**STROKE RATE PREDICTION USING DEEP NEURAL NETWORK: A COMPARATIVE STUDY**

**Submitted by**

**1. Name: Md. Redowan Chowdhury**

**(ID: CSE04180301288)**

**2. Name: Md. Abu Raihan**

**(ID: CSE04180101225)**

**3. Name: Raiyan Bin Noor**

**(ID: CSE04180301281)**

A Project Report Submitted in Partial Fulfillment of the Requirements for the Degree of Bachelor of Science in Computer Science & Engineering



DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

NOTHERN UNIVERSITY BANGLADESH

July 2022

APPROVAL

The Project Report “**STROKE RATE PREDICTION USING DEEP NEURAL NETWORK: A COMPARATIVE STUDY**” submitted by **Md. Redowan Chowdhury** (ID: **CSE04180301288**), **Md. Abu Raihan**(ID: **CSE04180101225**) and **Raiyan Bin Noor** (ID: **CSE04180301281)** to the Department of Computer Science and Engineering, Northern University Bangladesh, has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of Bachelor of Science in Computer Science and Engineering and approved as to its style and contents.

## Board of Examiners

1. Nazmin Islam (Supervisor)

1. (Examiner)
2. (Examiner)

---------------------------------------

Md. Raihan Ul Masood

Associate Professor and Head

Department of Computer Science and Engineering

Northern University Bangladesh

# DECLARATION

We, hereby, declare that the work presented in this Project report is the outcome of the investigation performed by us under the supervision of Mr. Md. Ruhul Amin, Assistant Professor, Department of Computer Science and Engineering, University of Dhaka. We also declare that no part of this Project has been or is being submitted elsewhere for the award of any degree or diploma.

Signature

..............................................

Member 1: Md. Redowan Chowdhury

(ID:CSE04180301288)

.............................................

Member 2: Raiyan Bin Noor

(ID:CSE04180301281)

.............................................

Member 3: Md. Abu Raihan

(ID: CSE CSE04180101225)

**Candidates**

# Table of Contents

Contents

ABSTRACT

ACKNOWLEGDEMENT

1. INTRODUCTION
   1. Background of Study
   2. Specific Objects
   3. Assumption of Study
2. LITERATURE RIVIEW
3. METHODOLOGY
   1. Deep Neural Network
4. PREPROCESSING
   1. Performance Estimation of Activation Functions
   2. Considering Parameters
   3. Environment
5. RESULT
6. DISCUSSION
7. CONCLUSION

REFERENCES

**ABSTRACT**

Due to data science's swift development and the expanding amount of research being done on it, the number of applications has been growing quickly in recent years. Medicine is one of the most benefited areas and data-based decisions are being trusted increasingly because of its efficiency and accuracy, where the decisions made by the medical staff are vital for patients [1]. A blood artery that is either blocked by a clot or bursts causes a stroke. A portion of the brain cannot receive the necessary blood when the clot or bursts happen, therefore the blood cells perish [2]. A variety of diseases can be correctly predicted by using electronic medical claims (EMCs), which can help with targeted medical interventions [3]. Given the high prevalence of stroke, prompt identification and treatment are essential [4]. The goal of this research article is to use Data Analytics and Machine Learning to create a model capable of predicting Stroke outcome based on an unbalanced dataset containing information about 5000 individuals whose Stroke outcome is known [5].

Keywords—Dataset, Data Science, disease prediction, Machine Learning, Stroke, Machine Learning, Decision Tree

**ACKNOWLEDGEMENTS**

First of all, we would like to thank the Almighty ALLAH. Today we are successful in completing our work with such ease because He gave us the ability, chance, and a cooperating supervisor.

We are indebted to a number of individuals in academic circles as well as in university faculties who have contributed to prepare the book. Their contributions are important in so many different ways that we find it difficult to acknowledge them in any other manner but alphabetically. In particular, we wish to extend our appreciation to our respected supervisor **Nazmin Islam**, Lecturer, Dept. of Computer Science & Engineering, Northern University Bangladesh, our honorable Coordinator, Faculty of Science **…………………,** and other respected faculties of Science Faculty Northern University Bangladesh, for their valuable suggestions on preparing and improving the presentation and the book.

