[1]Dr. Rohini B. Jadhav

[2]Dr. Milind Gayakwad

[3]Shital Pawar

[4]Dr. S. D. Jadhav

[5]Dr. Rahul Joshi

[6]Amruta V. Patil

[7]Hiren Dand

[8]Disha Sushant Wankhede

Analyzing Existing Algorithms and Identifying Gaps in Brain Stroke Detection

Abstract: - One of the global issues that matters most right now is healthcare. The primary cause of death globally is brain stroke. A valuable contribution to medicine is the early prediction and identification of brain stroke events. Numerous factors, including BMI (body mass index), age, sex, family history, gender, smoking status, hypertension, and so on, are linked to brain stroke deaths. While forecasting heart illness has received a great deal of interest in the medical community, predicting a brain stroke has received less attention. This study's primary goal is to evaluate various previously published research publications and select the most effective machine learning methods for brain stroke prediction for our next projects. It was shown that mortality rate and functional outcomes are the expected outcomes for the majority of the study work done after analyzing the various machine learning techniques used for stroke predictions and after accounting for the previously published studies. The techniques that were used most commonly were LR, DTC, RFC, SVM, KNN.

Keywords: Brain Stroke, LR, RF, DT, SVM, Ischemic and hemorrhagic

I. INTRODUCTION

One disorder known as brain stroke is brought on by inadequate blood supply to the brain, which results in the death of brain cells [1]. Bleeding is brought on by a limited blood supply to the brain. When a brain stroke is not identified and treated early, the patient passes away. Currently, brain stroke diagnosis is made using an MRI and a computed tomography (CT) scan. Although brain stroke can be effectively detected by CT and MRI scans, these tests are not able to identify strokes that become severe before bleeding begins. An analysis of supervised machine learning techniques evaluated the effectiveness of supervised learning in identifying and predicting the risk of brain stroke. The random forest and support vector machines are the best supervised learning techniques for brain stroke diagnosis, according to an assessment of machine learning techniques for identifying stroke risk. Numerous studies examined the efficacy of decision trees, logistic regression, radon forests, and Naïve Bayes in predicting

[1]Associate Professor, Department of Information Technology, Bharati Vidyapeeth (Deemed to be University) College of Engineering, Pune, India

[2]Bharati Vidyapeeth Deemed to be University College of Engineering, Pune, India

[3]Department of Computer Engineering, Bharati Vidyapeeth's College of Engineering for Women Pune, India

[4]Bharati Vidyapeeth (Deemed to be University), College of Engineering, Pune. India

[5]Symbiosis Institute of Technology, Symbiosis International (Deemed University), Pune, India

[6]Research Scholar, CSE, SJJTU, Rajasthan, India

[7]Research Guide, CSE, SJJTU, Rajasthan, India

[8]Assistant Professor, Vishwakarma Institute of Information technology Pune, India

Email- rbjadhav@bvucoep.edu.in, mdgayakwad@bvucoep.edu.in, shitalp16@gmail.com,

sdjadhav@bvucoep.edu.in, rahulj@sitpune.edu.in, amrutayadav2010@gmail.com, dandhiren@yahoo.co.in, disha.wankhede@viit.ac.in

the risk of stroke. Random forest outperforms logistic regression, Naïve Bayes, and decision tree methods in terms of accuracy, scoring a higher 98% in the assessment of these approaches for stroke risk prediction. Over the past few years, the use of supervised learning techniques in conjunction with predictive analytics has grown in importance in clinical decision-making due to the expansion of medical datasets and computational power.

