Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol 32 No. 2s (2025)

Hybrid Machine Learning Approach for Early Stroke Prediction in Elderly People

Dr.P.Divya[1], Dr.K.Lakshmi Prabha[2], Dr.B.Aruna Devi[3], V.Vinoth Kumar[4], K.Balaji[5], Dr. C.Ezhilazhagan[6], Dr.R.Senthil Ganesh[7]

[1]Assistant Professor, ECE, Sri Krishna College of Technology, Coimbatore, Email: p.divya@skct.edu.in

[2]Head & Associate professor, Department of ECE (ACS), Chennai institute of technology, Kundrathur, Email:lakshmisslp@gmail.com

[3]Professor, Department of ECE, Dr NGP Institute of Technology, Coimbatore, Email:arunadevi@drngpit.ac.in

[4]Assistant Professor, Department of ECE, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, Email:icevinoth@gmail.com

[5]Assistant professor, Department of AI&DS, Sri Krishna College of Engineering and Technology, Coimbatore balajisparklers@gmail.com

[6]Assistant Professor, Department of ECE, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, Email:ezhisang20@gmail.com

[7]Associate Professor, Department of ECE, Sri Krishna College of Engineering and Technology, Coimbatore drsenthilganesh@gmail.com

Article History:

Received: *11-09-2024*

Revised: *19-10-2024*

Accepted: *28-10-2024*

Abstract:

Stroke represents a significant global health challenge, often leading to severe disability and mortality. The timely prediction and intervention of stroke are paramount in enhancing patient outcomes and reducing healthcare burdens. This project proposes a machine learning-based approach to predict stroke risk using multi-modal biosignals, specifically electrocardiogram (ECG) data. By leveraging advanced algorithms, including Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks, the system aims to classify patient health data into critical risk categories such as Normal, Abnormal, Ischemic, and Hemorrhagic. The study utilizes a comprehensive dataset consisting of ECG signals and incorporates techniques for data preprocessing, class balancing, and feature extraction. The predictive model is trained and validated using robust evaluation metrics, including accuracy, precision, recall, and F1-score. The findings underscore the efficacy of the proposed system in providing real-time stroke risk assessments, offering a cost-effective alternative to traditional diagnostic methods. Furthermore, this research explores the integration of wearable technology with machine learning, highlighting its potential for continuous patient monitoring and early detection of stroke symptoms. By creating a user-friendly interface for healthcare professionals and patients, the system aims to facilitate prompt decision-making and intervention, ultimately improving the overall quality of care for individuals at risk of stroke

Keywords: Stroke, electrocardiogram (ECG), Convolutional Neural Networks (CNNs), Convolutional Neural Networks (CNNs), diagnostic methods.

1. INTRODUCTION

Stroke is a leading cause of death and long-term disability, making timely prediction and intervention crucial for improving patient outcomes. Advances in wearable technology and machine learning provide new opportunities for real-time health monitoring and predictive analytics. By leveraging these innovations, we can enhance the accuracy of stroke risk assessments and empower individuals to take proactive health measures. This project aims to develop a machine learning-based system for stroke prediction using ECG data from wearable devices, ultimately facilitating early detection and intervention [1]. The early prediction and intervention of stroke are essential for enhancing patient

Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol 32 No. 2s (2025)

outcomes and reducing healthcare burdens. Understanding the risk factors and early symptoms associated with strokes can lead to improved management and treatment outcomes. Recent studies indicate that strokes can manifest suddenly, often without prior warning. The ability to predict stroke risk accurately can facilitate timely medical intervention, potentially saving lives and reducing the long-term impact on patients' quality of life. Therefore, effective screening methods and predictive models are necessary to identify individuals at high risk and to implement preventive measures.

Understanding these classifications is essential for developing targeted interventions and predictive models. Each type of stroke may require different management strategies, and early identification of symptoms can lead to improved treatment efficacy. Recent advancements in healthcare technology have significantly transformed how we monitor and manage health conditions. In particular, the emergence of wearable devices has led to a substantial increase in the volume of health data available for analysis [2].

The integration of these wearable technologies with machine learning algorithms has opened new avenues for early detection and prevention of health issues, including strokes. For instance, devices that capture ECG data enable healthcare providers to monitor heart rhythms and detect irregularities that may indicate an increased risk of stroke. The use of advanced algorithms allows for the processing and analysis of this data, facilitating real-time alerts for potential health risks.

