iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** VI **Month of publication:** June 2023

DOI: https://doi.org/10.22214/ijraset.2023.53345

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue VI Jun 2023- Available at www.ijraset.com

Ensemble-Based AI System for Brain Stroke Prediction

Bindu Gottam[1], Leena Mandula[2], Amulya Kanaparthi[3], Dr K Kranthi Kumar[4], Ganesh B Chavan[5]
[1, 2, 3]Student, [4]Associate Professor, [5]Assistant Professor, SreeNidhi Institute of Science and Technology, Yamnampet, Ghatkesar, and 501301, India

***Abstract:** Timely detection and proactive measures to avert stroke are of utmost importance due to the significant likelihood of severe disabilities or fatal consequences associated with this condition. It is imperative to promptly administer appropriate thrombolytic or anticoagulant medications for both ischemic and hemorrhagic strokes. The pivotal and initial stage revolves around timely recognition of the initial indicators of a stroke, which may differ among individuals, and promptly seeking medical intervention within the prescribed treatment window. This research introduces a machine learning-based system that employs real-time measurements of electrocardiogram (ECG) and photoplethysmography (PPG) data to forecast and interpret stroke prognostic symptoms in a meaningful way. To achieve real-time stroke prediction, we have developed and implemented an ensemble structure voting classifier that combines SVM, Random Forest, and decision tree classifiers. This approach accurately forecasts stroke diagnosis in patients and can be easily implemented by utilizing a patient's ECG and PPG attribute data.*
***Keywords:** Deep learning, machine learning, electrocardiogram (ECG), photoplethysmography (PPG), real-time brain stroke prediction*

I. INTRODUCTION

Brain strokes, often referred to as strokes or cerebrovascular accidents (CVAs), are brought on by interruptions in the blood supply to a portion of the brain, which results in brain cell damage or death. The blood vessels supplying the brain may get blocked (ischemic stroke) or bleed (hemorrhagic stroke), which can result in this disturbance. Strokes can have major repercussions, such as cognitive and physical deficits, and they necessitate rapid medical care. Options for treatment and recovery vary on the kind and extent of the stroke. In order to enhance patient outcomes, stroke is a serious medical condition that requires prompt treatment.

The World Health Organization (WHO) released its 2019 Causes of Death Report in December 2020. It found that the top 10 causes of death accounted for 55% of all deaths reported in 2019 (or around 55.4 million people). According to [7], which is referring to the US, stroke rates are relatively high: A stroke occurs every 40 seconds, affects about 800,000 people annually, and is the largest cause of adult disability. Stroke is also the sixth-highest cause of mortality[7].

Additionally, recent research has linked COVID-19 to stroke, increasing the likelihood that people may die from strokes [3]. According to Kummer et al., patients with COVID-19 who had a history of stroke had a considerably higher mortality rate than those without a stroke history.

Stroke disease can be identified using imaging tests like Computed Tomography(CT), Magnetic Resonance Imaging(MRI), CT Angiography (CTA) Magnetic Resonance Angiography (MRA), blood tests, ECG, and Transcranial Doppler (TCD) ultrasound.

The most popular methods for diagnosing stroke are CT and MRI, but they come with concerns like radiation exposure or possible allergic reactions to the contrast chemicals. These methods have drawbacks in that they take a long time to test and are expensive to test, and they are also difficult to see in real-time at an early stage. Access to CT and MRI scans may be limited in certain healthcare settings, particularly in resource-limited areas or during emergencies. The availability of these imaging modalities can be a challenge, leading to delays in obtaining timely scans for stroke prediction and diagnosis.

Recent research has sought to forecast stroke issues using statistical or machine learning techniques while taking specific risk factors into account in order to get over these constraints.

A heart's electrical activity is captured on an ECG, which can provide vital details about the heart's rhythm and functionality. An elevated risk of stroke is linked to specific cardiac diseases, such as atrial fibrillation (an erratic pulse). ECG monitoring and detection of these irregular cardiac rhythms can help to identify those who are at increased risk of stroke.

PPG uses light-based technology to monitor changes in blood volume through the skin. Vascular dynamics, such as arterial stiffness and endothelial function, can be reflected by PPG signals.

©IJRASET: All Rights are Reserved

The integration of ECG and PPG signals can provide a comprehensive assessment of both cardiac and vascular health, enabling a more holistic approach to stroke prediction.

