International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

AI-BASED STROKE DISEASE PREDICTION SYSTEM USING ECG AND PPG BIO-SIGNALS

Madhu Bandari, Om Jhitri

1. Assistant Professor, Department of Data Science and Artificial Intelligence, IcfaiTech School, Telangana, India. madhu.bandari@ifheindia.org

2. Om Jhitri, Department of Computer Science and Engineering, IcfaiTech School, Telangana, India. omcool71@gmail.com

Abstract : Because stroke illness often results in death or major disability, aggressive primary prevention and early diagnosis of prognostic signs are critical. Stroke illnesses are classified as ischemic or hemorrhagic, and they should be treated as soon as possible with thrombolytic or coagulant therapy. First, it is critical to notice the precursor symptoms of stroke in real time, which vary by person, and to give professional treatment by a medical institution within the appropriate treatment window. Prior research, however, has concentrated on creating acute therapy or clinical treatment recommendations after the onset of stroke rather than identifying predictive indicators of stroke. Image analysis, such as magnetic resonance imaging (MRI) or computed tomography (CT), has been utilised extensively in recent research to identify and predict prognostic signs in stroke patients. These approaches are not only difficult to identify early in real-time, but they also have drawbacks in terms of extended test times and expensive testing costs. In this research, we present a machine learning-based method for predicting and semantically interpreting stroke prognostic symptoms in the elderly utilising multi-modal bio-signals of electrocardiogram (ECG) and photoplethysmography (PPG) recorded in real-time. We devised and deployed a stroke disease prediction system with an ensemble structure that integrates CNN and LSTM to predict stroke illness in real-time while walking. The suggested system takes into account the ease of wearing bio-signal sensors for the elderly, and bio-signals were captured while walking at a sample rate of 1,000Hz per second from the three electrodes of the ECG and the index finger for PPG. Real-time prediction of elderly stroke patients demonstrated good prediction accuracy and performance.

Index Terms : *Deep learning, machine learning, electrocardiogram (ECG), photo plethysmography (PPG), multi-modal bio-signal, real-time stroke prediction, stroke disease analysis.*

I. INTRODUCTION

Stroke is classified as either ischemic (a blood vessel delivering blood to a portion of the brain is blocked) or hemorrhagic (a blood vessel bursts). It is a neurological symptom and condition caused by injury to a specific region of the brain. Stroke is regarded as one of the most dangerous illnesses in contemporary civilization since it may result in mortality in extreme instances, as well as physical and mental impairments such as hemiparesis, speech impairment (aphasia), ataxia, vision impairment, consciousness impairment, and dementia. According to the World Health

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

Organization's (WHO) 2019 Causes of Death Report, issued in December 2020, the top ten causes of death accounted for 55% of all documented deaths in 2019. (about 55.4 million people). Six million of them died as a result of cerebrovascular illness, which was believed to be the second greatest cause of mortality. According to the United Nations (UN), a nation is categorised as an ageing society if the percentage of its people aged 65 and over in the total population is 7% or higher, an aged society if the proportion is 14% or higher, and a super-aged society if the proportion is more than 20%. As a result, the social difficulties of the ageing society are becoming visible enough that the ageing society may be segmented. Furthermore, according to an ageing study report published by Moody's, an international credit rating organisation, as of 2013, Japan, Germany, Italy, and other countries have become super-aged societies, with a percentage of senior people above 20%. According to reports, 34 nations would have become super-aged civilizations by 2030. The prognosis and health condition of stroke patients varies significantly depending on their age and place of onset. According to a prior stroke research, elderly persons 65 and older accounted for more than 66% of overall stroke incidence. Aside from these societal challenges, the incidence and mortality of stroke are likely to become significant social and economic issues.