By leveraging machine learning techniques, we can analyze the vast amounts of data generated by these wearable devices to develop sophisticated monitoring systems that alert healthcare providers and patients to potential stroke risks. This proactive approach not only enhances the accuracy of stroke prediction but also empowers individuals to take preventive measures based on timely information. Machine learning algorithms, particularly deep learning techniques such as Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks, have shown remarkable promise in analyzing complex medical data. These algorithms can learn intricate patterns from historical health data, making them suitable for predicting health events such as strokes based on ECG signals [3].

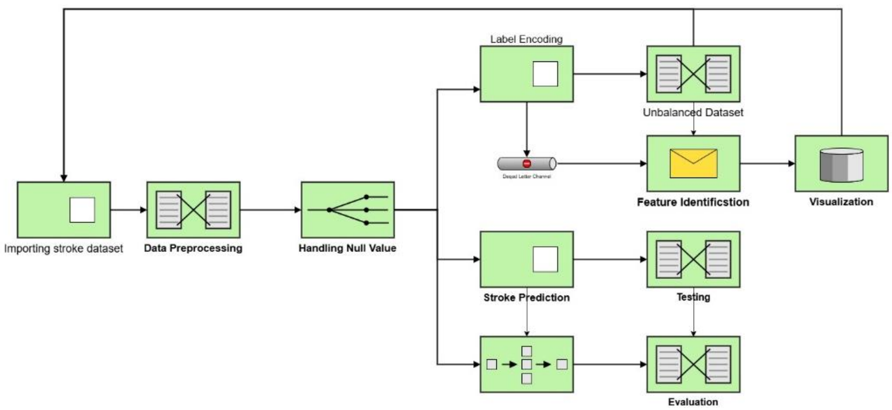

Figure 1: Architecture of the Stroke Prediction framework

Machine learning models can be trained to identify specific patterns associated with strokes, allowing for personalized health monitoring. This adaptability is particularly beneficial in healthcare settings, where individual patient characteristics can influence outcomes. Moreover, the increasing availability of large datasets enhances the training process, enabling more robust and accurate predictive models. In this paper, we aim to develop a machine learning-based system for stroke risk prediction that integrates ECG data from wearable devices. The model will classify patients into different risk categories and provide real-time alerts, ultimately aiding in timely medical interventions. By enhancing predictive accuracy and leveraging advancements in technology, this research strives to contribute to the field of preventive healthcare and improve the quality of care for individuals at risk of stroke. We hope to demonstrate the feasibility and effectiveness of integrating machine learning with wearable health technologies to create a proactive healthcare solution that addresses the critical issue of stroke prediction [4].

2. RELATED RESEARCH

The literature on this topic highlights various methodologies and approaches to analyzing physiological signals, particularly electrocardiogram (ECG) data. This survey reviews key studies that have contributed to the understanding of stroke prediction and related cardiovascular conditions through machine learning. By examining existing research, we can identify gaps in the current methodologies and highlight the importance of integrating advanced algorithms to enhance predictive accuracy [5]. The following summaries encapsulate the core findings and methodologies of six relevant studies that inform the framework of this paper. This research presents a wearable sensor-based stroke prediction model that identifies ischemic and hemorrhagic strokes through real-time data analysis [6]. The study highlights the importance of using multiple physiological signals to improve prediction accuracy, achieving a high reliability rate in stroke identification. The findings advocate for the integration of wearable technology with advanced machine learning algorithms to provide timely alerts for at-risk individuals [7].

The authors propose a model that combines features extracted from deep networks with machine learning classifiers [8]. The model showcasing the potential of machine learning in resource-constrained environments, particularly during the COVID-19 pandemic [9]. This study investigates sleep bruxism through the analysis of EEG signals using power spectral density. The research compares bruxism patients with normal subjects, finding significantly higher spectral density values in patients during various sleep stages [10]. The study also demonstrates the use of machine learning classifiers to improve the prognosis of sleep bruxism, highlighting the potential for accurate and timely diagnosis in clinical settings. The reviewed literature demonstrates a clear trend toward leveraging machine learning and deep learning algorithms for the analysis of physiological signals.