A supervised learning technique for brain stroke prediction has been developed by researchers using an enormous amount of medical datasets, and it shows promise in terms of accuracy and precision. In order to fill the research gap, this study will provide an experimental analysis of the state-of-the-art for seven well-known supervised machine learning techniques: logistic regression, decision trees, radial basis functions, k-nearest neighbors, radial support vector machines, linear support vector machines, and Naive Bayes. This study's contribution is to offer a concise overview of the literature about the effectiveness of current predictive models in predicting strokes. There are two types of strokes in the brain: hemorrhagic (caused by intracerebral hemorrhage) and ischemic (caused by intracerebral artery obstruction). Fig 1 and Fig 2 show types of BS .
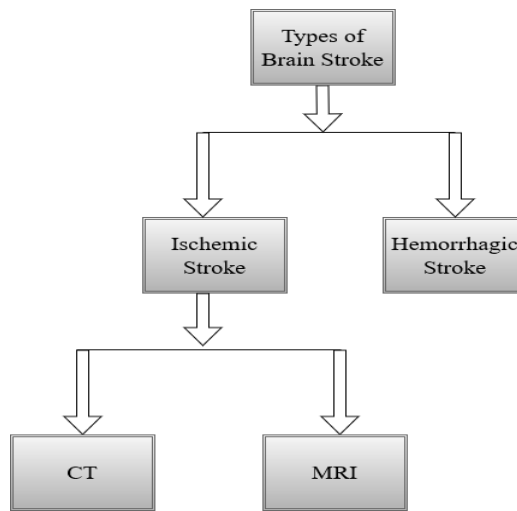
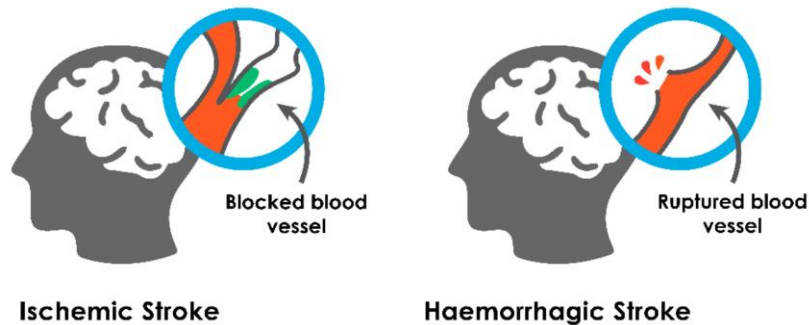

Figure.1- Shows Types of Brain Stroke

Figure. 2- Ischemic and hemorrhagic brain stroke.

1.1 Ischemic Heart Attack

Ischemic stroke is caused by a decrease in cerebral blood flow (CBF) brought on by blood vessel blockage. Large artery blockage can make these strokes deadly, even though ischemia tolerance varies throughout tissue types [28]. The damaged tissue starts to lose vital nutrients and oxygen and starts to expel toxins, which build up and interfere with normal function. Tissue infarction, or death, is a possible outcome if blood vessel recanalization is unsuccessful [28].
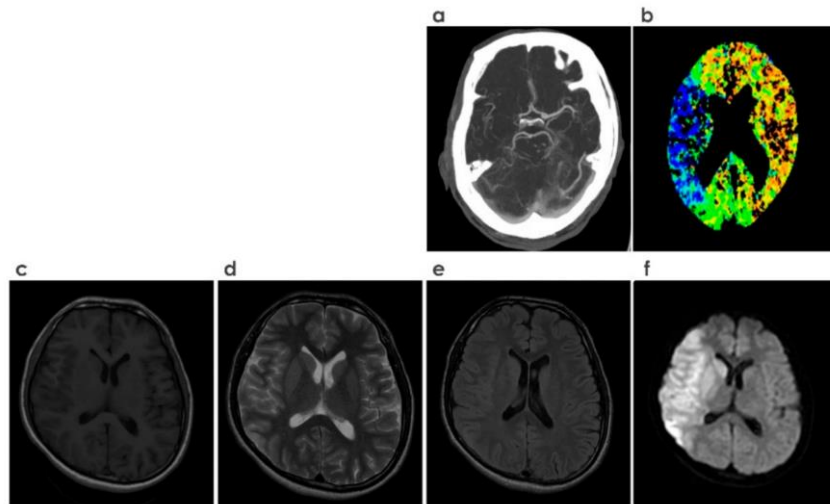
Figure 3- Various neuroimaging modalities

1.2.   Hemorrhagic Stroke

Hemorrhagic stroke is caused by spontaneous blood extravasation from a ruptured artery. The hemoglobin protein density (in relation to plasma concentrations) inside the hematoma determines how visible the bleeding appears on CT. A linear, quantitative measure of radio density, Hounsfield units are used to express the attenuation of CT in the immediate aftermath of a vascular rupture [33]. A'spot sign,' or active bleeding site within the hematoma, can be shown by a contrast-enhanced CT angiography (CTA), which can be used to identify patients at high risk of hemorrhage enlargement (HE). While cerebral microbleeds that are clinically quiet and previously resolved cannot be found on CT, they can be found on MRI.