We also want to convey our thanks to our fellow friends for their great encouragement and suggestions. We are thankful to Jon W. Mark and Weihua Zhuang for their great book “Wireless Communications and Networking”.

Again, we would like to give thanks to our honorable supervisor **Nazmin Islam** for his commitment to excellence in all aspect of the production of this book and completion of the work.

Last of all, we are grateful to our family members; who are, always with us in our every step of life.

***Dedicated to Our Parents***

**Chapter I**

**Introduction**

* 1. **Background of Study**

A stroke, often referred to as a brain attack, occurs when blood supply to the brain is interrupted, depriving it of oxygen and nutrients. As a result, brain cells start to degenerate within minutes [1]. After ischemic heart disease, it is the second leading cause of mortality globally, according to the World Health Organization (WHO) [2]. Paralysis, sluggishness, or visual loss are all possible symptoms of stroke.

* 1. **Specific Objects**

While certain stroke risk factors, such as age, gender, race, and family history of cerebrovascular illnesses, cannot be changed, others may and are thought to be responsible for between 60% and 80% of all stroke risks in the general population [3]. There were 6.55 million stroke-related fatalities and 12.2 million event cases in 2021 [4]. A stroke happens when a blood artery in the brain bursts and bleeds, preventing blood and oxygen from reaching the brain's tissues. Ischemic stroke (IS), transient ischemic stroke (TIA), and hemorrhagic stroke are three different forms of strokes (HE). The most frequent kind of stroke is ischemic. 87 percent of strokes are ischemic. Mine strokes, which account for 10 to 15 percent of strokes, are transient ischemic attacks. TIA Within 23 hours, the symptoms of a temporary obstruction will go away. TIA is an indication of a potential stroke. A brain artery may rupture and cause a hemorrhagic stroke.

* 1. **Assumption of Study**

We predicted the stroke in our model using a machine learning method. The patient may get medical care and reduce their risk of stroke with the aid of early stroke prediction. Less persons had strokes and died from them than had strokes and were still alive. In order to extract valuable information from the vast quantity of data, strong data tools are required. Machine learning is used to forecast illness in the healthcare industry. where patient information such as name, age, blood pressure, blood sugar, etc. are kept. Multiple characteristics are used by classification algorithms to identify the illness. The values will be properly predicted using machine learning. We may use a variety of machine learning methods, and we'll choose the one that will provide the highest level of accuracy.

**Chapter II**

# LITERATURE RIVIEW

To get the necessary understanding of numerous ideas connected to the current study of existing literature were examined. At some of the crucial from them, conclusions were drawn are mentioned here, and just a few scientists worked on Machine Learning for Stroke Prediction Some of them from recent years are detailed here.

In [1], five machine learning approaches were used to the Cardiovascular Health Study (CHS) dataset to predict strokes. The authors used a mix of the Decision Tree with the C4.5 method, Principal Component Analysis, Artificial Neural Networks, and Support Vector Machine to determine the ideal solution. However, the CHS Dataset used for this study contained fewer input parameters.

In [2], people's social media postings were used to predict strokes. The DRFS approach was used by the researchers in this study to identify the different signs and symptoms of stroke. Using Natural Language Processing (NLP) to extract text from social media postings increases the model's execution time, which is not ideal.

In [3], the authors used an adapted random forest algorithm to handle the job of stroke prediction. Analysing stroke risk levels was done this way. This strategy, according to the authors, outperformed the competition in terms of speed and accuracy. Only a small number of strokes can be studied in this way, and it cannot be applied to any additional strokes in the future.

Stroke prediction model developed using Decision Tree, Random Forest, and Multi-layer Perceptron, according to research article [4] The three approaches yielded similar results, with very minor changes in accuracies. Decision Tree had a calculated accuracy of 74.31%, Random Forest had a calculated accuracy of 74.53%, and Multi-layer perceptron had a calculated accuracy of 75.02%. According to this study, the Multi-layer Perceptron approach is the most accurate of the three. Using just one metric to measure performance, the accuracy score cannot always offer good results.

