A COMPARATIVE STUDY FOR BRAIN TUMOUR DETECTION USING MAGNETIC RESONANCE IMAGES

Thesis Submitted in Partial Fulfilment of the Requirements for the Degree of

MASTER OF TECHNOLOGY

in

Software Engineering

by

Smriti Srivastav

(2K22/SWE/19)

Under the Supervision of

Dr. Ruchika Malhotra

Head of Department (Software Engineering)

Department of Software Engineering

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Shahbad Daulatpur, Main Bawana Road, Delhi-110042, India

June,2024

DELHI TECHNOLOGICALUNIVERSITY

(Formerly Delhi College of Engineering)

Shahbad Daulatpur, Main Bawana Road, Delhi-42

CANDIDATE'S DECLARATION

I Smriti Srivastav hereby certify that the work which is being presented in the thesis entitled" A comparative study for brain tumour detection using magnetic resonance images" in partial fulfilment of the requirements for the award of the Degree of Master of Technology submitted by me to Department of Software Engineering, Delhi Technological University is an authentic record of my own work carried out during the period from 2022 to 2024 under the supervision of Dr. Ruchika Malhotra.

The matter presented in the thesis has not been submitted by me for the award of any other degree of this or any other Institute

Smriti

Candidate's Signature

This is to certify that the student has incorporated all the corrections suggested by the examiners in the thesis and the statement made by the candidate is correct to the best of our knowledge.

Signature of Supervisor

DELHI TECHNOLOGICALUNIVERSITY

(Formerly Delhi College of Engineering)

Shahbad Daulatpur, Main Bawana Road, Delhi-42

CERTIFICATE BY THE SUPERVISOR

Certified that **Smriti Srivastav**(2K22/SWE/19) has carried out their search work presented in this thesis entitled "**A comparative study for brain tumour detection using magnetic resonance images**" for the award of **Master of Technology** from Department of Software Engineering, Delhi Technological University, Delhi, under my supervision. The thesis embodies results of original work, and studies are carried out by the student herself and the contents of the thesis do not form the basis for the award of any other degree to the candidate or to anybody else from this or any other University/Institution.

Signature

Dr. Ruchika Malhotra

SUPERVISOR

HOD and Professor

(Department of Software Engineering)

Date:

ABSTRACT

Brain tumours are abnormal neural growths that can be fatal and develop an immense variety of symptoms. Accurate categorization of brain tumours is crucial for optimizing therapeutic approaches and improving overall patient outcomes. In this study, we propose a comparative brain tumour detection system using 6 machine learning algorithms which are Naïve Bayes algorithm, KNN, Random Forest, Ada Boost, SVM and CNN. In the study Experimental results demonstrate that deep learning models, particularly CNN outperforms the other algorithm with the accuracy of 98.16.

The suggested system is divided into three primary phases: feature extraction, classification, and picture pre-processing. In the pre-processing stage, noise in Magnetic Resonance Imaging (MRI) scans is reduced, and contrast is improved. This comparison analysis clarifies the advantages and disadvantages of each approach and offers guidance on which models to use depending on particular computational and clinical needs. Subsequent research endeavours will investigate the amalgamation of these models into a cohesive structure to optimise their mutual advantages for heightened precision in diagnosis. The practical ramifications of these findings for clinical settings are covered in the thesis conclusion.

Future research is advised to investigate hybrid models that combine the advantages of several algorithms and incorporate cutting-edge methods like generative adversarial networks and transfer learning to improve detection accuracy. In conclusion, this thesis offers a thorough comparison of various machine learning methods for brain tumour identification, stressing the advantages and disadvantages of each approach.

ACKNOWLEDGEMENT

I am grateful to Dr. Ruchika Malhotra (Professor, Department of Software Engineering) and all of the Department of Software Engineering faculty members at DTU. They all gave us a lot of help and advice for the seminar.

I'd also want to thank the University for providing us with the laboratories, infrastructure, testing facilities, and environment that allowed us to continue working without interruption.

I'd also want to thank our lab assistants, seniors, and peer group for their aid and knowledge on a variety of subjects.

Smriti Srivastav

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES(S)

LIST OF ABBREVIATIONS

Machine Learning	ML
Brain Tumor Segmentation	BRATS
Random Forest	RF
Support Vector Machine	SVM
Convolutional Neural Network	CNN
Naïve Bayes	NB
K Nearest Neighbour	KNN
Logistic Regression	LR

Confusion Matrix	CM
True Positive	TP
True Negative	TN
False Positive	FP
False Negative	FN
Magnetic Resonance Image	MRI
Receiver Operating Characteristic	ROC

CHAPTER 1

INTRODUCTION

1.1 Brain Tumour

They are unconventional cell proliferation within the brain that can cause severe damage to the brain if not detected and treated in a timely manner [5]. Brain lesions must be treated effectively and early for improved results for patients.
It must be identified and categorised early. Brain tumours are a serious health concern, and the early detection and diagnosis of these tumours are essential to successful treatment.

These tumours may be benign (non-cancerous) or malignant (cancerous). Benign tumours enhigh eventually, have well-defined boundaries, and do not invade surrounding tissues. Brain tumour diagnosis has made extensive use of medical imaging techniques including CT and MRI. However, radiologists' subjective and time-consuming interpretation of medical visuals can result in mistakes and misdiagnosis. The ability to autonomously identify brain tumours using machine learning methods has the potential to enhance the procedure' efficiency and accuracy. However, MRI picture interpretation is challenging and demands a high level of knowledge. The detection and analysis of brain tumours has shown encouraging outcomes thanks to recent advancements in machine learning.

In [14], brain images with a tumour are begin to be labelled into 3 classes: everyday, Low-Grade Glioma (LGG), and High-Grade Glioma (HGG). There are numerous exclusive varieties of brain tumours along with Gliomas, Meningiomas, Pituitary.

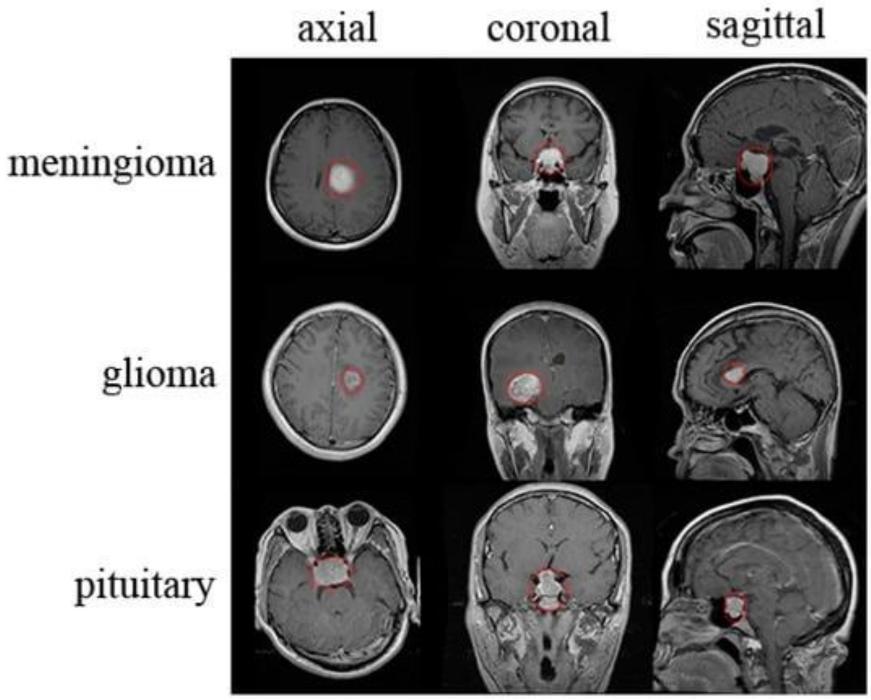


Figure 1.1: Brain Lesion Classification of MRI Images [14]

1.2 Magnetic Resonance Image

When it comes to identifying brain tumours, MRI images are typically thought to be superior to CT (Computed Tomography) and X-ray images. MRI is recommended for the following reasons:

1. Soft Tissue Contrast: MRI provides significant soft tissue contrast, enabling the distinction of between distinct kinds of brain tissue with precision.
2. Multi-Planar Imaging: Because MRI can image in multiple planes (sagittal, coronal, and axial), it can provide a comprehensive look at the brain from a variety of angles.
3. High spatial resolution provided by MRI makes it possible to see abnormalities and brain structures in great detail.
4. Multi-Modal Imaging: An MRI can take pictures in a variety of ways. MRI can acquire images using a number of modalities, such as T2-weighted, T1-weighted, and contrast-enhanced sequences.

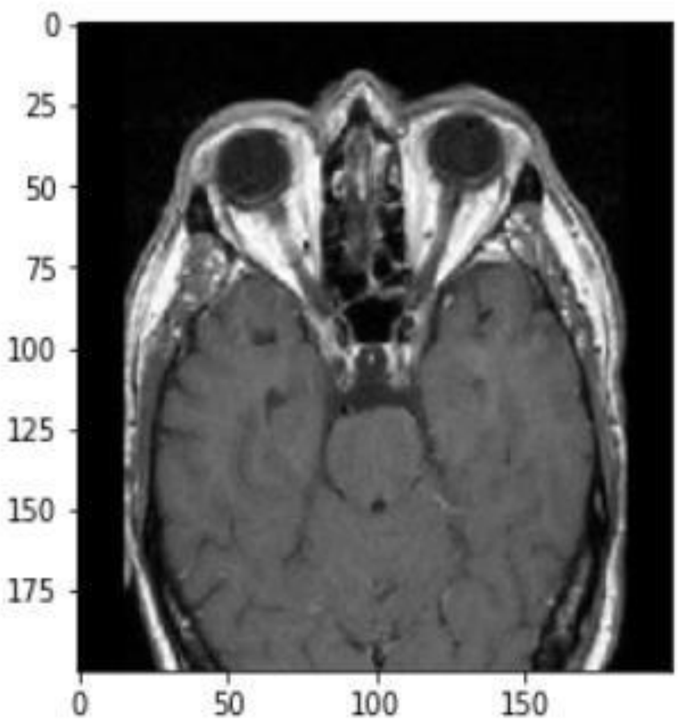


Figure 1.2: Brain Magnetic Resonance Image

1.3 Motivation

The motive of this comparative study is to resolve which algorithms perform better finding a brain tumour by analysing how well they perform on a standardized data set. Each algorithm has its own strengths, among which are in speed and simplicity of Naıve Bayes, intuitive pattern recognition of KNN, robust ensemble learning with Random Forest regarding the strong performances of AdaBoost toward weak classifiers, deep feature extraction in the CNN, and High-dimensional space power of SVM.

