International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

SPE: Ensemble Hybrid Machine Learning Model for Efficient Diagnosis of Brain Stroke towards Clinical Decision Support System (CDSS)

D Ushasree[1], AV Praveen Krishna[2], Ch Mallikarjuna Rao[3], D V Lalita Parameswari[4]

Submitted: 05/11/2022 **Revised:** 19/12/2022 **Accepted:** 02/01/2023

Abstract: As per World Health Organization (WHO) Brain stroke is the second underlying cause of death categories or major ICD(International Cause of Death) and disability across the globe . Artificial Intelligence (AI) enabled approaches using Machine Learning (ML) are widely used for stroke detection automatically in a non-invasive fashion with data-driven approach. However, from the literature, it is understood that there is need for improving quality of training and also find best classifiers to ensemble them to enhance prediction performance. In this paper, we proposed a framework known as Stroke Prediction Ensemble (SPE) which exploits a hybrid approach considering feature engineering and ensemble classification. From multiple brain stroke prediction models, best models that exhibit accuracy >90% are chosen for ensemble model. Two algorithms are proposed to realize the framework. They are known as "Hybrid Measures Approach for Feature Engineering (HMA-FE)" and Hybrid Ensemble and Feature Engineering for Stroke Prediction (HEFE-SP). The former is published in our prior work which is meant for finding best features from given dataset while the latter is meant for ensemble ML towards more efficient stroke prediction performance. Empirical study has revealed that our ensemble model showed highest accuracy with 97.93% while the average accuracy of all constituent base line models is 95.25%. Thus the ensemble model can be used for efficient brain stroke diagnosis as part of Clinical Decision Support System (CDSS).

Keywords: Brain Stroke Detection, Clinical Decision Support System (CDSS), Ensemble Learning, Feature Engineering, Machine Learning.

1. Introduction

Machine learning has capability to solve many problems in real world applications. This proposition is found to be interesting in the recent past due to the realization of different solutions in medical and other fields. Since brain stroke is one of the underlying cause of death and disabilities, ML which is part of Artificial Intelligence (AI) is considered to be useful research area for leveraging stroke diagnosis. This approach can be exploited in addition to the conventional methods [1]. ML models are widely used to detect brain stroke automatically using data-driven approach as explored in [2], [3], [4] and [5]. Shoily *et al*. [2] focused on different algorithms with specific functionality towards stroke detection. Javeria *et al*. [3] opined that features in dataset play crucial role in stroke detection. They proposed a feature selection and fusion approach to improve quality in training process. Sung *et al*. [4] focused on emergency triage model that helps in alerting patients in order to have medical intervention early. Shahriar [5] explored different ML techniques that paved way for improving detection of brain abnormalities.

There are also ensemble models found in the literature for stroke research as investigated in [11], [19] and [21]. Luca *et al*. [11] combined the techniques such as feature selection and ensemble learning towards diagnosis of medical diseases by exploiting known ML prediction models. Kumar *et al*. [19] proposed ensemble approach in order to detect brain abnormalities automatically. It is a learning based approach that makes use of multiple underlying baseline methods in the ensemble towards better accuracy. Agnes *et al*. [21] focused on ensemble learning towards detection of brain stroke for clinical diagnosis. From the literature it is observed that the existing methods strive to diagnose brain stroke using ML models. However, there is need for a more comprehensive framework that follows a systematic ensemble approach with efficient feature selection to form a hybrid candidate for stroke diagnosis. Towards this end, our contributions in this paper are as follows.

1. We proposed a framework known as Stroke Prediction Ensemble (SPE) which exploits a hybrid approach considering feature engineering and ensemble classification.

Department of CSE, Koneru Lakshmaiah Education Foundation[1],Department of CSE, Koneru Lakshmaiah Education Foundation[2], Department of CSE,GRIET[3], Department of CSE, GNITS[4]
dupakuntlausha@gmail.com[1], praveenkrishna@kluniversity.in[2], professorcmrao@gmail.com[3], lplalitap97@gmail.com[4]

2. Two algorithms are proposed to realize the framework. They are known as "Hybrid Measures Approach for Feature Engineering (HMA-FE)" and Hybrid Ensemble and Feature Engineering for Stroke Prediction (HEFE-SP). The former is published in our prior work which is meant for finding best features from given dataset while the latter is meant for ensemble ML towards more efficient stroke prediction performance.
3. Empirical study has revealed that our ensemble model showed highest accuracy with 97.93% while the average accuracy of all constituent base line models is 95.25%. Thus the ensemble model can be used for efficient brain stroke diagnosis as part of Clinical Decision Support System (CDSS).

The remainder of the paper is structured as follows. The different methods currently used by ML for stroke detection are described in Section 2. Section 3 presents materials and methods associated with the brain stroke research presented in this paper. Section 4 presents experimental results pertaining to our empirical study. Section 5 draws conclusions and provide future possibilities.

2. Materials And Methods

Our methodology for brain stroke detection is data-driven ensemble machine learning approach. Dataset is collected from [27]. It has 12 attributes pertaining to patients with 5110 observations. Out of the 12, there are 11 clinical features that reflect patient details including vitals and one attribute holds ground truth or class label.

2.1 The Framework

Our approach to efficient brain stroke detection is based on ensemble learning which is crucial for improving prediction accuracy. There are widely used ML classifiers available. However, towards Clinical Decision Support System (CDSS) there is need for highly accurate model to enable doctors to diagnose probability of brain stroke. Therefore, our approach has mechanism to identify best models and then make ensemble of them instead of choosing all classifiers as part of ensemble. Our approach is illustrated in Fig 1.

