1.Paper Name: ***machine learning approach to select features important to stroke prognosis***

In this paper [1], the authors proposed initially important features ranked by Shapiro-wilk algorithm where Recursive Feature Elimination was extensively used. ML algorithms used are Linear SVC, Random-Forest-Classifier, Extra Trees-Classifier, AdaBoost-Classifier, and Multinomial-Naive-Bayes-Classifier as estimator Recursive Feature Elimination with Cross-Validation (RFECV). All the features selection that was done were based on RFECV method and its related parameters. Relevant features best before six months were incorporated. Then RFECV algorithms for ML algorithms are plotted against No of features selected and score. Missing data and outliners were removed to test efficiency of RFEC method as aforesaid. The RFECV algorithmic approach was deviated with highest accuracy of 0.69 due to immature data and lack of other essential features.

2. Paper Name**: A machine-learning model for automatic detection of movement compensations in stroke patients**

In this paper [2], the authors proposed a ML based automatic model is used to gather precise info on the compensatory movements that a person makes. A total of 18 movements were considered for this Random Forest algorithm was used. Various Human movement data were collected that were considered for locomotion activities. There is presence of Motion capture system to record movements based on the data collection then comes motion segment parameter to analyse and record movement pace with respect to time, these multilabel classification is performed by using Random K- label sets (RAkEL) algorithm, further a Leave one out cross validation (LOOCV) is used thus 85% average precision overall is achieved.

3. Paper Name**: An Enhanced Stroke Prediction Scheme Using SMOTE and Machine Learning Techniques**

In this paper [3], the authors proposed SMOTE termed oversampling method with Logistic Regression, Random Forest and XGBoost is applied to the prescribed dataset. The outliners data present in the dataset which is a imbalanced data is handled using smote. Due to the presence of various missing values in the prescribed dataset one of the important handling techniques called “Mean Value Imputation” is applied and from the Scikit-learn library a module named Simple Imputer is accompanied along. Further Grid search CV algorithm is applied for finest parameters. Thus, the overall accuracy of 99.07% is achieved with the Random Forest and 99% Precision.

4. Paper Name**: An Integrated Machine Learning Approach to Stroke Prediction**

In this paper [4]**,** the authors proposed dataset emphasis is given over the cardiovascular health study where there is through study of Cox proportional hazards model. The model comprises of one of the automatic feature selector algorithms in order to select the best feature. Missing entries are calculated with the Imputation through linear regression and Regularized Expectation Maximization. The major ML algorithms used are Support Vector Machines and the Margin based Censored Regression. The aforesaid feature selection algorithm achieves high score under ROC curve when compared to the Cox model.

5. Paper Name: **Analysing the Performance of Stroke Prediction using ML Classification Algorithms**

In this paper [5]**,**  the authors proposed This paper encompasses various physiological factors in its prescribed dataset. Several ML algorithms like Logistic Regression, Decision Tree Classification, Random Forest Classification, K-Nearest Neighbours, Support Vector Machine and Naïve Bayes Classification are used, Further label encoding followed by Handling various imbalanced data are achieved. The ROC curve is drawn for all the aforesaid algorithms along with their recall score, Precision score, F1 score and accuracies of the above-mentioned algorithm by visualizing them in bar diagrams, thus a detailed overall comparative analysis. The naïve bayers algorithms performed the best with overall accuracy of 82%.

6.Paper Name: **Classification of stroke disease using machine learning algorithms**

In this paper [6], the authors proposed there is extensive use of data mining techniques in order to precisely track information’s. ML algorithms like Artificial neural networks, SVM, boosting and bagging and random forest were applied more emphasis on ANN was given which is constructed with around 22 inputs with one hidden layer. The AUC curve depicting area bounded by True positive rate and False positive rate is represented to each of the aforesaid listed ML algorithms. Further ensemble method is used for more precise accurate predictions Random Forest is used in bagging classifier. AdaBoost and random forest gave accuracy of 90.9% while the best performance was given by ANN algorithm with accuracy score of almost 95%

7.Paper Name: **Chi-Squared Based Feature Selection for Stroke Prediction using AzureML**

This paper [7], the authors proposed totally based on Server-side Resource utilization where the analysis and prediction is done using the Microsoft Azure ML a cloud based platform where the algorithms run also train, build and deployment of ML models takes place from the desired browser. ML algorithms used are Two class decision jungle which is derived from decision tree and the two-class boosted decision tree, Bar diagram results are illustrated before and after applying SMOTE. Further Confusion matrix and ROC graph is drawn for all the features. F1 score, precision and accuracy is calculated.