3. PROPOSED METHODLOGY

The proposed system is designed to predict stroke risk and provide semantic interpretation for elderly individuals by utilizing ECG-based multi-modal bio-signals. This system is capable of real-time detection and prediction of early symptoms related to stroke by gathering multi-modal bio-signals continuously. The study focused on participants aged 65 and older, collecting and storing ECG signals while they were engaged in walking activities. These bio-signals were segmented into specific waveform intervals to develop predictive models using machine learning, aiming to achieve more precise predictions and interpretations.

The research also demonstrated that deep learning models for time series analysis could accurately identify early stroke symptoms directly from raw data, eliminating the need for separate feature extraction. The multi-modal bio-signal-based predictive system introduced here shows potential for real-time detection and prediction of stroke symptoms, which have a high mortality and incidence rate among the elderly.

Neural networks, inspired by the structure of the human brain, consist of a vast array of interconnected neurons. These artificial neurons mimic biological ones, where each unit is linked to others, either promoting or inhibiting activation. The network processes inputs through a summation function, which combines the values received. Additionally, each connection and unit incorporates a threshold function that must be surpassed for signals to propagate to other neurons.

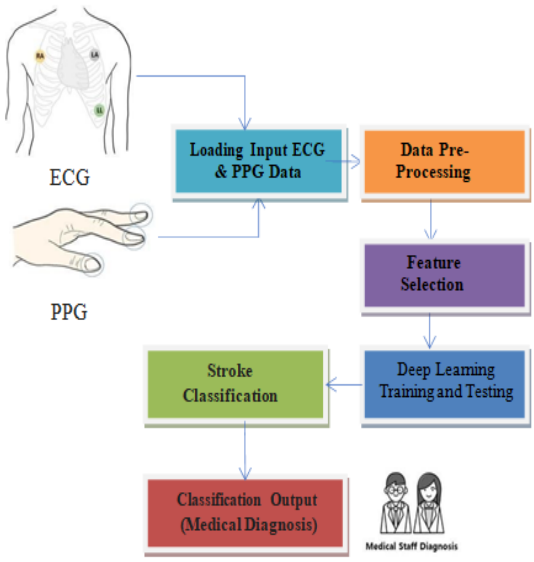

Figure 2: Proposed Block Diagram

In this study, we identified and introduced 29 novel attributes that have not been previously utilized in ECG-based multi-modal bio-signal studies involving machine learning and deep learning approaches. This contribution is significant as it enhances the ability to provide objective diagnoses and prognostic treatment, offering semantic analysis results that can support medical professionals in their decision-making. The proposed stroke prediction and monitoring system was experimentally validated to effectively predict early stroke symptoms in real-time. Additionally, it has the potential to be integrated into low-cost, everyday healthcare services for the elderly. The experimental data consisted of real-time ECG bio-signals collected from participants aged 65 and above while they were walking.

The system uses a layered feed-forward neural network, which is structured into multiple layers of processing elements. Each layer independently processes the input data and passes the results to subsequent layers. The next layer performs its own computations based on the received data and forwards the outcome to the next layer in the sequence. Ultimately, a final set of processing elements determines the network's output, providing a predictive result.

The hardware setup of the proposed system integrates various sensors and modules to collect and transmit physiological signals such as ECG, heart rate, and temperature. Each component is carefully chosen to ensure seamless data acquisition, processing, and communication, forming a robust foundation for the predictive model.

Table 1: Key Metrics of the Proposal

Hardware Requirements	Specifications
ECG Sensor	AD8232
Arduino NANO	ATmega328P
Temperature Sensor	DS18B20
Heart Rate Sensor	IR Based Sensor
Power Requirement(Node MCU)	ESP8266 WiFi Module

Internal Body Heat Level (IBL): The IBL is an approximation that shows the body's maximum internal temperature within the standard parameters. This rule becomes increasingly unusual at high or low elevations, as well as in other locations due to the extreme temperature fluctuations in the surrounding air. For example, vasodilation and drying out may result from high SH and low RH, which could quickly lead to decompensating.