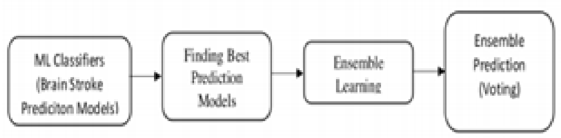

Fig. 1. Overview of our ensemble methodology

We begin by assessing each classifier that can be used to detect brain strokes using a data-driven approach.After feature engineering (based our proposed feature engineering algorithm) all prediction models are evaluated with the dataset [27]. We considered accuracy as criterion and threshold is set at >90%. This is the modus operandi to find out best prediction models. Only models that pass this test are used for ensemble learning. A weighted majority voting is the phenomenon used to determine final prediction for each test instance.

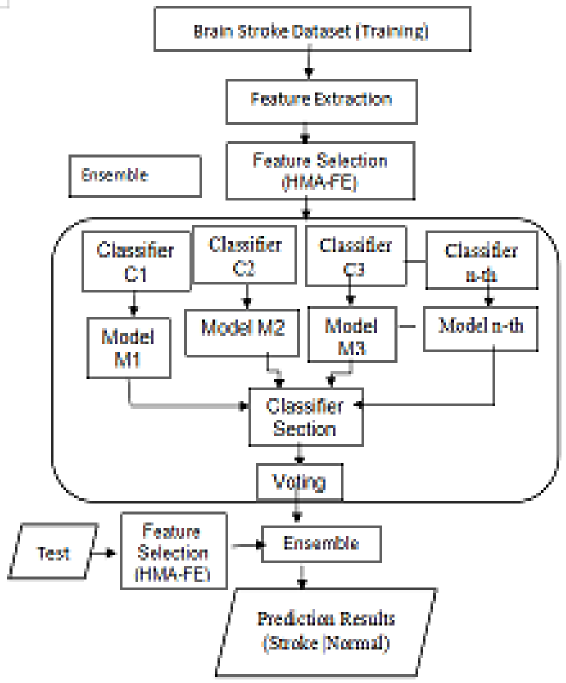

Fig. 2. Our framework known as Stroke Prediction Ensemble (SPE)

As presented in Fig 2, it reflects our framework meant for more efficient prediction of brain stroke. The given dataset is subjected to feature extraction which is meant for identifying all features from 11 attributes of the dataset. Then, the proposed "Hybrid Measures Approach for Feature Engineering (HMA-FE)" algorithm is used to determine the best characteristics that aid in class label prediction. More details of HMA-FE are available in our previous research paper. However, for completeness, we are including algorithm and some information in this paper. Therefore, the selected features are given to classifiers that are part of ensemble. Each classifier results in a knowledge model. Afterwards, as discussed earlier and illustrated in Fig 1, our approach uses only the selected ML classifiers that show accuracy >90% to be part of final voting method. We followed weighted majority voting method which is more useful in improving accuracy. Once ensemble model is built, it takes test dataset and detects brain stroke. Finally, the framework results in classification of given test instances into NORMAL and STROKE.

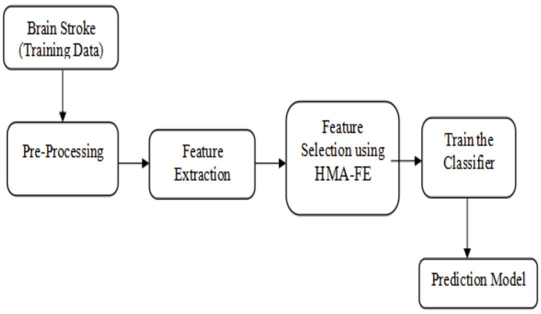

Fig. 3. Procedure in the training phase is illustrated

The training data is passed through pre-processing, which deals with missing values, as depicted in Fig 3. Feature extraction is the process in which features are identified from the 11 attributes of the dataset [27]. Afterwards, our feature engineering algorithm HMA-FE is applied to find best features. Instead of training a classifier with entire data, only selected features are considered to improve quality of training.

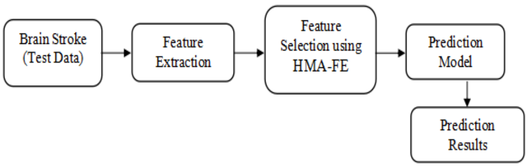

Fig. 4. Illustrates procedure in testing phase

As presented in Fig 4, the testing data is subjected to feature extraction. Feature extraction is the process in which features are identified from the 11 attributes of the dataset [27]. Afterwards, our feature engineering algorithm HMA-FE is applied to find best features. Instead entire data, only selected features are considered to improve quality of testing. A trained model is used to perform classification of test instances into NORMAL and STROKE.

2.2 Weighted Majority Voting Approach

In our framework SPE, weighted majority voting is followed as part of final determination of class label for given test instance. It is a soft approach that makes decisions based on probabilities to improve prediction accuracy. STROKE and NORMAL probabilities are weighted and final prediction is determined instead of making hard decisions. This approach is expressed in Eq. 5.

$$\sum_{t=1}^{T} w_t\, d_{t,c^*} = max_c \sum_{t=1}^{T} w_t\, d_{t,c} \qquad (5)$$

Where number of different prediction models is denoted by T. The t^{th} classifier decision is denoted as $d_{t,c}$. The t^{th} classifier weight is denoted as w_t. It chooses a class based on the weighted majority voting.