Using multi-modal bio signals based on ECG and PPG, we proposed a unique method in this research for the early diagnosis of stroke sickness. Seniors 65 years of age or older provided the ECG and PPG data that was measured and gathered for this study.

In order to evaluate deep learning models alongside machine learning models, the gathered ECG and PPG data were subjected to training and assessment processes. Our experiments showed high accuracy for Support Vector Machine(SVM), Voting Classifier, Decision Tree, and Random Forest.

In this paper, we used the Voting Classifier algorithm to predict stroke disease. The vote classification approach combines a number of different classifiers. We used SVM, decision tree, and Random Forest classifiers in the voting classification method because of the high accuracy rate. Furthermore, we verified that the strategy put forth in this study accurately predicts in real time the early warning signs of the stroke disease with exceptionally excessive fatality and resumption rates.

II. LITERATURE REVIEW

Both the diagnosis and management of strokes depend heavily on the quantitative analysis of brain MRI images. Deep neural networks with a high learning capacity enable lesion detection. Deep learning technology, which takes the form of a convolutional neural network (CNN), enables intelligent MRI interpretation. To create increasingly complex abstract features for classification, detection, and segmentation, CNN automatically gathers characteristic values from many samples.

A specific type of object identification job employing deep learning, such as Faster R-CNN, SSD, and YOLOv3, includes the detection of lesions in medical pictures. Creating anchor boxes and combining them to identify the picture using faster R-CNN takes a long time. The first convolution layer, conv4_3, in the single shot detection technique, has a spatial size of 38X38, which is a significant decrease from the input picture and makes prediction for smaller objects impossible. Because each grid cell in YOLO is intended for single-object identification, the YOLOv3 algorithm has difficulty detecting a single object inside a group of objects. With an accuracy of 89.77%, these three networks are used for automated lesion identification.

In summary, the diagnosis and management of strokes heavily rely on quantitative analysis of brain MRI images. Deep neural networks, particularly CNN, are employed for lesion detection, leveraging their ability to learn complex features. Faster R-CNN, SSD, and YOLOv3 are deep learning methods used for lesion detection, each with its own challenges and limitations. These networks have demonstrated high accuracy and are valuable tools for automated lesion identification[2].

Electroencephalography(EEG) data can be easily measured compared to other imaging techniques. The utilization of EEG can assist in the diagnosis of various conditions, including sleep disorders and brain tumors, among numerous other illnesses.

Yoon-A Choi claimed that our approach can forecast stroke illnesses in older people utilizing real-time EEG data. In order to test this technology, they developed a walking regimen that mimics the everyday activities of an old person. For this study, EEG data from Korean seniors 65 years of age or older were measured and collected. In order to contrast deep learning models with machine learning models, we divided the acquired EEG data into two sets: raw data and data extracted in the frequency domain. The CNN-Bidirectional LSTM models in this experiment demonstrated 94.0% accuracy. This shows tremendous trust in the results[1].

According to Ardabili et al. (2020), the SARS-CoV-2 (severe acute respiratory syndrome corona virus 2) viral strain is likely to be the source of COVID-19 (corona virus disease 2019). The World Health Organization (WHO) and international governments have categorized this sickness as being highly contagious.

Pre-processing, feature extraction, and classification are a few of the phases in the proposed COVID-19 prediction model. The COVID-19 dataset was gathered via Kaggle and consists of patients from Mexico. The input dataset is converted such that instances that are positive are shown as 1, and instances that are adverse are shown as 0. By calculating the mean of the entire dataset, the pre-processing stage eliminates the dataset's various unimportant characteristics as well as any missing values.

To minimize the number of features, GurjotKour and Pawanesh Abrol employed the PCA technique. The reduced features were then clustered using the K-mean method. The COVID-19 prediction's voting categorization uses the k-mean algorithm's result as its input. In this study, we merge Bernoulli Naive Bayes, SVM, Random Forest Classifier, and naive Bayes for COVID-19 prediction into a hybrid voting classification model[3].

III. METHODOLOGY

To conduct a performance analysis of algorithms, we evaluated our model using a set of 15 algorithms. Among these, five widely recognized algorithms are from the domain of deep learning and 10 from machine learning. We build our proposed system as an ensembled structured with the best classifiers according to their performance.

©IJRASET: All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue VI Jun 2023- Available at www.ijraset.com

A well-liked supervised machine learning approach for classification problems is the SVM classifier, commonly known as the Support Vector Machine classifier. It operates by identifying the appropriate hyperplane for dividing various data point classes. By using a set of labeled training examples, the SVM classifier learns to classify new, unseen data points into the appropriate categories.

