ISSN: 2350-0328
International Journal of AdvancedResearch in Science, Engineering and Technology
Vol. 11, Issue 6, June 2024

Optimizing Brain Tumor Diagnosis: A Stacking Ensemble Approach with Hyperparameter Tuning

A Anish, Aiswarya Anantharaman, Dr Sindhu R

B. Tech Scholar, Department of Electronics and Communication, NSS College Of Engineering, Palakkad
B. Tech Scholar, Department of Electronics and Communication, NSS College Of Engineering, Palakkad
Professor, Department of Electronics and Communication, NSS College Of Engineering, Palakkad, Kerala, India

ABSTRACT: A tumor is an abnormal growth resulting from the division of cells in an uncontrolled and disorderly manner. This research focuses on developing a Stacking Classifier model for automatic detection of brain tumors from MRI image data. The study begins with pre-processing the MRI dataset and analyzing the prediction performance of various Machine Learning models like Support Vector Machine (SVM), Decision Tree, Random Forest, Linear Discriminant Analysis (LDA), Gradient Boosting, Naive Bayes, Logistic Regression and K- Nearest Neighbors (KNN). The best-performing models were further tuned by tweaking their hyperparameters to improve the performance. These hyperparameter tuned models were further blended into a stack using a meta-learning algorithm. The Stacking classifier showed a significant improvement in performance with reduced false negative and false positive predictions. The findings suggest that the Stacking Classifier model can be highly beneficial for automated tumor detection from MRI input.

I. INTRODUCTION

Brain tumor is an abnormal clump of cells that grow uncontrollably in the brain. These abnormal cells form a lump or mass, which can be benign (non-cancerous) or malignant (cancerous).
Malignant tumors can spread to other parts of the Central Nervous System (CNS), making them challenging to treat. Secondary tumors, also known as Metastatic brain tumors, are of another type which originate from other parts of the body and then metastasize to the brain. An estimated 3,08,102 people were diagnosed with brain tumors in 2020 worldwide. The detection of brain tumors mainly involves the use of imaging techniques like Magnetic Resonance Imaging (MRI) and Computed Tomography (CT) scans. Another diagnosis method requires the use of Stereotactic Biopsy which is minimally invasive compared to its counterparts like Open-Surgical and Endoscopic Biopsies [1][2].

Manual detection of brain tumors from MRI images is an exhaustive process. It would be very challenging to identify subtle anomalies. Automated brain tumor detection allows faster diagnosis even at an early stage of tumor. This paper proposes an approach to optimize tumor detection using Stacking Classifier which is an ensemble learning technique to detect tumor presence effectively with high prediction performance.

II. LITERATURE REVIEW

There have been varied research approaches performed for brain tumor detection for achieving high accuracy and sensitivity. Some of these research approaches are elucidated below.

Shraddha S. More et al. [3], proposed Convolutional Neural Network based brain tumor detection which uses image augmentation and pre-processing techniques like grayscale conversion, image cropping, gaussian blurring and binary conversion. Further, Convolutional Neural Network (CNN) is used to train and predict from MRI images. The model provides an accuracy of 89%. However, the model uses a batch normalization mechanism for training the images in multiple batches. This results in computational overhead.

ISSN: 2350-0328
International Journal of AdvancedResearch in Science, Engineering and Technology
Vol. 11, Issue 6, June 2024

Harsh H. Patel et al. [4], studied the performance of Decision Tree algorithms ID3, C4.5 and CART on a dataset of cars, to understand the car acceptability and its technical characteristics. The results concluded that the CART algorithm had outperformed others with an accuracy and precision of 97.11% and 0.972 respectively. However, the time required is slightly higher compared with the other three algorithms.

R Venkadothu et al. [5], described the prediction of brain tumors from their CT scan images using SVM and Fuzzy classifier. The approach used image preprocessing, followed by feature extraction using GLCM (Gray-Level Co-occurrence Matrix), LGXP (Local Binary Pattern with Gray Level Differences) for texture analysis and Mutual Information data. The classification of images was performed with SVM. The predicted abnormal images were further segmented using Fuzzy C-Means clustering algorithm. The segmented images provide an accuracy of 98%.

Kashvi Taunk et al. [6], defined varied methodologies adopted with KNN to improve its accuracy scores including Adaptive KNN, KNN with Mahalanobis Metric, Informative KNN, Bayesian KNN and many more. Although KNN is known for its simplicity, comprehensibility, and scalability features, it can be expensive in determining K in larger datasets. It also details its applications in medical imaging, stock market forecasting, fault detection and surveillance systems. The research continues to improve the prediction performance of KNN.

Mihir Gada et al. [7], discussed the AutoML pipeline which uses feature engineering methodologies like removing highly correlated features, considering all features, Analysis of Variance (ANOVA), Principal Component Analysis (PCA), and Scikit-Learn Feature Selection. Multiple combinations of Hyperparameter optimization and feature processing techniques were evaluated to ensure optimal performance for classification and prediction tasks. The results concluded that ANOVA, PCA and Bayesian Optimization techniques proved very effective in the AutoML pipeline.

Swapnil R. Telrandhe et al. [8], proposed a method using the K-Means algorithm to detect brain tumors. Brain MR images were converted to grayscale, denoised and passed through a high pass filter mask to detect edges. Deep learning-based skull stripping techniques were used for smoothing the MR image. Image segmentation was done using K-Means and object labeling algorithms. feature extraction was carried out using HOG. SVM was used for detection.