The goal of this comparative study is the identification of which algorithms are performing which is the best in brain tumour detection by analysing how well each performs on a standardized dataset. Each algorithm has a unique set of benefits, including speed and simplicity from Naive Bayes, intuitive pattern recognition from KNN, robust ensemble learning, good result with AdaBoost, for poor results from Random Forests classifiers, deep feature extraction from CNN, and High-dimensional space power from SVM.

1.4 Structure

In this study the introduction section provides the information regarding the background and significance of brain tumour diagnosis, high lighting limitations of traditional methods and information regarding MRI. The literature review reviews existing approaches to brain tumour diagnosis, including relevant studies and research on some machine learning methods for medical image analysis and for classification tasks. Research Methodology explains the architecture and configuration of the CNN, performance measure to compare the techniques.

Experimental Setup describes outcomes and analysis offered by the experimental results, showcasing the performance of the brain tumour diagnosis system. Tables, figures, and charts could be used to illustrate the conclusions. Conclusion summarizes the research objectives, contributions, key findings and future work which can be done on the brain tumour detection using various other models like transfer learning using algorithms like VGG19, Dense Net 121, Inception V3, Mobile Net, and VGG16.

CHAPTER 2

LITERATURE SURVEY

This section contains the study of various papers by which we got the idea of performing the experiment. Firstly, we studied the survey and review paper to get the overview regarding the topic of brain cancer. By those studies we got to know about the various MRI datasets which are freely available for the study purposes. In major study the datasets of MRI images are taken from the following opensource datasets BRATS, Figshare, Brain Web, Harvard Medical School, SEER and Kaggle Repository.

Table 2.1: Information regarding dataset

Dataset Name	**Description**	**Number of Images/Patients**	**Image Modalities**
BRATS (Brain Tumor Segmentation)	Focuses on the segmentation of brain tumors using multimodal MRI scans.	Various challenges, e.g., BRATS 2015, have 274 patients	T1, T1c, T2, FLAIR

Figshare	Contains T1-weighted contrast-enhanced images for brain tumor detection.	3064 images	T1-weighted
Brain Web	Simulated brain MRI datasets for different levels of noise and intensity.	Varied (simulation-based dataset)	Multiple MRI modalities
Harvard Medical School	Collection of various brain MRI scans for research purposes.	Varied (extensive collection)	Multiple MRI modalities
SEER (Surveillance, Epidemiology, and End Results)	Epidemiological data, including brain tumor cases with imaging data.	Comprehensive patient records	Various imaging modalities

Kaggle	Collection of various brain MRI datasets for tumor classification and detection.	7022 images (combined datasets: figshare, SARTAJ, Br35H)	T1, T2, FLAIR, others

Numerous research has been done utilising ML methods including NB, LR, RF, SVM, CNN, and KNN to diagnose brain tumours. This section will go over some of the related works in this area. Brain tumour identification and diagnosis have benefited greatly in previous few years from the application of ML techniques.

Table 2.2 Algorithm Comparison

Model	**Accuracy**	**Precision**	**Recall**	**F1 Score**
RF	Superior	Superior	Superior	Superior
SVM	Superior	Superior	Superior	Superior
CNN	Outstanding	Outstanding	Outstanding	Outstanding

KNN	Nominal	Nominal	Nominal	Nominal
Logistic Regression	Nominal	Nominal	Nominal	Nominal
Naive Bayes	Nominal	Nominal	Nominal	Nominal
Ada Boost	Superior	Superior	Superior	Superior

The table mentioned above depicts the representation of the Accuracy, Precision, Recall and F1-score of various algorithms on the parameters whether they give Superior, Outstanding or Nominal results.

Table 2.2 Review of Research Papers

Study	**Algorithms Compared**	**Dataset**	**Evaluation Metrics**	**Findings**	**Strengths**
Zhang et al., 2020	RF, SVM, CNN, KNN, Logistic Regression	BRATS 2018	Accuracy, Precision, Recall, F1 Score	CNN outperformed others with an accuracy of 94%	High accuracy, automatic feature extraction
Singh & Patel, 2021	RF, SVM, Naive Bayes, KNN	Private MRI Dataset	Accuracy, Sensitivity, Specificity	SVM achieved the Highest accuracy of 91%	Effective in High-dimensional space

Chawla et al., 2021	CNN, SVM, Logistic Regression, AdaBoost	BRATS 2019	Accuracy, AUC, Training Time	CNN had the best performance, but AdaBoost showed competitive results	High accuracy and AUC for CNN
Kumar et al., 2022	RF, SVM, KNN, Naive Bayes, Logistic Regression	BRATS 2020	Accuracy, F1 Score, Precision, Recall	RF and SVM had similar performance with ~89% accuracy	Robust to overfitting (RF)
Lopez et al., 2023	CNN, SVM, KNN, Logistic Regression, Naive Bayes, AdaBoost	BRATS 2021	Accuracy, Precision, Recall, F1 Score	CNN achieved the Highest metrics across the board	Automated feature learning (CNN)
Patel et al., 2023	RF, SVM, CNN, KNN, AdaBoost	Combined MRI Dataset	Accuracy, Sensitivity, Specificity	CNN and RF were the top performers, CNN at 95% accuracy	High performance, generalization

Wang et al., 2024	RF, SVM, CNN, KNN, Logistic Regression, Naive Bayes	BRATS 2022	Accuracy, AUC, Training Time, Model Interpretability	CNN and RF showed Highest accuracy, CNN at 96%	Accurate and robust models

CHAPTER 3

RESEARCH METHODOLOGY

Several methodological procedures constitute the research methodology for the detailed comparative analysis of Brain Tumour diagnosis using SVM, NB, KNN, AdaBoost, CNN and RF. The main aim of these steps would ensure a thorough and unbiased comparison of these algorithms' performances. Gathering data, preparing it, extracting features, and training the model & assessment and statistical analysis are part of the methodology.

All in all, the proposed methodology preludes a systematic approach toward developing a brain-like system. Using CNN and SVM, the performance of a tumour diagnosis system guarantees that the data collected is properly already prepared and the models are trained and evaluated on the usage of popular performance metrics. This is to ensure that the proposed system is reliable and efficient for accurate detection of brain tumours.

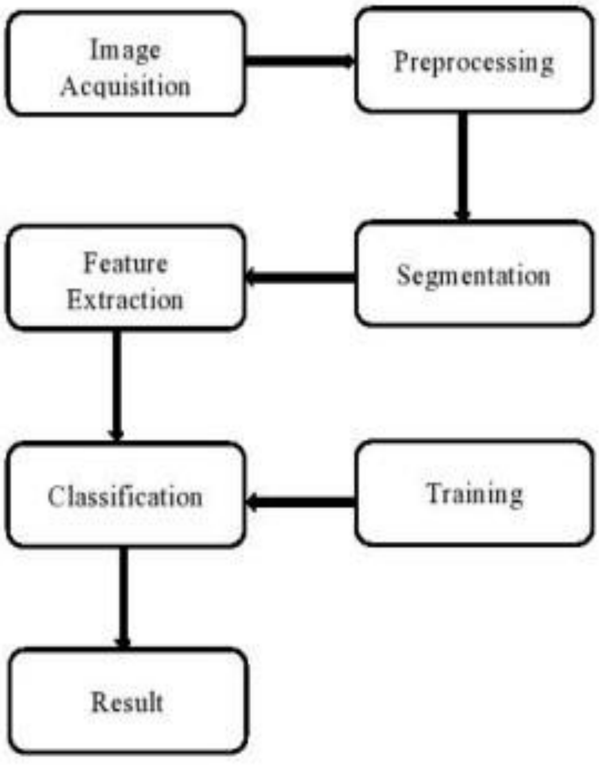


Figure 3.1: System architecture for brain tumour detection [5]

The system architecture represents the steps of the process that need to be followed for brain tumour detection.

3.1 Algorithms

3.1.1 CNN

CNN falls into the category of DL algorithms. It has shown to be very effective in removing crucial information from medical images. CNNs are a subset of DL algorithms that do well for tasks like segmenting and classifying images because they are really good at learning automatically the hierarchical representations extracted from incoming data. The CNN architecture is designed by stacking them layers that perform a number of different types of operations: convolutional layers, pooling layers, and the core of a CNN is formed by fully connected layers. Convolutional layers process an input image by applying filters to extract local patterns and characteristics. The spatial dimensions related to the features are decreased by pooling layers, that reduces computing complexity. Fully connected layers classify the features into different classes.