The decision tree classifier is a widely employed supervised machine learning technique, particularly favored for its effectiveness in classification tasks.

It constructs a tree-like model that makes decisions based on feature values to assign data points to different classes or categories. The decision tree classifier uses a series of if-else conditions based on the features to navigate through the tree and reach a final prediction. However, decision trees can be prone to overfitting, and ensemble methods like random forests or gradient boosting are often used to improve their performance.

A machine learning ensemble technique called Random Forest mixes different decision trees to generate predictions. It is an effective method that is used for both classification and regression applications. By utilizing a randomized subset of the training data, a Random Forest algorithm constructs an ensemble of decision trees. Additionally, each tree only considers a random subset of the input attributes while

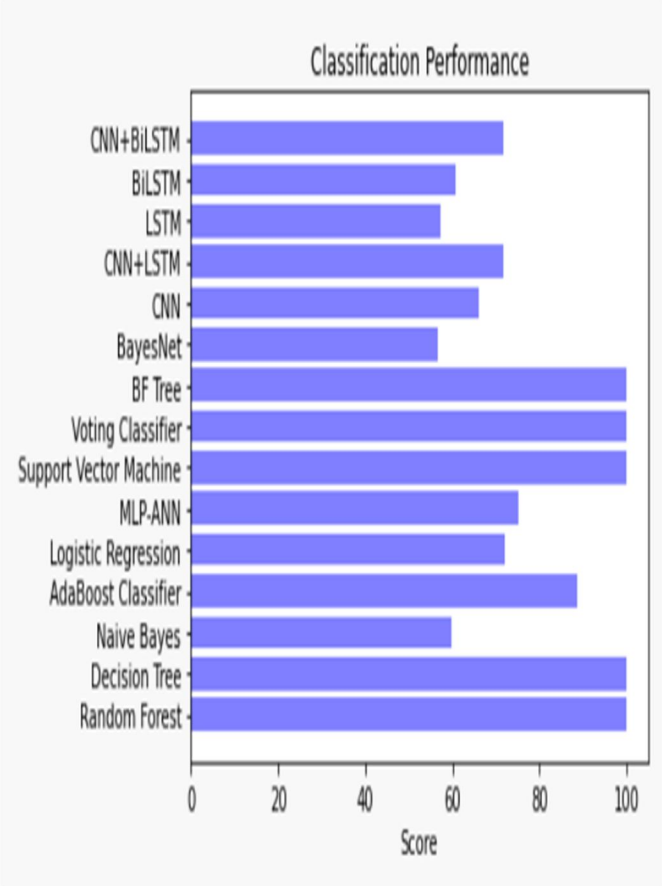


Figure 1. Algorithms Comparison

rendering a decision. This randomization enhances the model's capacity for generalization and helps to minimize overfitting.

Voting Classifier is a technique to machine learning that combines the output of several classifiers into a single prediction. In the system that was suggested, we used hard

voting, in which each classifier in the ensemble offers a

prediction, and the chosen prediction is determined by a majority vote. By incorporating a diverse range of base models into the voting classifier, any potential errors made by individual models can be effectively addressed and resolved.

To enhance the classification performance, we incorporated three classifiers—SVM, decision tree, and random forest into our ensemble voting classifier