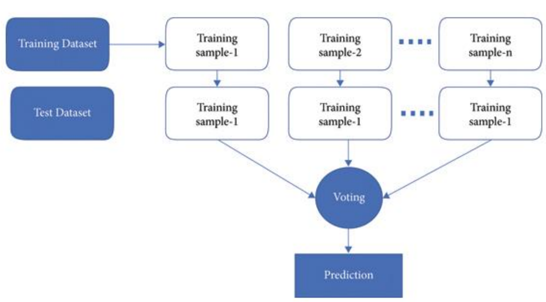

Fig.1: Example figure

The neurological diagnostic and severity information provided by a medical team are used to determine the diagnosis of stroke, which is represented by cerebrovascular illness [6], [10]-[12]. Brain MRI and CT are the most often used techniques for neurological diagnosis in stroke diagnosis, but other research have shown that bio-signals such as brain waves, muscle, and ECG may also be used to diagnose and prevent stroke disorders [13]-[15]. Furthermore, ultrasonography, echocardiography, cerebral angiography, and single photon emission computed tomography (SPECT) are being utilised to diagnose the most prevalent causes of stroke. Imaging methods like as CT and MRI have recently been popular for stroke detection, but they still have drawbacks in the examination and diagnostic process due to hypersensitive responses to contrast agent medication penetration, radiation exposure, and claustrophobia in a confined environment. Because test findings may include inaccuracies, medical staff judgement based on professional medical knowledge and factual evidence is regarded as critical.

II. RELATED WORK

Sylvie De Raedt , Aurelie De Vos , Jacques De Keyser : Impaired autonomic function, characterized by a predominance of sympathetic activity, is common in patients with acute ischemic stroke. This review describes methods to measure autonomic dysfunction in stroke patients. It summarizes a potential relationship between ischemic stroke-associated autonomic dysfunction and factors that have been associated with worse outcome, including cardiac complications, blood pressure variability changes, hyperglycemia, immune depression, sleep disordered breathing, thrombotic effects, and malignant edema. Involvement of the insular cortex has been suspected to play an important role in causing sympathovagal imbalance, but its exact role and that of other brain regions remain unclear. Although sympathetic overactivity in patients with ischemic stroke appears to be a negative prognostic factor, it remains to be seen whether therapeutic strategies that reduce sympathetic activity or increase parasympathetic activity might improve outcome.

Tapuwa D Musuka, Stephen B Wilton, Mouhieddin Traboulsi, Michael D Hill : Globally, stroke is the second leading cause of death.1 The estimated 62 000 strokes that occur each year in Canada affect all age groups, from neonates to elderly people, with occurrence rates rising by age. The lifetime risk of overt stroke is estimated at one in four by age 80 years, and the lifetime risk of silent or covert stroke is likely closer to 100%. Stroke affects men and women equally and causes major social and economic burdens to society, with direct costs above $3 billion annually in Canada.2 Acute stroke and acute coronary syndromes have many similarities. Here, we review the diagnosis and management of acute ischemic stroke and compare its treatment with that of acute coronary syndrome, to help illustrate how the rapid relief of arterial occlusion and restoration of normal blood flow can save lives and prevent disability. This narrative review is based upon a critical appraisal of relevant clinical trials (Box 1).

Qiaofeng Song, Xiaoxue Liu, Wenhua Zhou, Ling Wang, Xiang Zheng, Xizhu Wang & Shouling Wu: The objective of this study was to examine the relationship between sleep duration and ischemic and hemorrhagic stroke in a community-based cohort. The current analysis included 95,023 Chinese participants who were free of stroke at the baseline survey (2006–2007). Cox proportional hazards models were used to calculate hazard ratios (HRs) and their confidence intervals (CIs) for stroke, according to sleep duration. After a mean follow-up period of 7.9 years, 3,135 participants developed stroke (2,504 ischemic stroke and 631 hemorrhagic stroke). The full adjusted hazard ratio (95% CI) of total stroke (with 6–8 hours of night sleep being considered for the reference group) for individuals reporting greater than 8 hours was 1.29 (1.01–1.64). More significant association between long sleep duration and total stroke was found in the elderly (HR, 1.47; 95% CI, 1.05–2.07). Compared with participants getting 6–8 hours of sleep, only women who reported sleeping more than 8 hours per night were associated with hemorrhagic stroke (HR, 3.58; 95% CI, 1.28–10.06). This study suggested that long sleep duration might be a potential predictor/ marker for total stroke, especially in the elderly. And long sleep duration increased the risk of hemorrhagic stroke only in women.