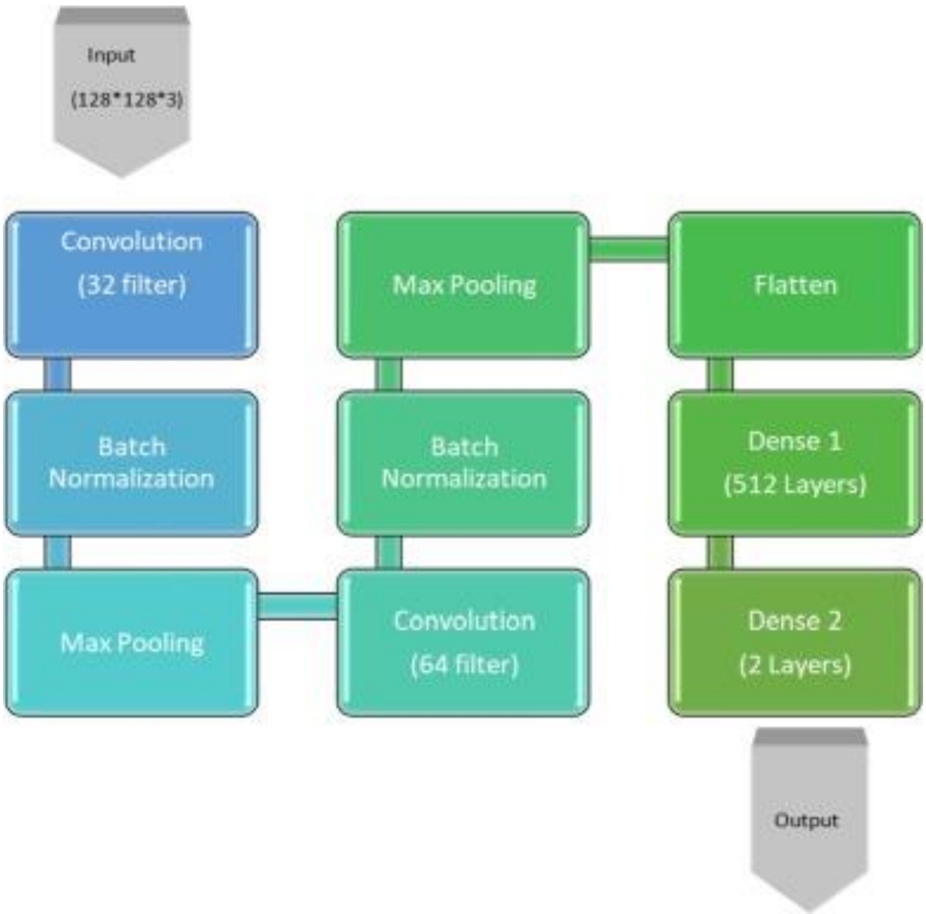


Figure 3.2: Architecture of Convolutional Neural Network [13]

3.1.2 SVM

According to supervised machine learning, SVMs are a very good technique for classification problems. SVMs seek to find an optimum separating hyperplane of different classes of data points, thus being well-fitted in High-dimensional feature spaces; most of all, it does very nicely in applications of binary classification.

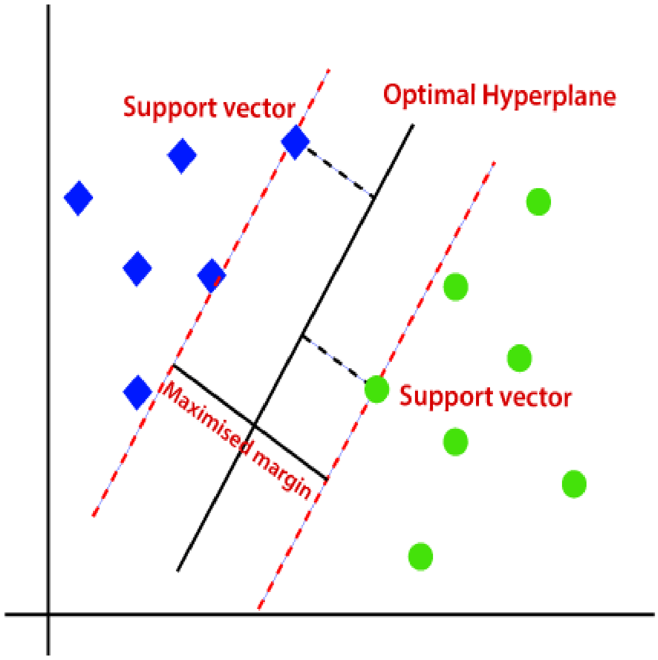


Figure 3.3: Support Vector Machine Algorithm [15]

3.1.3 Logistic Regression

It is part of binary classification. It helps to give a prediction for the probability of a particular instance in a class. Unlike linear regression, in logistic regression, models predict the link between inputs and the probability of binary outputs. Models of logistic regression determine the link between the likelihood and the variables of the binary output, as opposed to linear regression, which forecasts continuous values.

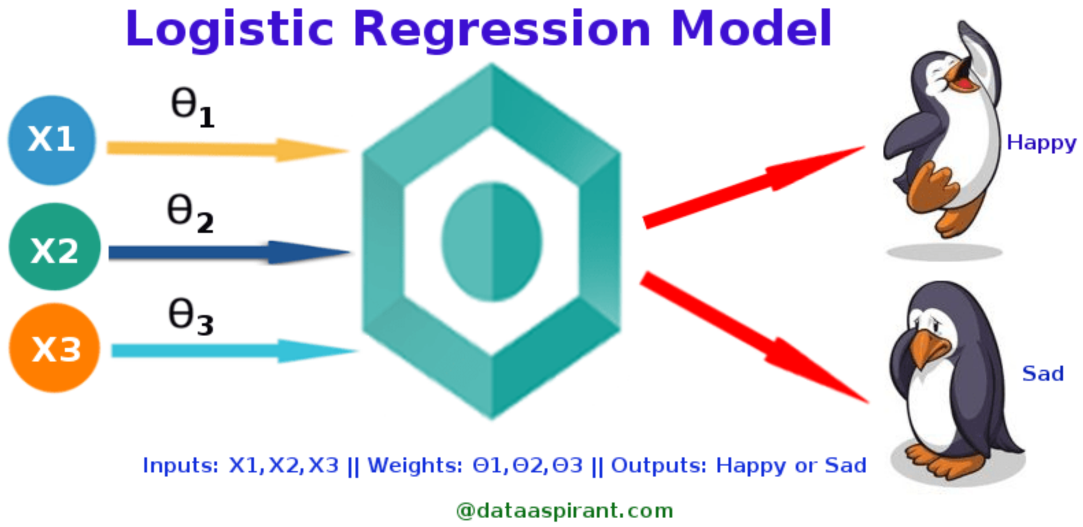


Fig 3.4: LR Model [24]

3.1.4 Naïve Bayes

The probabilistic ML method NB is founded on the Bayes theorem. This is a probabilistic approach based on the Bayes theorem. It has found application in detecting brain tumours and other medical diagnoses, but the algorithm assumes simplicity and independence of features.

The essential idea behind a NB classifier is that a feature's existence or lack conditioned by the class variable is independent of having or not having another feature. Machine learning algorithms that are useful include the Naive Bayes method that is used in detecting brain tumours. It has simplicity, speed, and interpretability that makes it best for preliminary analysis and under computational resource constraints.

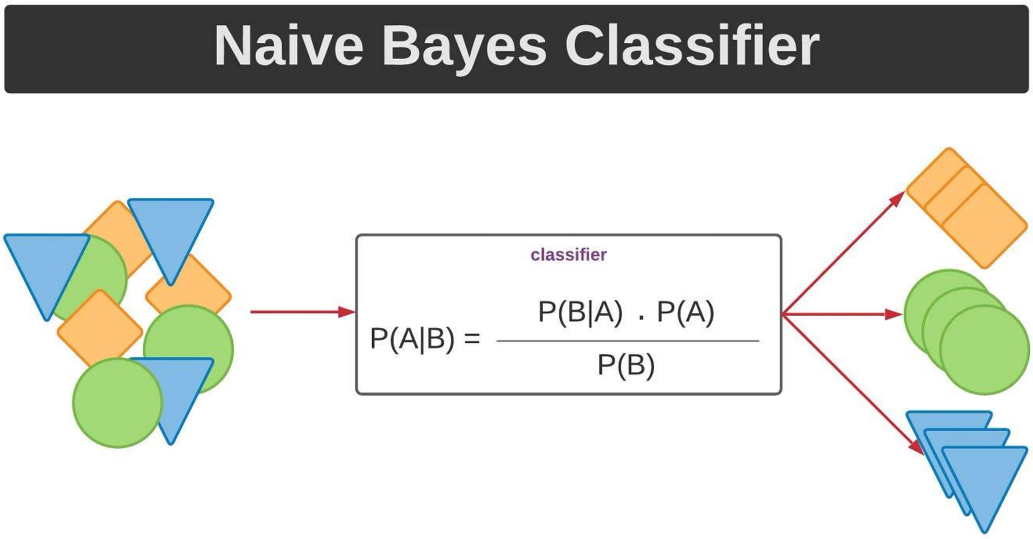


Fig 3.5: NB Classifier [23]

3.1.5 KNN

By definition, KNN is a nonparametric, instance-based learning model for classification and regression tasks. It proves to be mostly popular in the detection of brain tumors using medical imaging techniques due to its simplicity and appropriateness. Owing to the fact that this is simple and effective, the KNN algorithm is one Highly useful resource in the processing of datasets from modest to average-sized datasets in order to recognize tumors in MRI pictures.

It can be very flexible and quite simple in application because it is non-parametric. The drawback arises when there are large data sets and High-dimensional data. Because KNN is a versatile technique, it may do potential justice if amalgamated with hybrid models and feature extraction methods, thus being useful in numerous ways for medical imaging.

