Intelligent Systems with Applications 17 (2023) 200165

Contents lists available at ScienceDirect

Intelligent Systems with Applications

journal homepage: www.journals.elsevier.com/intelligent-systems-with-applications

# A stacked ensemble model for automatic stroke prediction using only raw electrocardiogram

Prashant Kunwar [a], Prakash Choudhary [*,b]

[a] Department of Computer Science and Engineering, National Institute of Technology Hamirpur, Hamirpur, Himachal Pradesh, 177005, India
[b] Department of Computer Science and Engineering, Central University of Rajasthan, Ajmer, Rajasthan, 305817, India

ARTICLE INFO

*Keywords:*
Stroke
CNN
Stacked ensemble model
ECG
Physiological data

ABSTRACT

Stroke is one of the main causes of death and disability in elderly people worldwide. Early diagnosis of this disease is desirable. An electrocardiogram (ECG) signal is a powerful tool in the diagnosis of stroke. Many recent studies are using ECG signals in the diagnosis of stroke. This study is done with the aim to develop an efficient classification model which can be used in the diagnosis of stroke disease using ECG features. This study uses ECG data collected from 71 different subjects of which 35 were stroke patients and 36 were normal patients. This study proposes a stacked ensemble model which is built by stacking three different convolutional neural networks (CNN) models. The raw ECG signals are used as input to the model for training and testing. The result shows that the proposed model is capable of predicting stroke with an accuracy of 99.7%. F1-score, precision and recall are 99.69%, 99.67% and 99.71% respectively. Hence, this study reports that with the proposed model, ECG can be used as an aid in the diagnosis of stroke disease with high efficiency.

## 1. Introduction

### 1.1. Background

Stroke is the second highest leading cause of death and the third leading cause of death and disability combined Acharya et al. (2017). According to a study in 2010, there were 16.9 million strokes reported, 5.9 million deaths from stroke and 33 million patients were survivors who experience this at least once in their lives (Fekadu et al., 2020). Stroke is a condition of medical emergency as it has a higher mortality rate. This is the reason why this disease puts demands significant attention from the healthcare sector. A stroke occurs when the blood supply to a part of the brain is interrupted or reduced, preventing brain tissue from getting oxygen and nutrients, this causes the brain cells to begin to die in minutes (Subudhi et al., 2020; Zhang et al., 2013). Stroke can be classified into two broad categories ischemic stroke and cerebral hemorrhagic stroke. Ischemic stroke (cerebral infarction) is caused by blockages in the blood vessels, and cerebral hemorrhagic stroke (cerebral haemorrhage) is caused by the rupturing of blood vessels in the brain (Sengupta et al., 2016). An ischemic stroke further can be classified into two types which are cerebrovascular thrombosis and cerebral embolism. Cerebral embolism is a condition in which a blood clot from the heart blocks the blood vessels, which is responsible for supplying the oxygen through the blood to the brain, whereas cerebral thrombosis is a condition in which blocking blood clots form in the brain due to arteriosclerosis or because of problems with the inner wall of the blood vessel. Now, there are two types of hemorrhagic stroke, intracerebral haemorrhage and subarachnoid haemorrhage. In an Intracerebral haemorrhage condition, weak blood vessels can burst if there is a sudden rise in blood pressure, which can cause damage to the surrounding cells. Subarachnoid haemorrhage can cause bleeding in the space between the brain and the surrounding membrane (subarachnoid space).

Aging is the most important independent risk factor for stroke but is obviously not a modifiable factor. For each successive 10 years after age 55 years, the stroke rate more than doubles in both men and women. Thus, the risk of stroke in elderly patients is significantly high (Olindo et al., 2003).

The plot of Electrical impulse (in voltage) generated by heart versus time is called an Electrocardiogram (ECG or EKG) (Lilly and S, 2012). The ECG is recorded by attaching electrodes to the skin of the patients. Electrocardiography is a process of recording the electrical activity of the heart (Feather et al., 2020). Electrodes attached to the skin of the patient detect the small electrical changes produced by cardiac muscle's depolarisation and repolarisation during each cardiac cycle i.e.

* Corresponding author.
*E-mail addresses:* kunwarprashant21@gmail.com (P. Kunwar), prakash.choudhary@curaj.ac.in, pc@nith.ac.in (P. Choudhary).

https://doi.org/10.1016/j.iswa.2022.200165
Received 25 May 2022; Received in revised form 20 November 2022; Accepted 9 December 2022
Available online 13 December 2022

heartbeat. An ECG is represented with its three main components, the first is the P-wave which represents the depolarization of the atria; the second is QRS-complex which represents the depolarization of the ventricles, and the third is T-wave which represents the repolarization of the ventricles (Lilly and S, 2012). This is well known that in numerous cardiac abnormalities, these ECG patterns tend to change from normal patterns but it shows a similar correlation in the case of stroke diseases also that we will see in a later section. A normal ECG plot is shown in Fig. 1.

