-6 Model Accuracy

## 5. CONCLUSIONS

The comparison of Naive Baye's, Random Forest and Support Vector machines is made and applied for testing the real-time bipolar disease classification and identification of bipolar disorder within the students or not is presented in this paper. The findings indicate that the Support Vector Machine classification method has achieve the highest (min 90.2%, max 92.1 %) classification accuracy in comparison with Naive Baye's and Random forest classification methods. On the contrary, Naive Baye's has got the lowest average accuracy values (min 83.1%, max 87.6%). The investigation indicates that increasing the size of the training data set per class leads to insignificant growth of the classification accuracy (1 – 2 %) of Naive Baye's, Random forest and Support Vector machines classifiers. These results show that a training set size 500 per class is sufficient for all analyzed classification methods, and classification accuracy relates more to the properties. The

Sequential Deep learning model is used for the prediction of bipolar disorder students and found that the accuracy is 99%.

## References

- [1] Renjit JA, Shanmuganathan KL, "Distributed and helpful multi-specialist based interruption discovery framework", *Indian J Sci Technology* 3(10): 2010, pp. 1070–1074
- [2] Priyadarshini R, Dash N, Mishra R, "An epic way to deal with anticipate diabetes mellitus utilizing changed limit learning machine". In: *International Conference on Electronics and Communication Systems (ICECS)*, 2014, pp 1–5.
- [3] Sankaranarayanan S, Perumal TP, "Diabetic guess through information mining techniques and techniques". In: *International Conference on Intelligent Computing Applications*, 2014, pp 162–166
- [4] Dahiwade D, Patle G, Meshram E, "Designing sickness forecast model utilizing AI approach.", In: *Third IEEE International Conference on Computing Methodologies and Communication (ICCMC)*, 2019.
- [5] Geetha R, Sivasubramanian S, Kaliappan M et al., "Cervical disease distinguishing proof with engineered minority oversampling strategy and PCA investigation utilizing irregular woodland classifier", *J Med Syst* 43:286. <https://doi.org/10.1007/s10916-019-1402-6>
- [6] Annamalai S, Udendhran R, Vimal S, "A wise framework network dependent on distributed computing infra structures". *Nov Pract Trends Grid Cloud Computing*. <https://doi.org/10.4018/978-1-5225-9023-1.ch005>
- [7] Wu H, Yang S, Huang Z, He J, Wang X (2018) Type 2 diabetes mellitus forecast model dependent on information mining. *Advise Med Unlocked* 10:100–1075218 V. Jackins et al. 1 3
- [8] Sarwar A and Sharma, "Intelligent Naïve Bayes way to deal with analyze diabetes type-2", In: *Special Issue of International Journal of Computer Applications on Issues and Challenges in Networking, Intelligence and Computing Technologies*, November 2012.
- [9] Pradeepa S, Manjula KR, Vimal S et al (2020) DRFS: "Recognizing hazard factor of stroke infection from online media utilizing AI strategies", *Neural Process Lett*. <https://doi.org/10.1007/s11063-020-10279-8>.
- [10] Kalaiselvi C, Nasira GM (2014) another methodology of conclusion of diabetes and expectation of malignancy utilizing ANFIS. In: *IEEE Computing and Communicating Technologies*, 2014, pp 188–190.
- [11] Robinson YH, Vimal S, Khari M, Hernández FCL, Crespo RG, "Tree-based convolutional neural organizations for object classification in portioned satellite pictures", *Int J High Perform Comput Appl*. <https://doi.org/10.1177/1094342020945026>.