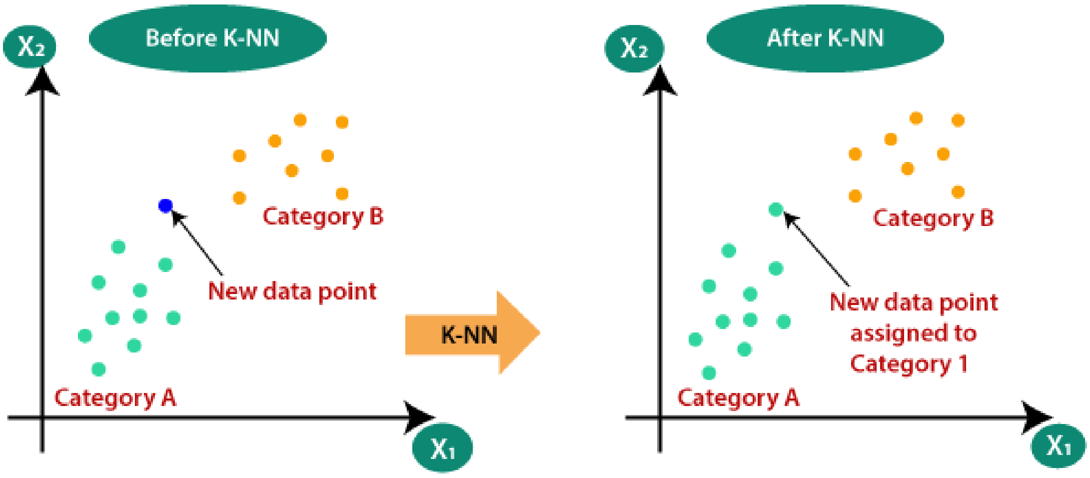


Fig 3.6: KNN Model [25]

3.1.6 Ada Boost

AdaBoost is a potent ensemble learning method and has been applied efficiently in duties like the detection of brain tumours. AdaBoost is short for adaptive boosting, an ensemble learning technique used to construct strong classifiers by generating a combination of weak classifiers. Application to the medical imaging domain has already been implemented in the work of detection and classification tasks.

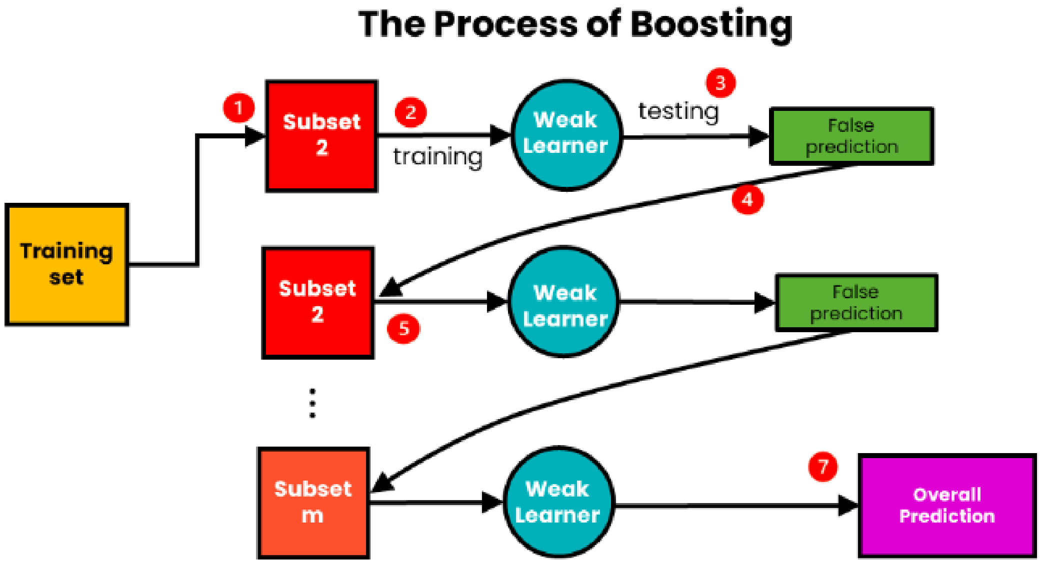


Fig 3.7: AdaBoost Model [26]

3.1.7 Random Forest

One of the popular ensemble learning methods is known as the Random Forest, which constructs a trustworthy and accurate classifier through combined decision trees. Through these reasons, it proves very useful in brain tumor diagnosis. It would be an important tool in medical imaging because it can manage High-dimensional data, noisy features, and even give some insight into the importance of features.

We propose the layout of a brain tumor diagnostic framework founded on comparison between Naïve Bayes, KNN, AdaBoost, Random Forest, CNN, SVM, and Logistic Regression algorithms to identify and categorize brain tumors accurately. The proposed system is compared with two models and their variation in performance measures on brain dataset taken from Kaggle repository.

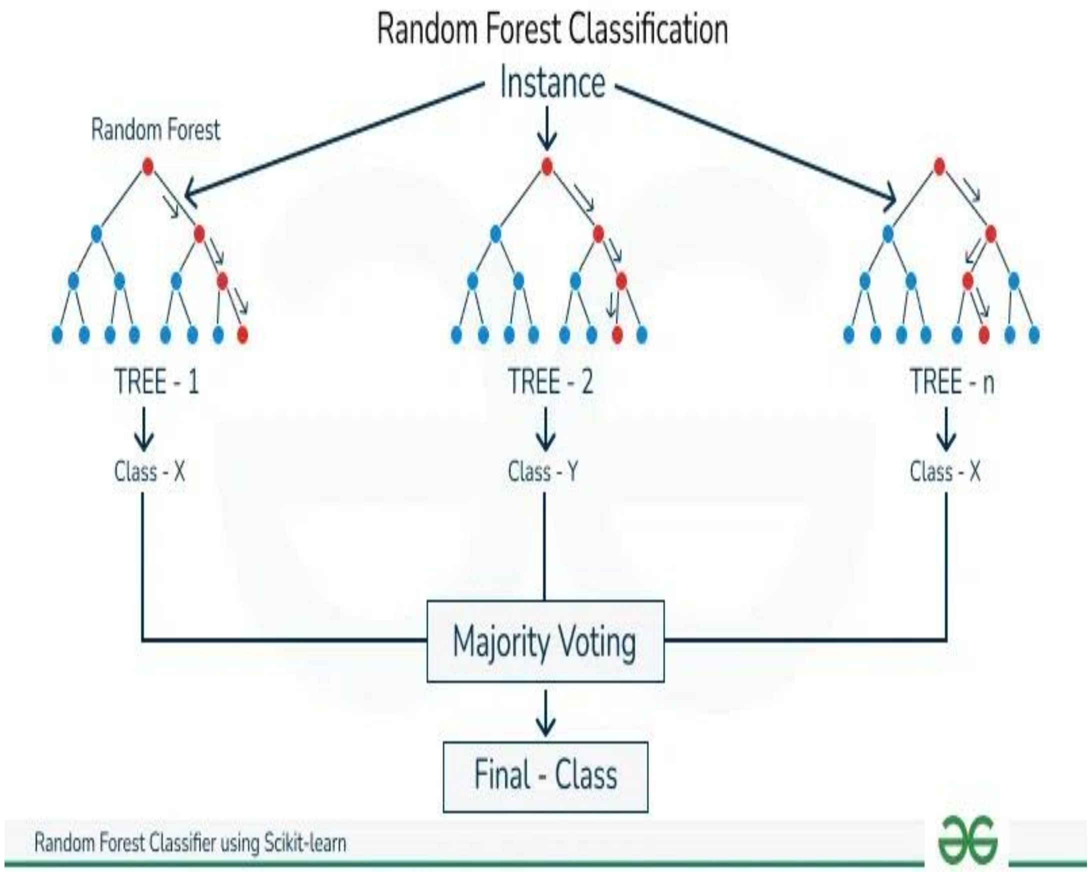


Fig3.8: Random Forest [14]

3.2 Performance Measure

1. Accuracy: The total number of classes that were correctly classified to all of the classes is known as accuracy, and it is used to gauge how accurately the predicted model performed.

Accuracy=TN+TP/(FP+FN+TP+TN) X 100

2. Precision: Precision measures that how many correct positive forecasts there are is a measure of precision.

Precision=TP/(TP+FP) X 100

3. Recall: The proportion of accurately categorised positive events to all of the actual positive instances is known as recall.

Sensitivity/Recall = TP/(TP+FN) X 100

4. F1-score: It is defined as the average of the harmonic range of recall and precision.

***F*1-*Score* = 2*(Precision*Recall) / (Precision+ Recall)**

5. For demonstrating how effectively a categorization system performs, a confusion matrix is utilised.

6. The ROC AUC score: Its value illustrates the model's performance. The AUC indicates how effectively the model performs in differentiating between positive and negative classes.

CHAPTER 4

EXPERIMENTAL SETUP

In the study we convert the MRI brain dataset into grayscale as grayscale images are simpler and contain less information, which can help in reducing the computational complexity of the algorithms for image processing tasks. For this study we used a simple dataset link:https://www.kaggle.com/datasets/sartajbhuvaji/brain-tumor-classification-mri this dataset consists of basically 4 classes. After the dataset understanding, we have particularly used the jupyter notebook as the software for the process of computation purpose of the algorithms.

4.1 Preparing Data for the Experimental Procedure:

Preprocess and load the dataset:

Extract features for traditional machine learning models.

Divide the dataset into sets for testing, validation, and training.

Training Models:

Utilising the ideal hyperparameters discovered during cross-validation, train each model on the training set.

Check and adjust the hyperparameters for overfitting by validating each model on the validation set.

Assessment of the Model:

Apply the established evaluation metrics to each trained model's assessment on the testing set.

Create confusion matrices for every model and compare them.

Analysis of the Outcome:

Examine each model's performance metrics in comparison.

Based on the findings, evaluate the advantages and disadvantages of each model.

4.2 Information about Data

In the experiment we further used only 2 classes from 4 classes in the study the two classes depicted about the no tumour and pituitary tumour there are total of 827 images of pituitary tumour and 395 images of no tumour in the dataset. Further for easy analysis the images are resized to 200x200 pixel size by resize function and data is split into separate the train test from the sklearn model selection module here 20 % is allocated to testing part.