II.   LITERATURE REVIEW

A support vector machine's ability to predict strokes was examined in another study [1]. Several kernel functions, including linear, radial, quadratic, and polynomial, were used to test the SVM model's performance. As evidenced by the findings, the linear kernel outperformed other kernel functions in terms of accuracy. With the SVM model and the linear kernel function, a 91% stroke prediction accuracy was attained. The 350 stroke samples used to train the model led the researchers to suggest testing the model's performance with a bigger dataset [2][3][1]. [7] examined the efficacy of several classifiers for stroke prediction, including LR,DT,RF and KNN. A total of 5,110 stroke samples with 12 attributes were utilised in the model's testing to initiate the classification model for stroke prediction. According to the results, the random forest model has an accuracy score of 96% [9]. Using the synthetic minority oversampling (SMOTE) method, the random forest model's performance was evaluated on a balanced stroke dataset. Similar to this, a different deep learning-based model for predicting stroke was created in [9]. Using MRI scan data, stroke data was used to train the deep learning model. The model might compete with human specialists, according to analysis of its deep learning capabilities.[5][4]. The corresponding label indicating whether or not a patient is suffering from a stroke is found by analyzing clinical cardiac stroke factors obtained from an MRI scan. An evaluation of the deep learning model's performance revealed that it achieves an accuracy score of 83.4% on the test data [6][8].An automated model for stroke prediction has been developed using the random forest technique [10].Even though the analysis revealed that random forest's performance may be enhanced, the outcome shown that random forest performs better when it comes to predicting strokes. On the prediction of stroke, the random forest model's accuracy was determined to be 90.4% .A convolution-based deep learning technique is developed in [11]to assess whether a patient is experiencing a stroke. A positive result with a 78% score is obtained with deep learning using the LASSO feature selection strategy when the model is evaluated for prediction accuracy [5].The bagging ensemble method is examined for stroke prediction in another study [12]. The analysis's conclusion demonstrates that the pruning technique enhances the bagging classifier's capacity to anticipate strokes. The bagging approach for stroke prediction has a 96% accuracy rating in its performance evaluation. Since a 96% accuracy score is a good result for stroke prediction, the bagging method is therefore thought to be helpful for classifying strokes[13].Furthermore, [14] examined the predictive power of different supervised learning models for strokes. The supervised models that were assessed were Naive Bayes, KNN, SVM,

LR, random forests, and decision trees (NB). When compared to other supervised learning models, the simulation result shows that the NB model achieves a superior accuracy score of 82%. The study was conducted using a highly unbalanced stroke dataset that is publicly available from the Kaggle data repository[28]. Despite the encouraging outcome of their research, they did not utilise preprocessing techniques like resampling. Similarly, using the stroke dataset, [15] evaluated various feature extraction techniques which is publically available from the Kaggle data repository[28].Similarly, using the stroke dataset, F. Calesella1et al. [16] evaluated various feature extraction techniques. As the model is trained for stroke prediction, the study showed that features that are employed for pattern matching have a big impact on how well the machine-learning model performs. It has been discovered that principal component analysis (PCA) is a useful tool for dimensionality reduction and for discovering important features that enhance model performance. In order to identify the most effective technique for hemorrhagic stroke identification, Phong et al. tested three different forms of CNN [156]. By calculating the mean output for input image rotations, Majumdar et al. trained a CNN with enhanced performance [17]. A predictive DL model that can identify ICH was proposed and evaluated by Arbabshirani et al. [18]. An end-to-end FCN model network that carries out combined classification and segmentation on CT images was used by Kuo et al. to address the difficulty of identifying minute and subtle abnormalities in a large 3D volume with higher sensitivity [19]. [31-35] Author work on Brain Tumor detection and classification using various CNN architecture.

Table 1- Summarized Literature Survey from 2018 to 2023

Reference	Stroke Type	Model Used	Accuracy	year
Garg et al. [22]	CT	(RF, GBM, KNN, XGBOOST, SVM	combined data = 0.57	2019
Peixoto et al. [23]	CT	CT SCM, SVM, MLP	97.1%	2018
Lin et al. [21]	CT	DBSCAN, hierarchical DBSCAN (HDBSCAN)	DBSCAN (Avg) IACC 96.9	2019
Pereira et al. [20]	MRI	RF classifier	0.84	2018
Wang et al., [27]	MRI	RF classifier	0.89	2019
Scrutinio et al.,	MRI	RF, ADA-B	089	2020
Karthik R[24]	MRI images	FCN	70%	2019
Lee BJ[25]	MRI images	LR, bagging and RF,CNN	LR on excessive data AUC = 0.569-0.731	2020
Bento M[26]	MRI	SVM	97.5%	2019
Kuang et al. [28]	CT	CNN	85.7%	2021
Barros et al.[29]	MRI	CNN	0.63	2020
Zhiliang Zhang et al. [40]	MRI	Masked RNN and ML	99.89	2023

Table 2 shows lists the inclusion and exclusion criteria.

Table 2. The exclusion and inclusion criteria.