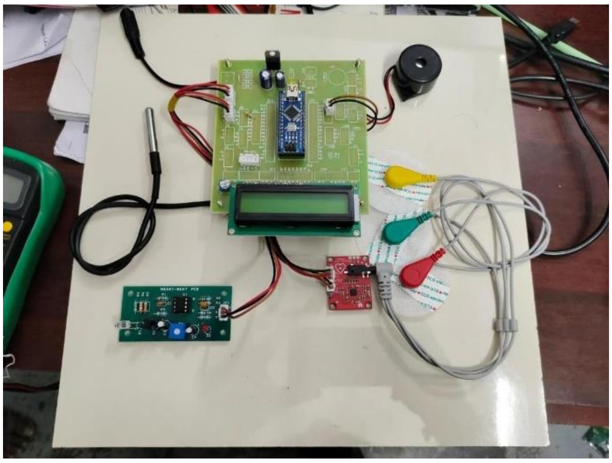

Figure 3: Hardware Setup of the proposed method

The hardware setup depicted in the diagram illustrates the integration of key components designed for real-time monitoring of physiological signals. The Arduino Nano serves as the central processing unit, efficiently managing data from various sensors, including the ECG Sensor (AD8232), Heart Rate Sensor, and Temperature Sensor. Each component plays a critical role in capturing vital health metrics, which are then processed and transmitted via the ESP8266 WiFi Module for real-time analysis. This configuration not only facilitates accurate stroke prediction but also enhances the system's overall functionality by providing essential alerts and visual feedback through the LCD Display and Buzzer.

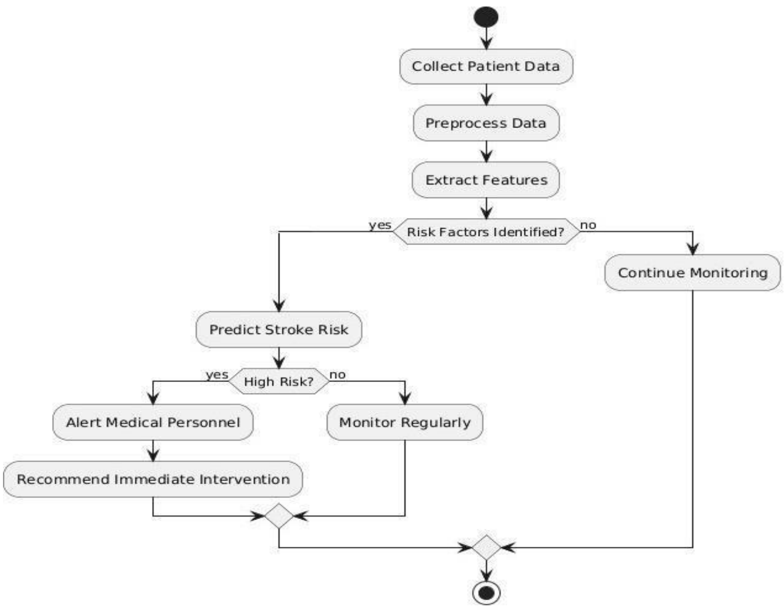

Figure 4: Software Model of the proposed method

This streamlined integration enables the system to efficiently collect, analyze, and respond to health data, supporting accurate and responsive stroke prediction.

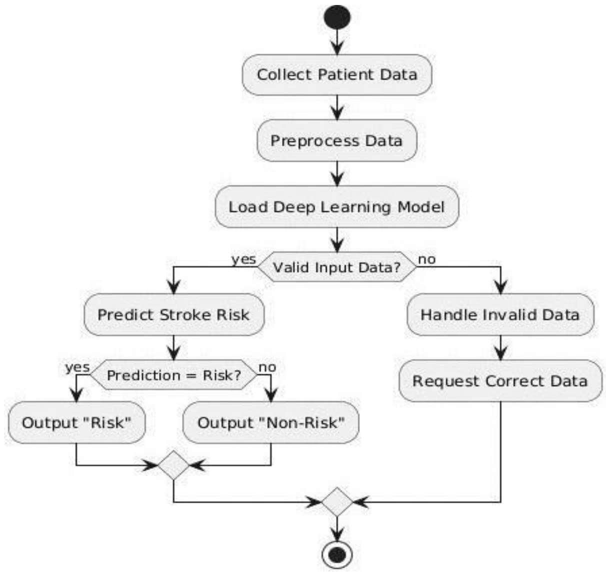

Figure 5 End-to-End Process Flow Diagram.