On the Cardiovascular Health Study (CHS) dataset, the researchers conducted stroke prediction in [5]. On the basis of their suggested conservative mean, they introduced a unique automated feature selection technique that chooses robust features. For further effectiveness, they paired this approach with the Support Vector Machine technique. However, this led to the creation of a number of vectors that have the tendency to make the model perform worse.

The prediction of thromboembolic stroke pathology using artificial neural networks is suggested by research in [6]. The Back-propagation algorithm was employed as the prediction approach. An accuracy of around 89 percent might be obtained with this approach. However, due of their complicated structure and growing neuronal population, neural networks take longer to train and analyse information.

Computer Techniques and Applications in Biomedicine - Bora Yoo, Kyunghee Cho, Dong-wook Kim, Soon-ae Shin, Jae-woo Lee, Hyunsun Lim, and This paper's objectives included calculating the 10-year stroke prediction probability and categorizing each user's personal stroke risk into five groups.

Stroke Risk Profile from the Framingham Study: Probability of Stroke - Albert J. Belanger, William B.Kannel, MD, Philip A. Wolf, DO, Ralph B. D'Agostino, PhD This study's Framingham Study cohort was used to build a health risk evaluation function for the prediction of stroke.

According to Tasfia Ismail Shoily et alcomparison .'s of the Naive Bayes, J48, k-NN, and Random Forest models, the former has higher accuracy. The dataset was compiled from a variety of medical records and cross-referenced by medical professionals using WEKA (Waikato Environment for Knowledge Analysis). The proposed model will assist patients in determining whether they are at risk of having a stroke or not. 4 distinct models, including Naive Bayes, J48, k-NN, and Random Forest, were trained. To verify the models, precision and accuracy were seen. The machine learning models are applied on the dataset [7].

We may learn more about potential limitations caused by stroke using Jaehak Yu et alfavored's C4.5 decision tree method, which employs the NIHSS score and real-time variables to classify stroke severity into four categories. This information aids in predicting the potential timing of a stroke and its associated handicap, allowing for the administration of additional drugs and the essential safety measures. Random Forest has a high accuracy of 88.9 percent, whereas Naive Bias has an accuracy rate of 85.4 percent [8].

A Bayesian model known as Bayesian Rule Lists (BRL) was predicted by Benjamin Letham et al., and it builds a distribution of permutations from a huge, processed collection of data. The approach scales with the least amount of the data set with numerous characteristics since the pre-processed data minimizes the model space for different sets of fragments. High accuracy, precision, and tractability may be attained with the aid of the BRL approach [9].

Pei-Wen Huang1 et al. used physiological data to predict stroke using the multimodal analysis approach. This information includes photoplethysmography, arterial blood pressure, and electrocardiography (EKG) (PPG). Each of these signals has been examined for accuracy. The three signals were combined, and they claimed that multi model analysis provides a greater accuracy for stroke prediction [10].

The information gathered from Sugam Multispeciality Hospital was used by Govindarajan et al. [11]. The dataset includes 507 patient records and 22 distinct class labels for the two main kinds of strokes. They used Decision Tree, Logistic Regression, Bagging, and Boosting, as well as Artificial Neural Networks (ANN), Support Vector Machines (SVM), and Decision Trees.

Data gathered from the physical therapy faculty was used by Singh and Choudhary [12]. It provides details about cardiovascular health research (CHS). 5,888 samples make up the dataset. The information was divided into three distinct medical terminologies: stroke and claudication, stroke and transient ischemic attack, and stroke and angioplasty. More than 600 qualities are also included. For dimensionality reduction, they employed principal component analysis. The feature selection algorithm is C4.5. They achieved 95, 95, and 97.7 percent accuracy with the use of ANN.

Clinical data from 739 patients' ischemic strokes were examined by Sung et al. in their study (Sung et al., 2013). The accuracy of the machine learning algorithm used to predict END will be evaluated using a total of 17 clinical variables, including past TIA history, risk factors for vascular illnesses, patient demographic data, stroke subtypes, and neuroimaging characteristics. They used four machine learning algorithms to validate their results: Bootstrap decision forest, deep neural network, boosted trees, and logistic regression. The accuracy score derived from the model is 0.966, 0.966, 0.966, and 0.946. Boosted Tree method has the greatest area under curve value (0.934) and accuracy (0.966) of all algorithms.