Inclusion	Exclusion
Studies pertaining to 1. Research on CT and MRI (including variations) 2. Heart attacks, both hemorrhagic and ischemic 3. Determining how severe the damage and infarction are 4. The likelihood of damage and the prognosis for	Studies pertaining to 1. Stroke treatment (exclusive) 2. Treatment approaches that are only statistical and biological in nature. 3.Algorithm development and technical functioning 4. non-stroke-related lesions

strokes 5. Identifying and dividing lesions into core and penumbra areas 6.Sectioning lesion locations using ML and DL approaches 7. The most recent architectures for feature-specific algorithms and factorization methods in deep learning.	

III.   ROPOSED METHODOLOGY USING MACHINE INTELLIGENCE STROKE DETECTION

Pre-processing data is required before building a model to remove noise and outliers from a dataset that could lead to the model deviating from its training. This stage addresses every problem preventing the model from performing better. Once the relevant dataset has been gathered, data needs to be cleaned and processed in order to build models. The dataset consists of twelve attributes, as was previously mentioned. First, the column id is disregarded since its presence has no effect on the construction of the model. Subsequently, any null values in the dataset are detected and filled in. The null values in the BMI column are filled in this case using the "most frequent" value from the data column. As a computer is usually trained on numerical data, it is required to convert the strings to integers. Five columns with data of the type string are present in the retrieved dataset: gender, ever married, work type, residence type, and smoking status. The whole dataset is transformed into a set of numbers during the label encoding process, which encodes all strings.

 The stroke prediction dataset is severely unbalanced. Recall and precision are inadequate measures of accuracy, even though utilizing such data to train a machine-level model may improve accuracy. If handling such uneven data is not done correctly, the forecast would be useless and the findings will be erroneous. Therefore, in order to create an effective model, it is necessary to address this uneven data first. The next step was to apply PCA, which determines the smallest number of principal components required to preserve 95% of the variation. The phase of building the model starts once the unbalanced dataset has been managed and the data prepared. To improve the accuracy and productivity of this task, the data is divided into training and testing segments using an 80/20 split. The model is partitioned and then trained with multiple classification approaches. The workflow is shown in Fig. 4
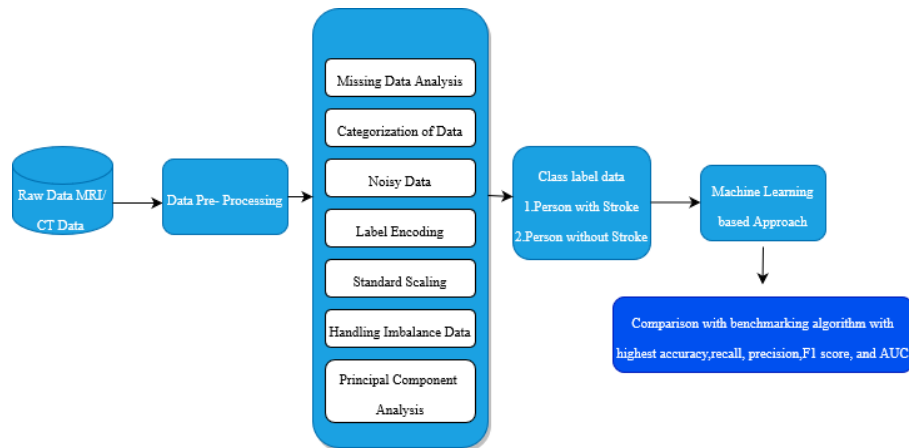

Figure 4- Proposed Workflow Process for Predicting Brain Stroke

IV.   ML BASED TECHNIQUES AVAILABLE

Three basic categories apply to publications in this field: (1) classification using discriminants (2) solely probabilistic types; and (3) hybrid types, where the outcomes are further enhanced using additional techniques. While the former is effective in decreasing processing time and complexity, it suffers from the necessity for specialized intervention to confirm the kind of stroke and its location. Additionally, adequate radiation attenuation is required for successful delineation, which presents a significant problem. The majority of texture-based algorithms suffer from significant false positive rates as a result of missing hardly noticeable lesions because to minute intensity variations, scan image. For network training, a system that provides medical expert advice to aid

in technical detection might be useful . Future research is therefore anticipated to focus on creating intensity-resistant algorithms, which may be combined with improved image enhancement methods.