2.3 Feature Engineering

In our prior work [26] we proposed a hybrid mehtod for feature engineering. It is known as “Hybrid Measures Approach for Feature Engineering (HMA-FE)”. This algorihtm is designed to have multiple filter methods combined. It exploits a composit metric that could improve performance in predicting feature imortance. Only the selected featuers are used to train classifiers that are part of ensemble in this paper. More technical details can be found in our published work [26].

2.4 Stroke Prediction

For efficient stroke prediction, we proposed the Hybrid Ensemble and Feature Engineering for Stroke Prediction (HEFE-SP) algorithm.It is a hybrid approach that considers feature engineering and ensemble of best classifiers. It exploits HMA-FE algorithm for selecting best features from the dataset. Figure 5 shows the selection of baseline classifiers used for ensemble method.

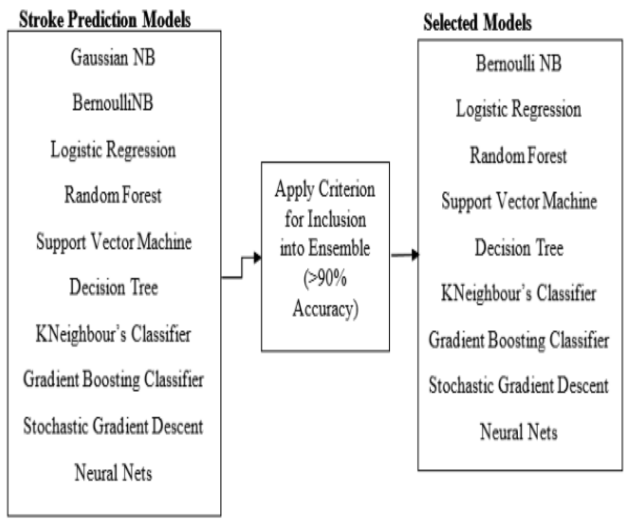

Fig. 5. Shows the baseline models chosen based on given criterion

The list of prediction models used for stroke prediction are provided. From the list of all models, accuracy based threshold (>90% accuracy) is applied to choose final list of classifiers to be used in the ensemble model.

Algorithm: Hybrid Ensemble and Feature Engineering for Stroke Prediction (HEFE-SP)
Inputs:
Patient Brain Stroke Dataset- *D*
Selected stroke prediction classifiers- *B*
Output:
Stroke Classification Results *R'*

1. Begin
2. (*T1*, *T2*)🡐SplitData(*D*)
3. *F*🡐Run HMA-FE(*T1*)

Model Training

4. For each model *b* in *B*
5. *t*🡐Train baseline model *b* using selected features *F*
6. Update models map M with *b* and *t* to M

```
7.    End For
Model Testing
8.    F←Run HMA-FE(T1)  //our algorithm published in [26]
9.    For each entry m in M
10.     Fit the model m.t for T2 using F
11.     Update results vector R with results of m.t
12.     Print results of confusion matrix
13.     Add R and m.t to ensemble map E
14.   End For
Ensemble Decision
15.   R'←ApplyVoting(E)
16.   Display results after ensemble decision R'
17.   End
```

Algorithm 1: Hybrid Ensemble and Feature Engineering for Stroke Prediction (HEFE-SP) algorithm

It uses the patient brain stroke dataset D and chosen stroke prediction classifiers B as inputs and outputs stroke classification results R', as shown in Algorithm 1. The given dataset is split into 80% and 20% training set and testing set respectively. Both training and test data are subjected to feature engineering using our proposed algorithm HMA-FE. The selected baseline models are trained and then they are used as ensemble model. The ensemble model results in final predictions based on weighted majority voting approach.

2.5 Evaluation Methodology

Performance of different prediction models is evaluated using confusion matrix based measures like precision, recall, F-measure and accuracy. These measures are based on the predictions of the models in terms of number of true positives (TP), number of true negatives (TN), number of false positives (FP) and number of false negatives (FN).

$$\text{Precision (p)} = \frac{TP}{TP+FP} \qquad (1)$$

$$\text{Recall (r)} = \frac{TP}{TP+FN} \qquad (2)$$

$$\text{F1-score} = 2 * \frac{(p*r)}{(p+r)} \qquad (3)$$

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \qquad (4)$$

Performance metrics is computed as in Eq.1,2,3 & 4. These metrics hold a value between 0.0 and 1.0 reflecting least and highest performance. Positive prediction capability is measured by precision while true positive rate is measured by recall. The harmonic mean of these two is measured by F1-scor. Accuracy is another measure used to know performance of the models.

3. Results And Discussion

This section presents experimental results and evaluate the performance of ensemble model and the constituent baseline models that are part of ensemble. Besides,Results of exploratory data analysis are also provided in this section.

3.1 Exploratory Data Analysis

The dataset [27] is subjected to exploratory data analysis to ascertain certain facts prior to using it for stroke prediction.

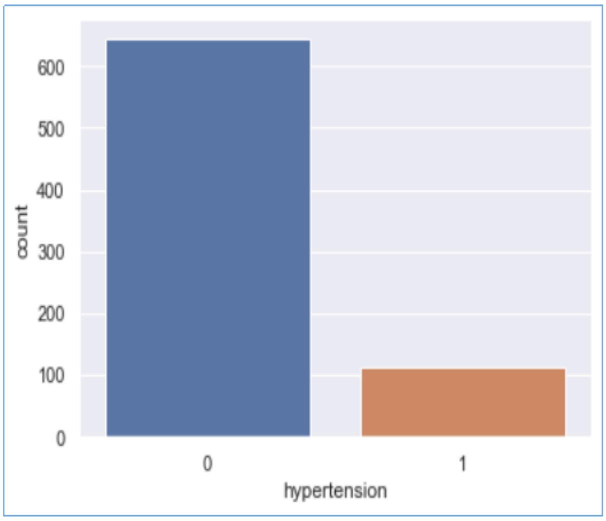

Fig. 6. Normal and stroke distribution dynamics among patients based on hypertension

Fig 6 illustrates the dynamics of normal and stroke distribution among patients based on the hypertension attribute.