Jaehak Yu ,Sejin Park ,Hansung Lee ORCID,Cheol-Sig Pyo andYang Sun Lee : Recently, with the rapid change to an aging society and the increased interest in healthcare, disease prediction and management through various healthcare devices and services is attracting much attention. In particular, stroke, represented by cerebrovascular disease, is a very dangerous disease, in which death or mental and physical aftereffects are very large in adults and the elderly. The sequelae of such stroke diseases are very dangerous, because they make social and economic activities difficult. In this paper, we propose a new system to prediction and in-depth analysis stroke severity of elderly over 65 years based on the National Institutes of Health Stroke Scale (NIHSS). In addition, we use the algorithm of

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
IJARST
ISSN: 2457-0362

decision tree of C4.5, which is a methodology of prediction and analysis of machine learning techniques. The C4.5 decision trees are machine learning algorithms that provide additional in-depth rules of the execution mechanism and semantic interpretation analysis. Finally, in this paper, it is verified that the C4.5 decision tree algorithm can be used to classify and predict stroke severity, and to obtain additional NIHSS features reduction effects. Therefore, during the operation of an actual system, the proposed model uses only 13 features out of the 18 stroke scale features, including age, so that it can provide faster and more accurate service support. Experimental results show that the system enables this by reducing the patient NIH stroke scale measurement time and making the operation more efficient, with an overall accuracy, using the C4.5 decision tree algorithm, of 91.11%.

Saerom Park, Min-Ji Yang, So-Nyeong Ha, Jeong-Sang Lee : The societies of the world in the 21(st) century have faced challenges arising from an aging population as the fertility rate has dropped dramatically and medical advances have extended the average human life span. The elderly aged 65 years or older make up at least 20% of the population in Korea, making the country a super-aging society as defined by the United Nations. The number of elderly women is higher than that of elderly men and women live longer than men. Based on the analysis of recent trends in previous studies, this study aimed to suggest practical strategies to utilize isoflavones, substances chemically similar to the female hormone estrogen, and to search for effective anti-aging strategies using this substance for women to be prepared to reach the elderly stage in good health.

III. ARCHITECTURE DIAGRAM

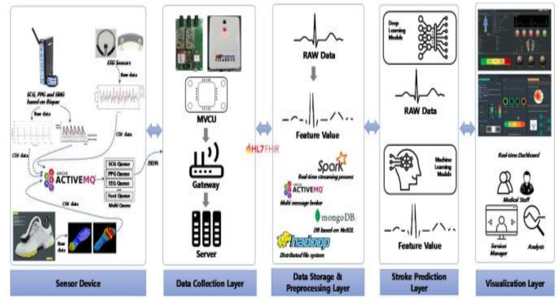

Fig 2 Proposed Architecture

In this research, we present a machine learning-based method for predicting and semantically interpreting stroke prognostic symptoms in the elderly utilising multi-modal bio-signals of electrocardiogram (ECG) and photoplethysmography (PPG) recorded in real-time. We devised and deployed a stroke disease prediction system with an ensemble structure that integrates CNN and LSTM to predict stroke illness in real-time while walking. The suggested system takes into account the ease of wearing bio-signal sensors for the elderly, and bio-signals were captured while walking at a sample rate of 1,000Hz per second from the three electrodes of the ECG and the index finger for PPG.

The benefits of this system is the real-time prediction of senior stroke patients shown good prediction accuracy and performance. It was experimentally shown that the prognostic symptoms of stroke patients may be predicted with more than 90% accuracy using just ECG and PPG obtained while walking.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
IJARST
ISSN: 2457-0362

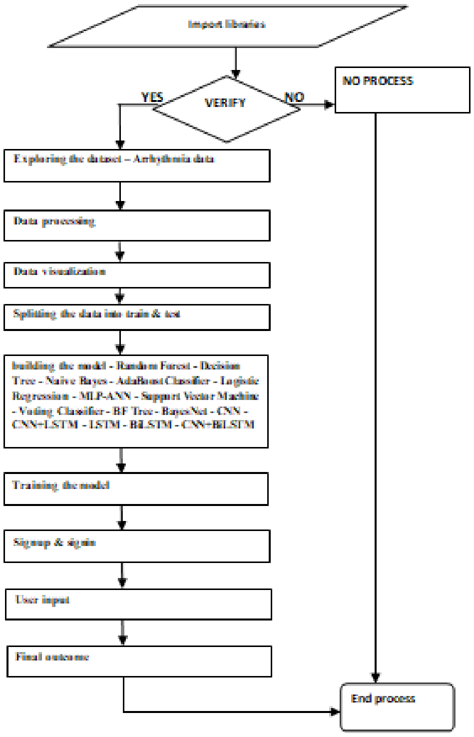

Fig 3 Work Flow of the proposed model

To carry out the aforementioned project, we created the modules listed below.