1)Logistic Regression(LR): The purpose of binary classification questions, such as yes/no or true/false, is to predict a binary outcome. A common machine learning approach for these types of issues is logistic regression [21]. The probability of a binary result based on input features is modelled using a logistic function. It is predicted that the positive class will occur if the logistic function's output is larger than 0.5, while the negative class will occur if it is smaller than 0.5. Logistic regression creates probabilistic predictions by modelling the connection between the dependent and independent variables as a logistic curve. Predictions of binary classes can then be made using these probabilities as thresholds. A lot of features can be handled by logistic regression, which is also computationally efficient and reasonably easy to apply. One-vs-all or softmax regression can be used to expand the logistic regression algorithm to address multiclass classification situations. It is extensively employed in numerous fields, including marketing, credit scoring, medical diagnosis, and picture categorization.

2) *Decision Tree(DT)*: A decision tree is a tree-based machine learning technique that divides the data into smaller subsets depending on the features that yield the most information gain, recursively, until there is no more data to be split. The data is split into distinct branches, each of which represents a distinct conclusion, and a choice is taken at each split depending on a particular attribute. The end product is a tree structure with predictions at the leaves and decision rules at each internal node. Decision trees handle both numerical and categorical data and are straightforward and simple to understand. They can be applied to challenges involving regression as well as classification. Through parameter adjustments, such as lowering the minimum number of samples needed to divide a node or form a leaf, they can deal with data that is not evenly distributed.

3) *Random Forest(RF)*: Several decision trees are created by Random Forest, an ensemble of decision trees, and their predictions are combined to arrive at a final choice. By lowering the variance and boosting the robustness of the model, several trees can cooperate to overcome the drawbacks of a single decision tree, such as overfitting. With its ability to handle missing values, big datasets, and both linear and non-linear correlations, Random Forest is a robust and adaptable machine learning technique. Numerous applications, including credit scoring, environmental monitoring, and medical diagnostics, make extensive use of it.

4) *Support Vector Machine (SVM)*: Classification and regression applications can both benefit from the application of SVM, a supervised learning method. SVM can deal with unbalanced data by changing the class weight parameter, which gives the minority class a larger weight. SVMs seek to identify the ideal hyperplane that divides the data points of various classes as much as possible. The closest data point from each class is separated by the maximum margin when the hyperplane is selected. Support vectors, as the name "support vector machines" suggests, are the data points that are closest to the hyperplane. Compared to other classification techniques, SVM offer a number of advantages, such as the capacity to manage high-dimensional data, cope with non-linear data by utilizing kernel functions, and manage imbalanced datasets. Bioinformatics, image recognition, text categorization, and other fields have all benefited from the effective application of SVMs.

5) *K-Nearest Neighbors (KNN)*: A machine learning algorithm based on instances is KNN. When using KNN, a new data point's categorization in the feature space is based on the class of its K closest neighbors. The hyper parameter K is one that the user sets in value. Generally speaking, decision borders with greater values of K are smoother, but data with smaller values of K can contain more local information. The number of nearest neighbors is adjusted to regulate the decision, allowing it to handle uneven data. Regression and classification problems are both compatible with it. Its drawbacks, however, are the curse of dimensionality—that is, the possibility of performance degradation in high-dimensional spaces—and its sensitivity to the selection of distance measure. Additionally, KNN requires a large amount of memory to store the entire training dataset for each prediction.

V. PUBLICLY AVAILABLE DATASETS

The researchers assess the suggested methodologies using a number of publicly accessible datasets. In this section, a few significant and difficult datasets are covered. The hardest MRI datasets are called BRATS. A greater number of BRATS Challenges with a resolution of 1 mm3 voxels are published in various years. Both Table 1 and Fig. 1 provide the dataset details.

Available Dataset For Experiment Purpose-

Modality	Database	Data Info	Website ink
MRI	ISLES-2015,2016,2017,2018	ISLES2015	http://www.isleschallenge.org/
MRI		ISLES2016	
MRI– DWI/ADC		ISLES2017	
CTP ,TMAX ,CBF, CBV ,MTT,			http://www.isles-challenge.org/

VI.   METRICS FOR MEASUREMENT OF PERFORMANCE

Performance evaluation measures are used to evaluate an ML model's effectiveness. We used five different classifiers to predict when a brain stroke will occur. We computed the performance assessment metrics to find out which classifiers performed well.