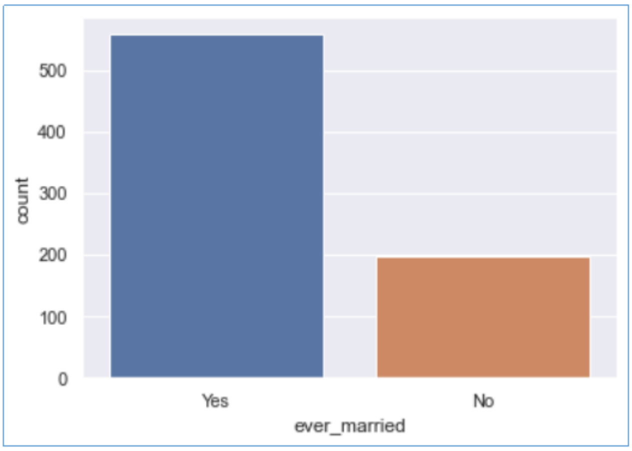

Fig. 7. Marital status distribution among the patients in given dataset

As presented in Fig 7, it considers the attribute ever_married and reflects on marital status distribution among the patients in given dataset.

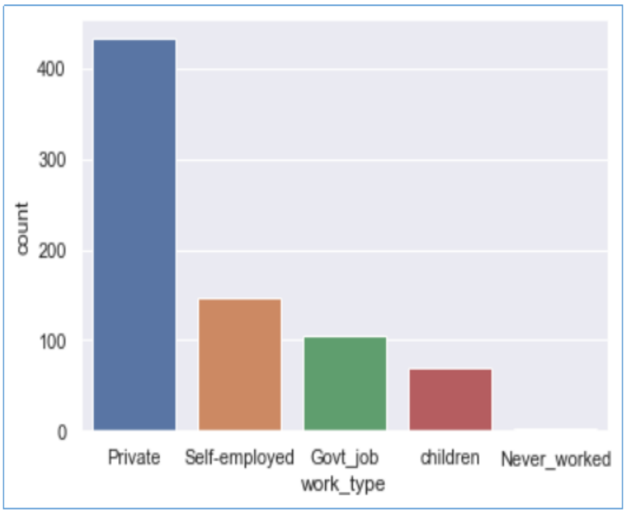

Fig. 8. Data analysis based on work_type attribute

As presented in Fig 8, it considers the attribute work_type and reflects on its values and their distribution statistics in the stroke dataset.

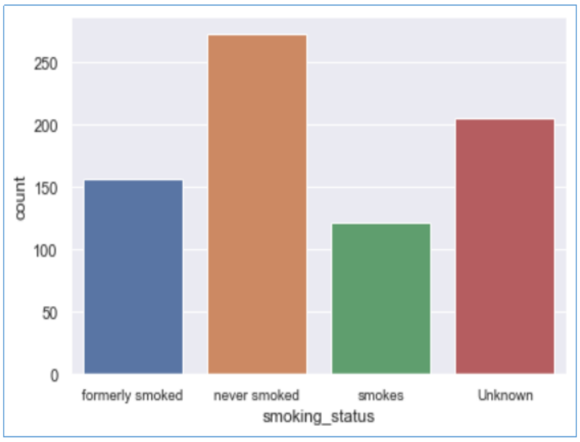

Fig. 9. Data analysis based on smoking_status attribute

As presented in Fig 9, it considers the attribute smoking_status and reflects on its values and their distribution statistics in the stroke dataset.

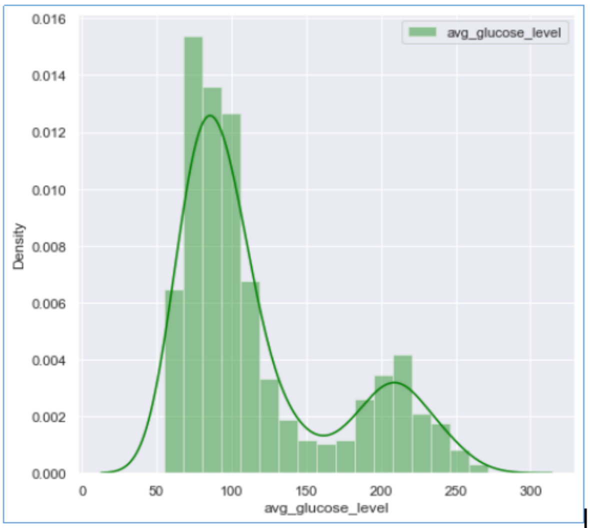

Fig. 10. Average glucose level and density distribution among the patients in given dataset

As presented in Fig 10, it considers the attribute avg_glucose_level and reflects on its values and their distribution statistics in the stroke dataset.

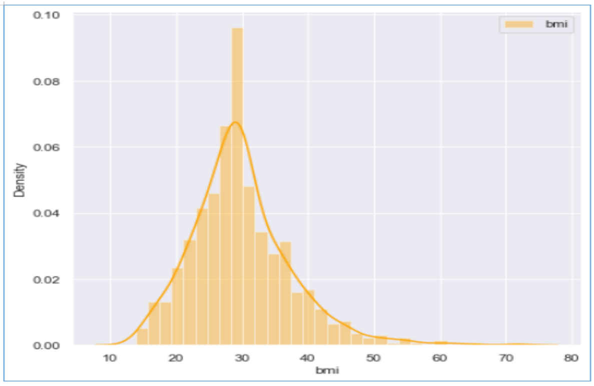

Fig. 11. Body mass index and density distribution among the patients in given dataset

As presented in Fig 11, it considers the attribute bmi and reflects on its values and their distribution statistics in the stroke dataset.