4. RESULTS AND DISCUSSION

The Results provides an in-depth analysis of the system's performance in predicting stroke and classifying ECG signals. Various techniques, including signal preprocessing, data augmentation, and machine learning model evaluation, have been employed to achieve accurate classification across categories such as Normal, Abnormal, Ischemic, and Hemorrhagic. Key results and metrics are discussed below.

This study used multimodal bio-signals, particularly real-time ECG data from elderly stroke patients and healthy individuals, to develop and validate the prediction model. The dataset includes signals from both stroke-diagnosed and non-stroke-diagnosed participants, providing a comprehensive basis for detecting abnormal heart conditions. Important data preparation steps included:Scaling the ECG signals to ensure consistency. Balancing the dataset through resampling to address class imbalances, this enhances model training. Techniques such as noise addition and signal rotation were applied to improve the model's robustness.

The ECG data was classified into the following categories, each representing different health conditions:

Normal: Signals with no abnormalities.

Abnormal: General irregularities that require further investigation.

Ischemic: Irregularities indicating restricted blood flow, often associated with ischemic strokes.

Hemorrhagic: Abnormalities related to brain bleeding, typical in hemorrhagic strokes.

False: Data flagged as unreliable or corrupted, often due to sensor noise.

This figure illustrates the distribution of classes (Normal, Abnormal, Ischemic, Hemorrhagic, False) used in the CNN classification model, showcasing the dataset's structure and balance after resampling. In this project, resampling is used to balance the dataset when there is an uneven distribution of stroke cases across different categories (Normal, Abnormal, Ischemic, Haemorrhagic). For example, if there are more data points for Normal cases and fewer for Ischemic or Haemorrhagic strokes, this imbalance can affect the model's ability to predict strokes accurately.

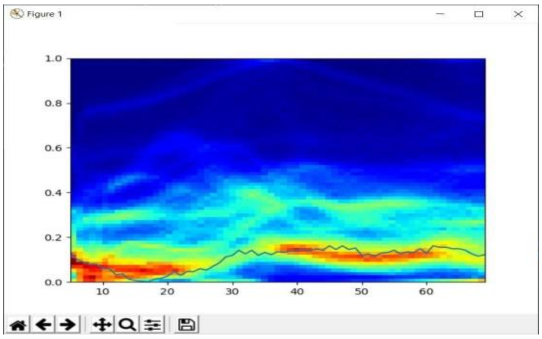

Figure 6 Signal heat map – Normal Beat

A signal heat map for a Normal Beat in the stroke detection project visually represents the ECG (electrocardiogram) or other bio- signal data across time. It uses color intensity to indicate the strength of the electrical activity of the heart during a normal heartbeat. For a Normal Beat, the heat map will show consistent, regular patterns with even spacing between heartbeats, reflecting a healthy heart rhythm without any abnormalities indicative of stroke or other issues.

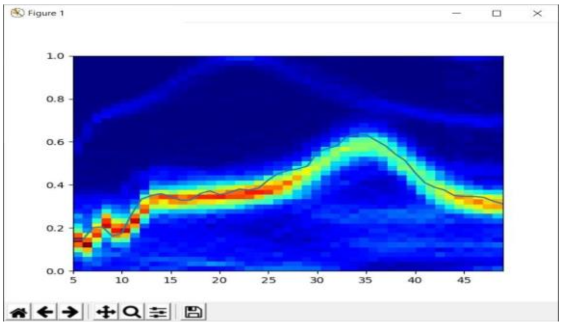

Figure 7 Signal Heat map - Abnormal Beat

A signal heat map for an Abnormal Beat in the stroke detection project visually highlights irregularities in the ECG or other bio-signal data. It uses colour gradients to reflect variations in the electrical activity of the heart. For an Abnormal Beat, the heat map will show irregular patterns, uneven intervals between beats, or sudden spikes or dips in signal strength. These variations may indicate arrhythmias, heart rate variability, or other early signs of potential stroke, distinguishing abnormal heart function from normal rhythms.