### 1.2. Related work

The fact that cardiac diseases lead to stroke is known for a very long time. Edwin Byer reported a stroke case with abnormal ECG patterns, which indicates the correlation of cardiac abnormalities with stroke (Czabanski et al., 2020). Thus, many studies have been done to show that diagnosis of stroke disease is possible through ECG tracings of patients. Qing Zang et. al. developed a regression based model to predict the stroke onset in a patient, which gives the highest precision rate of 90% by using some physiological variables and ECG tracings of patients (Zhang et al., 2013). Kalaivani Rathakrishnan et. al. used various machine learning techniques to classify ECGs of stroke patients and found the K-Nearest Neighbour model outperformed other machine learning techniques like the random forest, support vector machine (SVM), Naïve Bayes and logistic regression with accuracy and F-1 score of 96.6% (Rathakrishnan et al., 2021). An artificial neural network with three hidden layers was proposed by Pattanapong C. and Madhu Goyal to predict stroke. They used physiological data, medical history of patient and family and other habits like smoking, consumption of alcohol etc. as input features for the model; the model achieved a mean square error of 0.2596 (Chantamit-o-pas & Goyal, 2017). Songhee cheaon et. al. used medical history and health behaviour data of patients to predict mortality of stroke patients. They used principal component analysis (PCA) to extract dominant features and then used it to train and test a deep neural network; it achieved an accuracy of 83.48% (Cheon et al., 2019). Sushravya Raghunath et. al. proposed a deep neural network with four dense layers (hidden layers) to predict Atrial fibrillation related stroke. They used 12-lead digital ECG tracings as input data; model achieved area used receiver operator curve and area under precision recall curve as 0.85 and 0.22 respectively (Raghunath et al., 2021). Yoon-A Choi et. al. worked with EEGs collected from patients and experimented with deep learning models like LSTM, bi-directional LSTM, CNN-LSTM and CNN bi-directional LSTM to predict stroke. The experimental results showed that CNN bi-directional LSTM model is capable of predicting stroke with 94.0% accuracy and false positive rate (FPR) of 6.0% (Choi et al., 2021). Eric S. Ho et. al. used a 1-dimensional CNN model with Gradient-weighted Class Activation Mapping (GRAD-CAM) to predict stroke by using ECGs with an accuracy of 90% (Ho and Ding, 2021). Yifeng Xie et. al. proposed a CNN based model, which can take ECG tracing in form of an image and can predict the stroke with 85.82% accuracy. The accuracy of the model could have improved by using more examples, as only 98 examples were used in their study (Xie et al., 2021). Yan Liu et. al. have proposed a multi-neural network, which consists of a VGG-16 model and one 1-dimensional CNN based model stacked on top of a densely connected neural network. VGG-16 model is used to extract features from ECG tracing and CNN model was used to extract features from various physiological and behavioural data and then predicted the stroke based on those features with an accuracy of 98.53% (Liu et al., 2020). Noushin Rabinezhadsadatmahaleh et. al. proposed a stacked ensemble of deep and conventional machine learning classifiers (NRSE-DCML) for human bio-metric identification from ECG which achieves the Accuracy of 99.02, FAR of 0.95 and FRR of 1.02 (Rabinezhadsadatmahaleh and Khatibi, 2020). Matteo Bodini et. at. proposed an ensemble machine learning approach for classification of 27 cardiac abnormalities from 12-lead ECG signals (Bodini et al., 2020), the proposed model was presented in a PhysioNet 2020 Challenge, where it achieved an score of -0.179 on a data set which was private to organizer, it achieved 40th position in the challenge. Mehrdad Javadi et. al. proposed stacked ensemble model for ECG classification for the purpose of classifying normal heartbeats, Premature Ventricular Contraction (PVC) and other abnormalities which achieved an accuracy of 95.26% (Javadi et al., 2011).

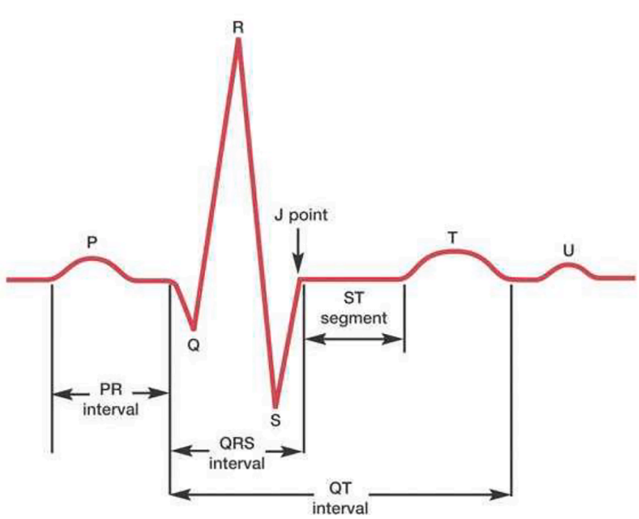


**Fig. 1.** A Plot showing normal ECG pattern and its different components. Image source- (Ashley and Niebauer, 2004).

### 1.3. Contributions

When it comes to health accuracy of diagnosis is the greatest concern, the prediction should be highly accurate and the false positive rate (FPR) and false negative rate (FNR) should be very less. The predictive model should use as less input data as possible to reduce the data collection time. Moreover, patients' personal data such as their medical history, personal habits and family members' medical history are sensitive data that creates data security concerns and patients might not be comfortable sharing that data. To overcome all those challenges, we propose a stacked ensemble model which shows very high accuracy and it does not need any personal data of patients, except ECG tracings. The proposed stacked ensemble model is built using three different CNN based models. First, three 1-dimensional CNN (conv1D) based models are trained and then they are combined and stacked on a meta-learner model. The meta-learner model is a densely connected neural network that tries to combine the predictions of all three conv1D based models in the best possible way. As result, by using just ECG data of patient, proposed model is able to predict stroke with accuracy and with a very less number of false positives and false negatives. It solves the problem of data security concerns as well as accuracy.