Ayesha Jabbar et al. [9], proposed a hybrid model that integrates CNN-based pre-trained architectures, CapsNet and VGGNet architectures. The model classifies pre-processed MRI images of brain tumors into categories: ordinary, pituitary, meningioma, and glioma by employing multiple layers. The model outperformed traditional models by providing faster detection of tumor and superior classification performance. However, the architectures used are complex, and the model was not evaluated on a larger dataset.

III.DATASET USED

The initial dataset comprises 3000 images. Due to challenges in computation of a large dataset, stratified sampling was applied. Stratified sampling is a technique used to reduce dataset size while preserving the ratio of the class labels.

The sampled dataset used to assess performance of machine learning models comprises 605 images, with 80% (484 images) designated for training and 20% (121 images) for validation. Lastly, a Stacking Classifier, formed by aggregating multiple ML models, was evaluated on 100 images of stratified sampled unseen testing dataset for performance analysis.

IV.METHODOLOGY

The main objective of the paper is to develop an algorithm which can predict brain tumors from MR images with an optimized approach with good prediction results.

The proposed method streamlines the analysis of brain MR images by initially subjecting them to crucial pre-processing steps which will be discussed in a separate section. Subsequently, feature extraction is performed which involves transforming 2D images into 1D arrays for efficient information extraction.

ISSN: 2350-0328
International Journal of AdvancedResearch in Science, Engineering and Technology
Vol. 11, Issue 6, June 2024

To bolster the model's training capabilities, data augmentation techniques are applied to expand the training dataset. This expanded dataset is then divided into an 80-20% split for training and result validation. Machine learning models undergo rigorous evaluation, and the top-performing ones are fine-tuned through hyperparameter optimization. These optimized models are ultimately combined into a stacking classifier, capitalizing on their individual strengths. These models' effectiveness is thoroughly assessed using an unseen testing dataset ensuring robust performance in real-world scenarios.

The following flow chart illustrates the different steps involved in the implementation.

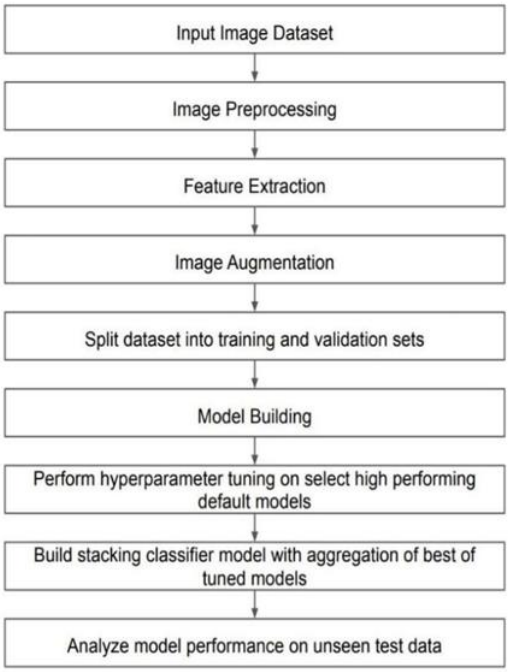

Fig. 1 Process diagram of proposed system.

A. **Image Pre-processing**

MR images may contain differentiated intensity, noise and unwanted data. This can be overcome using various preprocessing techniques including pixel resizing, image normalization and denoising. The various techniques used are listed below:

1. Pixel Resizing: It is the process of changing the size of an image. This can be done by increasing the number of pixels (upscaling) or decreasing the number of pixels (downscaling). This facilitates reduction in computational complexity. In this work, dataset images were downscaled to 200 x 200 pixels.
2. Image Normalization: It involves adjusting pixel values of an image to fit within a specified range. This is achieved by employing the contrast stretching technique to ensure that pixel intensities fall between 0 and 1.
3. Image Denoising: It is performed using a three-step approach that incorporates Median blurring, Gaussian blurring and Bilateral filtering. Median blurring replaces each pixel with the median value of its neighboring pixels, which is particularly effective in eliminating salt and pepper noise. This is followed by Gaussian blurring employs a weighted average for pixel values within a local neighborhood, effectively reducing Gaussian noise and smoothing out random high frequency variations in the image. Finally, the Bilateral filter is applied to denoise and preserve edges, achieving this by computing the weighted average that considers spatial distance and intensity differences with neighboring pixels. These techniques effectively provide a comprehensive denoising of image datasets [10].

ISSN: 2350-0328
International Journal of AdvancedResearch in Science, Engineering and Technology
Vol. 11, Issue 6, June 2024

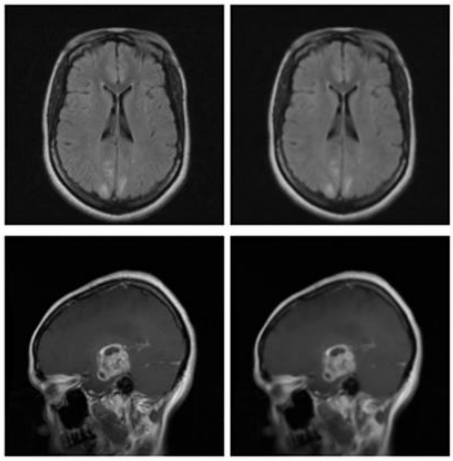

Fig. 2 Original (left) and denoised (right) MRI images.

B. Feature Extraction

Feature extraction is employed to transform 2D images to 1D arrays, essentially converting images into flattened feature vectors. This step helps in dimensionality reduction and noise removal. It reduces the complexity for training the model, saving time for classification and diagnosis [10][11].