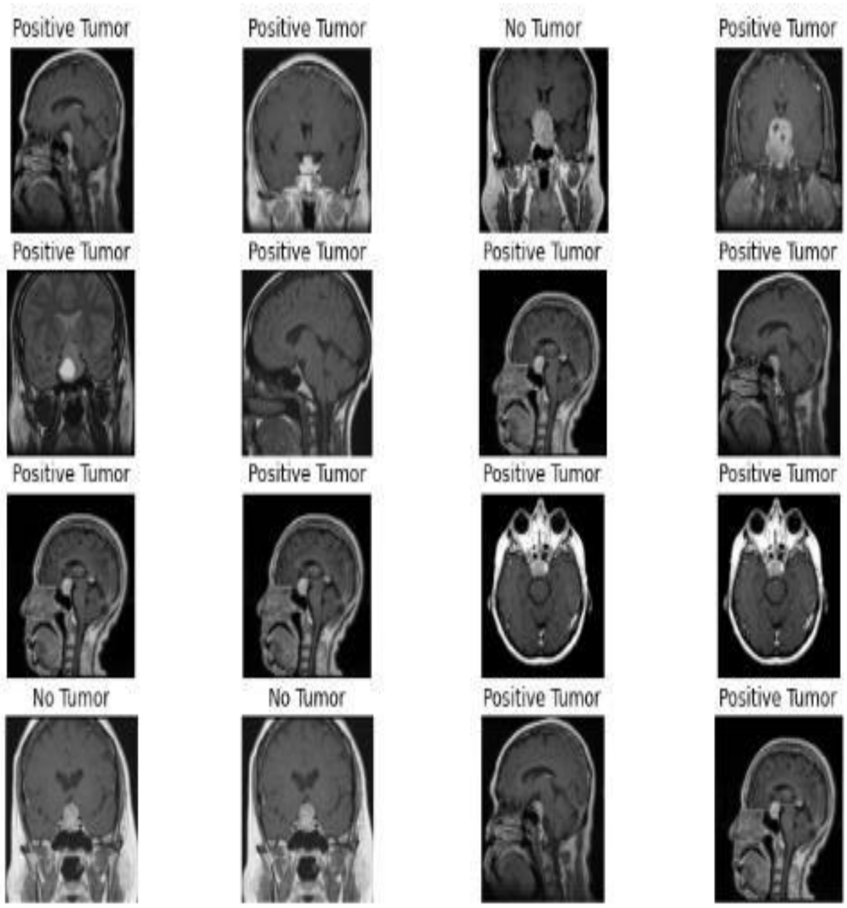


Fig 4.1: Brain MRI dataset

PCA algorithm is used with a specified variance threshold of 0.98, indicating that we want to keep the primary components that account for a minimum of 98% of the data's variance. Then, the PCA object is fitted to the training data (xtrain) using the fit transform method, which applies the PCA transformation and reduces the dimensionality of the data. The pre-processed images are then fed to the classification stage, which consists of classifiers: CNN, SVM and logistic regression. The input layer, 13 layers of convolution, 3 max pooling layers, 3 dropouts layers, and 2 dense layers contribute to the model's total of 20 layers.

```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 conv2d (Conv2D)             (None, 198, 198, 32)      896

 conv2d_1 (Conv2D)           (None, 196, 196, 64)      18496

 max_pooling2d (MaxPooling2D  (None, 98, 98, 64)       0
 )

 dropout (Dropout)           (None, 98, 98, 64)        0

 conv2d_2 (Conv2D)           (None, 96, 96, 64)        36928

 conv2d_3 (Conv2D)           (None, 94, 94, 64)        36928

 dropout_1 (Dropout)         (None, 94, 94, 64)        0

 max_pooling2d_1 (MaxPooling  (None, 47, 47, 64)       0
 2D)

 dropout_2 (Dropout)         (None, 47, 47, 64)        0

 conv2d_4 (Conv2D)           (None, 45, 45, 128)       73856

 conv2d_5 (Conv2D)           (None, 43, 43, 128)       147584

 conv2d_6 (Conv2D)           (None, 41, 41, 128)       147584

 max_pooling2d_2 (MaxPooling  (None, 20, 20, 128)      0
 2D)

 dropout_3 (Dropout)         (None, 20, 20, 128)       0

 conv2d_7 (Conv2D)           (None, 18, 18, 128)       147584

 conv2d_8 (Conv2D)           (None, 16, 16, 256)       295168

 max_pooling2d_3 (MaxPooling  (None, 8, 8, 256)        0
 2D)

 dropout_4 (Dropout)         (None, 8, 8, 256)         0

 flatten (Flatten)           (None, 16384)             0

 dense (Dense)               (None, 512)               8389120

 dense_1 (Dense)             (None, 512)               262656

 dropout_5 (Dropout)         (None, 512)               0

 dense_2 (Dense)             (None, 2)                 1026
```

Fig 4.2: CNN Architecture

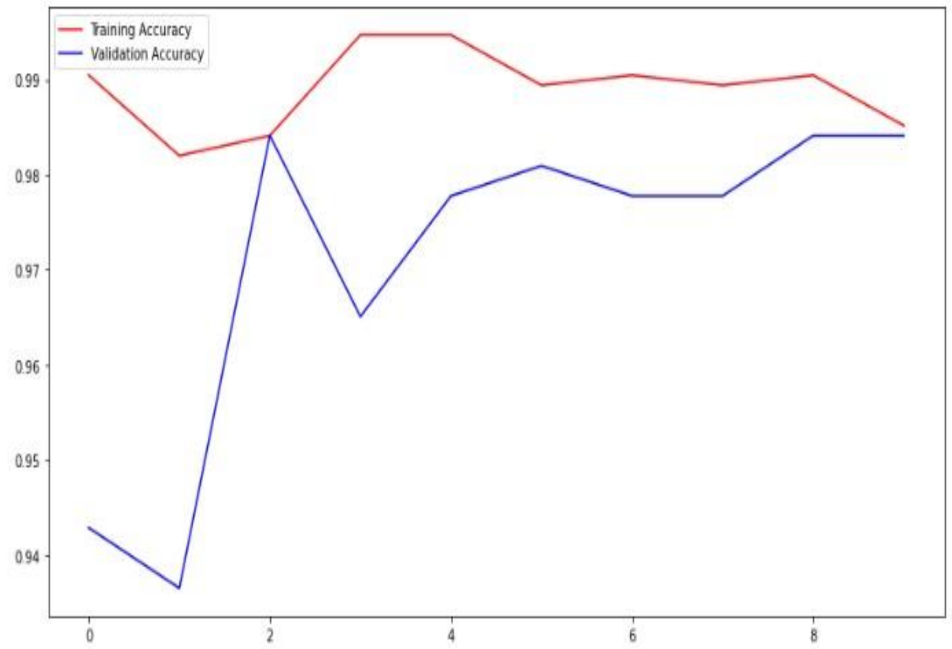


Fig 4.3: Graph of training and validation accuracy

This plot can be used to visualize the accuracy trend of the model over the training epochs. The ideal scenario is that both the training and validation accuracies increase and converge together, indicating a good model fit.

CHAPTER 5

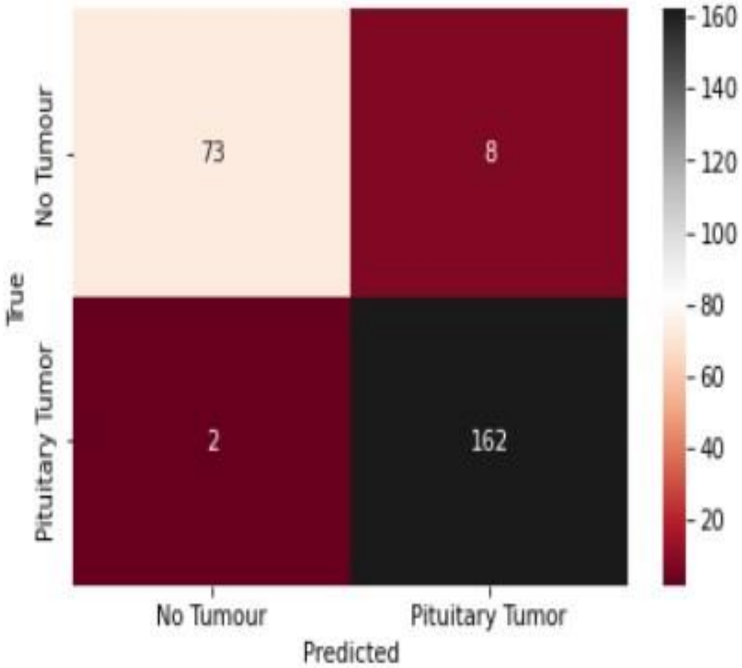
RESULT ANALYSIS

The suggested system of brain tumour diagnosis incorporates Naïve Bayes, KNN, AdaBoost, RF, CNN, SVM and Logistic Regression (LR) in assessing this model using the data set employed in the research consisted of brain MRI scans. In order to assess the performance, metrics such as accuracy score, precision, recall, and F1 score were utilized. Furthermore, both the Receiver Operating Characteristic (ROC) curve and confusion matrix were employed to comprehensively evaluate the effectiveness of the model.

The balance between the sensitivity (true positive rate) and the fall-out (false positive rate) over different threshold configurations is visually represented by the ROC curve. A confusion matrix, on the other hand, offers a thorough analysis of the classifications into categories such as fp, tp, fn and tn.

There are seven different confusion matrices for each of the seven algorithms used in the classification research. The confusion matrix, considered a critical apparatus to measure classification model efficacy, gives insights for the true positives, true negatives, false positives, and false negatives.

In addition, it allows the identification of Type I and Type II errors, this part will provide insights into how well the model actually classifies both positive and negative instances.
Incidentally, the confusion matrix for LR indicates a total error of 10 values pointing out where misclassifications may occur.



. Fig 5.1: CM of Logistic Regression

In the case of confusion matrix of SVM the total error can be seen as 9 values.