In the remainder of this paper, Section 2 discusses dataset and architecture of the proposed model. The performance evaluation metrics introduced in Section 3, and corresponding experimental results are presents in Section 4. Section 5 analyzed remarks based on empirical results drawn in Section 4. Section 6 conclude the study.

## 2. Materials and method

### 2.1. Dataset

The dataset used in this study is cerebral vasoregulation in elderly with stroke, it's a public dataset with open access, published on Oct 4, 2018 (Novak et al., 2010). This database has multi-model data collected in large scale studies examining the effects of the ischemic stroke on cerebral vasoregulation. The study compared 60 participants who suffered strokes, to 60 control participants, collecting the following data for each patient across multiple days: transcranial doppler of cerebral arteries, 24-hour blood pressure numerics, waveforms of high resolution (ECG, blood pressure, CO2, and respiration) during various movement tasks, 24-h ECG, EMG, and accelerometer recordings, and gait pressure recordings during a walking test. Other historical data of patients are also provided in the database related to their medical history and daily habits (such as consumption of alcohol and smoking etc.). However, 121 ECG tracings from 71 subjects of which 35 are stroke subjects and 36 control subjects are used in this study, other physiological data or data of the medical history of patients were not used. The data of other subjects is not used as it was incomplete. The ECG tracings have been collected at 500Hz for 10–12 minutes which has been broken into smaller windows in this study.

## 2.2. Experimental design

### 2.2.1. Overall architecture of proposed model

The neural network model proposed in this paper uses the concept of stacking ensemble models in which three different sub-models are combined and stacked on the top of a dense neural network (meta-learner). All the sub-models are built using Conv1D and dense layers. The output of these sub-models becomes the input for the meta-leaner model, which combines the results of individual models in the best possible way and predicts the occurrence probability of stroke. The model and the entire training process are explained in the detail in further sections. The entire network structure system is shown in Fig. 2.

### 2.2.2. Architecture of sub-models

Each sub-model was built using a one-dimensional convolutional neural network model to extract the features from the ECG tracings of different subjects. Each sub-model consists of three layers of one-dimensional convolutional layers containing 32, 64, and 128 neurons, each sub-models had a filter size of 16, 32, and 64 at subsequent layers, and ReLU is used as the activation function; three max-pooling layers and three dropout layers were also used. A flattening layer was used to flatten all extracted features and then fed into densely connected layers. Three fully connected layers containing 128, 256, and 1 neuron are used and Relu is used as the activation function except for the output layer where the sigmoid activation function is used. Table 1 presents some details about the training of sub-models.

**Table 1**
Detailed information on the architecture of the sub-models.

| Parameters | *Sub-model 1* | *Sub-model 2* | *Sub-model 3* |
|---|---|---|---|
| Input size | (2000,1) | (2000,1) | (2000,1) |
| No. of CNN layers | 3 | 2 | 3 |
| Filter Size | (16, 32, 64) | (16, 32, 64) | (16,32, 64) |
| No. of Dropout layers | 3 | 2 | 3 |
| No. of Pooling layers | 3 | 2 | 3 |
| No. of Dense Layers | 3 | 3 | 3 |
| No. of neurons/layer | (128, 256, 1) | (128, 128, 1) | (128, 256, 1) |