C. Image Augmentation

Image augmentation is a method used to increase the practical size of a dataset by creating modified versions of the original images. This involves applying random changes like rotation, flipping, scaling and shifting to the original image. This technique enables the model to train on a varied dataset, enhancing its effectiveness and robustness for tasks such as classification and prediction, even when dealing with low-quality images. Furthermore, it helps in preventing the model from overfitting [10].

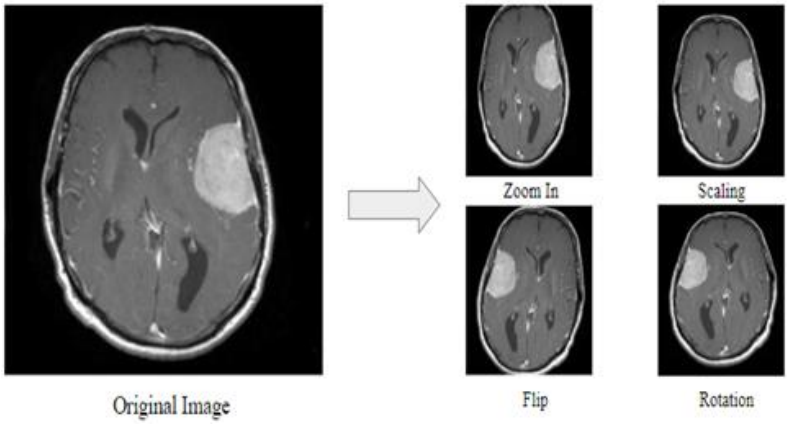

Fig. 3 Image augmentation.

D. Model Building

The models investigated include supervised learning techniques - Logistic Regression, Support Vector Machine (SVM), Decision Tree, Gradient Boosting, K-Nearest Neighbors (KNN), Random Forest, Linear Discriminant

ISSN: 2350-0328
International Journal of AdvancedResearch in Science, Engineering and Technology
Vol. 11, Issue 6, June 2024

Analysis (LDA), and Naive Bayes. A comprehensive analysis was conducted on these models, which underwent training and validation on the dataset. The analysis involved the examination of their respective confusion matrices for both training and validation datasets, which provided in-depth insights into the classification outcomes. The performance of these models was evaluated using key metrics such as Accuracy, Precision, Recall and F1 scores. The ML models studied, and the inferences obtained are detailed below:

Logistic Regression: Logistic Regression serves as a supervised learning technique designed for predicting the likelihood of a binary outcome. The model employs a sigmoid function to compute the probability associated with one of two possible outcomes, yielding a result bounded within the range of 0 to 1. By applying a predefined threshold value, the algorithm determines whether the output should be categorized as "Positive" or "Negative", making it a valuable tool for classification tasks [12][13]. The logistic regression model predicted validation data with an accuracy of 91% and a recall of 98%. An F1-score of around 92% was obtained.

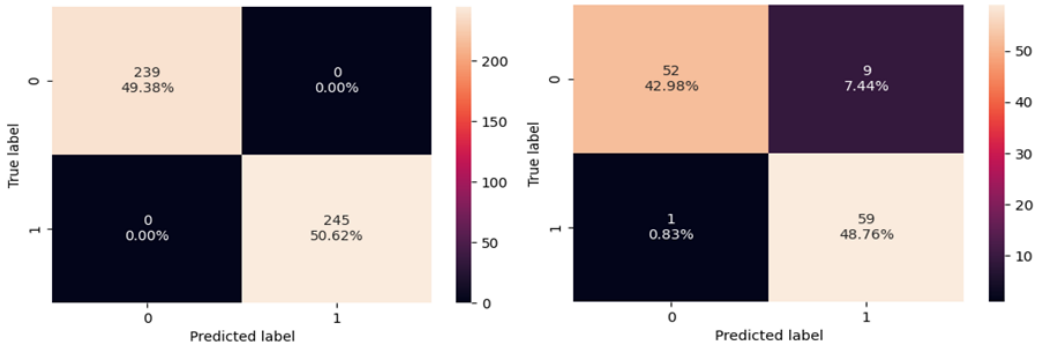

Fig. 4 Confusion matrix for training and validation datasets in Logistic Regression model.

Support Vector Machine (SVM): SVM is a supervised machine learning algorithm employed for solving classification and regression tasks. It operates by identifying a hyperplane or decision boundary, within a high dimensional space, with the objective of maximizing the separation margin between two classes. SVM is known for its robustness against noise and overfitting, and it is recognized for its efficiency in classification tasks [12][14]. The support vector machine model outperformed the logistic regression model. It was able to predict the validation data with an accuracy of 93% and a recall of 98%. An F1-score of around 93% was obtained.

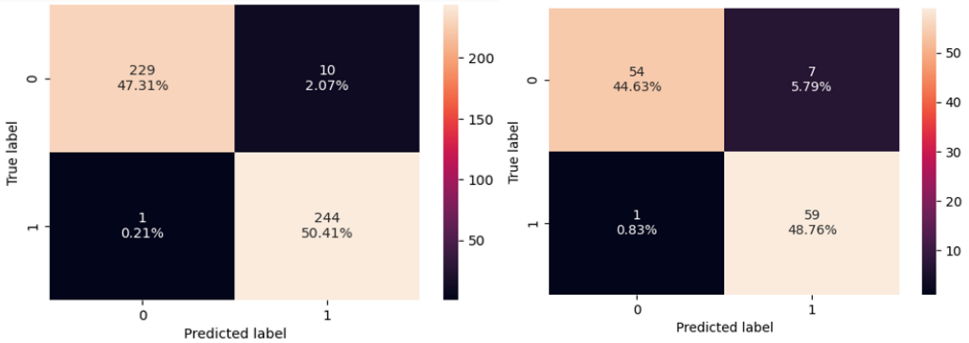

Fig. 5 Confusion matrix for training and validation datasets in SVM model.