8.Paper Name: **Early Detection of Brain Stroke using Machine Learning Techniques**

In this paper[8]**,** the author proposed This paper involves detailed comparative analysis of ML algorithms consisting K nearest neighbour, logistic regression, random forest classifier and SVC. Using these algorithms, a model is built which is further built a GUI using FLASK framework which becomes useful for any people (end user) to check predictions using the details of their health in the website. K Nearest Neighbour has highest accuracy of 99.35% followed by decision tree random forest, SVM then the least accurate being Logistic Regression and a bar diagram is plotted of all the aforesaid algorithms with respect to accuracy.

9.Paper Name: **Detection of Stroke Disease using Machine Learning Algorithms**

In this paper the [9], author proposed This paper encompasses dataset directly collected from hospital also based on various brain scans. The algorithms used are Naive Bayes, j48, k-NN, Random Forest, ANN, SVM and two ensemble methods. Datasets are cleaned using WEKA. The ANN algorithm proved to be best with highest accuracy of 95.3%. Further All the prediction parameters like recall, F1 score and precision are calculated, further for each of the above-mentioned algorithms a respective confusion matrix is calculated. In the given dataset the naïve Bayer couldn’t performed well as expected. Random forest was the only algorithm with constant score across all accuracy parameters.

10.Paper Name: **Prediction of Long-Term Stroke Recurrence Using Machine Learning Models**

In this paper the [10], author proposed This paper made the use of 6 ML algorithms like Logistic Regression, Extreme Gradient Boosting, Gradient Boosting Machine, Random Forest, SVM, Decision Tree for 5 years the predictions is tested. Area under Receiver Operating Characteristic (AUROC) curve was implemented. SMOTE was used for up-sampling of the data. A parameter grid was built. A model performance summary of the algorithms along with different features was plotted. The model’s metrics was much more improved after applying the sampling strategies

11.Paper Name: **Using Machine Learning to Improve the Prediction of Functional Outcome in Ischemic Stroke Patients**

In this paper the [11], authors proposed in this paper datasets is directly taken when the patient is admitted with the symptoms and history. Data cleaning was performed to remove some non-required data’s, further ASTRAL score, DRAGON score and the THRIVE score based on persons stroke related medical info of the patient is taken in account. Four classification algorithm is used i.e., the regularized Logistic Regression, Decision Tree, SVM, Random Forest and XGBoost, Further ROC curve for the Random Forest classifier was plotted because it displayed the overall best performance. Finally, a paired t-test result was drawn for each algorithm along with the medical score parameters as aforementioned

12. Paper Name: **Stroke Prediction Using SVM**

In this paper [12] the authors proposed a unique singleton ML algorithm Support Vector Machine where SVM is implemented with various Kernel mathematical functions. Here the Data are collected directly from international database of stroke diseases. The kernel functions of SVM used are Linear quadratic RBF and Polynomial. Further each of these kernel’s Performance metrices like Accuracy Precision Sensitivity specificity and F score are calculated and plotted. Amongst the Kernel used Linear Kernel gave best performance of 91% along with all other performance parameters much higher then rest whereas the RBF accountant to least accuracy of just close to 59%.

13. Paper Name: **Risk Detection of Stroke using a Feature Selection and Classification Method**

In this paper [13] the authors proposed the use of support vector machine which is further added with the glow-worm swarm Optimization Algorithm (GSO) which is totally based on the selected features of last five years based on several important medical reports and several essential scans carried out. Thus the above combinational of Algorithms fetched an accuracy of around 82.58% from 18 of the selected features form the given dataset. Feature selection model was constructed with the help of weighted voting form the initial given dataset records. First six feature weighted voting table was constructed. Further it was concluded that among the 18 features the first six features were most important in prediction.