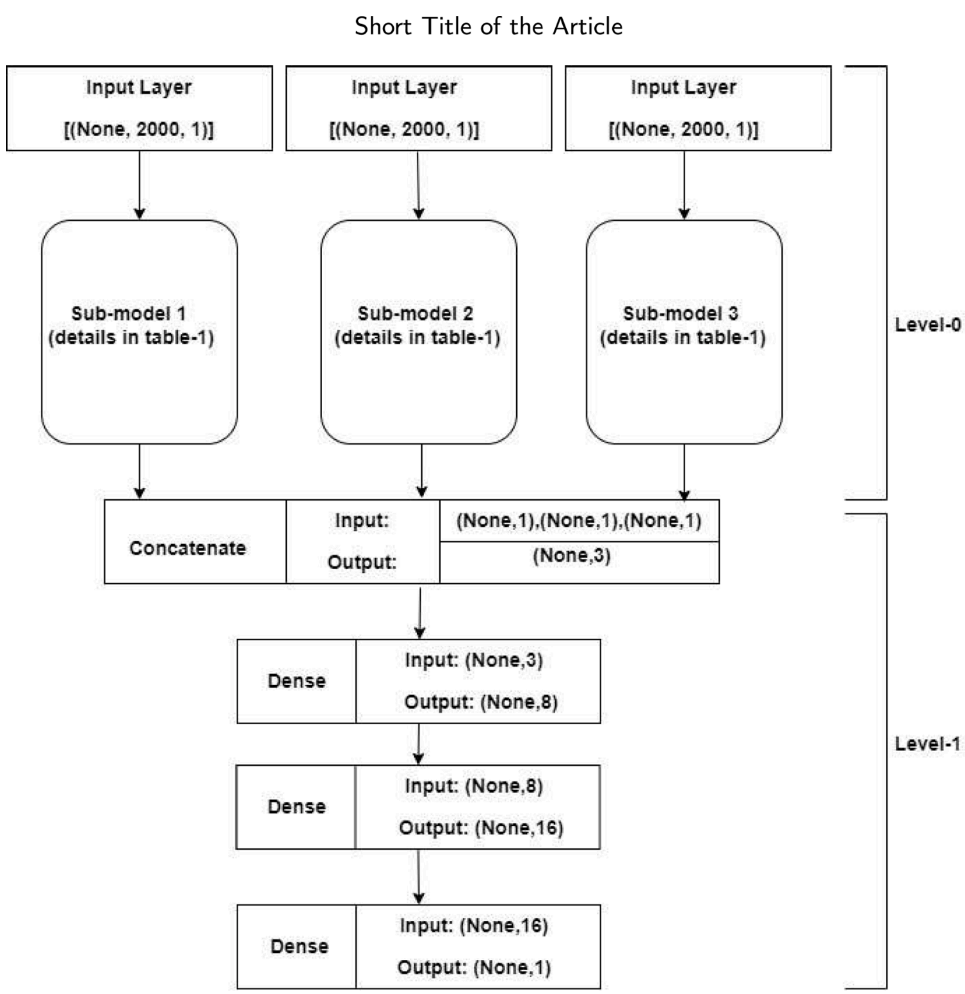


**Fig. 2.** The architecture of the proposed Stacked ensemble model: Three sub-models become the level-0, the upper part of the stacked model, below part of the model is called meta-learner which consists of three dense layers (level-1). This model has three input heads, takes the same input from all three channels, and sub-models process that input then probabilities are sent to meta-learner which will generate a final prediction.

#### 2.2.3. Stacked ensemble model and stacking ensemble technique

Ensemble learning combines several individual models to obtain better generalization performance. Currently, deep learning models with multilayer processing architecture is showing better performance as compared to the shallow or traditional classification models. Deep ensemble learning models combine the advantages of both the deep learning models as well as the ensemble learning such that the final model has better generalization performance (Ganaie et al., 2021). There are some variations of ensemble learning based on how it combines the predictions of sub-models. First, a model averaging ensemble combines the predictions from multiple trained models but this approach has the limitation that each sub-model contributes the same amount to the ensemble prediction, irrespective of how well that individual sub-model performed. Second, a variation of this approach is called a weighted average ensemble. It weighs the contribution of each sub-model by the expected performance of the model on a part of the test dataset. This ensures well-performing sub-models contribute more and less-well-performing sub-models contribute less. Third, a further generalization of this approach is to replace the linear weighted sum (e.g. linear regression) model used to combine the predictions of the sub-models with any deep-learning algorithm. This approach is called stacked generalization. In this stacking approach, an algorithm takes the outputs of sub-models as input and tries to learn how to best combine the input predictions to make a better output prediction. In this study, a stacked generalization ensemble is used in which three CNN based models have been used as sub-models and a dense neural network is used as a meta-learner which takes input from all three sub-models and tries to combine the predictions of them in order to outperform all of them if used individually.

#### 2.2.4. Training and testing of sub-models (level-0)

Training and testing in this study have been executed in two separate phases. In the first phase, sub-models have been trained and evaluated and in the second phase, the entire stacked ensemble model has been trained on a separate dataset to avoid overfitting and evaluation. As mentioned earlier, 121 tracings from 71 different subjects have been used in this study. Tracings were 10 to 12 minutes long, which have been broken down into smaller windows (4 sec) to decrease the input size to the model, in order to decrease the complexity. Other lengths of ECG tracings such as 1 sec, 2 sec and 3 sec, etc. have been tried but it turned out that sub-models performed best with 4 sec long ECG tracings. In this way, all tracings broke down in the smaller windows and resulted in nearly 18,000 examples. The labeled dataset is then split into four parts: training dataset (60%), validation dataset (15%), testing dataset (15%), and holdout dataset (10%). In order to ensure that the weights of the sub-models converge during the training process ECG signal values were normalized. Various parameters during training are given in Table 2 for sub-models. After training all the sub-models on the training dataset, all three sub-models have been evaluated using various performance evaluation metrics mentioned in a further section.