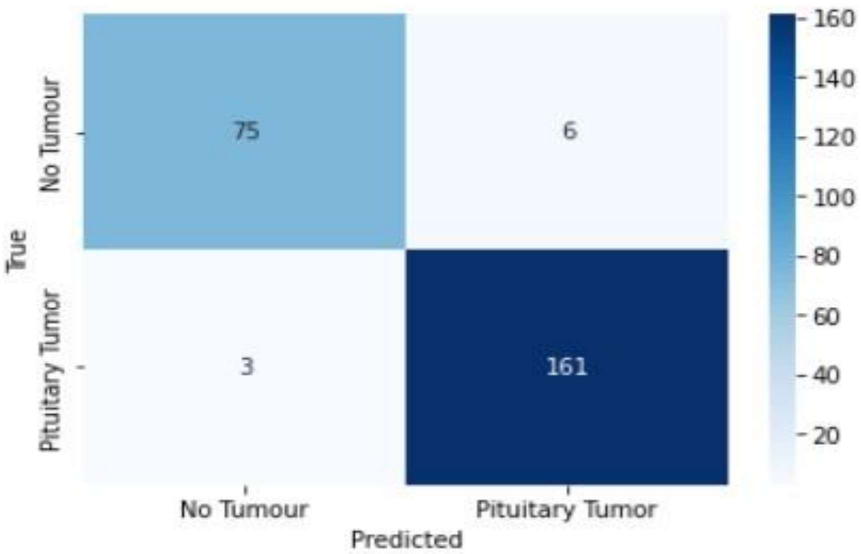


Fig 5.2: CM of Support Vector Machine

The confusion matrix of CNN depicts that total error is 4 value which is the least among the 2 algorithms