14. Paper Name: **Prognostic the Risk of Stroke Using Integrated Supervised Machine Learning Techniques**

In this paper the [14]**,** authors proposed model comprising of ML algorithms used Support Vector Machine, Logistic Regression, Gaussian NB, K-Nearest Neighbours, Bernoulli and Linear regression. Pre-processing the data was achieved and further label encoding was performed. The Decision Tree Classifier was best achieved Accuracy score of 93%, the linear Regression performed worst with accuracy of just 0.95%. Further comparative study of same algorithms with other datasets were analysed. The SVC algorithm used uses feature of Multidimensional Kernel for enhancing better performance

**15**. Paper Name: **Predicting Stroke Risk With an Interpretable Classifier**

In this paper[15]**,** the authors proposed use of ML Algorithms like Support Vector Machine, Naive Bayes, Decision tree and Multilayer Perceptron which is an abstract algorithm form the Artificial Neural Networks (ANN). All the results form the model so obtained there is a comparison of Dempster-Shafer method with the aforementioned algorithms. Dempster-Shafer method is used with Gradient Descent Classifier. A conclusion Table was drawn for Each models representing aforesaid algorithms with Accuracy and AUC ROC value and ROC curve for each of the mentioned algorithm. Dempster-Shafer using Gradient Descent Classifier (DSGD) was the best performer while KNN and SVM performance was worst amongst all

16. Paper Name: **Performance Analysis of Machine Learning Approaches in Stroke Prediction**

In this paper [16]**,** the authors proposed training of different classifiers by selecting best features classification algorithms used are Stochastic Gradient Descent, XGBoost, Multilayer Perceptron, AdaBoost ,K- Neighbours and Gradient Boosting classifiers were taken into account. Further a heatmap was plotted for Target variable against features present in the dataset. Amongst all the weighted voting classifier performed the best with measured an accuracy score of 97%. Further a table was concluded with ML classifier with each algorithm with performance metrics. XGB and Gradient Boosting classifier resultant in 96% accuracy score while Gaussian accountant for least accuracy

17. Paper Name: **Machine Learning–Based Model for Prediction of Outcomes in Acute Stroke**

In this paper[17]the authors proposed model comprising of ML algorithms like Deep Neural Network Random Forest and Logistic Regression. For the given Neural network of Artificial intelligence consisted of three hidden layers along with the fifteen ANN units for the decision tree algorithm approximately three hundred decision trees were used. Further ASTRAL score was taken into account to evaluated the accuracy of the model. Further a ROC curve was drawn for all above mentioned algorithms it was concluded that the area under the ROC curve for Deep Neural Network was higher then its competitor algorithms

18. Paper Name: **Machine Learning in Acute Ischemic Stroke Neuroimaging**

In this paper [18], the authors proposed, some certain importance arena of Machine learning in neuroimaging specially for the region of Acute Ischemic in Stroke. The data set consists of Practical parameters on hospital database of Various brain scans like CT and MRI scan. Support Vector Machine is used. Various Data was collected from articles on the basis of Machine learning algorithm used with their Feature selected to achieve best result under selected Algorithm. Further a Conclusion table was drawn featuring the aforesaid. The authors have expressed current as well as future challenges in Machine learning for the treatment of Acute Ischemic stroke.

19. Paper Name: **Automated Ischemic Stroke Subtyping Based on Machine Learning Approach**

In this paper [19] the authors proposed model consisting of ML algorithms like SVM, Random Forest, AdaBoost and the Multinomial Naïve Bayes. Further some important features were taken in account which were further indexed using the Shapiro-Wilk algorithm. Recursive Feature Elimination with Cross-Validation (RFECV) was used that consisted all the aforementioned algorithms. Further data pre-processing with label encoding was performed. The Performance of RFECV with selected feature with each of the mentioned algorithm was plotted based on the features that were indexed by Shapiro Ranking. It was concluded that the RFECV method was better in accuracy score of 81.5%

20. Paper Name: **Machine Learning for Brain Stroke**

In this paper [20]the authors proposed model consisting of ML algorithms were SVM, Deep learning (DL), k-nearest neighbours, ANN, Logistic Regression and the Random Forest. From the given dataset based on different tests and scans some objectives were concluded based on Machine learning algorithm approach including the samples takes and the contrasting feature for enhancing the accuracy based on all these an optimal result that included ROC-AUC score the recall value and the precision with accuracy and error value was observed further on this basis some clinical implications were drawn