- Data exploration: we will put data into the system using this module.
- Processing: we will read data for processing using this module.
- Data splitting into train and test: Using this module, data will be split into train and test.
- Model construction: Random Forest, Decision Tree, Naive Bayes, AdaBoost Classifier, Logistic Regression, MLP-ANN, Support Vector Machine, Voting Classifier, BF Tree, BayesNet, CNN, CNN+LSTM, LSTM, BiLSTM, and CNN+BiLSTM. Calculated algorithm accuracy.
- User registration and login: Using this module will result in registration and login.
- Using this module will provide input for prediction.
- Prediction: final predicted shown

IV. IMPLEMENTATION

Using the ECG , PPG and Bio – Signals dataset, this code aims to predict the stroke risk disease.

The code starts by importing libraries. Now the dataset is visualizing and generating the modules in preprocessing method. After that dataset will split into training and test dataset. Now the model is building for algorithms and predicting the result for algorithms. The final result is shown in the graph.

Algorithms

Random Forest: A Supervised Machine Learning Algorithm that is commonly utilised in Classification and Regression applications. It constructs decision trees from several samples and uses their majority vote for classification and average for regression.

Decision Tree: Decision trees use numerous methods to determine whether or not to divide a node into two or more sub-nodes. The development of sub-nodes promotes the homogeneity of the sub-nodes that arise. In other words, the purity of the node rises in relation to the target variable.

Naive Bayes: A probabilistic classifier, the Naive Bayes classification technique. It is based on probability models with high independence assumptions. The independence assumptions often have little effect on reality. As a result, they are seen as naïve.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

AdaBoost Classifier: An AdaBoost classifier is a meta-estimator that starts by fitting a classifier on the original dataset and then fits additional copies of the classifier on the same dataset with the weights of incorrectly classified instances adjusted so that subsequent classifiers focus more on difficult cases.

Logistic Regression: Logistic regression is a statistical analytic approach that uses past observations of a data set to predict a binary result, such as yes or no. A logistic regression model forecasts a dependent variable by examining the connection between one or more existing independent variables.

MLP-ANN: A multilayer perceptron (MLP) is a kind of feedforward artificial neural network that is completely linked (ANN). The word MLP is used ambiguously, sometimes to refer to any feedforward ANN, and sometimes to networks built of many layers of perceptrons (with threshold activation); see Terminology. Multilayer perceptrons are commonly referred to as "vanilla" neural networks, particularly when just one hidden layer is present.

SVM: Support Vector Machine (SVM) is a supervised machine learning technique that may be used for both classification and regression. Though we call them regression issues, they are best suited for categorization. The SVM algorithm's goal is to identify a hyperplane in an N-dimensional space that clearly classifies the input points.

Voting classifier: A voting classifier is a machine learning estimator that trains numerous base models or estimators and predicts based on the results of each base estimator. Aggregating criteria may be coupled voting decisions for each estimator output.

BF Tree: The breadth-first search (BFS) method searches a tree or graph data structure for nodes that fulfil a set of criteria. It explores all nodes at the current depth level before going on to nodes at the next depth level, starting at the root of the tree or graph.

Bayesian Net:Bayesian networks are a sort of Probabilistic Graphical Model that may be used to create models based on data and/or expert opinion. They may be used for a variety of activities like as prediction, anomaly detection, diagnostics, automated insight, reasoning, time series prediction, and decision making in the face of ambiguity.

CNN: A CNN is a kind of network architecture for deep learning algorithms that is primarily utilised for image recognition and pixel data processing jobs. There are different forms of neural networks in deep learning, but CNNs are the network design of choice for identifying and recognising things.

LSTM: Long short-term memory (LSTM) is a kind of artificial neural network used in artificial intelligence and deep learning. Unlike traditional feedforward neural networks, LSTM has feedback connections. A recurrent neural network (RNN) of this kind may analyse not just single data points (such as photos), but also complete data sequences (such as speech or video).

BiLSTM: BiLSTM stands for Bidirectional Long Short-Term Memory (BiLSTM) In general, LSTM ignores future information in time series processing. BiLSTM processes series data in forward and reverse directions on the basis of LSTM, linking the two hidden layers.