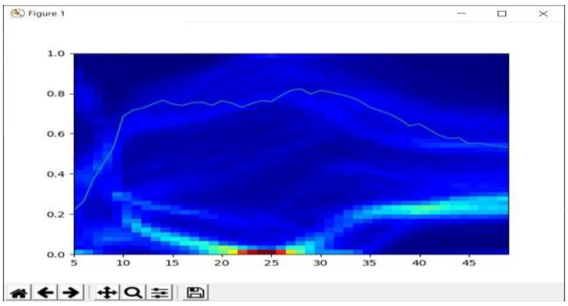

Figure 8 Signal Heat map - Ischemic Beat

A signal heat map for an Ischemic Beat in the stroke detection project visually captures the irregularities in heart activity caused by ischemia, which is typically a blockage in blood flow to the brain. The heat map uses color intensity to indicate electrical activity from ECG data over time. For an Ischemic Beat, the heat map will often display: ST-segment depression or elevation (a key marker of ischemia).Prolonged interval between heartbeats or erratic patterns in signal strength. Inconsistent coloring, indicating a reduction in blood flow to the heart, causing abnormal electrical activity. These abnormal patterns are significant for detecting ischemic strokes, providing visual clues of compromised heart function.

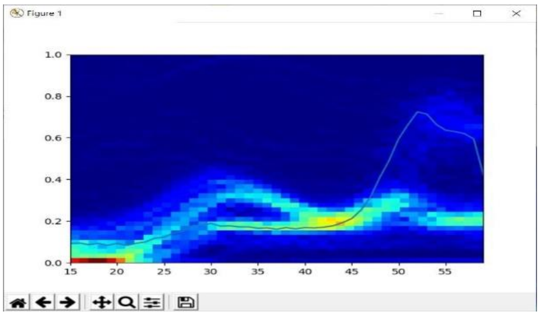

Figure 9 Signal Heat map - Haemorrhagic Beat

A signal heat map for a Haemorrhagic Beat in the stroke detection project visually represents the irregularities in heart activity that may arise due to haemorrhagic stroke, which involves bleeding in the brain. The map uses color intensity to show variations in the electrical activity of the heart over time. For a Haemorrhagic Beat, the heat map may show the Disrupted or erratic patterns in signal strength, reflecting the brain's response to internal bleeding.Abnormal fluctuations in heart rhythm or sharp shifts in signal intensity. Irregular coloring that suggests a disruption in the brain-heart communication.

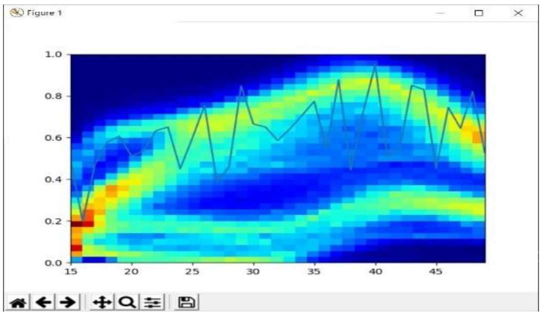

Figure 10 Signal Heat map - False Data

A signal heat map for False Data in the stroke detection project visually represents corrupted or misleading data that could arise from sensor errors, environmental interference, or incorrect data recording. This type of map helps identify when the data captured does not accurately reflect the patient's physiological state. For False Data, the heat map may show the Random or nonsensical patterns, with no consistent rhythm or relationship between time and signal strength This kind of heat map helps in identifying when the input data is unreliable, enabling the system to flag or filter out such essential to clean the signal and ensure accurate analysis. A spike in ECG data can represent noise or artifacts (such as power line interference, electrode issues, or motion artifacts) that need to be treated.

Table 2: Confusion Matrix Layout of the Proposal

	Predicted: Normal	**Predicted: Abnormal**	**Predicted: Ischemic**	**Predicted: Hemorrhagic**	**Predicted: False**
True: Normal	TP(Normal)	FP(Abnormal)	FP(Ischemic)	FP(Hemorrhagic)	FP(False)
True: Abnormal	FN(Normal)	TP(Abnormal)	FP(Ischemic)	FP(Hemorrhagic)	FP(False)
True: Ischemic	FN(Normal)	FN(Abnormal)	TP(Ischemic)	FP(Hemorrhagic)	FP(False)
True: Hemorrhagic	FN(Normal)	FN(Abnormal)	FN(Ischemic)	TP(Hemorrhagic)	FP(False)

True: False	FN(Normal)	FN(Abnormal)	FN(Ischemic)	FN(Hemorrhagic)	TP(False)

AI techniques can be applied to create personalized stroke risk prediction models that adapt based on individual health data over time, continuously learning from the user's habits, health patterns, and trends to improve prediction accuracy. As wearable devices, such as smartwatches and fitness trackers, become more sophisticated, the system can be optimized to collect real-time data seamlessly. This integration of AI- driven personalization and advanced wearables would enable continuous monitoring, providing more practical and accessible health management solutions. The system would not only deliver personalized stroke risk assessments but also offer proactive health insights tailored to the user's evolving needs.