**Table 2**
Detailed information on training of the models.

| Parameters | *Sub-models* | *Stacked ensemble model* |
|---|---|---|
| No. of examples | 10600 | 1766 |
| No. of epochs | 20 | 15 |
| Batch size | 32 | 32 |
| Optimizer | Adam | Adam |
| Learning rate | 0.001 | 0.001 |
| Total parameters | (4113825),(2221921),(4064929) | 10400868 |
| Trainable parameters | All | 193 |

#### 2.2.5. Training and testing of stacked ensemble model

After successful training and testing of sub-models, the stacked ensemble model has been trained on holdout dataset, which was different from the training dataset previously used, this was done to avoid overfitting. This time, weights of sub-models have been made non-trainable, only the weights of the stacked ensemble model have been trained. The Various parameters during the training process are given in Table 2.

## 3. Performance evaluation metrics

The effectiveness of any deep learning model is determined by its capacity to give accurate results. To evaluate the performance of the proposed model, various performance evaluation metrics are used as follows.

1. Accuracy: Accuracy shows how often the proposed model classifies the unseen data instance correctly.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

2. Specificity: Specificity is a measure of the capacity of the model to determine the true negatives of each class correctly.

$$Specificity = \frac{TN}{TN + FP} \tag{2}$$

3. Precision: Precision shows that what proportion of positive predictions is actually belongs to positive class.

$$Precision = \frac{TP}{TP + FP} \tag{3}$$

4. Recall: Recall shows what proportion of correct positive predictions made out of all positive predictions that could have been made.

$$Recall = \frac{TP}{TP + FN} \tag{4}$$

5. F1-score:It is the harmonic mean of precision and recall.

$$F1 - score = \frac{2 * Precision * Recall}{Precision + Recall} \tag{5}$$

In the equation given above, TP, TN, FP, and FN represent True Positives, True negatives, False positives, and False negatives, respectively. True positives and true negatives represent the correct prediction of whether a subject has stroke disease or not. However, false positives and false negatives determine the number of wrong predictions made by the proposed model.

## 4. Results

In this section, we will discuss the obtained results of experiments conducted. Firstly, all three sub-models were trained and tested on 60% and 15% of the dataset respectively then the stacked ensemble model was trained and tested on 10% and 15% datasets to predict the probability of stroke in a given subject.

#### 4.1.1. Results of training the sub-models

Our sub-models were trained on 60% dataset, validated, and tested on 15% and 15% dataset respectively. The optimizer used was Adam with a learning rate was 0.001 and the binary cross-entropy loss function used. Then, Sub-models were successfully trained for 15 epochs. Training and validation accuracy of sub-models after completion of training phase in given in Table 3. The curves of accuracy and loss rate of each sub-model are shown in Figs. 3–5.

#### *4.1.2. Results of training the stacked ensemble model*

The stacked ensemble model was trained on a holdout dataset (10%) to train the meta-learner neural network so that it can combine the prediction made by sub-models in the best possible way. The optimizer used was Adam's learning rate was 0.001 and the binary cross-entropy loss function was used. Then, a stacked ensemble model was successfully trained for 15 epochs. The metrics like training accuracy and validation accuracy, etc. are given in Table 3. The curves of accuracy and loss rate of stacked ensemble model are shown in Fig. 6.

#### *4.1.3. Results of testing the sub-models*

The test dataset containing 2650 samples was used to evaluate the sub-models. The index values of the sub-models evaluation (precision, recall, accuracy and f1-score) are shown in Table 4. The precision, recall and f1-score were obtained by Equations (1)–(5). Now, these three models will become the base models for the stacked model, more accurate prediction these sub-models will give more accurate the whole stacked model becomes.

#### *4.1.4. Results of testing the stacked ensemble model*

The stacked ensemble model has been tested on 2650 samples of the test dataset and it was capable of classifying the samples in stroke and normal (or Control) groups with 99.70% accuracy. Various performance evaluation metrics have been listed in Table 4. The proposed model has shown the capability of predicting stroke with a precision and recall score of 0.996737 and 0.99712 respectively. The confusion matrix depicting the true positive, true negative, false positive, and false negative is shown in Fig. 7. These metrics also show that the ensemble model easily outperforms all the sub-models. The proposed can accept the ECG tracing of a single patient also and can predict the probability of occurrence stroke with accuracy which can act as a pre-stroke warning which will help medical practitioners take appropriate preventive measures.

## 5. Discussion

In this study, the proposed model uses just ECG data, and based on those tracings it is capable of predicting stroke with 99.7% accuracy. Also, no noise removal techniques is used in this study which means the proposed model takes in the raw ECG readings which shows the robustness of the proposed model. Hence, it can also be concluded that the sub-model of the proposed model can extract the features even from noisy ECG beats. Thus, we might be able to accurately classify the new noisy ECG beat with our proposed model. The stacked ensemble model has been previously used in a study to predict post-COVID-19 complications and has been found very useful (Gupta et al., 2021). The summary of various studies done by researchers to predict stroke using ECG

**Table 3**
Training results evaluation metrics of the models.

| Metric | *Sub-model 1* | *Sub-model 2* | *Sub-model 3* | *Stacked Ensemble Model* |
|---|---|---|---|---|
| Training accuracy | 99.45% | 98.47% | 99.46% | 99.89 |
| Training loss | 0.0107 | 0.0211 | 0.0144 | 0.0077 |
| Validation accuracy | 99.23% | 0.9913% | 99.21% | 99.78 |
| Validation loss | 0.0206 | 0.0348 | 0.0263 | 0.0083 |
| Training time | 465 sec | 425 sec | 420 sec | 50 sec |