V. PSEUDO CODE

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
IJARST
ISSN: 2457-0362

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from sklearn.preprocessing import
LabelEncoder, StandardScaler
from imblearn.over_sampling import
SMOTE
from sklearn.model_selection import
train_test_split
from sklearn.svm import SVC
from sklearn.naive_bayes import
GaussianNB
from sklearn.linear_model import
LogisticRegression
from sklearn.tree import
DecisionTreeClassifier
from sklearn.ensemble import
RandomForestClassifier
from lightgbm import LGBMClassifier
from xgboost import XGBClassifier
from sklearn.metrics import
classification_report,accuracy_score
,confusion_matrix
from sklearn.metrics import
auc,roc_auc_score,roc_curve,precisio
n_score,recall_score,f1_score
import time as timer
from sklearn.inspection import
permutation_importance
# import package
# open dataset
filename = "train_2v.csv"
data = pd.read_csv(filename)
with
pd.option_context('expand_frame_repr
', False):
    print(data.head())
print("Data shape:
{}".format(data.shape))
miss_val =
data.isnull().sum()/len(data)*100
print(miss_val)
print("# Missing values in variable
bmi\t\t:
{:.2f}%".format(miss_val['bmi']))
print("# Missing values in variable
smoking_status\t:
{:.2f}%".format(miss_val['smoking_st
atus']))
print("Data shape:
{}".format(data.shape))
# Safely disable new warning with
the chained assignment.
pd.options.mode.chained_assignment =
None  # default='warn'
# replace missing values in variable
'bmi' with its mean
data['bmi']=data['bmi'].fillna(data[
'bmi'].mean())
# remove (drop) data associated with
missing values in variable
'smoking_status'
clean_data =
data[data['smoking_status'].notnull(
)]
# drop variable 'id'
clean_data.drop(columns='id',axis=1,
inplace=True)
# validate there's no more missing
values
miss_val =
clean_data.isnull().sum()/len(clean_
data)*100
print(miss_val)
print("# Missing values in variable
'bmi'\t\t:
{}".format(miss_val['bmi']))
print("# Missing values in variable
'smoking_status'\t:
{}".format(miss_val['smoking_status'
]))
print("Shape of data without missing
values:
{}".format(clean_data.shape))
print("Unique 'gender':
{}".format(clean_data['gender'].uniq
ue()))
print("Unique 'ever_married':
{}".format(clean_data['ever_married'
].unique()))
print("Unique 'work_type':
{}".format(clean_data['work_type'].u
nique()))
print("Unique 'Residence_type':
{}".format(clean_data['Residence_typ
e'].unique()))
print("Unique 'smoking_status':
{}".format(clean_data['smoking_statu
s'].unique()))
# create encoder for each
categorical variable
label_gender = LabelEncoder()
label_married = LabelEncoder()
label_work = LabelEncoder()
label_residence = LabelEncoder()
label_smoking = LabelEncoder()
```

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
IJARST
ISSN: 2457-0362