For the categorization of strokes, Sudha et al. [14] employed a Bayesian classifier, a decision tree, and a neural network. The medical institution provides the stroke dataset. Their patient's details and history make up their dataset are designed to be error-free. There are 1000 entries in the dataset. They had utilized PCA to lessen the dimensionality. With 92 percent for the neural network, 91 percent for the naive Bayes classifier, and 94 percent for the decision tree, they have the highest accuracy.

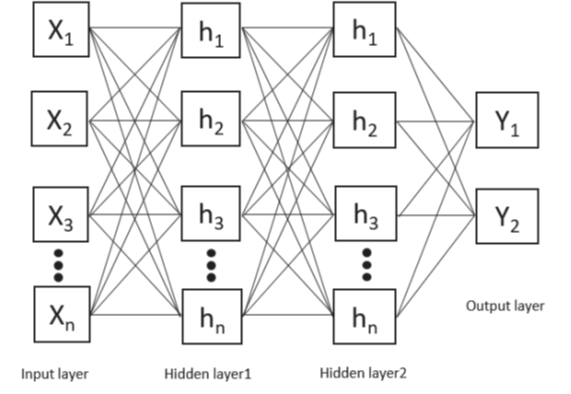
**Chapter III**

**METHODOLOGY**

The Deep Neural Network (DNN) approach, which is shown below, was employed in this suggested system to predict strokes using various activation functions.

* 1. **Deep Neural Network**

DNN has become as one of the most debated and studied topics in the modern world. This artificial neural algorithm, which was inspired by the human brain, is in charge in several fields. This model starts with multiple layers of weighted input and produces output. The relevance of the input data is dominated by weight [13]. For initializing the weight, Glorot uniform initializer has been introduced. Iterating through the data during forward propagation results in a cost function that specifies the distinction between genuine input and artificial intelligence. Cycle relapse modification of the weights, sometimes referred to as backpropagation, is used to reduce the loss Gradient batch descent.



*Figure 1: Deep Neural Network Structure*

**Chapter IV**

**PREPROCESSING**

* 1. **Performance Estimation of Activation Functions**

The output of a node is described by an input or group of inputs in its activation function [15]. The output of a neural network model is derived by a numerical identification. Each neuron in the model is subjected to the function, which determines whether to activate it or not by computing a weighted sum depending on whether the input from the neuron is relevant to the system's forecast. The activation function's objective is to introduce irregularity into the output. Five activation functions, including the hyperbolic tangent, sigmoid, soft plus, rectified linear unit, and swish, were assigned to the DNN model in this study. The investigated outcome shows that the swish activation function, with 10,000 epochs and 0.001 learning rate, reveals the maximum accuracy. With the same epochs and learning rate, the hyperbolic tangent (tanh) and sigmoid activation function exhibit the lowest accuracy, which is 67.929 percent. Table 1 compares the effectiveness of several activation functions as Table 1

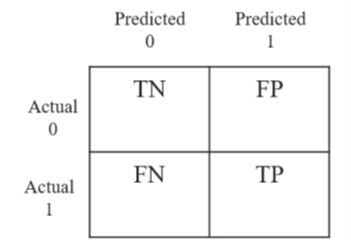
*Table 1: Comparison of Activation Functions*

|  |  |  |
| --- | --- | --- |
| **Activation Function** | | **Accuracy (for 10,000 epochs and 0.001 learning rate)** |
| Hidden Layer | **Output Layer** |
| Hyperbolic Taangent(tanh) | Hyperbolic Taangent(tanh) | 69.34% |
| Sigmoid | Sigmoid | 67.92% |
| Softplus | Softplus | 97.05% |
| Rectified Linear Unit (relu) | Rectified Linear Unit (relu) | 67.24% |
| Swish | Swish | 79.72% |

* 1. **Considering Parameters**

The mapping of formal and actual parameters was used to evaluate the stroke dataset using the evaluation parameter. A confusion matrix can be seen as a table that shows how a categorization model operates. Based on the test dataset, the true values are examined as Table 2.

Table 2: Confusion Matrix



It offers metrics parameters for recall, F-measure, accuracy, and precision. Traditionally, these variables provide the final performance evaluation result [16].



