Received: 2025-08-08
Accepted: 2025-08-25
Published Online: 2025-09-20
DOI: 10.70454/JRIST.2025.10101

Journal of Recent Innovation in Science and Technology
E-ISSN: 3117-3926

sA Comparative Study of Machine Learning Algorithms for Brain Tumor Detection

Vimmi Kochher *[1]

[1]Assistant Professor, CSE, Guru Tegh Bahadur Institute of Technology, Delhi, India.
***Corresponding Author**: v.kochher1990@yahoo.com

Abstract

This study examines brain imaging to identify areas with tumors and categorizes these regions into three distinct types: meningioma, glioma, and pituitary tumors. This paper also compares different machine learning algorithms for the identification of brain tumors. The term "brain tumor" describes the excessive growth of cells in the brain, that can be either benign or malignant.

In this study, machine learning algorithms have been utilized to both identify what kind of brain tumor in patients who have been diagnosed with one using multi-class classification and to detect whether brain tumors are present or absent by binary classification. There are two types of brain MRI pictures in the dataset used for the binary classification challenge: those containing tumors and those without. Several machine learning algorithms, including Support Vector Machine (SVM), Logistic Regression, K-Nearest Neighbor (KNN), Naïve Bayes (NB), Decision Tree (DT) classifier, and Random Forest classifier, have been applied to classify the MRI images.

Accuracy, recall, precision, and F1-score were among the performance indicators used in a comparative analysis of machine learning algorithms. With an accuracy of 90.8%, recall of 95.2%, precision of 81.1%, and F1-score of 87.6% across 1,475 brain tumor images, the Random Forest classifier outperformed the other techniques.

Keywords: brain tumor classification, artificial intelligence, machine learning, MRI image

Introduction

The application of machine learning techniques in various radiological imaging tasks has seen a substantial rise. This advancement has largely facilitated the discovery of concealed insights that enhance clinical decision-making. Clinical uses of machine learning encompass digital pathology as well as imaging of the chest, brain, heart, and abdomen. Given that brain tumors are the primary cause of cancer-related deaths among children and adults under the age of 40, it is crucial to promote early diagnosis. Consequently, there is a need to develop methods that expedite the early detection of brain tumors. Early detection of brain cancers improves patient survival chances by enabling a quicker response to treatment. A technology that could automatically identify and locate brain tumors would be the perfect cure. Brain tumors can be successfully identified and categorized using machine learning, which has gained popularity in almost every decision-making arena.

Received: 2025-08-08
Accepted: 2025-08-25
Published Online: 2025-09-20
DOI: 10.70454/JRIST.2025.10101

Journal of Recent Innovation in Science and Technology
E-ISSN: 3117-3926

This research aims to investigate the use of machine learning methods for brain tumor identification. Data collection, data pre-processing (including data labeling and image pre-processing), classification using sophisticated machine learning algorithms, and a comparative analysis of the implemented models are all phases of the suggested framework.

Related Work

Machine learning is critical in the development of tumor detection systems based on large-scale medical imaging and diagnostic data sets. It identifies patterns and anomalies in data such as MRI scans, CT pictures, and histopathological characteristics. Algorithms like regression and classification models leverage this information to detect or predict the presence of malignancies. Important factors—such as tumor size, form, and texture—are quantified and studied utilizing these methodologies, resulting in more accurate and objective diagnostic assistance. New developments in deep learning and machine learning have completely changed the process of detecting and classifying tumors, improving patient outcomes, treatment planning, and early diagnosis accuracy. These techniques have been used in numerous research on a variety of datasets, with notable outcomes.

In Lisa M. DeAngelis [1], The word "brain tumor" refers to a variety of growths that occur inside the skull, each having unique traits, prognoses, and methods of treatment. Despite not coming from brain tissue, several of these tumors—like meningiomas and lymphomas—are nevertheless categorized as intracranial neoplasms. Many of these cancers have similar symptoms, diagnostic techniques, and first treatment approaches, despite their variances. This review focuses on the general presentation, diagnosis, and specific treatment options for these conditions. New England Journal of Medicine

In terms of prevalence, the American Cancer Society estimated that in 1999, there were 16,800 new cases of intracranial tumors diagnosed. This figure is more than half of the leukemia cases and more than twice the Hodgkin's disease cases that were recorded in that year.