The formula for those measures looks like this:

$$Acc = ((TP + TN)/(TP + FP + TN + FN)) \times 100\% \qquad \ldots\ldots\ldots\ldots(1)$$

$$Sen = ((TP) / (TP + FN)) \times 100\% \qquad \ldots\ldots\ldots\ldots(2)$$

$$Spe = ((TN) / (TN + FP)) \times 100\% \qquad \ldots\ldots\ldots\ldots(3)$$

$$FPR = ((FP) / (TN + FP)) \times 100\% \qquad \ldots\ldots\ldots\ldots(4)$$

$$FNR = ((FN) / (TP + FN)) \times 100\% \qquad \ldots\ldots\ldots\ldots(5)$$

VII.   CONCLUSION

Stroke is a medical disorder that can be fatal and requires immediate attention to avoid further complications. Early stroke detection and subsequent mitigation of its severe repercussions may be facilitated by the development of machine learning (ML) and deep learning models. Based on a variety of biological parameters, this study investigates the predictive power of several machine learning (ML) and boosting algorithms for stroke. Our study examined stroke detection and segmentation trends, status, and future directions. It is evident that prompt neuroimaging analysis requires a quick, flexible method, which is essential for managing stroke patients. This is because neuroimaging plays a major part in the best possible diagnosis and treatment of various stroke types. The creation of automated diagnostic tools is distinctly feasible given current developments in neuroimaging, machine learning, artificial intelligence, and computing power. Patient morbidity and death will decrease as a result of developments in this area and their application to clinical practice. We suggested employing machine learning (ML) to improve the functionality of automated diagnostic approach and create a more complete automated system based on the results of this systematic literature review.

REFERENCE

[1]   Virani, S.S.; Alonso, A.; Aparicio,. Heart Disease and Stroke Statistics—2021 Update: A Report from the American Heart Association.Available online: https://www.heart.org/-/media/PHD-Files-2/Science-News/2/2021-Heart-and-Stroke-Stat-Update/2021Stat_Update_factsheet_Global_Burden_of_Disease.pdf (accessed on 27 January 2021).

[2]   GBD 2016 Stroke Collaborators. Global, regional, and national burden of stroke, 1990–2016: A systematic analysis for the GlobalBurden of Disease Study 2016. Lancet 2019, 18, 439–458.

[3]   Disha Sushant Wankhede, R. Selvarani, Dynamic architecture based deep learning approach for glioblastoma brain tumor survival prediction, Neuroscience Informatics, Volume 2, Issue 4, 2022, 100062, ISSN 2772-5286, https://doi.org/10.1016/j.neuri.2022.100062. (https://www.sciencedirect.com/science/article/pii/S2772528622000243)

[4]   O'Donnell, M.J.; Xavier, D.; Liu, L.; Zhang, H.; Chin, S.L.; Rao-Melacini, P.; Rangarajan, S.; Islam, S.; Pais, P.; McQueen, M.J.; et al. Risk factors for ischaemic and intracerebral haemorrhagic stroke in 22 countries (the INTERSTROKE study): A case-control study. Lancet 2010, 376, 112–123.

[5]   Campbell, B.C.V.; De Silva, D.A.; MacLeod, M.R.; Coutts, S.B.; Schwamm, L.H.; Davis, S.M.; Donnan, G.A. Ischaemic stroke. Nat. Rev. Dis. Primers 2019, 5, 70.

[6]   Switzer, A.J.; Hess, D.C.; Nichols, F.T.; Adams, R.J. Pathophysiology and treatment of stroke in sickle-cell disease: Present and future. Lancet Neurol. 2006, 5, 501–512.

[7]   Krishnamurthi, R.V.; Feigin, V.L.; Forouzanfar, M.H.; Mensah, A.G.; Connor, M.; Bennett, A.D.; Moran, E.A.; Sacco, R.L.; Anderson,L.M.; Truelsen, T.; et al. Global and regional burden of first–ever ischaemic and haemorrhagic stroke during 1990–2010: Findingsfrom the Global Burden of Disease Study. Lancet Glob. Health 2013, 1, e259–e281.

[8] Sagar, R.; Dandona, R.; Gururaj, G.; Dhaliwal, R.S.; Singh, A.; Ferrari, A.; Dua, T.; Ganguli, A.; Varghese, M.; Chakma, J.K.; et al.The burden of neurological disorders across the states of India: The Global Burden of Disease Study 1990–2019. Lancet Psychiatry2020, 7, 148–161.

[9] D. Wankhede, V. Mishra, M. Karnik, A. Kekane and A. Shukla, "The Impact of the Latest Technology on Healthcare and how can it be leveraged to improve patient outcomes and reduce Healthcare costs," 2023 4th IEEE Global Conference for Advancement in Technology (GCAT), Bangalore, India, 2023, pp. 1-6, doi: 10.1109/GCAT59970.2023.10353516.

[10] Donkor, E.S. Stroke in the21st Century: A Snapshot of the Burden, Epidemiology, and Quality of Life. Stroke Res. Treat. 2018, 2018,3238165.

[11] Saver, J.L. Penumbral salvage and thrombolysis outcome: A drop of brain, a week of life. Brain 2017, 140, 519–522.