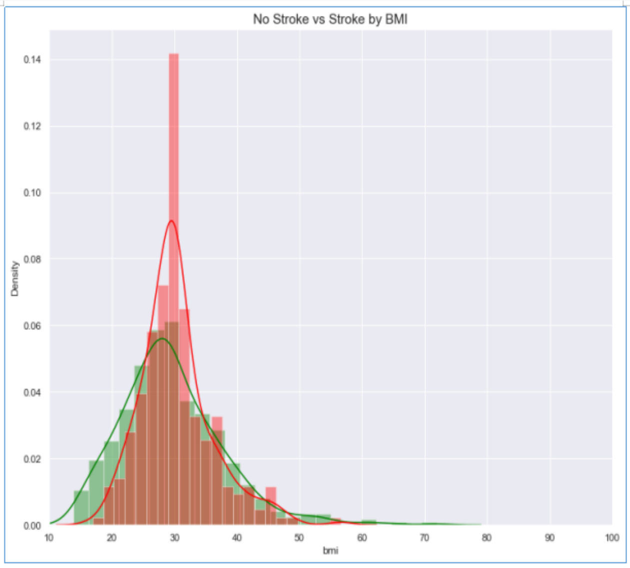

Fig. 12. Body mass index and density distribution for stroke and normal patients

As presented in Fig 12, it considers the attribute bmi and reflects on its values and their distribution statistics in the stroke dataset against stroke and no stroke class label.

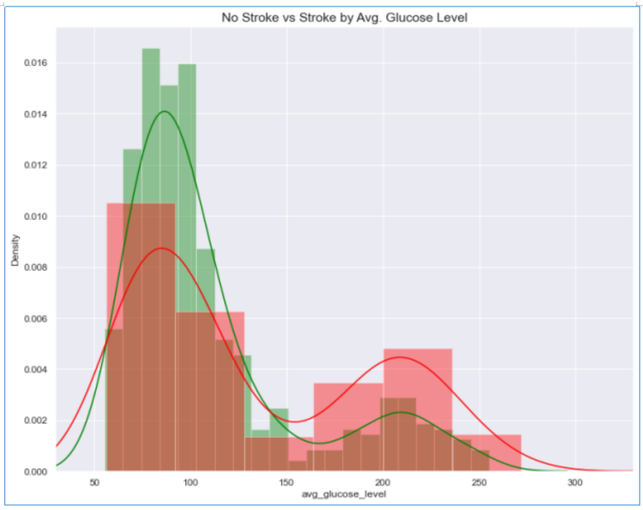

Fig. 13. Average glucose level and density distribution for stroke and normal patients

As presented in Fig 13, it considers the attribute avg_glucose_level and reflects on its values and their distribution statistics in the stroke dataset against stroke and no stroke class label.

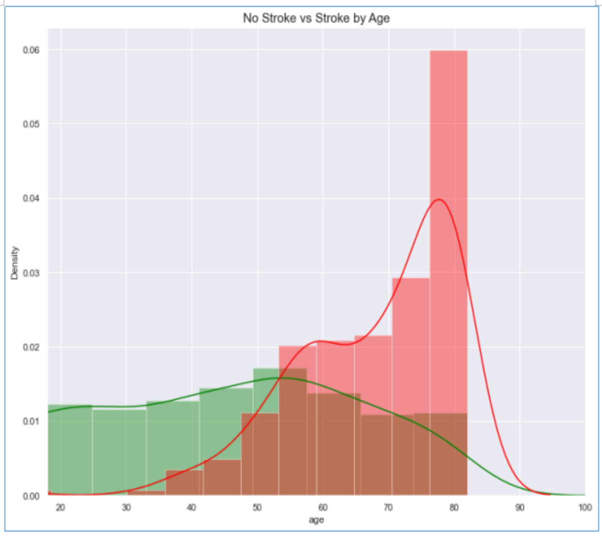

Fig. 14. Age and density distribution for stroke and normal patients

As presented in Fig 14, it considers the attribute age and reflects on its values and their distribution statistics in the stroke dataset against stroke and no stroke class label.

3.2 Evaluation of Results

The proposed ensemble model and its constituent models are evaluated to ascertain their performance statistics. Since all models are evaluated along with feature selection algorithm known as HMA-FE, feature extraction and computation of feature important play significant role. Therefore, feature importance and performance statistics of ensemble model and its constituent models are presented in this section.

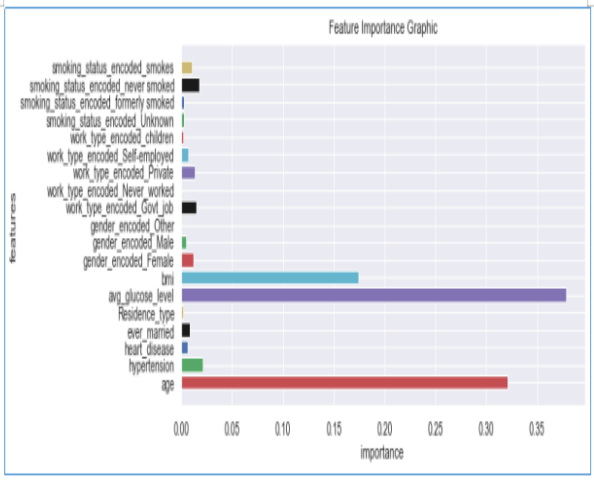

Fig. 15. Reflects extracted features from dataset and importance of each feature computed by HMA-FE

Fig 15 shows the features extracted from the dataset [27]. It also provides the computed feature impotence scores. These scores are used to choose only the best performing features that contribute to the class label prediction.