5. CONCLUSION

The proposed system successfully integrates hardware and software components to create an efficient, real-time stroke prediction model. By leveraging IDLE 3.7 for the software setup and combining powerful machine learning techniques like CNN and LSTM, the system can accurately analyze ECG signals and predict stroke risk. The hardware components, including ECG sensors and microcontrollers, provide reliable data acquisition and transmission, ensuring seamless communication between the physical sensors and the prediction algorithm.This project demonstrates how AI and machine learning can be effectively applied to healthcare, particularly in monitoring critical health parameters and alerting users to potential risks. The system's ability to operate with minimal computational resources makes it especially suitable for use in low-resource environments, where early detection of stroke can significantly improve patient outcomes. Overall, this project paves the way for developing more accessible and portable health-monitoring systems, highlighting the potential of smart healthcare solutions to prevent life-threatening conditions in real time.

References

[1] R. Mehrrotraa, M. A. Ansari, R. Agrawal, P. Tripathi, M. B. B. Heyat, M. Al- Sarem, A. Y. M. Muaad, W. A. E. Nagmeldin, A. Abdelmaboud, and F. Saeed, ''Ensembling of efficient deep convolutional networks and machine learning algorithms for resource effective detection of tuberculosis using thoracic (chest) radiography,'' IEEE Access, vol. 10, pp. 85442–85458, 2022

[2] M. Bahrami and M. Forouzanfar, ''Sleep apnea detection from single-lead ECG: A comprehensive analysis of machine learning and deep learning algorithms,'' IEEE Trans. Instrum. Meas., vol. 71, pp. 1–11, 2022

[3] D. Lai, Y. Zhang, X. Zhang, Y. Su, and M. B. Bin Heyat, ''An automated strategy for early risk identification of sudden cardiac death by using machine learning approach on measurable arrhythmic risk markers,'' IEEE Access, vol. 7, pp. 94701–94716, 2019

[4] G. Xu, X. Shen, S. Chen, Y. Zong, C. Zhang, H. Yue, M. Liu, F. Chen, and W. Che, "A deep transfer convolutional neural network framework for EEG signal classification," IEEE Access, vol. 7, pp. 112 767–112 776, 2019

[5] D. Lai, ''Prognosis of sleep bruxism using power spectral density approach applied on EEG signal of both EMG1-EMG2 and ECG1- ECG2 channels,'' IEEE Access, vol. 7, pp. 82553–82562, 2019,

[6] J. Yu, D. Kim, H. Park, S. Chon, K. H. Cho, S. Kim, S. Yu, S. Park, and S. Hong, ''Semantic analysis of NIH stroke scale using machine learning techniques,'' in Proc. PlatCon, Jeju, South Korea, Jan. 2019, pp. 82–86

[7] T.-Y. Kim and S.-B. Cho, ''Predicting residential energy consumption using CNN-LSTM neural networks,'' Energy, vol. 182, no. 1, pp. 72–81, Sep. 2019

[8] F. Rundo, S. Conoci, A. Ortis, and S. Battiato, ''An advanced bio-inspired photoplethysmography (PPG) and ECG pattern recognition system for medical assessment,'' Sensors, vol. 18, no. 2, pp. 1–22, Jan. 2018.

[9] J. M. Bumgarner, C. T. Lambert, A. A. Hussein, D. J. Cantillon, B. Baranowski, K. Wolski, and K. G. Tarakji, ''Smartwatch algorithm for automated detection of atrial fibrillation,'' J. Amer. College Cardiol., vol. 71, no. 21, pp. 2381–2388, Mar. 2018

[10] G. Sannino and G. De Pietro, ''A deep learning approach for ECG-based heartbeat classification for arrhythmia detection,'' Future Gener. Comput. Syst., vol. 86, pp. 446–455, Sep. 2018.