The phrase "brain tumor" refers to a variety of anatomical growths, each having unique traits, prognoses, and therapeutic modalities. Some of these tumors, including lymphomas and meningiomas, are categorized as intracranial neoplasms even though they don't start in brain tissue. Many of these cancers, in spite of their variances, have similar symptoms, diagnostic techniques, and initial treatment plans. This review focuses on the general presentation, diagnosis, and specific treatment options for these conditions. New England Journal of Medicine

In terms of prevalence, the American Cancer Society estimated that in 1999, there were 16,800 new cases of intracranial tumors diagnosed. This number is more than double the cases of Hodgkin's disease and over half the number of cases of leukemia reported that year.

Amin, Javaria, et al.[2] rain tumors are dangerous because they grow from abnormal cells in the brain and can be deadly if not detected early. This research focuses on using machine learning and image processing techniques to spot tumors in MRI brain scans as early as possible.

The authors cleaned up the MRI images using a filter to reduce noise, then used a clustering method to find which parts of the image might contain a tumor. After that, they applied mathematical techniques to highlight and separate the tumor areas, especially in specific types of MRI scans (called FLAIR and T2-weighted).

Received: 2025-08-08
Accepted: 2025-08-25
Published Online: 2025-09-20
DOI: 10.70454/JRIST.2025.10101

Journal of Recent Innovation in Science and Technology
E-ISSN: 3117-3926

To make the detection more accurate, they combined two different ways of pulling out important features from the images—one based on texture (Local Binary Patterns) and another based on frequencies (Gabor Wavelet Transform). These features were then used to classify the tumor areas.

They tested their method using both private datasets and well-known public datasets (BRATS 2013 and 2015), and it worked really well—achieving high accuracy and clear separation between tumor and non-tumor areas. This approach shows promise for helping doctors detect brain tumors more effectively using AI.

Painuli et al [5], In recent years, machine learning (ML) and deep learning (DL) technologies have made huge strides in improving how we diagnose cancer. This review looks at the latest developments in using these technologies to detect and diagnose six major types of cancer: breast, lung, liver, skin, brain, and pancreatic cancer.

The paper explores how different types of data (like medical images, patient records, etc.) are used, as well as the methods used to pull out key features from that data. It also goes over the pros and cons of various ML and DL techniques and how well they perform at early detection and diagnosis.

The authors highlight the promising impact these technologies could have on improving the accuracy of cancer diagnosis and, ultimately, patient care and treatment outcomes.

Hatami, Toktam, et al. [6]in this research, the authors introduced a method to separate (or "segment") brain tumors from MRI scans using a machine learning technique called an adaptive random forest algorithm. They tested this approach on brain MRI images and checked its accuracy with measures like the Dice Similarity Coefficient (DSC) and accuracy (ACC).The results were promising, with the algorithm achieving a DSC of 98.38% and an accuracy of 97.65%. This shows that the method works very well for segmenting brain tumors and might even be better than other methods currently in use.

Solanki, Shubhangi, et al. [7]This study offers a summary of the various artificial intelligence (AI) techniques for identifying and categorizing brain tumors. The authors talk about the application of AI methods, including deep learning (DL), machine learning (ML), and other intelligent algorithms, to medical imaging in order to analyze brain tumors. These techniques are essential since brain tumors can greatly enhance patient outcomes if detected early.

The study examines a number of methods that aid in differentiating between benign and malignant tumors, such as image processing, feature extraction, and classification algorithms. Along with highlighting developments that have increased diagnosis accuracy, it also looks at the difficulties with present approaches, such as the intricacy of brain imaging. By evaluating the state-of-the-art methodologies, the authors seek to highlight how AI is revolutionizing the way brain cancers are diagnosed and classified, ultimately leading to more effective and precise therapies.

Singh, Yogesh, Pradeep Kumar Bhatia et al. [8]The application of machine learning techniques in software development, specifically for expert estimation, is thoroughly reviewed in this study. By "learning" from historical data, machine learning has demonstrated significant promise in producing predictions that are more accurate. Helping academics comprehend how machine learning can be used to enhance expert estimates in software projects is the aim of this paper.

Several machine learning techniques that are frequently used to forecast software development are covered in the study, including neural networks, genetic algorithms, decision trees, and case-based reasoning. The authors point out that the kind of data and the field in which these methods are used determine how effective they are. The research concludes that although these approaches can compete with conventional estimating methods, their accuracy may differ depending on the dataset used for training.