Decision Tree: The Decision Tree algorithm, employed in supervised learning for classification and regression tasks, features a tree-like structure comprising root, internal, and leaf nodes. It starts at the root node with the entire dataset and makes decisions via internal nodes and branches until it reaches a leaf node for the final classification and regression label. Data splitting in Decision Trees relies on optimal attribute selection, often determined by metrics like Gini impurity and Entropy. Gini impurity gauges misclassification probability, while entropy assesses dataset disorder. Decision Trees can be computationally expensive on large datasets and prone to overfitting. However, they offer robustness against noise [12][15]. The decision tree model made better predictions on the validation data. An accuracy of 97%, recall of 98% and an F1-score of around 97% was achieved.

ISSN: 2350-0328
International Journal of AdvancedResearch in Science, Engineering and Technology
Vol. 11, Issue 6, June 2024

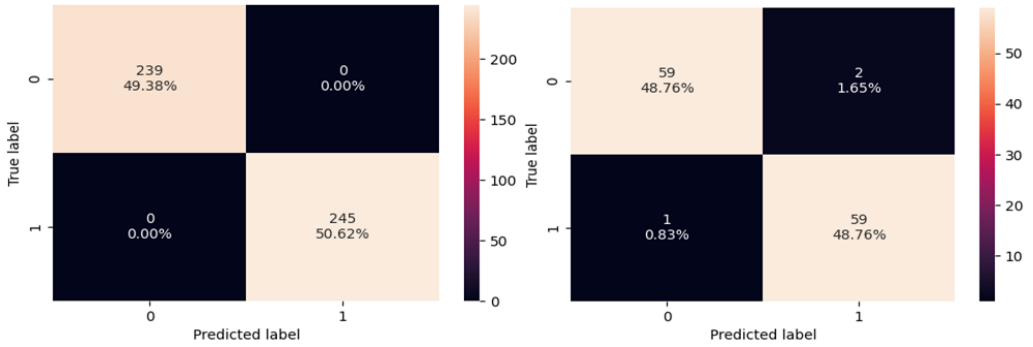

Fig. 6 Confusion matrix for training and validation datasets in Decision Tree model.

Gradient Boosting Classifier: Gradient Boosting is a supervised machine learning algorithm that builds a strong predictive model by combining multiple weak learners, typically Decision Trees. Unlike gradient descent, which is used for optimizing parameters in some machine learning algorithms, Gradient Boosting optimizes the loss function directly. It does this by iteratively adding weak learners to the ensemble, with each new learner focusing on the mistakes made by the previous ones. These mistakes are quantified using the gradient of the loss function, and the algorithm aims to minimize this gradient, thereby reducing the prediction errors in the ensemble model [12]. The Gradient Boosting classifier performed well with the validation data with an accuracy of 96% and a recall of 98%. The F1 score obtained was about 97%, albeit at the expense of computational resources.

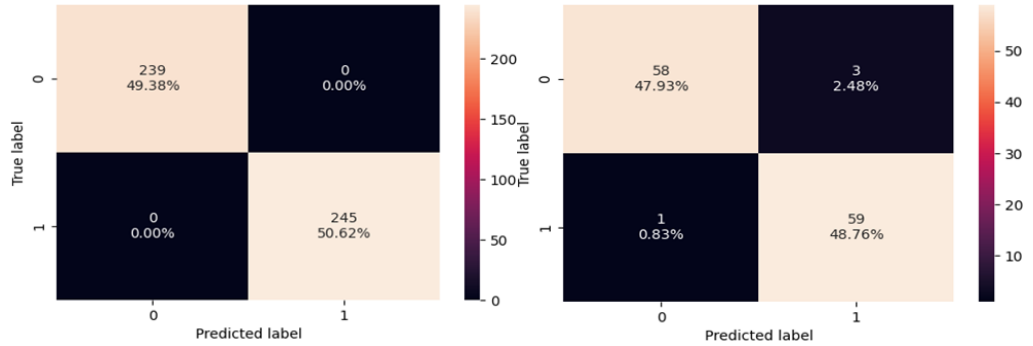

Fig. 7 Confusion matrix for training and validation datasets in Gradient Boosting model.

K-Nearest Neighbors (KNN): K-Nearest Neighbor (KNN) is a non-parametric supervised learning classifier employed for both classification and regression tasks. Its operational principle involves the computation of K nearest neighbors, which are determined based on the Euclidean, Manhattan, or Hamming distance metrics between data points. Subsequently, it tallies the data points within each category and assigns new data points to the category exhibiting the highest neighbor count. KNN is renowned for its simplicity and straightforward training process, as it retains the entire training dataset in memory. However, it exhibits susceptibility to suboptimal performance, particularly when dealing with imbalanced datasets. Additionally, the choice of the K-parameter necessitates careful consideration to mitigate the risks of overfitting or underfitting [12]. The model produced an accuracy of 89% and a recall of 96% with the validation data. The F1 score of about 90% was achieved.

