International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 03 | Mar 2020 www.irjet.net p-ISSN: 2395-0072

DIGITAL ASSISTANCE: A NEW IMPULSE ON STROKE PATIENT HEALTH CARE USING DEEP LEARNING

Narmada B[1], Abhineya A N[2],Dharsini B[3] ,Gayathri M S[4],Farzana Thuskin S E[5]

[1]Assistance Professor, Dept. of Computer Science and Engineering, Dhirajlal Gandhi College of Technology, Tamil Nadu, India
[2,3,4,5]Student, Dept. of Computer Science and Engineering, Dhirajlal Gandhi College of Technology, Tamil Nadu, India

---***---

Abstract - Stroke is one of the leading causes of death worldwide. Apart from risk factors like hypertension, diabetes, heart diseases, and positive family history, other lifestyle-related factors such as unhealthy diet, obesity, lack of physical activity, stress and tobacco use account for its occurrence. Stroke patients have longer hospital stays and higher readmission rates and medical costs than patients with other chronic diseases. However, the conventional predictive models or techniques are still not effective enough in capturing the underlying knowledge because it is incapable of simulating the complexity of feature representation of the medical problem domains. To overcome this process predict and diagnose stroke using a deep learning algorithm applied to heart disease datasets. Deep learning is a powerful tool to deal with large-scale data and provides an end-to-end solution to multi-class classification including dimensionality reduction. The outcomes of this process are more accurate than medical scoring systems currently in use for warning heart patients if they are likely to develop stroke. The deep learning technique is using Artificial Bee Colony (ABC) which is a new swarm-based optimization algorithm. It is used for feature extraction from a given dataset. Adding to this Neural Networks (NN) are used to classify the dataset.

***Key Words*: Stroke, Prediction, Deep Learning, Classification, Feature Extraction, Neural Network, Time Series Analysis, Datasets, Artificial Bee Colony (ABC).**

1. INTRODUCTION

Stroke is the leading cause of death and long term disability. Stroke occurs when an artery in your brain is blocked and leaks. It is not at all an unpredictable disease. People who receive treatment within 3 hours of having stroke are less likely to have disability. Taking this into consideration, the concept of artificial intelligence (AI) has recently permeated various sectors of life, including rapidly evolving healthcare systems. As electronic diagnoses, therapies, and record-keeping expand, it is essential to leverage, integrate, and optimize these advances. In the field of medicine, patient data are massively available in distributed electronic health record (EHR) databases and voluminous clinical, imaging, and laboratory datasets, among others. Such data can be utilized to predict disease incidence and prognosis.

Recent nationwide efforts seek to use big data to expand precision medicine to many other medical areas. Precision medicine is broadly defined as patient-specific diagnosis and therapy. Deep learning using big data has been employed to predict disease. Deep learning is actively used in many fields, yielding satisfactory results when conventional analyses are not appropriate. With these deep learning model we are also planned for developing a digital assistance which can be used by patient doctors and inmates of any emergency situation. This will allow healthcare policymakers to improve the quality of medical care, evaluate its appropriateness, and employ diagnostic resources efficiently. Our research was performed to predict stroke mortality using large-scale electronic health records. This study is expected to expand the research that can prescreen diverse diseases in e-health field in future.

1.1 Related Work

Several studies have used deep learning methods to solve various problems. In particular, there have been many computer-aided diagnosis systems using deep learning for detecting diverse diseases. Machine-learning/deep learning has been employed to detect or predict certain diseases using various approaches and datasets.

In recent years, there has been massive progress in artificial intelligence (AI) with the development of deep neural networks, natural language processing, computer vision and robotics. These techniques are now actively being applied in healthcare with many of the health service activities currently being delivered by clinicians and administrators predicted to be taken over by AI in the coming years. However, there has also been exceptional hype about the abilities of AI with a mistaken notion that AI will replace human clinicians altogether. These perspectives are inaccurate, and if a balanced perspective of the limitations and promise of AI is taken, one can gauge which parts of the health system AI can be integrated to make a meaningful impact. The four main areas where AI would have the most influence would be: patient administration, clinical decision support, patient monitoring and healthcare interventions. This health system where AI plays a central role could be termed an AI-enabled or AI-augmented health system. In this article, we discuss how this system can be developed based on a realistic assessment of current AI technologies and predicted developments.[3]

In order to support the stroke system of care in the age of big data, predictive analytics can be applied to forecast what will happen in the future. For example, we can predict post-stroke discharge disposition at an acute stroke admission, which will facilitate optimizing acute treatment and planning post-acute rehabilitation with the desired outcomes. In this preliminary study, we will explore deep learning for post-stroke discharge disposition prediction and evaluate prediction performance using hospital discharge data provided by Tennessee Department of Health. Deep learning is a powerful tool to deal with large-scale data and provides an end-to-end solution to multi-class classification including dimensionality reduction. Our preliminary results will demonstrate the effectiveness of deep learning and suggest the further exploration for performance improvement and clinical.[1]