Table 1: Presents performance comparison between base classifiers of ensemble model and the proposed ensemble mode

Brain Stroke Prediction Model	Performance (%)			
	Precision	Recall	F-measure	Accuracy
Decision Tree Classifier	0.05882	0.05	0.05405	0.91576
Neural Nets	0.05882	0.05882	0.05882	0.92298
Gradient Boosting Classifier	1	0.96	0.98	0.95788
Logistic Regression	1	0.96	0.98	0.95908
Random Forest Classifier	1	0.96	0.98	0.95908
Support Vector Machine	1	0.96	0.98	0.95908
Stochastic Gradient Descent	1	0.96	0.98	0.95908
KNeighbors Classifier	0.02941	1	0.05714	0.96028
Proposed Ensemble Model	0.8823	0.9393	0.9099	0.9793

As listed in Table 1, the performance comparison of the proposed ensemble model and its underlying constituent models are provided.

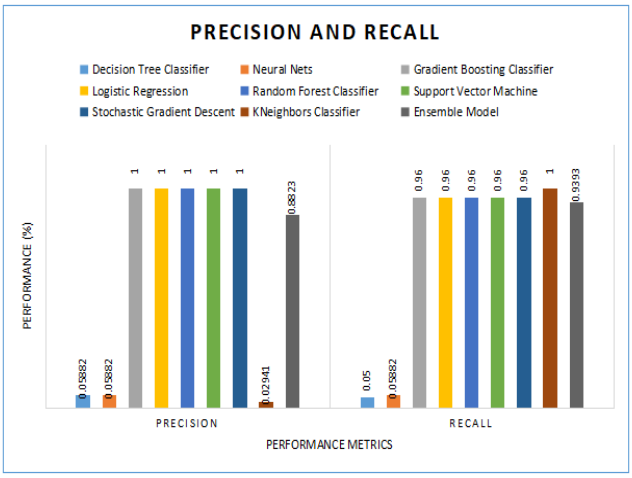

Fig. 16. Precision and recall comparison among stroke prediction models

As presented in Fig 16, the performance metrics of all stroke prediction models and proposed ensemble model are provided. Higher in precision or recall represents better performance. Due to their internal functionality, the prediction models exhibited different levels of performance . All models do have their approach in learning from dataset and knowledge representation.

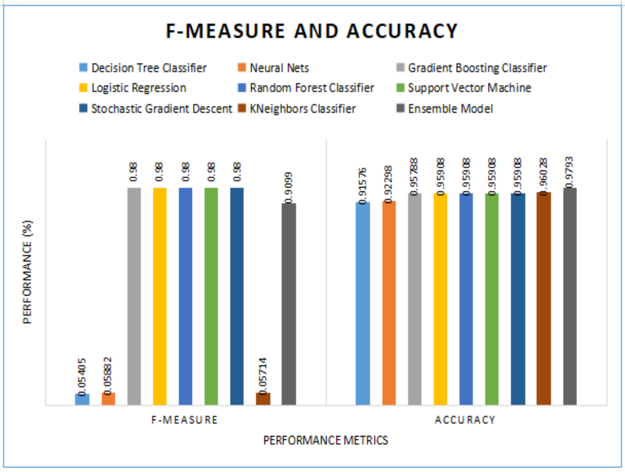

Fig. 17. F1-score and accuracy comparison among stroke prediction models

As presented in Fig 17, F1-score and accuracy of all stroke prediction models and the proposed ensemble model are provided. Higher in F1-score or accuracy indicates better performance. The prediction models showed different levels of performance due to their internal functionality. All models do have their approach in learning from dataset and knowledge representation.

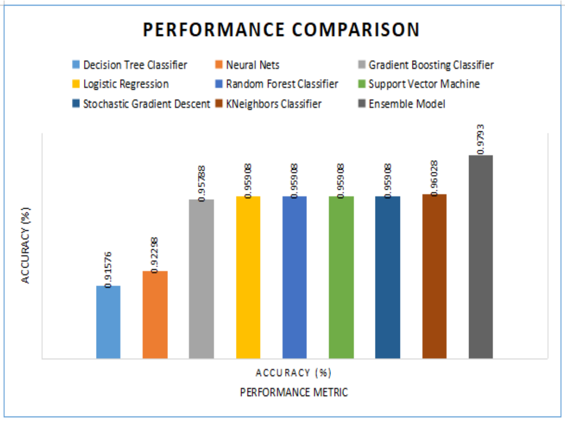

Fig. 18. Performance comparison among stroke prediction models in terms of accuracy

As shown in Fig 18, accuracy is the measure used to evaluate performance of different stroke prediction models. The prediction modes showed varied performance. Decision Tree (DT) exhibited 91.57% accuracy in stroke prediction which is the least performance among all models. Neural Nets showed better performance over DT with 92.29% accuracy. Interestingly there are many models that showed similar accuracy with 95.90%. They include KNeighbours, Random Forest (RF), Support Vector Machine (SVM), Stochastic Gradient Descent (SGD), and Gradient Boosting (GD). Highest performance is shown by the proposed ensemble model with 97.93% accuracy.