ISSN: 2350-0328
International Journal of AdvancedResearch in Science, Engineering and Technology
Vol. 11, Issue 6, June 2024

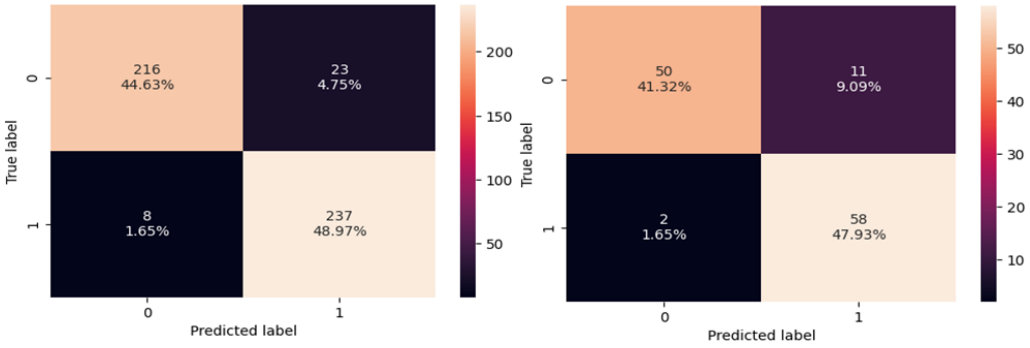

Fig. 8 Confusion matrix for training and validation datasets in KNN model.

Random Forest: Random Forest is a supervised ensemble learning algorithm employed for classification and regression tasks. This algorithm leverages a multitude of Decision Trees, each trained on a distinct random subset of the training dataset in an iterative manner. The ultimate prediction for unseen data by the Random Forest is derived from aggregating the majority predictions generated by individual Decision Trees. Random Forest is renowned for its resistance to overfitting and its adeptness at managing substantial datasets efficiently [12][13][14]. The model produced an accuracy of 97% and recall of 98%. The F1 score obtained was nearly 97%.

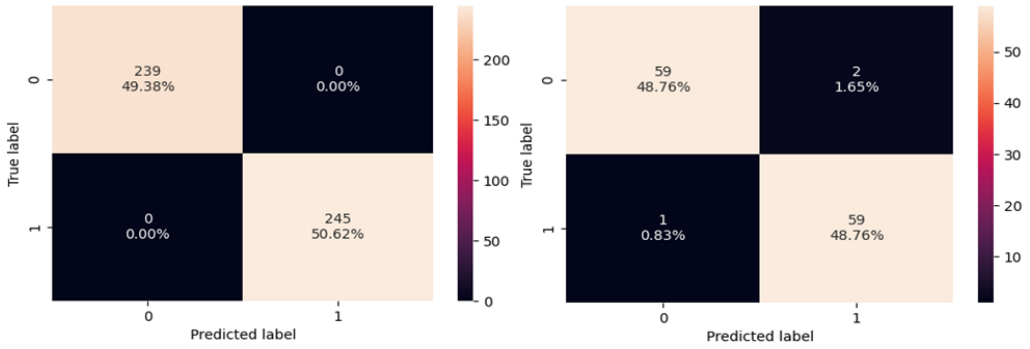

Fig. 9 Confusion matrix for training and validation datasets in Random Forest model.

Linear Discriminant Analysis (LDA): Linear Discriminant Analysis (LDA) is a supervised machine learning algorithm primarily employed for classification and dimensionality reduction purposes. LDA involves several key steps, including the computation of class means and covariances. Following this, it calculates the Fisher discriminant, which is a linear combination of features designed to maximize the ratio of between-class variance to within-class variance. The data is subsequently projected onto the Fisher discriminant space, effectively reducing the dimensionality of the data to facilitate more straightforward classification. LDA can be susceptible to overfitting, particularly when dealing with small datasets, and may exhibit bias towards the majority class in imbalanced datasets. The model achieved an accuracy of approximately 91% and a recall of about 98%. The F1 score obtained was about 91%.

ISSN: 2350-0328
International Journal of AdvancedResearch in Science, Engineering and Technology
Vol. 11, Issue 6, June 2024

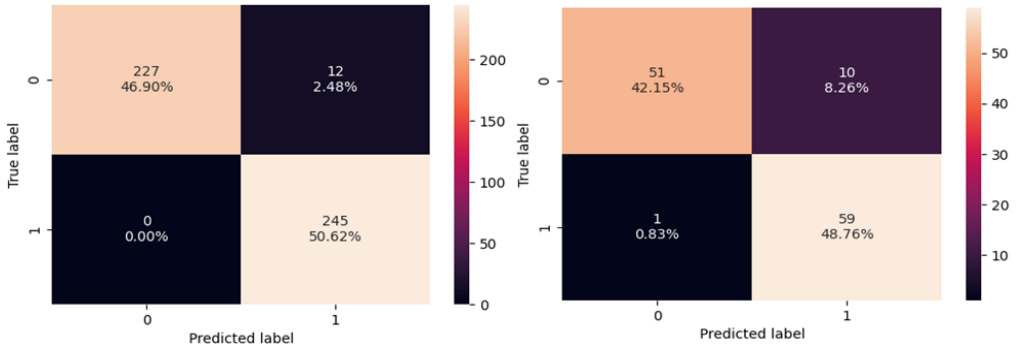

Fig. 10 Confusion matrix for training and validation datasets in LDA model.