* 1. **Environment**

The highest priority for medical technology that protects people's health and lives is accurately prescribing patients' conditions. For this, it is anticipated that prediction accuracy will fit into the model. A python-based open source program containing pandas, matplotlib, numpy, sci-kit learn, and other tools were introduced as a platform for model configuration in order to create the best environment for training the dataset. For faster and simpler network development, TensorFlow has been employed for mathematical computing. PowerPoint software has been utilized for design and evolution. MathType was assigned to evaluate parameters and mathematical expressions. The Google Collaboratory was used to run the entire programming.

**Chapter V**

**RESULT**

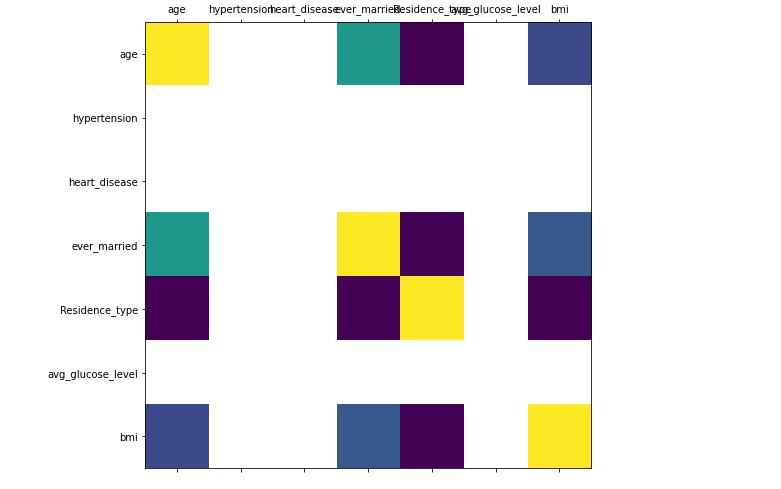
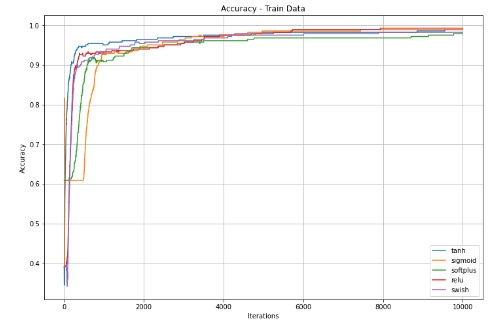
****According to Deep Neural Network, using this model would allow for higher precision and more accuracy. The classification and diagnostic accuracy with swish activation function as hidden layer and output layer were shown to be more acceptable than others through this investigation. Its accuracy of 97.05 percent was higher than that of earlier studies [17]. The correlation between the input columns is shown in Figure 2.

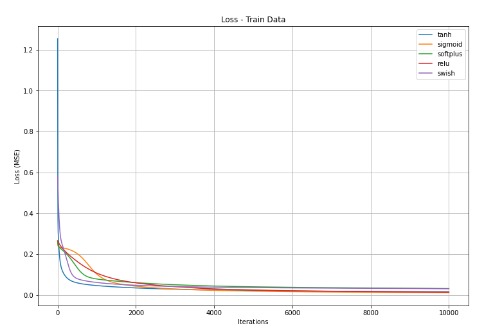
Figure 2: Correlation Between Inputs

**The obtained confusion matrix, which illustrates how well the model performed with this accuracy obtained from 0.001 learning rate and 50,000 epochs, is**

**Accuracy, precision, F1-measure, and recall assessment parameters were calculated to have values of 0.929, 0.9318, 0.953, and 0.976, respectively.**