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

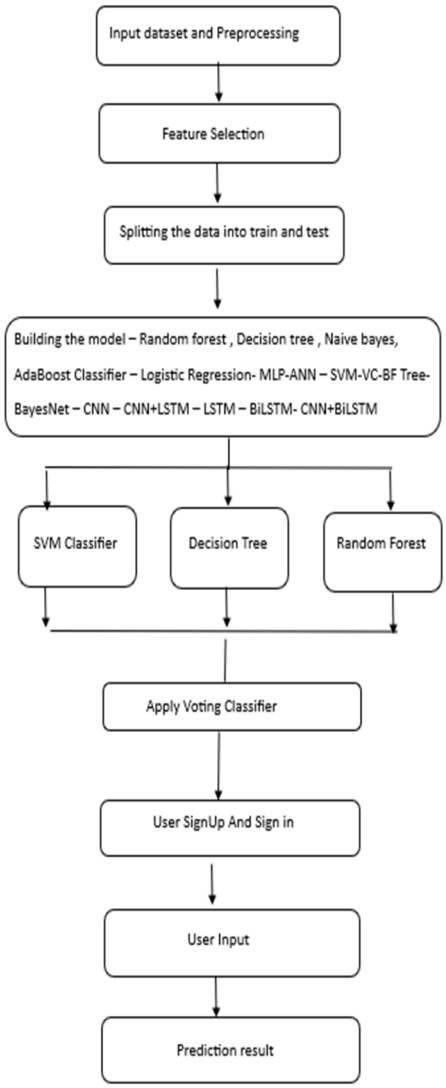


Figure 2. Proposed System Architecture

IV. IMPLEMENTATION

1) Using this module, we will import each and every package.
2) This module will be used to upload the dataset for the arrhythmia data analysis.
3) Processing of data: This module will be used to read data for processing.
4) Using this module, you may visualize data and information using Seaborn and Matplotlib.
5) Separating the dataset into a train and a test for processing: The train and test portions of the dataset will be separated using this module.
6) Creation of the model: We'll create all algorithms with the help of this module.
7) constructing the model: Since the Voting Classifier provides superior accuracy when compared to other models, the model was built using algorithms trained for processing and prediction using this module.
8) For registration and authentication, the Flask Framework uses SQLite Users may sign up, log in, and import packages using this module.
9) When utilizing this module, the user offers prediction input as feature values, which are subsequently preprocessed.
10) Using this module, the final outcome is shown through the frontend. A trained model is utilized for prediction.