Naive Bayes model: Naive Bayes model is a supervised machine learning algorithm used for classification problems. It is based on Bayes theorem, which is used to predict an event given that the other event has already occurred. It makes a naive assumption that the features are independent of one another in each class, even though this is not true in real world scenarios. Nonetheless, the model exhibits a fairly good performance in classification tasks. The model provided an accuracy of 92% and a recall of 96%. The F1 score achieved is about 93%.

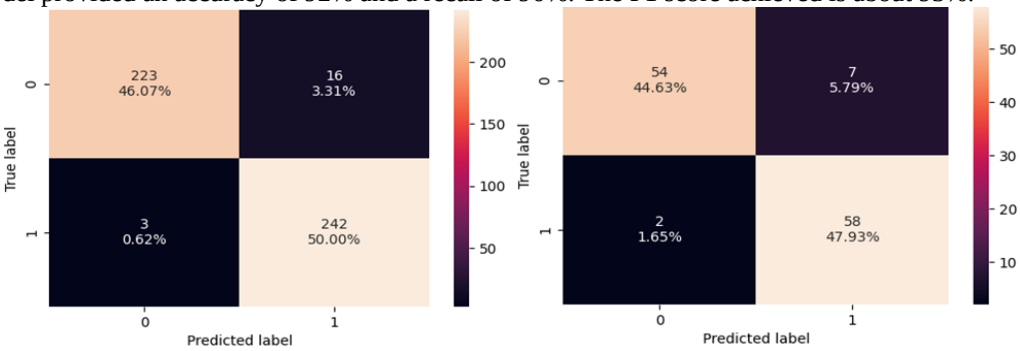

Fig. 11 Confusion matrix for training and validation datasets in Naive Bayes model.

E. Hyper-parameter tuning

In medical imaging and analysis, datasets are mostly imbalanced compared to other scenarios. This may cause bias in predictions of certain ML models towards the majority class. Accuracy, Precision, Recall and F1 scores are commonly used matrices to evaluate these models. Accuracy can be calculated using,

$$Accuracy = \frac{Number\ of\ correctly\ predicted\ instances}{Total\ number\ of\ instances}$$

High accuracy values can be misleading as the count of negative cases outnumber that of positive cases of brain tumor. The identification of cases in the minority class is crucial in medical applications. False Negatives, i.e. model's failure to identify tumor cases can be potentially life-threatening on patients. Hence, its count should be significantly reduced. Recall (Sensitivity) reduces this risk and is calculated using,

$$Recall = \frac{True\ Positives}{True\ Positives + True\ Negatives}$$

ISSN: 2350-0328
International Journal of AdvancedResearch in Science, Engineering and Technology
Vol. 11, Issue 6, June 2024

Precision measures the ability of the model to identify False Positive cases, i.e. incorrect predictions of cases as positive. Precision is calculated using,

$$Precision = \frac{True\ Positives}{True\ Positives + False\ Positives}$$

F1 score is the harmonic mean of precision and recall. It aims to strike a balance between these two parameters. F1 score is calculated using,

$$F1\ Score = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

Healthcare decisions based on machine learning predictions have real world consequences. The metrics F1 and recall provide better insights for proper diagnosis [16]. Machine learning models generated were evaluated based on their performance on the validation set. Models with higher recall and F1 scores were filtered. Randomized search with cross-validation was used to obtain the best set of hyperparameters for the models. The best set of hyperparameters were used to tune the model and optimize the scores on the training dataset. The performance of the tuned models was evaluated against the validation dataset.

Machine learning models generated were evaluated based on their performance on the validation set. Models with higher recall and F1 scores were filtered. Randomized search with cross validation was used to obtain the best set of hyperparameters for the models. The best set of hyperparameters were used to tune the model and optimize the scores on the training dataset. The performance of the tuned models was evaluated against the validation dataset.

Tuned decision tree: A set of hyperparameters for decision tree viz. maximum depth of tree, minimum samples per leaf, maximum leaf nodes and minimum decrease in impurity were randomly searched. The default model was tuned using the best set obtained, trained and its performance was evaluated [17]. It was found that the tuned model was performing well and was not overfitting the training dataset. A recall score of 98% and an F1-score of 92% were obtained.

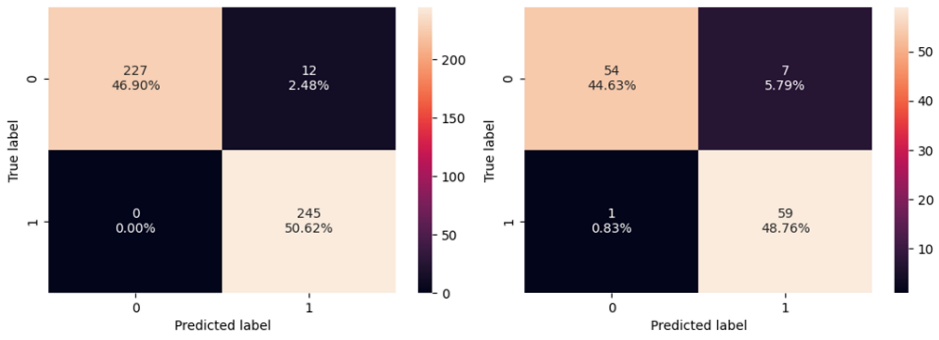

Fig. 12 Confusion matrix for training and validation datasets in tuned decision tree model.

Tuned random forest: The best set of hyperparameters consisting of number of trees, minimum samples required per leaf, maximum number of features considered, and maximum samples were utilized to tune the default random forest model, trained and its performance was evaluated [17, 18]. It was found that the tuned model was not overfitting the training dataset. A recall score of 98% was obtained. However, the F1-score increased to 97%.