Principal component analysis (PCA) featuring quantile scaling was used to extract relevant background features from medical records; we used these to predict stroke. We compared our method (a scaled PCA/deep neural network [DNN] approach) to five other machine-learning methods. The area under the curve (AUC) value of our method was 83.48%; hence; it can be used by both patients and doctors to prescreen for possible stroke.[6]

Stroke data analysis through a HVN visual mining platform. The visualization platform uses a hierarchical clustering algorithm to aggregate the data and map coherent groups of data-points to the same visual elements. The stroke data is collected from an open source Healthcare Dataset stroke data at Kaggle.com. Data dictionary is the data format used for this analysis. By using this data hierarchical virtual node is developed. The result of this research outcome is used as a new HVN visual data mining platform for interactively and transparently analyzing the stroke data.[4]

2. PROPOSED WORK

2.1. *Subject*

We have used data from Kaggle.com. The data has been collected from various hospitals around the world like Hungarian Institute of Cardiology. The subjects were about 343 stroke patients. The database used contains 76 attributes, but all published experiments refer to using a subset of 14 of them. The "goal" field refers to the presence of heart disease in the patient. It is integer valued from 0(no presence) to 4.Analysis shows presence (values 1,2,3,4) from absence(value 0).The mean age of the patients were 57 years. Of the patients 68.32% were male, 31.68% were female.

Table -1: Distribution of subjects by general character

Variables	N(%)
Mean age	57%
Gender Male Female	 68.32% 31.68%

2.2. *Principle Variables*

The dataset characteristics is Multivariate, the attributes used are of type categorical, integer and real. The dependent variable is associated with the mortality rate of stroke patients. Independent variables that reflected the social status were age and gender. Medical variables included thalassemia, chest pain, resting blood pressure, fasting blood sugar, cholesterol level, resting electrocardiographic, exercise induced angina, ST depression, the number of major vessels. Admission mode was emergency, ambulatory and others. The stroke type were classified into ischemic and hemorrhagic. Chart (1) shows the probability for disease based on the age attribute.

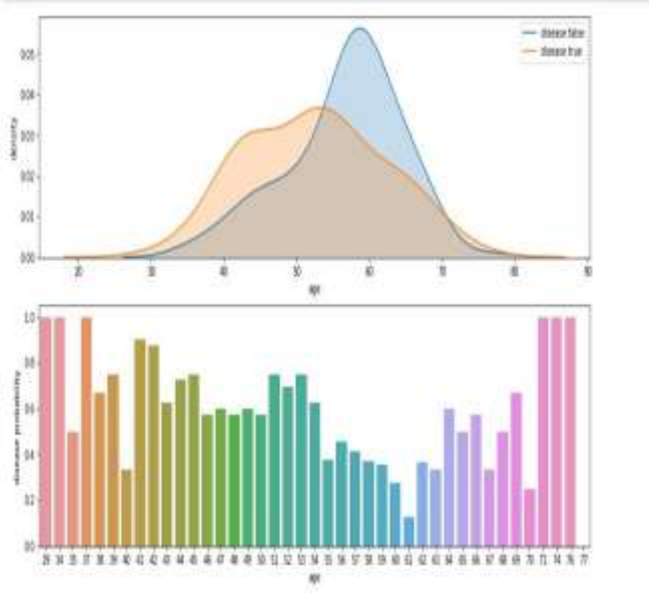

Chart -1: Probability of Stroke

2.3. *Methods*

Our deep learning model consists of the following attributes: Age, gender, included thalassemia, chest pain, resting blood pressure, fasting blood sugar, cholesterol level, resting electrocardiographic, exercise induced angina, ST depression, the number of major vessels. We used DNN and Artificial Bee Colony Algorithm (ABC) to automatically generate features from the data and identify risk factors for stroke. Figure 1 shows the system architecture: The dataset we used included 13 variables, where the data were classified into training and testing data. We used both ABC and scaler to convert categorical variable into continuous variable, and to generate models for testing. We then trained the DNN using ABC variables and compared to predict the results. The training and testing data did not overlap with one another.