©IJRASET: All Rights are Reserved

V. EXPERIMENTAL RESULTS

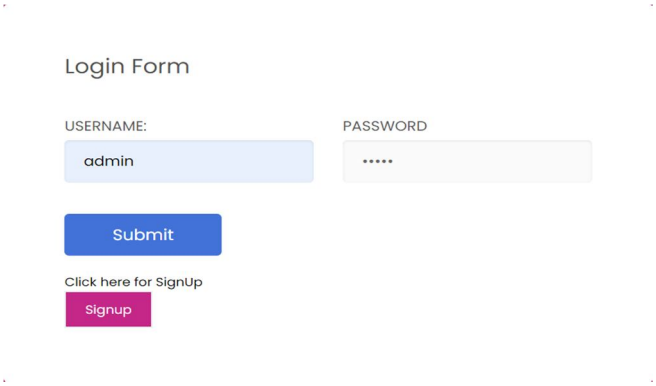


Figure 3. Sign in or Sign Up in

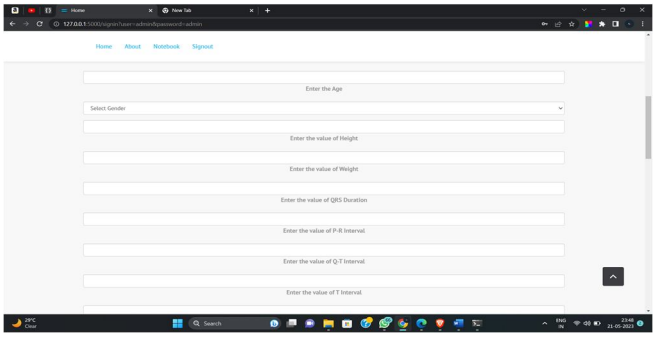


Figure 4. Welcome Screen

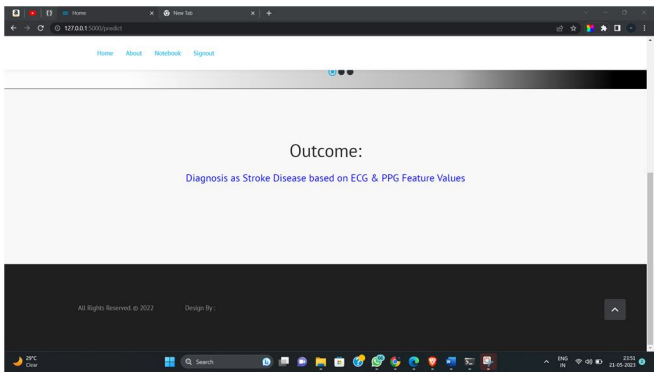


Figure 5. Enter the Values of the Patient

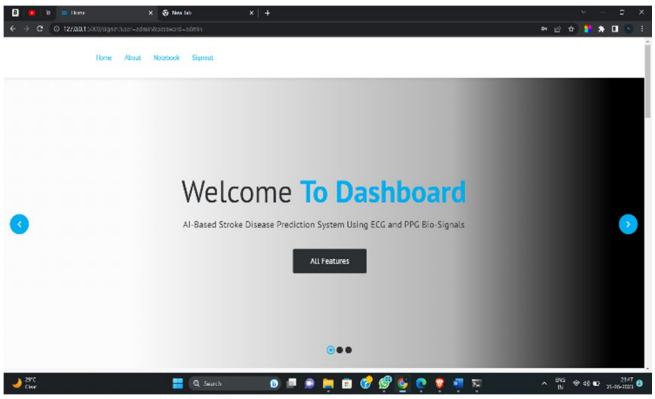


Figure 6. Diagnostic prediction of stroke

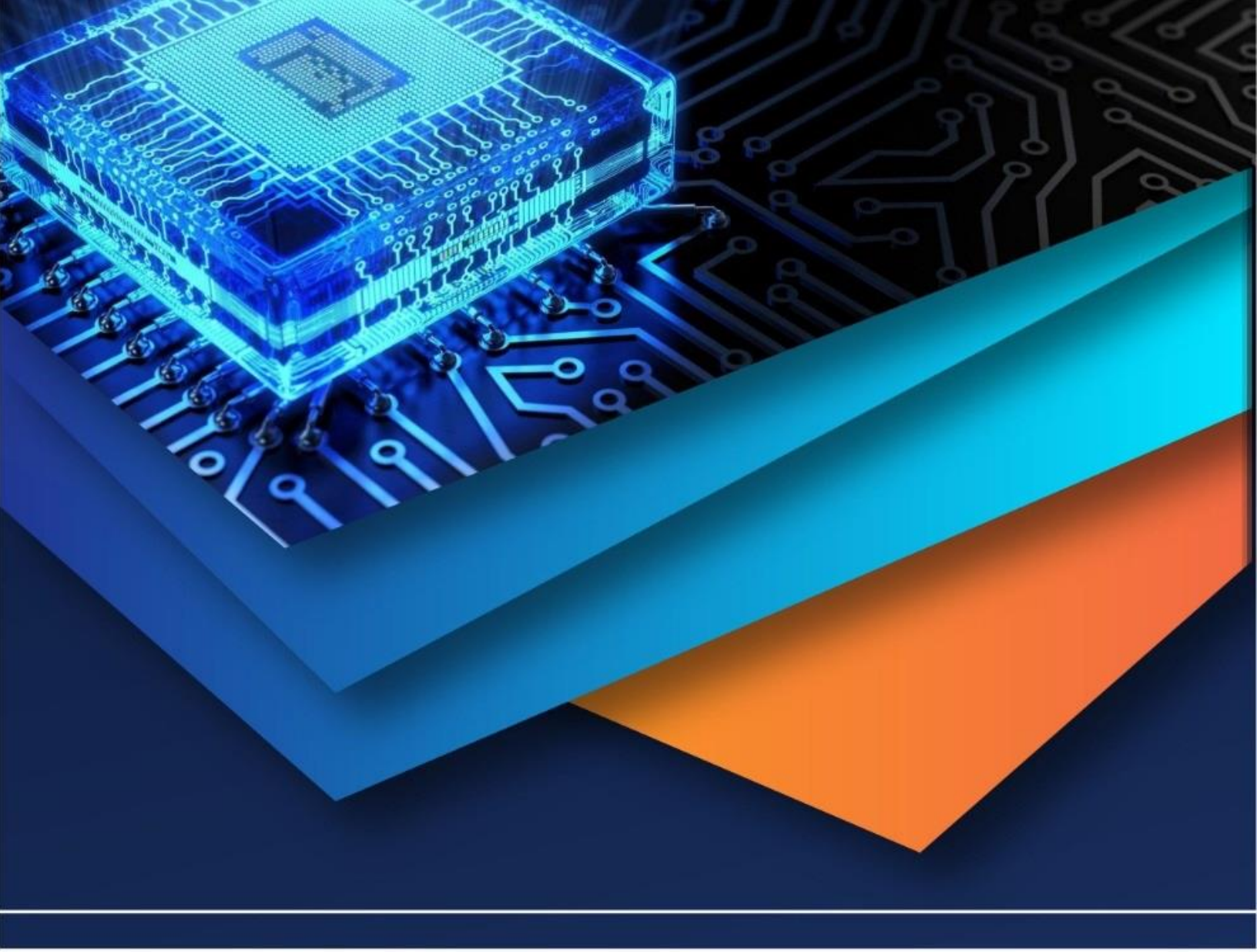
VI. CONCLUSION

In this research paper, a novel system is introduced that harnesses biological signals, such as ECG and PPG, collected during the regular activities of elderly individuals. The suggested technique enables quick recognition and prognostication of prognosis symptoms related to stroke illness by collecting real-time bio-signals as ECG and PPG. A machine learning-based prediction model is being used in the study that makes use of several bio signals. As a result of the signals' division into discrete components, prediction accuracy is increased and semantic interpretations are made easier.