ISSN: 2350-0328
International Journal of AdvancedResearch in Science, Engineering and Technology
Vol. 11, Issue 6, June 2024

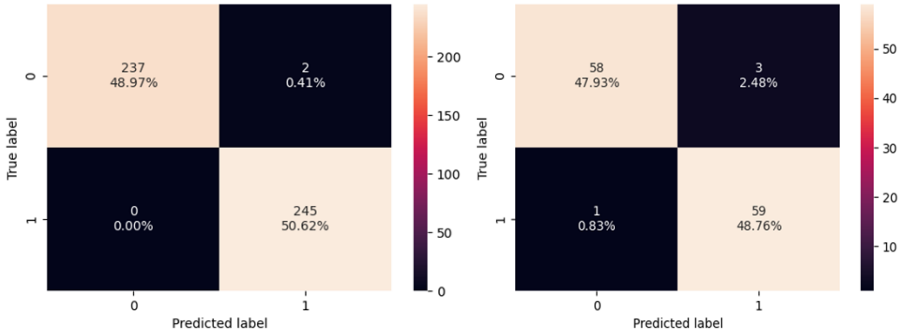

Fig. 13 Confusion matrix for training and validation datasets in tuned random forest model.

Tuned support vector machine: The default support vector machine model was tuned to achieve the best recall and F1-score. The hyperparameters considered for tuning were the regularization parameter and the kernel type for the algorithm [19]. The model's performance was evaluated, and it was found that the tuned model was not overfitting the training dataset. A recall score of 96% was obtained but the F1-score dipped to 81%.

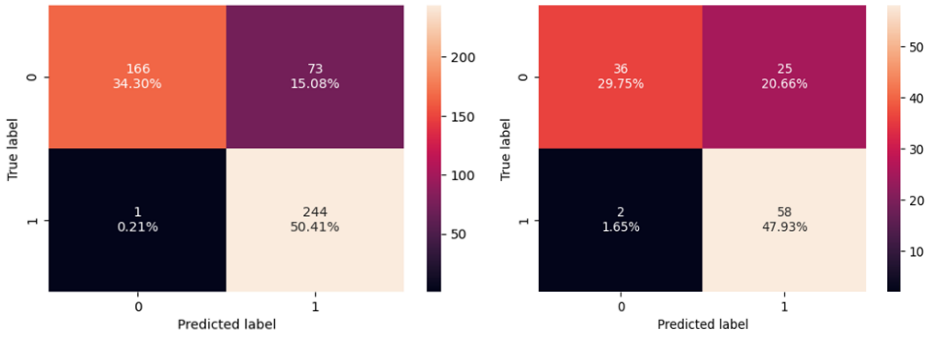

Fig. 14 Confusion matrix for training and validation datasets in tuned support vector machine model.

F. Stacking classifier

Stacking is an ensemble method that uses a meta-learning algorithm to combine the predictions of two or more models. This technique can harness the capabilities of the base models used and provides better predictions than any single model in the ensemble [20].

The tuned models were analyzed based on their predictions on the validation set. The concept of stacking was applied on the best performing models. The tuned decision tree and random forest models were set as the baseline models. These models predicted the output for the test dataset. The predictions were fed as input to the second layer - tuned support vector machine model and final predictions were obtained.

ISSN: 2350-0328
International Journal of AdvancedResearch in Science, Engineering and Technology
Vol. 11, Issue 6, June 2024

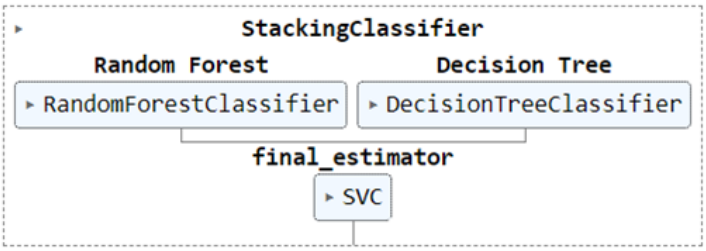

Fig. 15 Stacking Classifier model.

The stacking classifier outperformed other models and was able to predict the validation dataset with a recall score of around 98%. An F1-score of 97.5% was obtained.

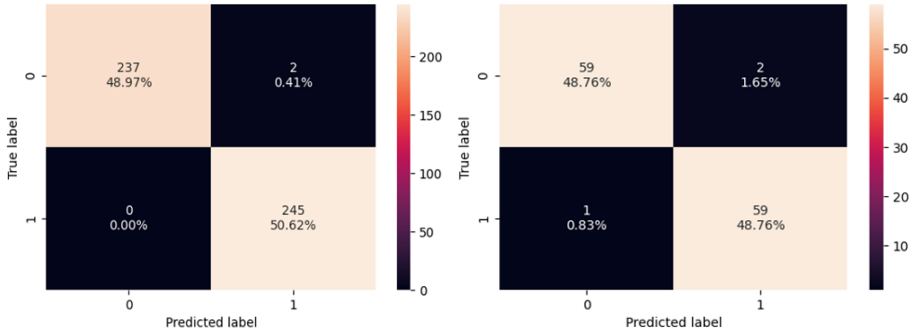

Fig. 16 Confusion matrix for training and validation datasets in tuned stacking classifier model.