[12] Meretoja, A.; Keshtkaran, M.; Tatlisumak, T.; Donnan, G.A.; Churilov, L. Endovascular therapy for ischemic stroke: Save aminute-save a week. Neurology 2017, 88, 2123–2127.

[13] Powers,W.J.; Rabinstein, A.A.; Ackerson, T.; Adeoye, O.M.; Bambakidis, N.C.; Becker, K.; Biller, J.; Brown, M.; Demaerschalk,B.M.; Hoh, B.; et al. 2018 Guidelines for the Early Management of Patients with Acute Ischemic Stroke: A Guideline for HealthcareProfessionals from the American Heart Association/American Stroke Association. Stroke 2018, 49, e46–e99.

[14] Lindsay, P.; Furie, K.L.; Davis, S.M.; Donnan, G.; Norrving, B. World Stroke Organization Global Stroke Services Guidelines andAction Plan. Int. J. Stroke 2014, 9, 4–13.

[15] Brazzelli, M.; Sandercock, P.A.; Chappell, F.M.; Celani, M.G.; Righetti, E.; Arestis, N.; Wardlaw, J.M.; Deeks, J.J. Magneticresonance imaging versus computed tomography for detection of acute vascular lesions in patients presenting with stroke

[16] symptoms. Cochrane Database Syst. Rev. 2009, 4, CD007424.

[17] Clinical Guidelines for Stroke Management. Available online: https://informme.org.au/en/Guidelines/Clinical-Guidelines-for-Stroke-Management (accessed on 3 October 2021).

[18] Wankhede, D.S., Shelke, C.J. (2023). An Investigative Approach on the Prediction of Isocitrate Dehydrogenase (IDH1) Mutations and Co-deletion of 1p19q in Glioma Brain Tumors. In: Abraham, A., Pllana, S., Casalino, G., Ma, K., Bajaj, A. (eds) Intelligent Systems Design and Applications. ISDA 2022. Lecture Notes in Networks and Systems, vol 715. Springer, Cham. https://doi.org/10.1007/978-3-031-35507-3_19

[19] Rowley, H.; Vagal, A. Stroke and Stroke Mimics: Diagnosis and Treatment. In Diseases of the Brain, Head and Neck, Spine 2020–2023:Diagnostic Imaging; IDKD Springer Series; Springer: Berlin/Heidelberg, Germany, 2020; pp. 25–36.

[20] Allen, M.L.; Anton, N.; Hasso, J.; Handwerker, J.; Hamed, F. Sequence-specific MR Imaging Findings that Are Useful in DatingIschemic Stroke. RadioGraphics 2012, 32, 1285–1297.

[21] Pereira, S.; Meier, R.; McKinley, R.;Wiest, R.; Alves, V.; Silva, C.A.; Reyes, M. Enhancing interpretability of automatically extractedmachine learning features: Application to a RBM-Random Forest system on brain lesion segmentation. Med. Image Anal. 2018, 44,228–244.

[22] Lin, C.-H.; Hsu, K.-C.; Johnson, K.R.; Luby, M.; Fann, Y.C. Applying density-based outlier identifications using multiple datasetsfor validation of stroke clinical outcomes. Int. J. Med. Inform. 2019, 132, 103988.

[23] Garg, R.; Oh, E.; Naidech, A.; Kording, K.; Prabhakaran, S. Automating Ischemic Stroke Subtype Classification Using Machine

[24] Learning and Natural Language Processing. J. Stroke Cerebrovasc. Dis. 2019, 28, 2045–2051.

[25] Peixoto, S.A.; Filho, P.P.R. Neurologist-level classification of stroke using a Structural Co-Occurrence Matrix based on thefrequency domain. Comput. Electr. Eng. 2018, 71, 398–407.

[26] Karthik R, Gupta U, Jha A, Rajalakshmi R, Menaka R. Adeep supervised approach for ischemic lesion segmentationfrom multimodal MRI using fully convolutional network.Applied Soft Computing 2019;84:105685

[27] Lee BJ, Kim KH, Ku B, Jang J-S, Kim JY. Prediction ofbody mass index status from voice signals based onmachine learning for automated medical applications.Artificial intelligence in medicine 2013;58(1):51-61.

[28] Bento M, Souza R, Salluzzi M, Rittner L, Zhang Y, FrayneR. Automatic identification of atherosclerosis subjects ina heterogeneous MR brain imaging data set. Magneticresonance imaging 2019.

[29] Wang, C.-W.; Lee, J.-H. Stroke Lesion Segmentation of 3D Brain MRI Using Multiple Random Forests and 3D Registration. Adv.Data Min. Appl. 2016, 9556, 222–232.