We intend to investigate the stroke illness in-depth and do research to forecast it in the future. Numerous bio-signals will need to be examined, including EEG, EMG, foot pressure, mobility data, electronic medical records (EMRs), and MRI image data. We want to get comprehensive knowledge on stroke sickness and enhance our comprehension of it through a multi-modal approach.

REFERENCES

[1] Yoon-A Chai, Se-Jin Park “Deep Learning-Based Stroke Disease Prediction System Using Real-Time Bio Signals” Chungnam National University College of Medicine,korea, Published online 2021 Jun 22. doi: 10.3390/s21134269.
[2] Shujun Zhang , Shuhao Xu, Liwei Tan, Hongyan Wang “Stroke Lesion Detection and Analysis in MRI Images Based on Deep Learning” Qingdao University of Science and Technology, Published 2021 Apr 10.
[3] GurjotKoura , Pawanesh Abrolb , Namrata Kalrupiac “Hybrid Voting Classifier Model for COVID-19 Prediction by Embedding Machine Learning Techniques” Mahant Bachittar Singh College of Engineering & Technology-Vol. 13 No. 2 (2022)
[4] Jeena R S, Dr.Sukesh Kumar “Stroke Prediction Using SVM” Rajiv Gandhi Institute of Development Studies. Published in 2016, **doi:** 10.1109/ICCICCT.2016.7988020
[5] Yuheng Liu, Chenxuan Zhang, Xiaoyang Zheng, Yuhan Liu, Jiangping He “Stroke prediction model based on decision tree” Chongqing University of Technology, Published on 2023 Mar 7, doi : 10.37394/23208.2023.20.3
[6] Michael Wiryaseputra’s “Stroke Prediction Using Machine Learning Classification Algorithm-Random Forest” International Journal of Scientific & Engineering Research-2017
[7] Jaehak Yu , Sejin Park , Soon-Hyun Kwon , Kang-Hee Cho , And Hansung Lee “AI-Based Stroke Disease Prediction System Using ECG and PPG Bio-Signals” Published on 2021 Apr 21
[8] Alessandro Fedeli , Igor Bisio “Mobile Smart Helmet for Brain Stroke Early Detection through Neural Network-Based Signals Analysis” Published on 2018 Jan 15, doi: 10.1109/GLOCOM.2017.8255029
[9] Sailasya, G., & Kumari, G. L. A. “Analyzing the Performance of Stroke Prediction using ML Classification Algorithms ” International Journal of Advanced Computer Science and Applications, 12(6). doi: 10.14569/IJACSA.2021.0120662
[10] Tazin, T., Alam, N. H., Dola, N., Bari, M., Bourouis, S., & Khan, M. M. “Stroke Disease Detection and Prediction Using Robust Learning Approaches” Journal of Healthcare Engineering, 2021, 1–12.
[11] Widodo, A., & Yang, B. “Support vector machine in machine condition monitoring and fault diagnosis. Mechanical Systems and Signal Processing” 21(6), 2560–2574. 2007
[12] Heo, J., Yoon, J. G., Park, H., Kim, Y. H., Nam, H. S., & Heo, J. H. “Machine Learning–Based Model for Prediction of Outcomes in Acute Stroke” 50(5), 1263–1265. 2019
[13] Wu, Y., Zhu, M., Li, D., Zhang, Y., & Wang, Y. “Brain stroke localization by using microwave-based signal classification” 2016 Nov 03. doi:10.1109/ICEAA.2016.7731527
[14] Sirsat, M. S., Fermé, E., & Câmara, J “Machine Learning for Brain Stroke: A Review”. Journal of Stroke and Cerebrovascular Diseases, 29(10), 105162. 2020 Aug 12
[15] Fang, G., Huang, Z., & Wang, Z. “Predicting Ischemic Stroke Outcome Using Deep Learning Approaches” Frontiers in Genetics, 12. 2022 Jan 24.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)