**LITERATURE SURVEY**

In the research conducted by Manisha Sirsat, Eduardo Ferme, Joana Camara, the main aim of the research was to classify state-of-arts on ML techniques for brain stroke into 4 categories based on their functionalities or similarity, and then review studies of each category systematically. The study further discusses the outcomes and accuracies obtained by using different Machine Learning models using text and image-based datasets.

In this study, the authors discussed many stroke related problems from the state-of-art. The reviewed studies were grouped in several categories based on their similarities. The study notes that it is difficult to compare studies as they employed different performance metrics for different tasks, considering different datasets, techniques, and tuning parameters. Hence, it only mentions the research areas which were targeted in more than one study and the studies which report highest classification accuracy in each section

Harish Kamal, Victor Lopez, Sunil A. Sheth, in their study discus how Machine Learning (ML) through pattern recognition algorithms is currently becoming an essential aid for the diagnosis, treatment, and prediction of complications and patient outcomes in several neurological diseases. The evaluation and treatment of Acute Ischemic Stroke (AIS) have experienced a significant advancement over the past few years, increasingly requiring the use of neuroimaging for decisionmaking. This study offers an insight into the recent developments and applications of ML in neuroimaging focusing on acute ischemic stroke. The implementations of machine learning are numerous, from early identification of imaging diagnostic findings, estimating time of onset, lesion segmentation, and fate of salvageable tissue, to the analysis of cerebral edema, and predicting complications and patient outcomes after treatment.

The paper finally concludes by discussing how Machine learning applications are expanding in the medical field for diagnostic and therapeutic purposes, and the rapidly expanding and increasingly neuro-imaging reliant field of AIS is proving to be fertile ground. There is a particular need for ML solutions in this field, which is faced with the challenge of increasingly complex data, with limited human expert resources. Future directions in ML for AIS may require collaborative approaches across multiple institutions to build a robust dataset for efficient training of ML networks [2].

In the research conducted by Chuloh Kim, Vivienne Zhu, Jihad Obeid and Leslie Lenert, they have assessed performance of natural language processing (NLP) and machine learning (ML) algorithms for classification of brain MRI radiology reports into acute ischemic stroke (AIS) and non-AIS phenotypes. The method followed included All brain MRI reports from a single academic institution over a two-year period were randomly divided into 2 groups for ML: training (70%) and testing (30%). Using “quanteda” NLP package, all text data were parsed into tokens to create the data frequency matrix. Ten-fold crossvalidation was applied for bias correction of the training set. Labelling for AIS was performed manually, identifying clinical notes. They applied binary logistic regression, naïve Bayesian classification, single decision tree, and support vector machine for the binary classifiers, and we assessed performance of the algorithms by F1-measure. They also assessed how n-grams or term frequency-inverse document frequency weighting affected the performance of the algorithms.

The paper concluded with the understanding how supervised ML based NLP algorithms are useful for automatic classification of brain MRI reports for identification of AIS patients. Single decision tree was the best classifier to identify brain MRI reports with AIS [3].

In the research conducted by R. Punitha Lakshmi et al [4], they put forward their work on SVM Classifier Based On Otsu Thresholding For Ischemic Stroke Detection. The dataset used in order to train the algorithms/models were a set of 32 different types of brain MRI images which were in JPEG format. Both the classifiers i.e. the Random Forest Classifier and SVM Classifier were trained with the help of these images but with different procedure. All the MRI Images were first transformed using the wavelet transformation and the segmentation of those images were carried out by Otsu Thresholding.

The images are obtained from Open Access Series of Imaging Studies (OASIS) which makes the MRI data sets of the brain which is available for the research purpose. Noise reduction of these images were done in pre-processing so as to get accurate results. After that, the data was then fed as input to the SVM Classifier. Thus, the maximum accuracy was given by SVM Classifier being 88% and Random Forest Classifier being at 81%.

The paper concluded with the understanding of how maximum accurate segmentation of brain and brain lesions is achieved with the help of SVM Classifier based on Otsu Thresholding and the dataset with scattering lesion tissues can also help to improve further accuracy rates of this Classification [4].

In the research conducted by Jaehak Yu et al, an implementation of system for semantic analysis of early detection of stroke and also the recurrence of stroke in Koreans over the age of 65 years based on the National Institute of Health (NIH) Stroke Scale was done by the researchers. The research was made possible with the help of data which was collected from the emergency medical center of the Chungnam National University Hospital consisting of 287 stroke patients out of which 16 patients, which had no stroke symptoms were excluded. Final NIHSS Data consisted of 227 patients, excluding the 60 patients whose data included missing values or outlier values among the NIHSS questionnaires. Patient subjects were the elderly over 65 years old, and consisted of 117 men and 110 women.

The Machine Learning Algorithm which was used was C4.5 Decision Tree Algorithm. The researchers found out that it is the most advanced algorithm and its function of classification and prediction is already proven. The proposed system in this experiment classifies and predicts stroke severity score into four classes using representative classification and prediction models of machine learning and data mining methodology. To measure the experiment accuracy of the proposed system, the recall and precision are used as the measurements. The experiment resulted in faster and more accurate predictions of stroke severity and efficient system operation with the help of various Machine Learning algorithm used and C4.5 decision tree and Random Forest classified and predicted the performance with high accuracy.

The paper concluded with the understanding of how efficient use of Machine Learning Methodologies and a proper dataset to build a model to predict Brain Stroke and also assess the severity of symptoms to predict results with high precision can be implemented to build a system providing an alarm service to visit a medical centre or hospital in real-time [5].

Gangavarapu Sailasya, Gorli L. Aruna Kumari, in their study discuss how Brain Stroke, which is the fouth leading cause of death in India, can be predicted with the help of trained Machine Learning Models so as to minimize risk of death due to Brain