**The accuracy and loss curve of the model are shown in Figure 3 and Figure 4.**

***Figure 3: Accuracy Curve for The Model***

****

***Figure 4: Loss Curve for The Model***

**Chapter VI**

**DISCUSSION**

In Bangladesh, stroke prevalence among the elderly has been exceptionally high. As a result of regular exposure to risk factors including hypertension [20], Bangladesh will soon face a significant challenge in the prevention and treatment of stroke. Between the two classes, the study's data were very unbalanced (stroke vs. non-stroke), As might be predicted, machine learning techniques performed poorly with the unbalanced data set. Meanwhile, the combined effectiveness of various prediction techniques increased. The data balancing process, demonstrating that it is dangerous to do prediction using inaccurate data. Using data balancing methods, excessively unbalanced classes might be successfully remedied. Avoidance is essential for precise prediction [19]. In the balanced ROS, RUS, and SMOTE/SVM data, the AUC for RLR, SVM, and RF compared to that in the unbalanced data set, respectively, sets improved significantly. The emphasis here is the significance of data balancing methods. only in the unbalanced data set's AUC for RF and Compared to RLR, the SVM's AUC in the ROS-balanced data set exhibited improvement. The remaining models and RLR were found to vary significantly. As shown by demonstrated the application context greatly influenced how well machine learning techniques performed [21]. RLR, a traditional machine learning approach, demonstrated impressive results given its simplicity and versatility [22]. excellent results in our research. SVM is another effective machine learning technique, and it also shown good performance in our study's ROS-balanced data set. RF is an example of a representation. when it comes to ensemble learning, which primarily combines the output from many classifiers to produce more precise projections [18]. RF functioned despite the fact that the initial data set was highly unbalanced. improved over RLR. Our study's key factors were consistent with those previously reported in other studies. The three machine learning models all included age, hypertension, and bmi as common characteristics methods. Age was recognized as the most significant stroke risk factor in our research into demographic factors, as shown by a prior research [23]. An essential factor was hypertension. predictor [24, 25], and considering how common hypertension is, prevention of it is therefore a crucial endeavour. and detrimental impact on stroke in Bangladesh. Bmi was discovered to be a standalone predictor of early mortality in stroke victims [26]. An essential component shared by RLR and RF was avg glucose. Previous research revealed demonstrated a deficit in the management of Avg glucose was a crucial indicator of the prognosis of stroke [27]. The following are some benefits of this research.

Our research also had a number of drawbacks. The study's outcome variable was self-reported stroke; As a result, there can be some personal prejudice. Additionally, the population's access to data is restricted, and included in our analysis was not sufficiently big. At the same time, over 50% of the participants were dropped. Considering the high percentage of missing outcome and predictive factors, this might possibly have some unpredictability in our findings. Furthermore, as data balancing methods have advanced, more approaches are developing, but we just covered the three that are most often utilized (ROS, SMOTE/SVM, and RUS) in this investigation. Last but not least, we merely carried out internal validation of our procedures, and future research need external validation in big populations.

**Chapter VII**

**CONCLUSION**

Deep neural networks are advantageous when detection or prediction with the best outcomes is the main objective in a changing environment. Stroke may be challenging to predict due to the intricacy of the symptoms. As a result, the study's main goal is to use artificial intelligence to deliver an effective and quick therapy. For this unbalanced system, a deeper neural network prediction technique has been devised. The suggested strategy had a 97.05 percent success rate in producing the intended result. The use of this algorithm to forecast strokes in the medical field will be extremely advantageous to doctors. This study on stroke prediction is generally superior to prior ones. Further analysis of this model and comparisons with other sickness prediction algorithms will be conducted. The model's low loss and high accuracy make it more useful and accurate than others for predicting the likelihood of a stroke occurring early.