[30] Kuang, H.; Menon, B.K.; Sohn, S.I.; Qiu, W. EIS-Net: Segmenting early infarct andscoring ASPECTS simultaneously onnon-contrast CT of patients with acute ischemic stroke. Med.Image Anal. 2021, 70, 101984.

[31] Barros, S.R.; Tolhuisen, M.L.; Boers, A.M.; Jansen, I.; Ponomareva, E.; Dippel, D.W.J.; Lught, A.V.D.; Oostenbrugge, R.J.V.; Zwam,W.H.V.; Berkhemer, O.A.; et al. Automatic segmentation of cerebral infarcts in follow-up computed tomography images with convolutional neural networks. J. NeuroInterv. Surg. 2020, 12, 848–852.

[32] Barman, A.; Inam, M.E.; Lee, S.; Savitz, S.; Sheth, S.; Giancardo, L. Determining Ischemic Stroke from CT-AngiographyImaging Using Symmetry-Sensitive Convolutional Networks. In Proceedings of the 2019 IEEE 16th International Symposium onBiomedical Imaging (ISBI 2019), Taichung, Taiwan, 8–10 November 2019; pp. 1873–1877.

[33] Mrs. Disha Sushant Wankhede, Dr. Selvarani Rangasamy,"REVIEW ON DEEP LEARNING APPROACH FOR BRAIN TUMOR GLIOMA ANALYSIS" Journal of Information Technology in Industry, VOL. 9 NO. 1 (2021) pp. 395 - 408 , DOI: https://doi.org/10.17762/itii.v9i1.144

[34] Feigin, V.L.; Norrving, B.; Mensah, G.A. Global Burden of Stroke. Circ. Res. 2017, 120, 439–448.

[35] Kamalakannan, S.; Gudlavalleti, A.; Gudlavalleti, V.S.M.; Goenka, S.; Kuper, H. Incidence & prevalence of stroke in India: Asystematic review. Indian J. Med. Res. 2017, 146, 175–185.

[36] Bivard, A.; Churilov, L.; Parsons, M. Artificial intelligence for decision support in acute stroke Current roles and potential. Nat.Rev. Neurol. 2020, 16, 575–585.

[37] Wankhede, D.S., Pandit, S., Metangale, N., Patre, R., Kulkarni, S., Minaj, K.A. (2022). Survey on Analyzing Tongue Images to Predict the Organ Affected. In: Abraham, A., et al. Hybrid Intelligent Systems. HIS 2021. Lecture Notes in Networks and Systems, vol 420. Springer, Cham. https://doi.org/10.1007/978-3-030-96305-7_56

[38] D. Wankhede, V. Mishra, M. Karnik, A. Kekane and A. Shukla, "The Impact of the Latest Technology on Healthcare and how can it be leveraged to improve patient outcomes and reduce Healthcare costs," *2023 4th IEEE Global Conference for Advancement in Technology (GCAT)*, Bangalore, India, 2023, pp. 1-6, doi: 10.1109/GCAT59970.2023.10353516.
Sutar, S., Jose, K., Gaikwad, V., Mishra, V., Wankhede, D., & Karnik, M. (2023). Enhancing Data Management: An Integrated Solution for Database Backup, Recovery, Conversion, and Encryption Capabilities. International Journal of Intelligent Systems and Applications in Engineering, 12(6s), 720–734. Retrieved from https://ijisae.org/index.php/IJISAE/article/view/4011

[39] Sakhare, N. N., Bangare, J. L., Purandare, R. G., Wankhede, D. S., & Dehankar, P. (2024). Phishing Website Detection Using Advanced Machine Learning Techniques. International Journal of Intelligent Systems and Applications in Engineering, 12(12s), 329 –. Retrieved from https://ijisae.org/index.php/IJISAE/article/view/4519

[40] P. M, D. S. Wankhede, R. Kumar, G. Ezhilarasan, S. Khurana and G. S. Sahoo, "Leveraging AI-Driven Systems to Advance Data Science Automation," *2023 International Conference on Emerging Research in Computational Science (ICERCS)*, Coimbatore, India, 2023, pp. 1-7, doi: 10.1109/ICERCS57948.2023.10434009.

[41] Zhiliang Zhang, Zhongxiang Ding, Fenyang Chen, Rui Hua, Jiaojiao Wu, Zhefan Shen, Feng Shi & Xiufang Xu (2024) Quantitative Analysis of Multimodal MRI Markers and Clinical Risk Factors for Cerebral Small Vessel Disease Based on Deep Learning, International Journal of General Medicine, 17:, 739-750, DOI: 10.2147/IJGM.S446531