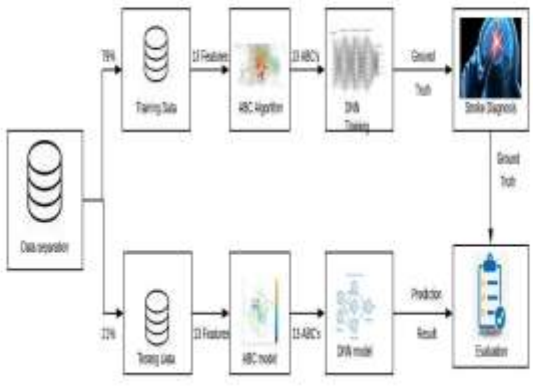

Fig- 1: System Architecture for ABC and DNN

2.4. *Preprocessing*

Artificial Bee Colony (ABC) is an optimization algorithm that adapts the foraging behavior of honey bees. ABC is one of the swarm intelligence based algorithm. This technique is being widely used in various problems in-order to extract meaningful insights from huge datasets. In general, ABC preprocessing efficiently defines new features, reducing dimensions to find hidden or simplified structures for inclusion in classification algorithms. Graph (2) depicts the type of angina and the level of fasting blood sugar level (>or<120 mg) based on the count of stroke patients.

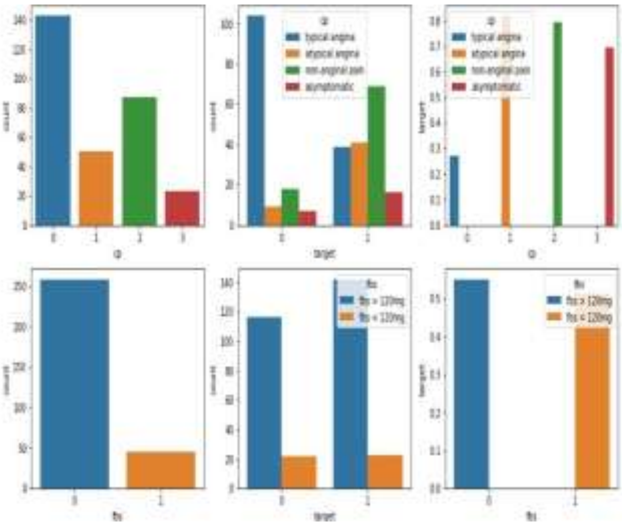

Chart -2: Plotting of angina type and fasting blood sugar level

2.5. *NN Architecture*

A simple feed-forward neural networks is used in our deep (four hidden layers) learning models. Training of datasets is carried out by following a standard backpropagation algorithm. For each DNN, we adjusted several hyper parameters, including the number of hidden layers, the number of neurons in each layer, the activation function, the optimization method, and the regularization technique.

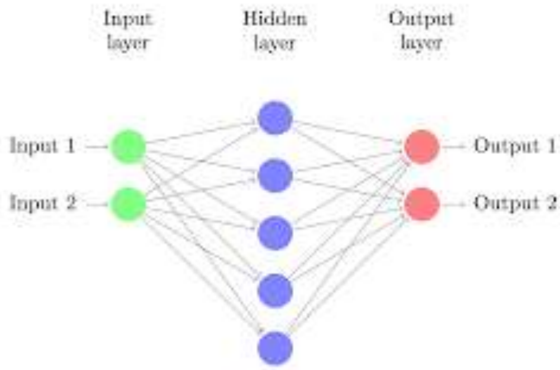

Fig -2: Neural Network (NN) Architecture

All models were implemented using **Anaconda** with a **Jupyter** notebook backend. Figure (3) shows the accuracy score achieved while comparing the training score and the cross validation score from our testing data used.

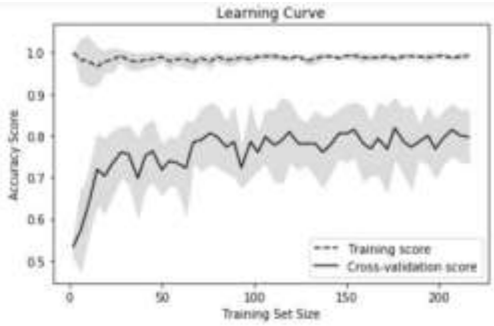

Fig -3: Learning Curve

2.6. *Result and Discussion*

This technique is suitable to use predictive neural networks or characteristic data as infectio event or non-event binomial effects. Figure (4) summarize the correlation values of all important attributes used.

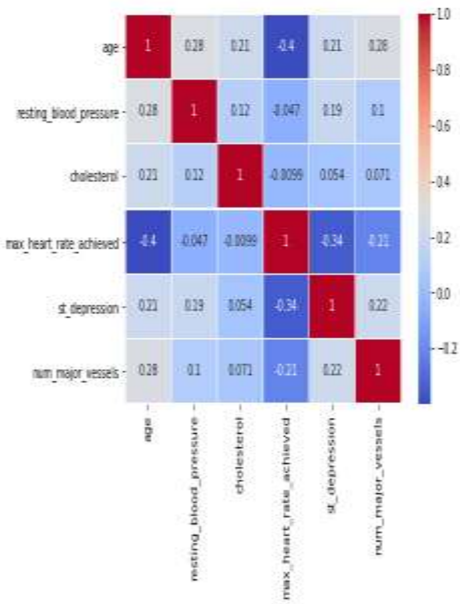

Fig -4: Correlation Values

Table (2) shows the comparison of various algorithms by considering the factors precision, recall value, false negative (FN), true positive (TP) and false negative rate (FNR).