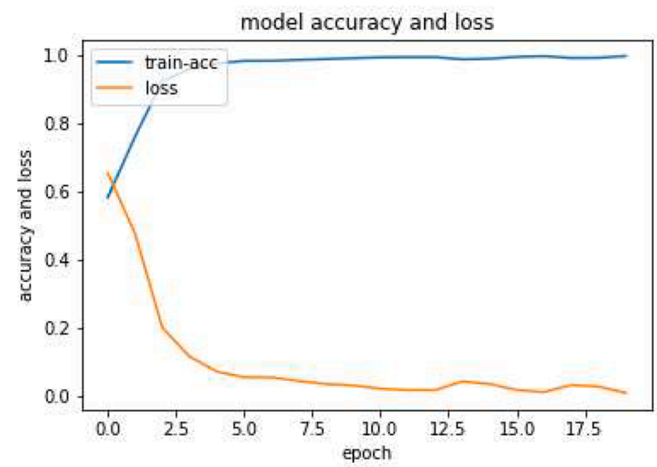


**Fig. 3.** Sub-model 1, training accuracy and loss progress with no. of epochs.

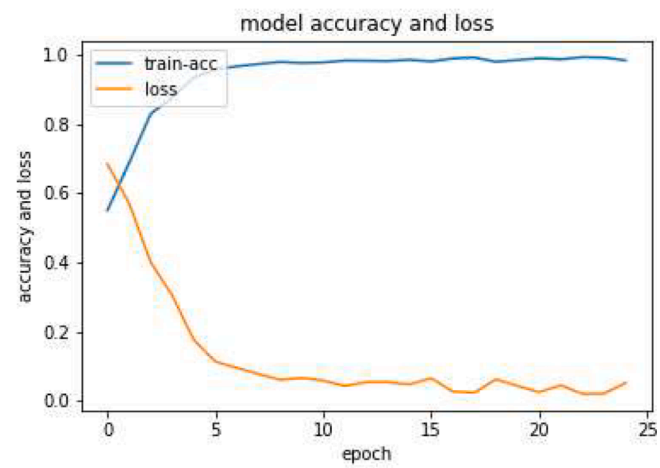


**Fig. 4.** Sub-model 2, training accuracy and loss progress with no. of epochs.

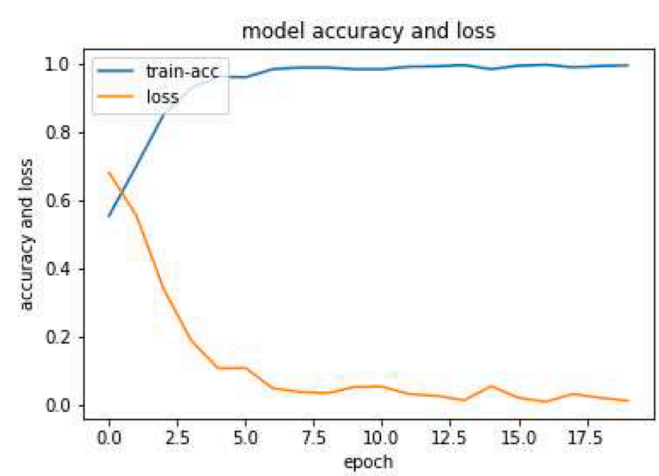


**Fig. 5.** Sub-model 3, training accuracy and loss progress with no. of epochs.

data is given in Table 5. However, some studies have used various physiological and medical history data besides ECG also (Liu et al., 2020; Zhang et al., 2013). The proposed model has shown better performance by just using ECG signal data. The fact that the ensemble model easily outperforms the individual CNN sub-models can be

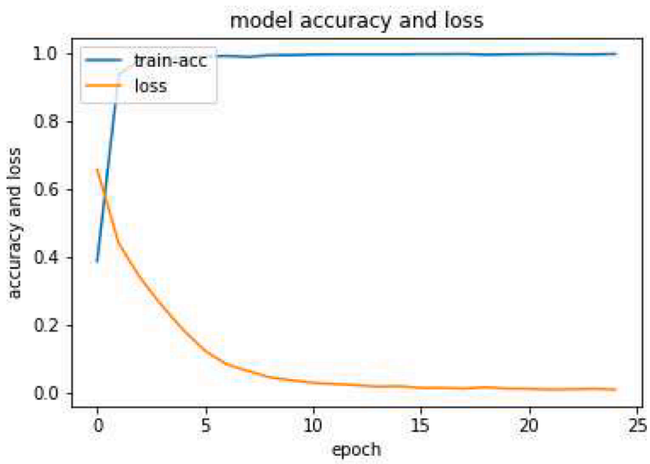


**Fig. 6.** For proposed model, training accuracy and loss progress with no. of epochs.

**Table 4**
Testing results evaluation metrics of models.

| Metric | *Sub-model 1* | *Sub-model 2* | *Sub-model 3* | *Stacked Ensemble Model* |
|---|---|---|---|---|
| Test accuracy | 99.47% | 98.72% | 99.32% | 99.70 |
| Test loss | 0.023 | 0.0389 | 0.0234 | 0.0130 |
| F1-score | 0.99462 | 0.98695 | 0.99310 | 0.99693 |
| Precision | 0.99443 | 0.98667 | 0.99239 | 0.996737 |
| Recall | 0.99482 | 0.98724 | 0.99388 | 0.99712 |