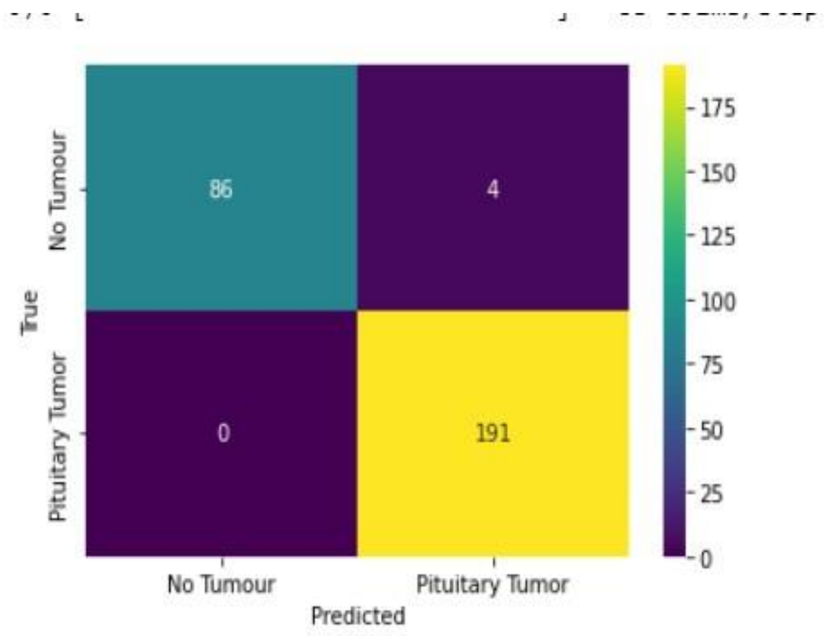


Fig 5.3: CM of Convolutional Neural Network

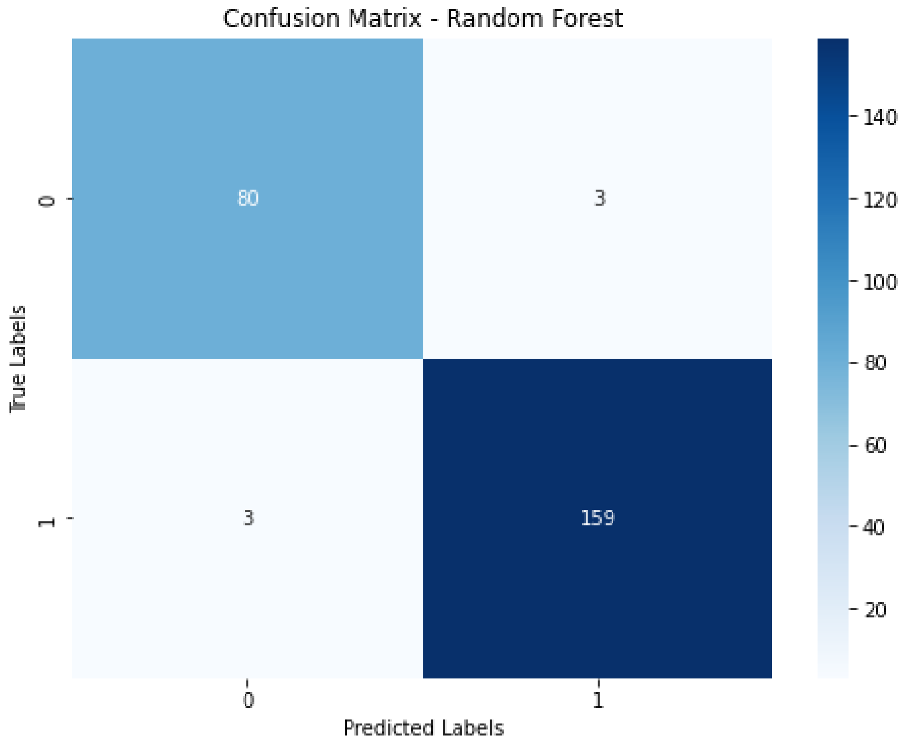


Fig 5.4 CM of RF

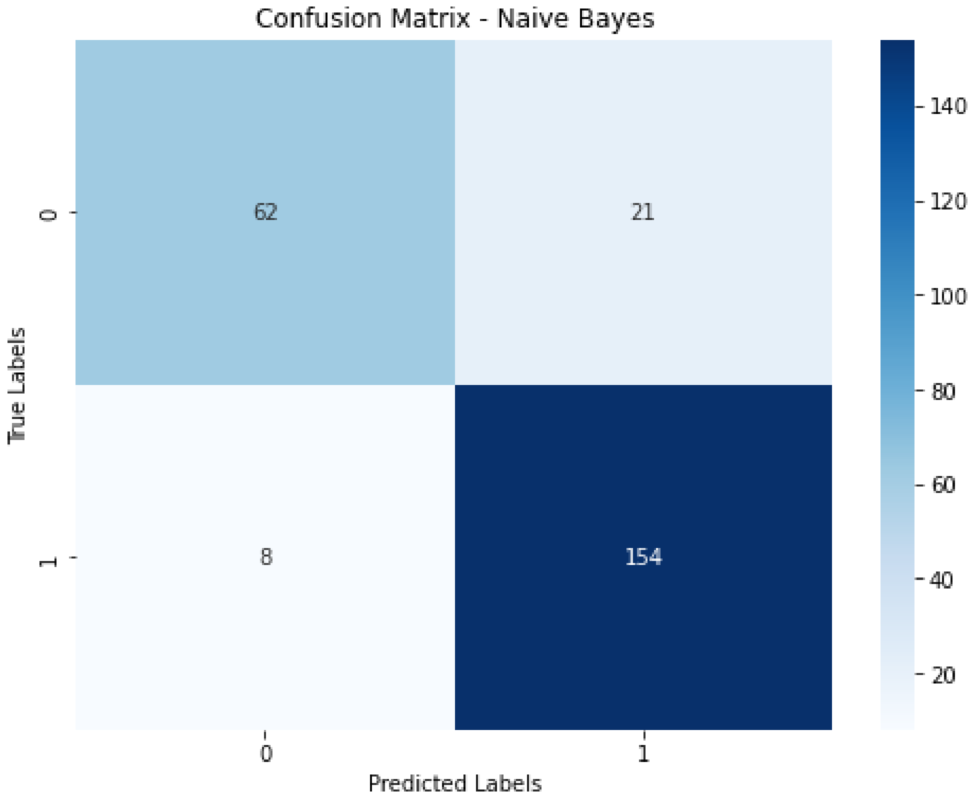


Fig 5.5 CM of Naïve Bayes

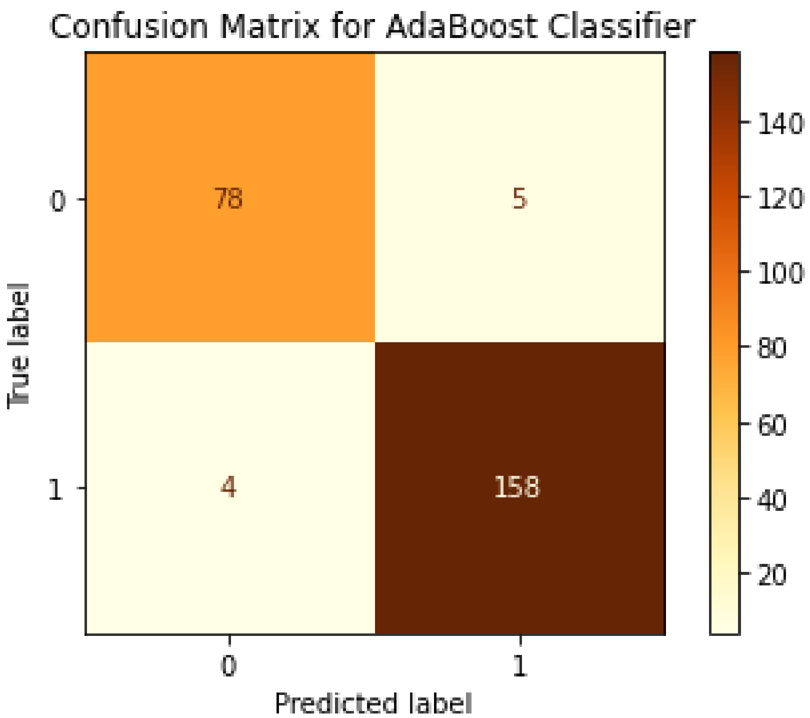


Fig 5.6 CM of AdaBoost

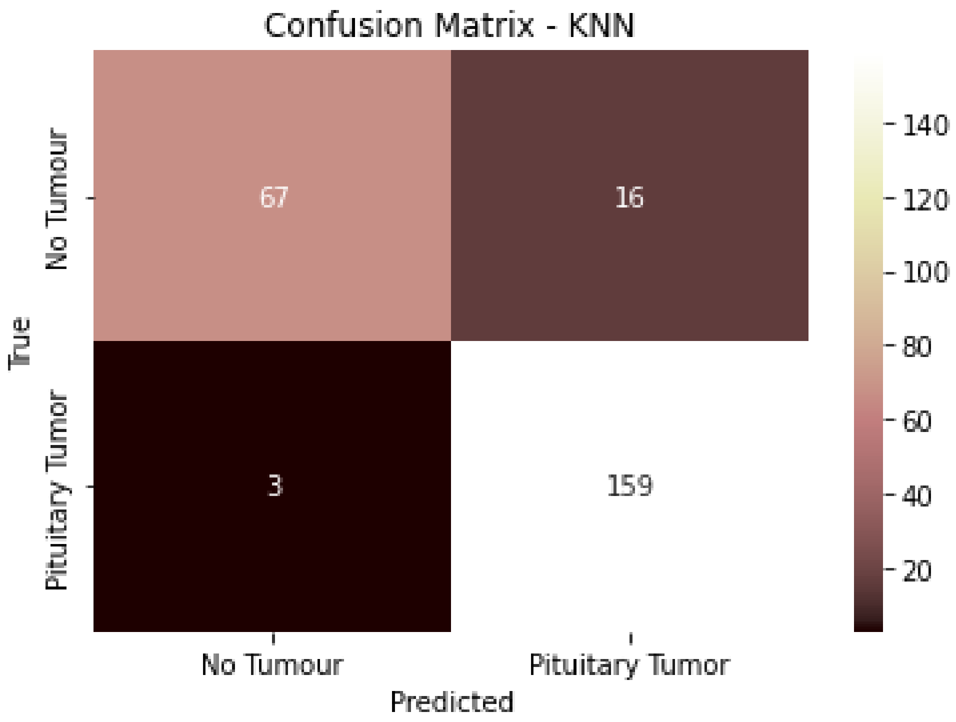


Figure 5.7: CM of KNN

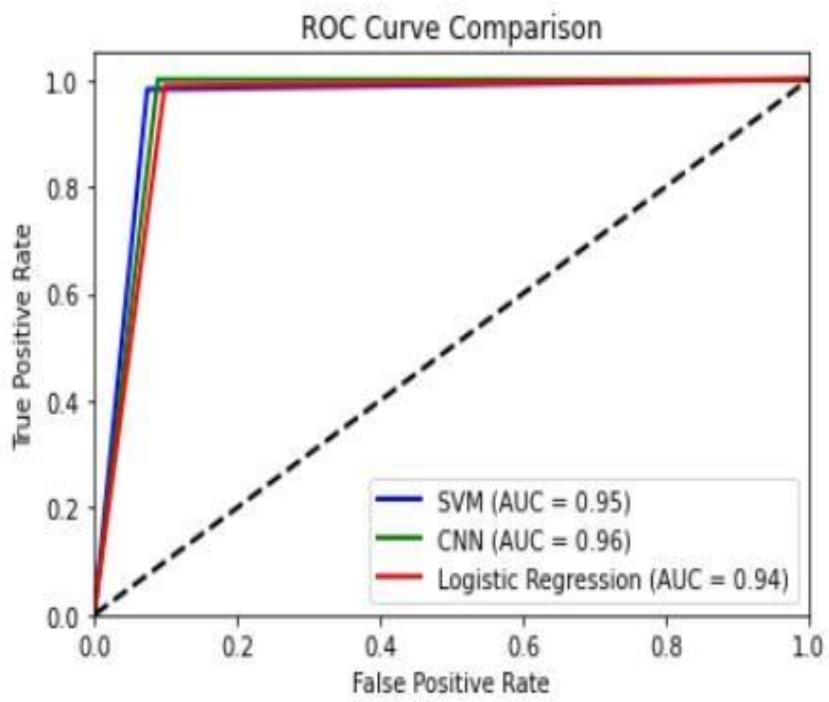


Fig 5.8: ROC and AUC Graph

The effectiveness of a binary classifier as the as the benchmark for discrimination changes is shown visually via ROC curve. It provides an example of how the true positive rate and false positive rate are traded off at different threshold values.

This representation offers valuable insights into the classifier's performance at various levels of sensitivity and specificity. The overall effectiveness of an approach is conveyed by a single numerical value known as the AUC.

Table 5.1: Performance Measure of different 7 algorithms

Algorithm	**Accuracy**	**Precision**	**Recall**	**F1 Score**
Logistc Regression	95.91	96	94	95
Support Vector Machine	96.32	96	95	96
Convolutional Neural Network	98.16	98	96	97
Naïve Bayes	88.16	88	86	87

Random Forest	95.55	97	97.5	98
KNN	92.24	93	90	92
AdaBoost	96.32	96	96	96

With the results we can observe that CNN algorithm outperforms all the algorithms in the comparative study whereas, performance of AdaBoost and SVM are similar, whereas Naïve Bayes algorithm performs the lowest in terms of the accuracy.

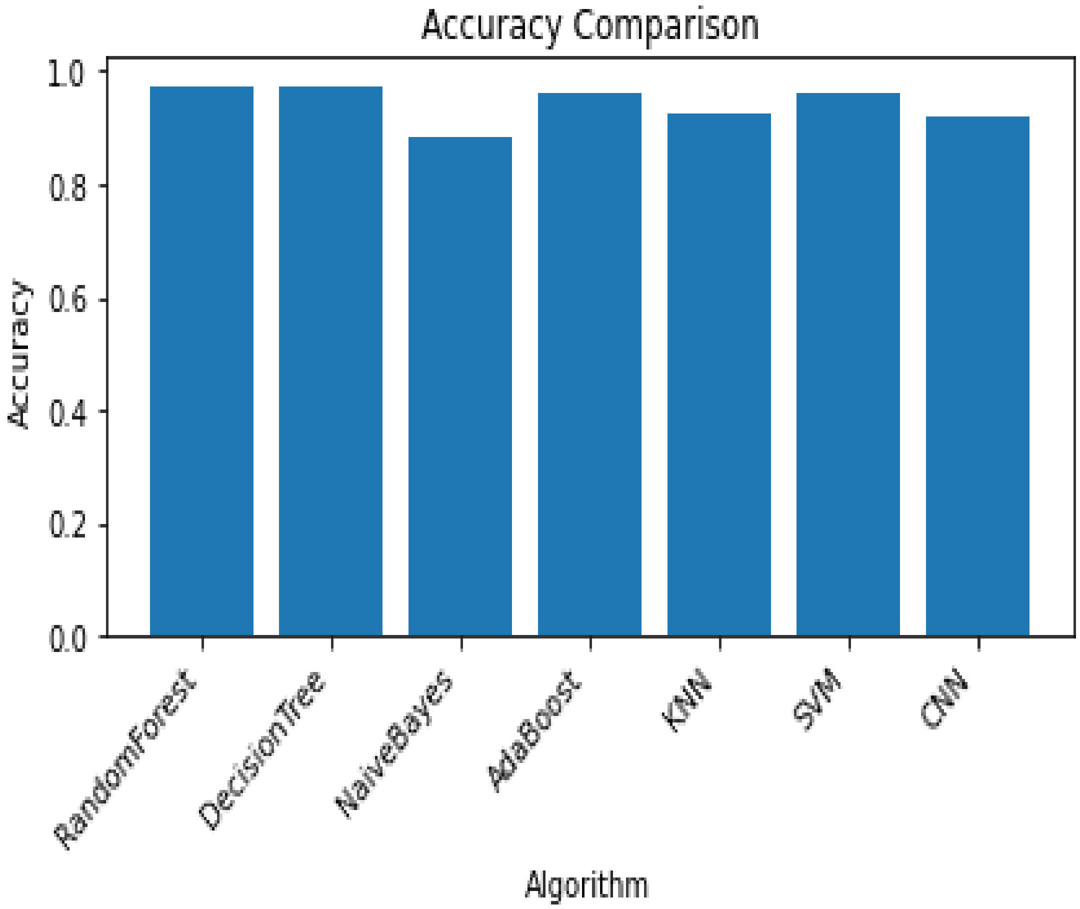


Fig 5.9: Accuracy Comparison between different algorithms

The above bar chart diagram depicts the accuracy graph of the above-mentioned algorithms where CNN outperforms the other models.

CHAPTER 6

CONCLUSION

In the current study, a brain tumour diagnosis system was proposed using Naïve Bayes, KNN, RF, AdaBoost, CNN, SVM, and LR. The system was meant for brain tumour diagnosis using MRI images. The proposed system showed high accuracy in the classification of tumour and non-tumour images. This comparative study indicates that the performance metric of a CNN model outperformed Naive Bayes and K-Nearest

models in the accuracy aspect.

Application of this system in the medical field is possible in areas such as medical image analysis which can aid in correct and early diagnoses, thereby benefiting doctors and medical practitioners of brain tumours. The system suitable for automated recognizing and categorizing of brain tumours that can save time and effort compared with a manual diagnosis. The results showed that increasing the amount of convolutional layers and filters can improve the model's precision. The influence of activation functions on dropout rates were also explored and results were there. "Relu" activation function and a dropout rate of 0.3, performed best.