Table -2: Comparison of Various Algorithm

PR RE FS FN TP FNR
LR 0.857 0.882 0.869 22 30 11.764
RF 0.861 0.911 0.886 22 31 8.823
NB 0.837 0.911 0.873 21 31 8.823
KNN 0.741 0.676 0.707 19 23 32.353
DT 0.870 0.794 0.830 23 27 20.518

LR, Logistic regression; RF, Random forest; NB, Naive Bayes; KNN , K-Nearest Neighbour algorithm; DT, Decision Tree.

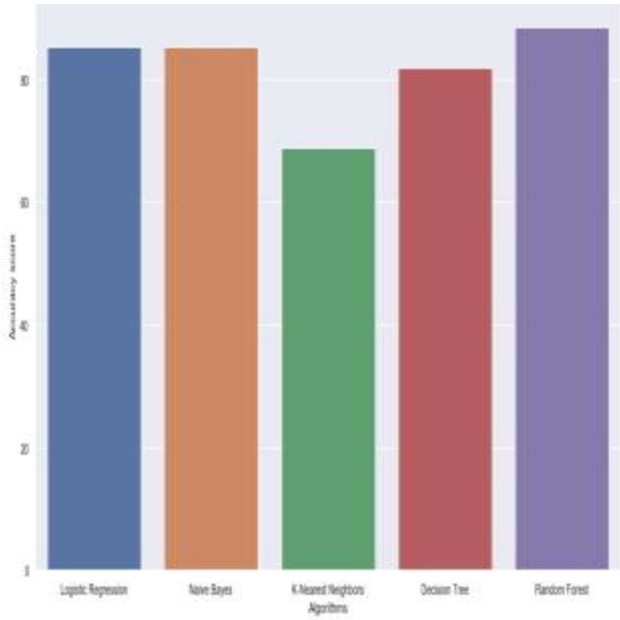

Chart -3: Accuracy Comparison

From chart (3), the algorithm with maximum accuracy score for prediction is Random Forest having more than 80% of accuracy.

Table -3: Accuracy of Different Algorithms in probability

Algorithms Accuracy
KNN 0.688525
Decision trees 0.819672
Logistic Regression 0.852459
Naive Bayes 0.852459
Random Forest 0.885246

We considered tuning hyper parameters, such as the number of nodes and depth of the DNN to improve stroke detection. Our method predicts stroke using indirect or limited data, such as medical service use history and health behavior. Our results, apply only to patients who are suffering from stroke and the patients receiving treatment are not considered, which may affect the overall accuracy of our model.

3. CONCLUSIONS

Stroke is the leading cause of death and long-term disability in the world. We are going to develop a deep learning model that predict the stroke based on the medical utilization history and health behavior. Our work allow early detection of patient at the high risk of stroke who needs additional checkups and appropriate treatments prior disease to exacerbation. As the input data are simple (albeit of low resolution, that is, binary or with a limited number of choices), we used a DNN to study the variables of interest and scaled PCA to generate improved continuous inputs for the DNN. The sensitivity, specificity and AUC value of our method were 64.32%, 85.56% and 83.48%, respectively. Our method can be used not only to predict stroke using limited data, but also other diseases.

In future, we will modify and apply our method for the analysis of other medical service use and health behavior datasets on conditions such as dementia. We will also use detailed indices and physiological signals as input data to achieve more meaningful DNN results. Finally, we will employ auto-fine tuning methods to reduce training time and improve performance.

REFERENCES

[1] Jin Cho*, Zhen Hu*, and Mina Sartipi* .Post-stroke Discharge Disposition Prediction using Deep Learning,2019.

[2] Shuo Tian [1] , Wenbo Yang [1] , Jehane Michael Le Grange , Peng Wang , Wei Huang * , Zhewei Ye. Smart healthcare: making medical care more intelligent, 2019.

[3] Sandeep Reddy[1], John Fox[2] and Maulik P Purohit[3] ·Artificial intelligence-enabled healthcare delivery ,2019.

[4] CERNER Kansas City INNOVATION , KS (US) , INC . , (72) Inventor : DOWNINGTOWN RICHARD MATTHEW , PA (US BALIAN).ASSISTANT VIRTUAL HEALTHCARE PERSONAL,2018.

[5] Suchitra Kataria[1] · Vinod Ravindran[2].Digital health: a new dimension in rheumatology patient care, 2018.

[6] Songhee Cheon [1], Jungyoon Kim [2] and Jihye Lim [3,*]The Use of Deep Learning to Predict Stroke Patient Mortality, 2019.