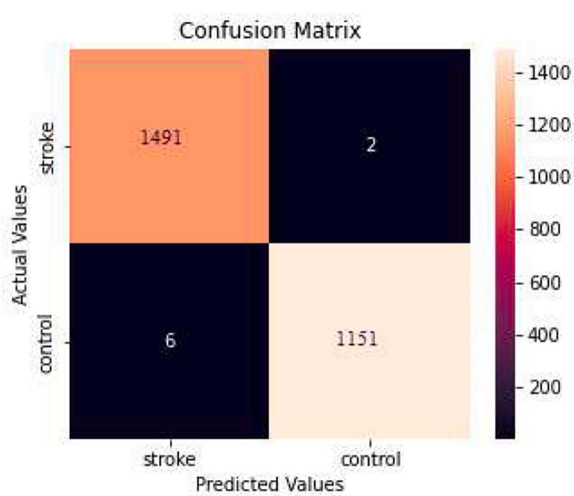


**Fig. 7.** Confusion matrix after testing the proposed stacked ensemble model, where stroke class represents subjects having stroke disease and control class represents normal patients, the model fails on only 8 examples out of 2650 examples.

observed by comparing the results of Table 4.

This study also has some limitations that open new dimensions of further studies. The data used in this study is collected from just 71 patients. The effectiveness of the proposed model will increase if we can train and test it on a huge scale. Further, the geographical location and genetic aspect of patients are not taken into consideration in this study the more diverse data will ensure a highly generalized model. This study does not consider the effect of an ongoing medication like anti-arrhythmic drugs on patients, whether it can change the ECG signal during data collection. The data set used in this study was obtained from

**Table 5**
Comparative analysis of proposed method with existing similar work on stroke prediction using electrocardiogram.

| Author, year | *Data used* | *Classifier* | *Performance* |
|---|---|---|---|
| Kalaivani Rathakrishnan et. al. 2020 (Rathakrishnan et al., 2021) | ECG | K-Nearest Neighbours(KNN) | Accuracy: 96.6%, Precision: 94.3% |
| Yifeng Xie et. al. 2020 ( Xie et al., 2021) | ECG | CNN | Accuracy: 85.82% |
| Eric S. Ho et. al. 2021 ( Ho and Ding, 2021) | ECG | CNN with Gradient-weighted class activation mapping | Accuracy: 90%, AUROC: 0.95 |
| Proposed model | ECG | Stacked ensemble method, CNN | Accuracy: 99.70%, F1-score: 99.69%, Precision: 99.67%, Recall: 99.71% |

small group that's why we increased the data samples by spliting the ECG tracings in smaller fragments, it might cause the overestimation of performance of the model, it leaves us with scope for further studies and validation of performance of the proposed model. The proposed model should be tested on larger ECG data sets in further studies, so that it's usefulness can be validated.

## 6. Conclusions

In summary, this study has analyzed the presence of signs of stroke in ECGs of patients. Thus, the proposed model becomes able to diagnose the stroke disease using ECG tracings of patients. Further, this study attempts to classify the stroke patients using ECGs only with high accuracy of 99.7%. This will save the time spent and cost in collecting and maintaining other physiological data like blood pressure, oxygen level, heart rate, etc., medical history, and daily habits like consumption of alcohol, use of tobacco, etc. Our findings show that the stacking ensemble technique seems to be useful in building automatic diagnostic models to be used in the real world by medical practitioners. Our proposed model can be applied on larger data set, with binary and multi classes in future studies, on a larger data set, we will be able to observe how stacking multiple models achieves better performance than an individual model by significant amount of accuracy. Our proposed model may assist in the diagnosis of stroke by using the ECG of patients.

## CRediT authorship contribution statement

**Prashant Kunwar:** Conceptualization, Methodology, Software, Validation, Resources, Data curation, Writing – original draft, Visualization. **Prakash Choudhary:** Formal analysis, Investigation, Writing – review & editing, Visualization, Supervision.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data availability

Data will be made available on request.

## References

Acharya, U. R., Fujita, H., Lih, O. S., Hagiwara, Y., Tan, J. H., & Adam, M. (2017). Application of deep convolutional neural network for automated detection of myocardial infarction using ecg signals. *Information Sciences, 415*, 190–198.