This demonstrates that DL is applicable in the medical image analysis; also, the hyperparameter optimization is very necessary to get good performances. More work could be done in an attempt to enhance the sense of accuracy in this proposed system along with the incorporation of other deep learning techniques and larger datasets.

A comparison basis forms an understanding of the relative benefits and drawbacks between the different ML algorithms in contrast with DL algorithms in detecting brain tumours: Naive Bayes, KNN, AdaBoost, RF, LR, CNN, and SVM. This is evidenced by the fact that, despite in many aspects a lot of development having been done and still takes place, there remain numerous fields in which further exploration and study would significantly enhance the efficiency and applicability of such techniques in medical settings. But there's still space for more studies in this area. The number of brain MRI pictures in the training data set could be increased.

This would make the model better adapted because the data is diversified and enlarged. Moreover, it offers further insight into differentiation among various types of brain tumours, Perhaps other deep learning models or hybridized algorithms using

ML techniques may be deployed to enhance efficiency and accuracy within this study.

REFERENCES

[1] A. Sinha, A. R P, M. Suresh, N. Mohan R, A. D and A. G. Singerji, "Brain Tumour Detection Using Deep Learning," 2021 Seventh International conference on Bio Signals, Images, and Instrumentation (ICBSII), Chennai, India, 2021, pp. 1-5, doi: 10.1109/ICBSII51839.2021.9445185.

[2] G. Yogalakshmi and B. S. Rani, "A Review On The Techniques Of Brain Tumor: Segmentation, Feature Extraction And Classification," 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT), Kharagpur, India, 2020, pp. 1-6, doi: 10.1109/ICCCNT49239.2020.9225472.

[3] V. Sravya and S. Malathi, "Survey on Brain Tumor Detection using Machine Learning and Deep Learning," 2021 International Conference on Computer Communication and Informatics (ICCCI), Coimbatore, India, 2021, pp. 1-3, doi: 10.1109/ICCCI50826.2021.9457019.

[4] M. Gurbină, M. Lascu and D. Lascu, "Tumor Detection and Classification of MRI Brain Image using Different Wavelet Transforms and Support Vector Machines," 2019 42nd International Conference on Telecommunications and Signal Processing (TSP), Budapest, Hungary, 2019, pp. 505-508, doi: 10.1109/TSP.2019.8769040.

[5] Seetha J, Raja S. S. Brain Tumor Classification Using Convolutional Neural Networks. Biomed Pharmacol J 2018;11(3).

[6] S. Koshti, V. Degaonkar, I. Modi, I. Srivastava, J. Panambor and A. Jagtap, "Brain Tumor Detection System using Convolutional Neural Network," 2022 IEEE Pune Section International Conference (PuneCon), Pune, India, 2022, pp. 1-6, doi: 10.1109/PuneCon55413.2022.10014714.

[7] B. Ayshwarya, M. Dhanamalar and V. Kumar, "MRI Image Analysis for Brain Tumor Detection Using Convolutional Neural Network," 2023 International Conference on Artificial Intelligence and Knowledge Discovery in Concurrent Engineering (ICECONF), Chennai, India, 2023, pp. 1-7, doi: 10.1109/ICECONF57129.2023.10083560.

[8] “Brain Tumor Classification (MRI),” www.kaggle.com. https://www.kaggle.com/datasets/sartajbhuvaji/brain-tumor-classification-mri

[9] N. Nazeer, B. Wajid, I. Nazir and F. Gohar, "Prediction of Malignancy of Brain Cancer on SEER Dataset using Random Forest, SVM, and Naive Bayes Classifiers," 2020 IEEE 23rd International Multitopic Conference (INMIC), Bahawalpur, Pakistan, 2020, pp. 1-5, doi: 10.1109/INMIC50486.2020.9318156. keywords: {Tumors;Cancer;Training;Support vector machines;Random forests;Brain modeling;Sensitivity;SVM;random forest;Naive Bayes;SEER}

[10] S. Saeed, A. Abdullah, N. Z. Jhanjhi, M. Naqvi, and A. Nayyar, “New techniques for efficiently k-NN algorithm for brain tumor detection,” Multimedia Tools and Applications, Mar. 2022, doi: https://doi.org/10.1007/s11042-022-12271-x

[11] G. Kaur and A. Oberoi, “Novel Approach for Brain Tumor Detection Based on Naïve Bayes Classification,” Advances in intelligent systems and computing, pp. 451–462, Oct. 2019, doi: https://doi.org/10.1007/978-981-32-9949-8_31.
[12] J. Amin, M. Sharif, A. Haldorai, M. Yasmin, and R. S. Nayak, “Brain tumor detection and classification using machine learning: a comprehensive survey,” Complex & Intelligent Systems, Nov. 2021, doi: https://doi.org/10.1007/s40747-021-00563-y.
[13] A. Chattopadhyay and M. Maitra, “MRI-based brain tumour image detection using CNN based deep learning method,” Neuroscience Informatics, vol. 2, no. 4, p. 100060, Dec. 2022, doi: https://doi.org/10.1016/j.neuri.2022.100060. M. M. Badža and M. Č. Barjaktarović, “Classification of Brain Tumors from MRI Images Using a Convolutional Neural Network,” Applied Sciences, vol. 10, no. 6, p. 1999, Mar. 2020, doi: https://doi.org/10.3390/app10061999.
[14] “Random Forest Classifier using Scikit-learn,” GeeksforGeeks, Sep. 04, 2020. https://www.geeksforgeeks.org/random-forest-classifier-using-scikit-learn/
[15] JavaTPoint, “Support Vector Machine (SVM) Algorithm - Javatpoint,” www.javatpoint.com. https://www.javatpoint.com/machine-learning-support-vector-machine-algorithm
[16] J. Zhang, Z. Jiang, J. Dong, Y. Hou and B. Liu, "Attention Gate ResU-Net for Automatic MRI Brain Tumor Segmentation," in IEEE Access, vol. 8, pp. 58533-58545, 2020, doi: 10.1109/ACCESS.2020.2983075.
keywords: {Tumors;Image segmentation;Logic gates;Task analysis;Magnetic resonance imaging;Feature extraction;Brain;MRI;brain tumor segmentation;U-Net;attention gate;residual module},
[17] S. Gaikwad, S. Patel and A. Shetty, "Brain Tumor Detection: An Application Based on Machine Learning," 2021 2nd International Conference for Emerging Technology (INCET), Belagavi, India, 2021, pp. 1-4, doi: 10.1109/INCET51464.2021.9456347. keywords: {Deep learning;Visualization;Shape;Magnetic resonance imaging;Transfer learning;Surgery;Medical services;Machine Learning;Deep Learning;MRI;Convolutional Neural Networks;Transfer Learning},
[18] S. Kumar and D. Kumar, “Human brain tumor classification and segmentation using CNN,” Multimedia Tools and Applications, Aug. 2022, doi: https://doi.org/10.1007/s11042-022-13713-2.
[19] C. L. Choudhury, C. Mahanty, R. Kumar and B. K. Mishra, "Brain Tumor Detection and Classification Using Convolutional Neural Network and Deep Neural Network," 2020 International Conference on Computer Science, Engineering and Applications (ICCSEA), Gunupur, India, 2020, pp. 1-4, doi: 10.1109/ICCSEA49143.2020.9132874. keywords: {Machine Learning;Brain Tumour;Biomedical Informatics;CNN},
[20] R. Jahan and M. M. Tripathi, "Brain Tumor Detection Using Machine Learning in MR Images," 2021 10th IEEE International Conference on Communication Systems and Network Technologies (CSNT), Bhopal, India, 2021, pp. 664-668, doi: 10.1109/CSNT51715.2021.9509695. keywords: {Machine learning algorithms;Communication systems;Conferences;Cerebrum;Machine learning;Functional magnetic resonance imaging;Tumors;MRI;brain tumor;machine learning;pattern recognition},

[21] R. R. Singh, S. Vijh and D. Chaubey, "An Efficient Brain Tumor Detection Using Modified Tree Growth Algorithm and Random Forest Method," 2021 Sixth International Conference on Image Information Processing (ICIIP), Shimla, India, 2021, pp. 1-6, doi: 10.1109/ICIIP53038.2021.9702692. keywords: {Sensitivity;Binary trees;Prediction algorithms;Feature extraction;Classification algorithms;Random forests;Tumors;Medical image processing;Tree Growth algorithm;Gray level co-occurrence matrix;Random Forest},

[22] H. T. Zaw, N. Maneerat and K. Y. Win, "Brain tumor detection based on Naïve Bayes Classification," 2019 5th International Conference on Engineering, Applied Sciences and Technology (ICEAST), Luang Prabang, Laos, 2019, pp. 1-4, doi: 10.1109/ICEAST.2019.8802562. keywords: {Tumors;Magnetic resonance imaging;Image segmentation;Feature extraction;Cancer;Classification algorithms;Databases;Glioblastoma Multiforme (GBM);Magnetic resonance imaging (MRI);Naïve Bayes classification;maximum entropy threshold;statistical features extraction},

[23] ajaymehta, "'Demystifying Naïve Bayes: Simple yet Powerful for Text Classification,'" Medium, Nov. 06, 2023. https://medium.com/@dancerworld60/demystifying-na%C3%AFve-bayes-simple-yet-powerful-for-text-classification-ad92b14a5c7

[24] S. Swaminathan, "Logistic Regression — Detailed Overview," Medium, Mar. 15, 2018. https://towardsdatascience.com/logistic-regression-detailed-overview-46c4da4303bc

[25] JavaTpoint, "K-Nearest Neighbor(KNN) Algorithm for Machine Learning - Javatpoint," www.javatpoint.com, 2021. https://www.javatpoint.com/k-nearest-neighbor-algorithm-for-machine-learning

[26] N. Vikas, "ADABOOST ALGORITHM," Medium, Apr. 16, 2024. https://medium.com/@nehra.vikas95/adaboost-algorithm-ef043e6104e0 (accessed May 21, 2024).