**REFERENCES**

1. NHLBI, “What Is a Stroke?” 2014. https://web.archive.org/web/20150218230259/http://www.nhlbi.nih.gov/health/health-topics/topics/stroke/# (accessed Jun. 06, 2021).
2. WHO, “The top 10 causes of death.” https://www.who.int/newsroom/fact-sheets/detail/the-top-10-causes-of-death (accessed Jun. 09, 2021).
3. M. A. Moskowitz, E. H. Lo, and C. IA DeCola, “The science of stroke: Mechanisms in search of treatments,” Neuron, vol. 67, no. 2, pp. 181–198, 2010, doi: 10.1016/j.neuron.2010.07.002.
4. https://www.thelancet.com/journals/laneur/article/PIIS1474-4422(21)00252-0/fulltext
5. 13. Sammut, C., and Webb, G. I. (Eds.). (2010). “Encyclopedia of Machine Learning.” doi:10.1007/978-0-387-30164-8
6. 15. K. Shankar, P. Manickam, G. Devika and M. Ilayaraja,” Optimal Feature Selection for Chronic Kidney Disease Classification using Deep Learning Classifier,” 2018 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), Madurai, India, 2018, pp. 1-5, doi: 10.1109/ICCIC.2018.8782340.
7. 16. Kumar, C Kumar M, Varun Gayathri, T Kumar, S. (2014). Data Analysis and Prediction of Hepatitis Using Support Vector Machine (SVM). International Journal of Computer Science and Information Technologies. 5. 2235-2237.
8. 17. N Komal Kumar, D Vigneswari. Hepatitis- Infectious Disease Prediction using Classification Algorithms. Research J. Pharm. and Tech 2019; 12(8): 3720-3725. doi: 10.5958/0974-360X.2019. 00636.X
9. 18. Breiman, L. Random Forests. Mach. Learn. 2001, 45, 5–32. [CrossRef]
10. 19. He, H.; Garcia, E.A. Learning from Imbalanced Data. IEEE Trans. Knowl. Data Eng. 2009, 21, 1263–1284.[CrossRef]
11. 20. Jia, Q.; Liu, L.;Wang, Y. Stroke in China. Clin. Exp. Pharmacol. Physiol. 2010, 37, 259–264. [CrossRef]
12. 21. Kearns, M.; Mansour, Y.; Ng, A.Y.; Ron, D. An Experimental and Theoretical Comparison of Model SeIection Methods. Conf. Learn. Theory 1995, 27, 21–30.
13. 22. Li, W.; Liu, H.; Yang, P.; Xie, W. Supporting Regularized Logistic Regression Privately and E\_ciently. PLoS ONE 2016, 11, e0156479. [CrossRef]
14. 23. Hametner, C.; Kellert, L.; Ringleb, P.A. Impact of sex in stroke thrombolysis: A coarsened exact matching study. BMC Neurol. 2015, 15, 10. [CrossRef] [PubMed]
15. 24. Alter, M.; Friday, G.; Lai, S.M.; Oconnell, J.; Sobel, E. Hypertension and Risk of Stroke Recurrence. Stroke 1994, 25, 1605–1610. [CrossRef]
16. 25. Linfante, I.; Starosciak, A.K.;Walker, G.R.; Dabus, G.; Castonguay, A.C.; Gupta, R. Predictors of poor outcome despite recanalization: A multiple regression analysis of the NASA registry. J. Neurointerv. Surg. 2016, 8, 224–229. [CrossRef]
17. 26. Karagiannis, A.; Mikhailidis, D.P.; Tziomalos, K.; Sileli, M.; Savvatianos, S.; Kakafika, A. Serum uric acid as an independent predictor of early death after acute stroke. Circ. J. 2007, 71, 1120–1127. [CrossRef]
18. 27. Jia, Q.; Zheng, H.; Zhao, X.; Wang, C.; Liu, G.; Wang, Y. Abnormal glucose regulation in patients with acute stroke across China: Prevalence and baseline patient characteristics. Stroke 2012, 43, 650–657. [CrossRef] Int. J. Environ. Res. Public Health 2020, 17, 1828 11 of 11
19. 28. Amarenco, P.; Goldstein, L.B.; Szarek, M.; Sillesen, H.; Rudolph, A.E.; Callahan, A. E\_ects of Intense Low-Density Lipoprotein Cholesterol Reduction in Patients with Stroke or Transient Ischemic Attack: The Stroke Prevention by Aggressive Reduction in Cholesterol Levels (SPARCL) Trial. Stroke 2007, 38, 3198–3204.[CrossRef]
20. 29. Dawood, F.Z.; Judd, S.; Howard, V.J.; Limdi, N.A.; Meschia, J.F.; Cushman, M. High-Sensitivity C-Reactive Protein and Risk of Stroke in Atrial Fibrillation (from the Reasons for Geographic and Racial Di\_erences in Stroke Study). Am. J. Cardiol. 2016, 118, 1826–1830. [CrossRef] [PubMed]

30. Sundell, L.; Salomaa, V.; Vartiainen, E.; Poikolainen, K.; Laatikainen, T. Increased stroke risk is related to a binge-drinking habit. Stroke 2008, 39, 3179–3184. [CrossRef] [PubMed]