```
clean_data['gender'] =
label_gender.fit_transform(clean_dat
a['gender'])
clean_data['ever_married'] =
label_married.fit_transform(clean_da
ta['ever_married'])
clean_data['work_type']=
label_work.fit_transform(clean_data[
'work_type'])
clean_data['Residence_type']=
label_residence.fit_transform(clean_
data['Residence_type'])
clean_data['smoking_status']=
label_smoking.fit_transform(clean_da
ta['smoking_status'])
with
pd.option_context('expand_frame_repr
', False):
    print(clean_data.head())
fig, ax =
plt.subplots(figsize=(8,6))
im = ax.matshow(clean_data.corr())
ax.set_xticks(np.arange(clean_data.s
hape[1]))
ax.set_yticks(np.arange(clean_data.s
hape[1]))
ax.set_xticklabels(clean_data.column
s,rotation=90)
ax.set_yticklabels(clean_data.column
s)
# Create colorbar
cbar = ax.figure.colorbar(im, ax=ax)
cbar.ax.set_ylabel("Correlation",
rotation=-90, va="bottom",
fontsize=12)
fig.tight_layout()
plt.show()
fig =
clean_data.hist(figsize=(10,8))
plt.tight_layout()
plt.show()
print("Classification report for
SVM:
\n{}".format(classification_report(y
_test,y_svm)))
print("Confusion matrix for SVM:
\n{}".format(confusion_matrix(y_test
,y_svm)))
print("Accuracy score for SVM:
{:.2f}".format(accuracy_score(y_test
,y_svm)))
# calculate precision, recall, and
f1 scores
prec_svm =
precision_score(y_test,y_svm)
rec_svm = recall_score(y_test,y_svm)
f1_svm = f1_score(y_test,y_svm)
print("Precision score for SVM:
{:.2f}".format(prec_svm))
print("Recall score for SVM:
{:.2f}".format(rec_svm))
print("F1 score for SVM:
{:.2f}".format(f1_svm))
# calculate sensitivity,
specificity, and auc
sens_svm,spec_svm =
calc_sens_spec(y_test,y_svm)
fpr, tpr, _ = roc_curve(y_test,
y_svm_prob[:,1])
auc_svm = roc_auc_score(y_test,
y_svm_prob[:,1])
print("Sensitivity score for SVM:
{:.2f}".format(sens_svm))
print("Specitivity score for SVM:
{:.2f}".format(spec_svm))
print("AUC score for SVM:
{:.2f}".format(auc_svm))
fig, ax = plt.subplots()
ax.plot(fpr, tpr, color='blue',
label='ROC curve (area = %0.2f)' %
auc_svm)
ax.plot([0, 1], [0, 1],
color='green', linestyle='--')
ax.set_xlim([-0.05, 1.0])
ax.set_ylim([0.0, 1.05])
ax.set_xlabel('False Positive Rate')
ax.set_ylabel('True Positive Rate')
ax.set_title('Receiver Operating
Characteristic (SVM)')
ax.legend(loc="lower right")
plt.show()
```

VI. EXPERIMENTAL RESULTS

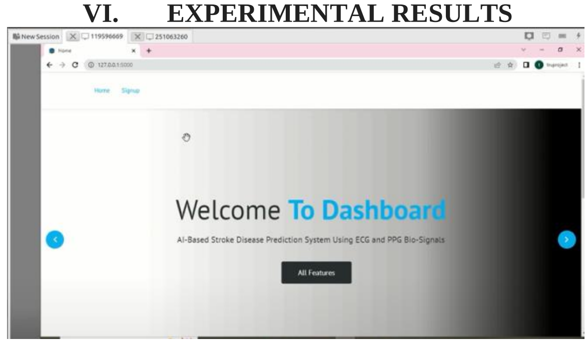

Fig.4: Home screen

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
IJARST
ISSN: 2457-0362

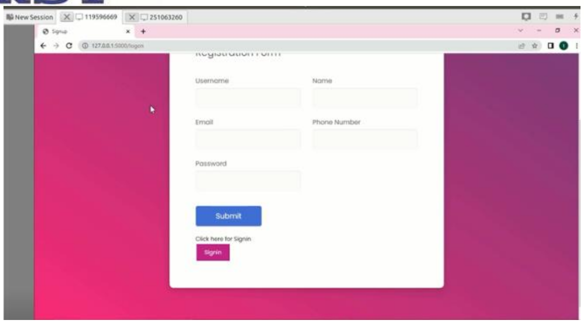

Fig.5: User registration

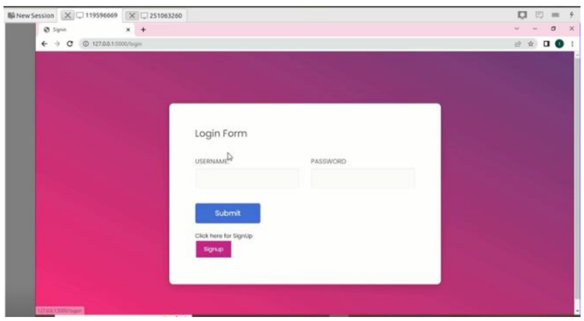

Fig.6: user login

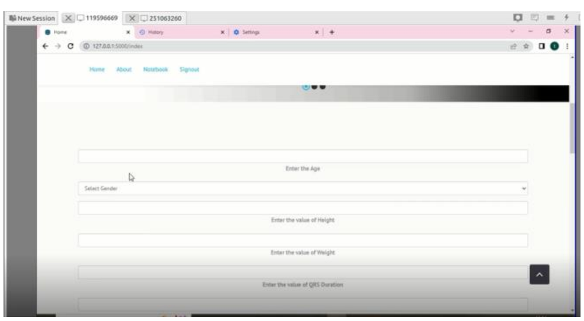

Fig.7: Main screen

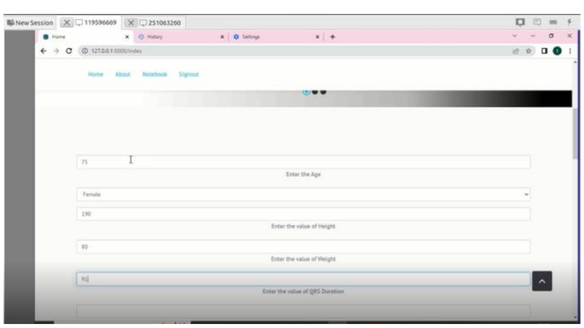

Fig.8: User input

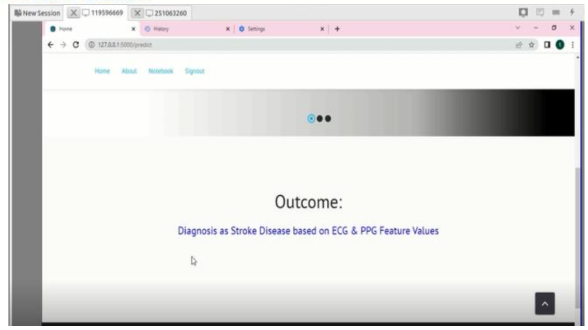

Fig.9: Prediction result

VII. CONCLUSION AND FUTURE SCOPE