Ashley, E. A., & Niebauer, J. (2004). Arrhythmia. *Cardiology explained*. Remedica.
Bodini, M., Rivolta, M. W., & Sassi, R. (2020). Classification of 12-lead ECG with an ensemble machine learning approach. *2020 computing in cardiology* (pp. 1–4). IEEE.
Chantamit-o-pas, P., & Goyal, M. (2017). Prediction of stroke using deep learning model. *International conference on neural information processing* (pp. 774–781). Springer.
Cheon, S., Kim, J., & Lim, J. (2019). The use of deep learning to predict stroke patient mortality. *International Journal of Environmental Research and Public Health, 16*(11), 1876.
Choi, Y.-A., Park, S.-J., Jun, J.-A., Pyo, C.-S., Cho, K.-H., Lee, H.-S., & Yu, J.-H. (2021). Deep learning-based stroke disease prediction system using real-time bio signals. *Sensors, 21*(13), 4269.
Czabanski, R., Horoba, K., Wrobel, J., Matonia, A., Martinek, R., Kupka, T., Jezewski, M., Kahankova, R., Jezewski, J., & Leski, J. M. (2020). Detection of atrial fibrillation episodes in long-term heart rhythm signals using a support vector machine. *Sensors, 20*(3), 765.
Feather, A., Randall, D., & Waterhouse, M. (2020). *Kumar and clark's clinical medicine E-Book*. Elsevier Health Sciences.
Fekadu, G., Adola, B., Mosisa, G., Shibiru, T., & Chelkeba, L. (2020). Clinical characteristics and treatment outcomes among stroke patients hospitalized to nekemte referral hospital, western ethiopia. *Journal of Clinical Neuroscience, 71*, 170–176.
Ganaie, M., Hu, M., et al. (2021). Ensemble deep learning: A review. *arXiv preprint arXiv: 2104.02395*.
Gupta, A., Jain, V., & Singh, A. (2021). Stacking ensemble-based intelligent machine learning model for predicting post-covid-19 complications. *New Generation Computing*, 1–21.
Ho, E. S., & Ding, Z. (2021). Electrocardiogram analysis of post-stroke elderly people using one-dimensional convolutional neural network model with gradient-weighted class activation mapping. *medRxiv*.
Javadi, M., Ebrahimpour, R., Sajedin, A., Faridi, S., & Zakernejad, S. (2011). Improving ECG classification accuracy using an ensemble of neural network modules. *PloS one, 6*(10), e24386.
Lilly, & S, L. (2012). *Pathophysiology of heart disease: a collaborative project of medical students and faculty*. Lippincott Williams & Wilkins.
Liu, Y., Yin, B., & Cong, Y. (2020). The probability of ischaemic stroke prediction with a multi-neural-network model. *Sensors, 20*(17), 4995.
Novak, V., Hu, K., Desrochers, L., Novak, P., Caplan, L., Lipsitz, L., & Selim, M. (2010). Cerebral flow velocities during daily activities depend on blood pressure in patients with chronic ischemic infarctions. *Stroke; A Journal of Cerebral Circulation, 41*(1), 61–66.
Olindo, S., Cabre, P., Deschamps, R., Chatot-Henry, C., Rene-Corail, P., Fournerie, P., … Smadja, F. M. D. (2003). Acute stroke in the very elderly: Epidemiological features, stroke subtypes, management, and outcome in martinique, French West Indies. *Stroke; A Journal of Cerebral Circulation, 34*(7), 1593–1597.
Rabinezhadsadatmahaleh, N., & Khatibi, T. (2020). A novel noise-robust stacked ensemble of deep and conventional machine learning classifiers (NRSE-DCML) for human biometric identification from electrocardiogram signals. *Informatics in Medicine Unlocked, 21*, 100469.
Raghunath, S., Pfeifer, J. M., Ulloa-Cerna, A. E., Nemani, A., Carbonati, T., Jing, L., vanMaanen, D. P., Hartzel, D. N., Ruhl, J. A., Lagerman, B. F., et al. (2021). Deep neural networks can predict new-onset atrial fibrillation from the 12-lead ECG and help identify those at risk of atrial fibrillation–related stroke. *Circulation, 143*(13), 1287–1298.
Rathakrishnan, K., Min, S.-N., & Park, S. J. (2021). Evaluation of ECG features for the classification of post-stroke survivors with a diagnostic approach. *Applied Sciences, 11*(1), 192.
Sengupta, A., Rajan, V., Bhattacharya, S., & Sarma, G. (2016). A statistical model for stroke outcome prediction and treatment planning. *2016 38th annual international conference of the IEEE engineering in medicine and biology society (EMBC)* (pp. 2516–2519). IEEE.
Subudhi, A., Dash, M., & Sabut, S. (2020). Automated segmentation and classification of brain stroke using expectation-maximization and random forest classifier. *Biocybernetics and Biomedical Engineering, 40*(1), 277–289.
Xie, Y., Yang, H., Yuan, X., He, Q., Zhang, R., Zhu, Q., … Yan, C. (2021). Stroke prediction from electrocardiograms by deep neural network. *Multimedia Tools and Applications, 80*(11), 17291–17297.
Zhang, Q., Yang, X., Pengjie, Y., & Chaoyi, P. (2013). Acute ischaemic stroke prediction from physiological time series patterns. *The Australasian Medical Journal, 6*(5), 280.