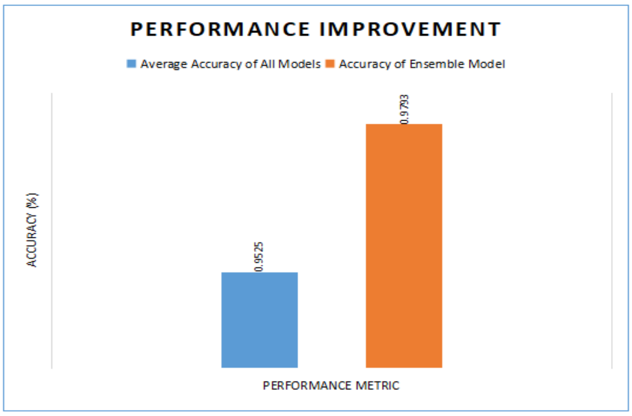

Figure 19: Overall performance improvement of ensemble model over average accuracy of constituent base models

Figure 19 illustrates how the ensemble model outperforms the average accuracy of its constituent base models in terms of overall performance. Our empirical study has revealed that ensemble model showed highest accuracy with 97.93% while the average accuracy of all constituent base line models is 95.25%. Proposed framework SPE witnesses 2.68% overall improvement in accuracy when compared with average accuracy of all base models used in ensemble.

4. Conclusion And Future Work

A framework known as Stroke Prediction Ensemble (SPE) is designed and implemented in this paper. The constituent base prediction models are empirically selected from many available models. The models that showed accuracy >90% in brain stroke detection are chosen for base classifiers in the ensemble. Our framework SPE reuses the hybrid approach for feature engineering published in our prior work [26]. An ensemble architecture is used along with weighted majority voting approach to have soft phenomenon in choosing final class label for each patient instance in the test dataset. We proposed an algorithm called Hybrid Ensemble and Feature Engineering for Stroke Prediction to realize the framework (HEFE-SP). It is meant for ensemble ML towards more efficient stroke prediction performance. Our empirical study has revealed that ensemble model showed highest accuracy with 97.93% while the average accuracy of all constituent base line models is 95.25%. Proposed framework SPE witnesses 2.68% overall improvement in accuracy when compared with average accuracy of all base models used in ensemble. Thus the ensemble model can be used for efficient brain stroke diagnosis as part of Clinical Decision Support System (CDSS).

References

[1]. Sirsat, Manisha Sanjay; Ferme, Eduardo; Camara, Joana (2020). Machine Learning for Brain Stroke: A Review. *Journal of Stroke and Cerebrovascular Diseases*, 29(10), 105162.

[2]. Tasfia Ismail Shoily, Tajul Islam, , Sumaiya Jannat, Sharmin Akter Tanna,Taslima Mostafa Alif, Romana Rahman Ema.. (2020). Detection of Stroke Disease using Machine Learning Algorithms. *IEEE*, p1-6.

[3]. Amin, Javeria; Sharif, Muhammad; Raza, Mudassar; Yasmin, Mussarat (2018). Detection of Brain Tumor based on Features Fusion and Machine Learning. Journal of Ambient Intelligence and Humanized Computing.

[4]. Sung, S.-F., Hung, L.-C., & Hu, Y.-H. Developing a stroke alert trigger for clinical decision support at emergency triage using machine learning. *International Journal of Medical Informatics*, 152, 104505, p1-8, 2021.

[5]. Kaisar, Shahriar . Developmental dyslexia detection using machine learning techniques : A survey. ICT Express, p1-4, 2020.

[6]. Uy, Sean Rich U.; Abu, Patricia Angela (2020). [IEEE 2020 *International Conference on Artificial Intelligence in Information and Communication (ICAIIC)* - Fukuoka, Japan (2020.2.19-2020.2.21)] 2020 International Conference on Artificial Intelligence in Information and Communication (ICAIIC) - Analysis of Detecting Compensation for Robotic Stroke Rehabilitation Therapy using Imbalanced Learning and Outlier Detection. , p432–437.

[7]. Md. Shafiul Azam, Md. Habibullah and Humayan Kabir Rana. (2020). Performance Analysis of Various Machine Learning Approaches in Stroke Prediction. *International Journal of Computer Applications* . 175 (21), p1-5.

[8]. Hussain, Lal . Detecting epileptic seizure with different feature extracting strategies using robust machine learning classification techniques by applying advance parameter optimization approach. Cognitive Neurodynamics, p1-24, 2018.

[9]. N. S. M. Noor1 , N. M. Saad2 , A. R. Abdullah3 , N. M. Ali4. (2019). Automated segmentation and classification technique for brain stroke. *International Journal of Electrical and Computer Engineering (IJECE)*. 9 (3), p1832-1841.

[10]. Cai, Siqi; Li, Guofeng; Zhang, Xiaoya; Huang, Shuangyuan; Zheng, Haiqing; Ma, Ke; Xie, Longhan (2019). Detecting compensatory movements of stroke survivors using pressure distribution data and machine learning algorithms. *Journal of Neuro Engineering and Rehabilitation*, 16(1), 131, p1-11.

[11]. Brunese, Luca; Mercaldo, Francesco; Reginelli, Alfonso; Santone, Antonella (2019). An Ensemble Learning Approach for Brain Cancer Detection exploiting Radiomic Features. Computer Methods and Programs in Biomedicine, 105134, p1-45.

[12]. Kamal, Haris; Lopez, Victor; Sheth, Sunil A. (2018). Machine Learning in Acute Ischemic Stroke Neuroimaging. Frontiers in Neurology, 9, 945, p1-11.