In this research, we offer a system that enables semantic analysis of illnesses in the elderly by using numerous biological signals of ECG and PPG recorded while walking in the elderly's everyday lives. The proposed system captures numerous ECG and PPG biosignals in real time and may identify and predict prognostic signs of stroke illness in the elderly. Using numerous biosignal data, a machine learning-based prediction model research was undertaken, which included separating the signal waveform into particular portions, and reasonably accurate prediction results and semantic interpretations were produced using this model. In this work, it was experimentally demonstrated that utilising the suggested features, the prognostic symptoms of stroke patients may be reliably predicted by more than 90% based purely on ECG and PPG obtained while walking. To summarise the experimental and verification findings, we demonstrated that by splitting stroke and general elderly into 10-folder CV datasets, we can properly forecast 91.56% C4.5 Decision Tree, 97.51% RandomForest, and 99.15% CNN-LSTM models for deep learning. The method presented in this work has high academic value since it can reliably predict prognostic symptoms and the development of stroke by recording ECG and PPG at a reasonable cost and with minimal discomfort throughout everyday life. With a high repetition probability, various bio-signal data gathered in everyday life may give objective interpretation information to stroke patients or medical professionals. The trial findings demonstrated that this technology may be utilised for practical healthcare services such as reducing stroke aftereffects and preventing emergency situations via continuous monitoring. We will perform in-depth analyses and predicting experiments of stroke illness in the future by assessing numerous bio-signals such as EEG, EMG, foot pressure, and motion, as well as electronic medical records (EMRs) and MRI image data.

REFERENCES

[1] S. De Raedt, A. De Vos, and J. De Keyser, ''Autonomic dysfunction in acute ischemic stroke: An underexplored therapeutic area?'' J. Neurol. Sci., vol. 348, nos. 1–2, pp. 24–34, Jan. 2015.

[2] C. P. Warlow, ''Epidemiology of stroke,'' Lancet, vol. 352, pp. 1–4, Oct. 1998.

[3] K.-D. Seo, M. J. Kang, G. S. Kim, J. H. Lee, S. H. Suh, and K.-Y. Lee, ''National trends in clinical outcomes of endovascular therapy for ischemic stroke in South Korea between 2008 and 2016,'' J. Stroke, vol. 22, no. 3, pp. 412–415, Sep. 2020.

[4] T. D. Musuka, S. B. Wilton, M. Traboulsi, and M. D. Hill, ''Diagnosis and management of acute ischemic stroke: Speed is critical,'' Can. Med. Assoc. J., vol. 187, no. 12, pp. 887–893, Sep. 2015.

[5] Q. Song, X. Liu, W. Zhou, L. Wang, X. Zheng, X. Wang, and S. Wu, ''Long sleep duration and risk of ischemic stroke and hemorrhagic stroke: The Kailuan prospective study,'' Sci. Rep., vol. 6, no. 1, pp. 1–9, Sep. 2016.

[6] J. Yu, S. Park, H. Lee, C. S. Pyo, and Y. S. Lee, ''An elderly health monitoring system using machine learning and in-depth analysis techniques on the NIH stroke scale,'' Mathematics, vol. 8, no. 7, pp. 1–16, Jul. 2020.

[7] World Health Organization. The Top 10 Causes of Death. Accessed: Apr. 22, 2022. [Online]. Available: https://www.who.int/newsroom/fact-sheets/detail/the-top-10-causes-of-death

[8] UNDESA World Social Report 2020. Accessed: Apr. 22, 2022. [Online]. Available: https://www.un.org/development/desa/dspd/world-socialreport/2020-2.html