EVALUATING MODEL PERFORMANCE ON TEST DATASET

A stratified sample of 100 MRI images were labeled as the test dataset. These images were kept isolated from the model training process. Since models tend to perform better on observations on seen data, use of unknown dataset can help evaluate the model performance by closely mimicking the real-time scenario.
From our extensive training and validation on various models, it was evident that the stacking classifier outperformed all models. The stacking classifier was used to predict the test dataset. The model performed well on the unseen data. A recall score of 98% and F1-score of 96% were obtained.

CONCLUSION

The paper focuses on building a machine learning model that can identify brain tumor cases from input MRI images. Since the medical dataset is imbalanced in nature, recall and F1-scores are used to compare and evaluate the model performance. The stacking classifier algorithm is highly efficient to detect the presence of a tumor.

Future work will focus on categorizing the type of brain tumor as well as its characterization. This can help doctors make informed decisions regarding the treatment process. Another area of research is the size of the dataset used. Medical data is highly imbalanced and is challenging to work with. Machine learning models perform even better when exposed to various types and huge sizes of datasets which at present are unavailable. With the advancements in medical technologies to capture finer details and more machine learning techniques, there is hope for better diagnosis and treatment of such life-threatening conditions.

ISSN: 2350-0328
International Journal of AdvancedResearch in Science, Engineering and Technology
Vol. 11, Issue 6, June 2024

REFERENCES

John Hopkins medicine. Brain tumors and brain cancers. [Online]. Available: https://www.hopkinsmedicine.org/health/conditions-and-diseases/brain-tumor

Cancer.Net. *Brain tumor*. [Online]. Available: https://www.cancer.net/cancer-types/brain-tumor

S. S. More, M. A. Mange, M. S. Sankhe and S. S. Sahu, "Convolutional Neural Network based Brain Tumor Detection," 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, 2021, pp. 1532-1538, doi: 10.1109/ICICCS51141.2021.9432164.

Patel, Harsh & Prajapati, Purvi. (2018). Study and Analysis of Decision Tree Based Classification Algorithms. International Journal of Computer Sciences and Engineering. 6. 74-78. 10.26438/ijcse/v6i10.7478.

Ramdas Vankdothu, Mohd Abdul Hameed, Brain tumor segmentation of MR images using SVM and fuzzy classifier in machine learning, Measurement: Sensors, Volume 24, 2022, 100440, ISSN 2665-9174

K. Taunk, S. De, S. Verma and A. Swetapadma, "A Brief Review of Nearest Neighbor Algorithm for Learning and Classification," 2019 International Conference on Intelligent Computing and Control Systems (ICCS), Madurai, India, 2019, pp. 1255-1260, doi: 10.1109/ICCS45141.2019.9065747.

M. Gada, Z. Haria, A. Mankad, K. Damania and S. Sankhe, "Automated Feature Engineering and Hyperparameter optimization for Machine Learning," 2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 2021, pp. 981-986, doi: 10.1109/ICACCS51430.2021.9441668.

Telrandhe, Swapnil & Kendhe, Ankita & Pimpalkar, Amit. (2016). Detection of Brain Tumor from MRI images by using Segmentation &SVM.

Jabbar, Ayesha, Shahid Naseem, Tariq Mahmood, Tanzila Saba, Faten S. Alamri, and Amjad Rehman. Brain Tumor Detection and Multi-Grade Segmentation Through Hybrid Caps-VGGNet Model. IEEE, 2023.

S. S. Hunnur, A. Raut and S. Kulkarni, "Implementation of image processing for detection of brain tumors," 2017 International Conference on Computing Methodologies and Communication (ICCMC), Erode, India, 2017, pp. 717-722, doi: 10.1109/ICCMC.2017.8282559.

Image Feature Extraction | Feature Extraction Using Python by Aishwarya Singh Available: https://www.analyticsvidhya.com/blog/2019/08/3-techniques-extract-features-from-image-data-machine-learning-python/

Ghosh A, Kole A. A comparative study of enhanced machine learning algorithms for brain tumor detection and classification. Authorea Preprints. 2023 Oct 30.

Sperandei S. Understanding logistic regression analysis. Biochem Med (Zagreb). 2014 Feb 15;24(1):12-8. doi: 10.11613/BM.2014.003. PMID: 24627710; PMCID: PMC3936971.

Hyun-Chul Kim, Shaoning Pang, Hong-Mo Je, Daijin Kim, Sung Yang Bang, Constructing support vector machine ensemble, Pattern Recognition, Volume 36, Issue 12, 2003, Pages 2757-2767, ISSN 0031-3203

Naik, Janki and Sagar Patel. "Tumor Detection and Classification using Decision Tree in Brain MRI." (2013).

Various ways to evaluate a machine learning model's performance by Kartik Nighania, Available: https://towardsdatascience.com/various-ways-to-evaluate-a-machine-learning-models-performance-230449055f15

Hyperparameter Tuning in Decision Trees and Random Forests by Himani Gulati, Section, Available: https://www.section.io/engineering-education/hyperparmeter-tuning/

A Beginner's Guide to Random Forest Hyperparameter Tuning by Sharoon Saxena, Available: https://medium.com/analytics-vidhya/a-beginners-guide-to-random-forest-hyperparameter-tuning-4692c3ab7537

Kaibo Duan, S.Sathiya Keerthi, Aun Neow Poo, Evaluation of simple performance measures for tuning SVM hyperparameters, Neurocomputing, Volume 51, 2003, Pages 41-59, ISSN 0925-2312

Seewald, Alexander K. "Towards a theoretical framework for ensemble classification." In IJCAI, vol. 3, pp. 1443-1444. 2003.