- [12] Undre P, Kaur H, Patil P, "Improvement in expectation rate and precision of diabetic determination framework utilizing fluffy rationale mixture blend", In: International Conference on Pervasive Computing (ICPC), 2015, pp 1–4
- [13] Yi Y, Wu J, Xu W, "Incremental SVM dependent on saved set for network interruption recognition", Elsevier J Expert Syst Appl 38(6):7698–7707
- [14] Ramamurthy M, Krishnamurthi I, Vimal S, Harold Y, "Robinson profound learning based genome examination and NGS-RNA LL identification with a novel crossover model", 2020, 197: 104211. <https://doi.org/https://doi.org/10.1016/j.biosystems.2020.104211>
- [15] Pradeepa S, Gayathri P, Nishmitha P, Vimal S, Oh-Young S, Usman T, Raheel N (2020) IoT based wellbeing related point acknowledgment from arising on the web wellbeing local area: drug help utilizing AI method. Gadgets 9(9):1469.
- [16] Babu S, Vivek EM, Famina KP, Fida K, Aswathi P, Shanid M, Hena M (2017) Heart sickness determination utilizing information mining procedure. In: International Conference on Electronics, Communication, and Aero space Technology, ICECA2017
- [17] Sampaul TGA, Robinson YH, Julie EG, Shanmuganathan V, Nam Y, Rho S, "Diabetic retinopathy diagnostics from retinal pictures dependent on profound convolution networks", Preprints. <https://doi.org/organization/10.20944/preprints2020.05.0493.v1>
- [18] Vimal S et al, "Deep learning-based dynamic with WoT for brilliant city advancement", In: Jain A, Crespo R, Khari M (eds) Smart development of web of things, CRC Press, Boca Raton, 2020, pp 51–62. , <https://doi.org/10.1201/9780429298462>
- [19] Kumari M, Vohra R, Arora A, "Prediction of diabetes utilizing Bayesian organization", Int J Comput Sci Inf Technol (IJCSIT) 5(4):2014, 5174–5178
- [20] Krishnaiah V, Narsimha G, Chandra NS, "Diagnosis of cellular breakdown in the lungs expectation framework utilizing information mining classification strategies", Int J Comput Sci Inf Technol 4(1): 2013, 39–45
- [21] C. Li Sa, D. H. b. Abang Ibrahim, E. Dahliana Hossain and M. canister Hossin, "Understudy execution investigation framework (SPAS)," The fifth International Conference on Information and Communication Technology for The Muslim World (ICT4M), Kuching, 2014, pp. 1-6.
- [22] G. Barata, S. Gama, J. Jorge and D. Gonçalves, "Early Prediction of Student Profiles Based on Performance and Gaming Preferences," in IEEE Transactions on Learning Technologies, vol. 9, no. 3, 2015, pp. 272-284.
- [23] Singh, A. S. Sabitha and A. Bansal, "Understudy execution investigation utilizing a bunching calculation," 2016 sixth International Conference -

Cloud System and Big Data Engineering (Confluence), Noida, 2016, pp. 294-299.

- [24] Duru, G. Dogan and B. Diri, "An outline of learns about understudies' presentation investigation and learning examination in MOOCs," 2016 IEEE International Conference on Big Data (Big Data), Washington, DC, 2016, pp. 1719-1723.
- [25] M. Agaoglu, "Anticipating Instructor Performance Using Data Mining Techniques in Higher Education," in IEEE Access, vol. 4, 2016, pp. 2379-2387.

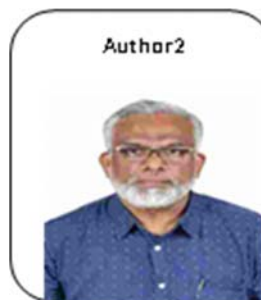
## AUTHORS PROFILE



**Mr. S. Peerbasha**, a Research Scholar at Jamal Mohamed College (Autonomous) has cleared his National Level Eligibility Test conducted by UGC NTA NET in the year 2020 and has cleared his State Level Test (TNSET-2018) conducted by Mother Teresa University, Kodaikanal. He received his M.Tech at School of Computer Science & Engineering, Bharathidasan University in 2012.

He received his M.B.A at Alagappa University in 2011. He has completed his Master of Philosophy (C.S) in 2008 and Master of Computer Applications in 2007. He is currently working as an Assistant Professor of Computer Science at Jamal Mohamed College for 10 Years. He was also working as a Senior Lecturer for Southern Cross University, Australia in 2016. He is acting as a Member in Guidance and Counseling Centre, Jamal Mohamed College (Autonomous). His areas of research include Data mining, Machine learning and Deep Learning.

Communication mail id: [bashapeer2003@gmail.com](mailto:bashapeer2003@gmail.com)



**Dr. M. Mohamed Surputheen**, Working as a Associate Professor in Jamal Mohamed College with more than 20yrs experience. Education Qualification is MSc., Mphil and PhD. Guiding Mphil Scholars and around eight PhD Scholars. Published and Presented more than twenty five research papers. Currently acting as a Controller of Examinations at Jamal Mohamed College

(Autonomous), Trichy. Communication mail id: [msurfudeen@yahoo.com](mailto:msurfudeen@yahoo.com)