[9] S. Park, M. J. Yang, S. N. Ha, and J. S. Lee, ''Effective anti-aging strategies in an era of super-aging,'' J. Menopausal Med., vol. 20, no. 3, pp. 85–89, Dec. 2014.

[10] B. C. Meyer and P. D. Lyden, ''The modified national institutes of health stroke scale: Its time has come,'' Int. J. Stroke, vol. 4, no. 4, pp. 267–273, Aug. 2009.

[11] D. Schlegel, S. J. Kolb, J. M. Luciano, J. M. Tovar, B. L. Cucchiara, D. S. Liebeskind, and S. E. Kasner, ''Utility of the NIH stroke scale as a predictor of hospital disposition,'' Stroke, vol. 34, no. 1, pp. 134–137, Jan. 2003.

[12] 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.

[13] J. Yu, S. Park, S. H. Kwon, C. M. B. Ho, C. S. Pyo, and H. Lee, ''AIbased stroke disease prediction system using real-time electromyography signals,'' Appl. Sci., vol. 10, no. 19, pp. 1–19, Sep. 2020.

[14] Y. A. Choi, S. Park, J. A. Jun, C. M. B. Ho, C. S. Pyo, H. Lee, and J. Yu, ''Machine-learning-based elderly stroke monitoring system using electroencephalography vital signals,'' Appl. Sci., vol. 11, no. 4, pp. 1–18, Feb. 2021.

[15] M. Lee, J. Ryu, and D. Kim, ''Automated epileptic seizure waveform detection method based on the feature of the mean slope of wavelet coefficient counts using a hidden Markov model and EEG signals,'' ETRI J., vol. 42, no. 2, pp. 217–229, Apr. 2020.

[16] Y. A. Choi, S. J. Park, J. A. Jun, C. S. Pyo, K. H. Cho, H. S. Lee, et al., "Deep learning-based stroke disease prediction system using real-time bio signals", Sensors, vol. 21, no. 13, pp. 1-17, Jun. 2021.

[17] S.-M. Lai, P. W. Duncan and J. Keighley, "Prediction of functional outcome after stroke: Comparison of the orpington prognostic scale and the NIH stroke scale", Stroke, vol. 29, no. 9, pp. 1838-1842, Sep. 1998.

[18] S. Shrestha, R. S. Poudel, D. Khatiwada and L. Thapa, "Stroke subtype age and baseline NIHSS score predict ischemic stroke outcomes at 3 months: A preliminary study from Central Nepal", J. Multidisciplinary Healthcare, vol. 8, pp. 443-448, Oct. 2015.

[19] E. A. Sartor, K. Albright, A. K. Boehme, M. M. Morales, A. Shaban, J. C. Grotta, et al., "The NIHSS score and its components can predict cortical stroke", J. Neurol. Disorders Stroke, vol. 2, no. 1, pp. 73-78, Sep. 2013.

[20] R. G. Hart and J. L. Halperin, "Atrial fibrillation and stroke: Concepts and controversies", Stroke, vol. 32, no. 3, pp. 803-808, Mar. 2001.

[21] G. Y. H. Lip and H. S. Lim, "Atrial fibrillation and stroke prevention", Lancet Neurol., vol. 6, no. 11, pp. 981-993, 2007.

[22] J. A. Reiffel, "Atrial fibrillation and stroke: Epidemiology", Amer. J. Med., vol. 127, no. 4, pp. 15-16, Apr. 2014.

[23] P. A. Wolf, R. B. D'Agostino, A. J. Belanger and W. B. Kannel, "Probability of stroke: A risk profile from the Framingham study", Stroke, vol. 22, no. 3, pp. 312-318, Mar. 1991.

[24] S. H. Jee, I. Suh, I. S. Kim and L. J. Appel, "Smoking and atherosclerotic cardiovascular disease in men with low levels of serum cholesterol: The Korea medical insurance corporation study", Jama, vol. 282, no. 22, pp. 2149-2155, Dec. 1999.

[25] P. E. McSharry, G. D. Clifford, L. Tarassenko and L. A. Smith, "A dynamical model for generating synthetic electrocardiogram signals", IEEE Trans. Biomed. Eng., vol. 50, no. 3, pp. 289-294, Mar. 2003.