[13]. van Os, Hendrikus J. A.; Ramos, Lucas A.; Hilbert, Adam; van Leeuwen, Matthijs; van Walderveen, Marianne A. A.; Kruyt, Nyika D.; Dippel, Diederik W. J.; Steyerberg, Ewout W.; van der Schaaf, Irene C.; Lingsma, Hester F.; Schonewille, Wouter J.; Majoie, Charles B. L. M.; Olabarriaga, Silvia D.; Zwinderman, Koos H.; Venema, Esmee; Marquering, Henk A.; Wermer, Marieke J. H. Predicting Outcome of Endovascular Treatment for Acute Ischemic Stroke: Potential Value of Machine Learning Algorithms. Frontiers in Neurology, 9, p784, p1-8, 2018.

[14]. Liu, Tianyu; Fan, Wenhui; Wu, Cheng (2019). A hybrid machine learning approach to cerebral stroke prediction based on imbalanced medical dataset. Artificial Intelligence in Medicine, 101723, p1-31.

[15]. Brunese, Luca; Mercaldo, Francesco; Reginelli, Alfonso; Santone, Antonella (2019). An Ensemble Learning Approach for Brain Cancer Detection exploiting Radiomic Features. Computer Methods and Programs in Biomedicine, p1-45.

[16]. Ong, Charlene Jennifer; Orfanoudaki, Agni; Zhang, Rebecca; Caprasse, Francois Pierre M.; Hutch, Meghan; Ma, Liang; Fard, Darian; Balogun, Oluwafemi; Miller, Matthew I.; Minnig, Margaret; Saglam, Hanife; Prescott, Brenton; Greer, David M.; Smirnakis, Stelios; Bertsimas, Dimitris; Peng, Yifan (2020). Machine learning and natural language processing methods to identify ischemic stroke, acuity and location from radiology reports. PLOS ONE, 15(6), e0234908, p1-16.

[17]. Ho, King Chung; Speier, William; Zhang, Haoyue; Scalzo, Fabien; El-Saden, Suzie; Arnold, Corey W. (2019). A Machine Learning Approach for Classifying Ischemic

Stroke Onset Time from Imaging. *IEEE Transactions on Medical Imaging*, p1–1.

[18]. C, B. Machine Learning for Interpretation of Brain Images: A Detailed Analysis through survey.International Conference on Computing, Communication, and Intelligent Systems (ICCCIS 2021).

[19]. Patra, Amit Kumar; Ray, Ratula; Abdullah, Azian Azamimi; Dash, Satya Ranjan (2019). Prediction of Parkinsonâ€™s disease using Ensemble Machine Learning classification from acoustic analysis. Journal of Physics: Conference Series, 1372(), 012041, p1-8.

[20]. Lin, Ching-Heng; Hsu, Kai-Cheng; Johnson, Kory R.; Fann, Yang C.; Tsai, Chon-Haw; Sun, Yu; Lien, Li-Ming; Chang, Wei-Lun; Chen, Po-Lin; Lin, Cheng-Li; Hsu, Chung Y. (2020). Evaluation of Machine Learning Methods to Stroke Outcome Prediction Using a Nationwide Disease Registry. Computer Methods and Programs in Biomedicine, 105381, p1-37.

[21]. Gyorfi, Agnes; Kovacs, Levente; Szilagyi, Laszlo . IEEE International Conference on Systems, Man and Cybernetics (SMC) - Bari, Italy (2019.10.6-2019.10.9)] 2019 *IEEE International Conference on Systems, Man and Cybernetics* (SMC) - Brain Tumor Detection and Segmentation from Magnetic Resonance Image Data Using Ensemble Learning Methods. ,IEEE 2019,p909–914.

[22]. Ashwini Kadam1. (IJRAR) Detection and Classification of Brain Hemorrhage Using Ensemble Learning. *International Journal of Research and Analytical Reviews* . (2020) 7 (1), p520-521.

[23]. Fulop, Timea; Gyorfi, Agnes; Csaholczi, Szabolcs; Kovacs, Levente; Szilagyi, Laszlo (2020). *[IEEE 2020 IEEE 15th International Conference of System of Systems Engineering* (SoSE) - Budapest, Hungary (2020.6.2-2020.6.4)] *2020 IEEE 15th International Conference of System of Systems Engineering (SoSE)* - Brain Tumor Segmentation from Multi-Spectral MRI Data Using Cascaded Ensemble Learning, p531–536.

[24]. Suk, Heung-Il; Lee, Seong-Whan; Shen, Dinggang . Deep ensemble learning of sparse regression models for brain disease diagnosis. Medical Image Analysis, 37, p101–113,(2017).

[25]. Serbes, Gorkem; Sakar, Betul Erdogdu; Gulcur, Halil Ozcan; Aydin, Nizamettin (2015). An emboli detection system based on Dual Tree Complex Wavelet Transform and ensemble learning. Applied Soft Computing, p1-8.

[26]. Ushasree, D., Krishna, A.V.P., Rao, C., Parameswari, D.V.L. A Hybrid Feature Selection for Improving Prediction Performance with a Brain Stroke Case Study. In: Shukla, S., Gao, XZ., Kureethara, J.V., Mishra, D. (eds) Data Science and Security. *Lecture Notes in Networks and Systems*,2022, vol 462. Springer, Singapore. https://doi.org/10.1007/978-981-19-2211-4_33.

[27]. Stroke Prediction-Dateset. Retrieved from https://www.kaggle.com/fedesoriano/stroke-prediction-dataset