Received: 2025-08-08
Accepted: 2025-08-25
Published Online: 2025-09-20
DOI: 10.70454/JRIST.2025.10101

Journal of Recent Innovation in Science and Technology
E-ISSN: 3117-3926

Datasets

This work leverages a dataset for the categorization of brain tumors, supplied from the Kaggle website. There are 493 MRI pictures of patients without tumors and 982 pictures of patients with tumors in the dataset. Images that show a brain tumor are labeled '1', while those that don't are labeled '0'.In total, there are 1,475 images included in the dataset. Figure 1 displays a selection of 18 brain tumor images from Dataset-A.

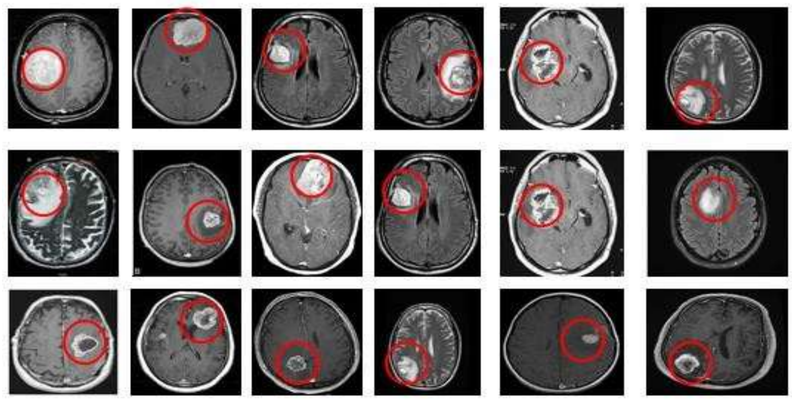

Fig.1.Collectionof18braintumorMRIimagesfromDataset-A

Types of Machine Learning Algorithms

In this context, several advanced machine learning (ML) algorithms that are effective in detecting brain tumors have been analyzed. 1. The Support Vector Machine (SVM) is a supervised ML classification algorithm that utilizes a hyperplane model for separation. The primary objective of SVM is to accurately distinguish between two classes of data [24]. 2. Logistic Regression is a supervised ML approach applied to anticipate a binary outcome based on a collection of independent variables. Finding the best model to depict the link between the outcome and the predictor variables is its primary goal[25]. 3. For binary classification tasks, the K-Nearest Neighbor (KNN) algorithm is a supervised machine learning technique. By calculating the distance between a particular data point and other points, KNN assigns it to the class of the closest data points. The number of surrounding points taken into consideration is indicated by the "K" in KNN.

[26]. 4. Naïve Bayes (NB) is a supervised ML algorithm primarily utilized for binary classification. It uses the Bayes theorem and assumes that the predictors are independent[27].According to the NB classifier, a feature's existence within a class has no bearing on the existence of any other features. 5. A supervised learning model called the Decision Tree (DT) was created for applications involving binary classification. DTs use the data's properties to generate basic decision rules that forecast a target variable's value[28]. 6. Random Forest is an ensemble ML algorithm that constructs multiple decision trees and combines their outputs [29]. This approach enhances accuracy and mitigates the risk of overfitting, which can occur with large datasets in decision trees. Random Forest is applicable for both classification and regression tasks.

Received: 2025-08-08
Accepted: 2025-08-25
Published Online: 2025-09-20
DOI: 10.70454/JRIST.2025.10101

Journal of Recent Innovation in Science and Technology
E-ISSN: 3117-3926

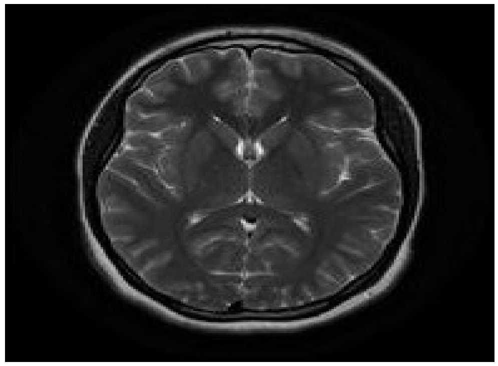

Fig.2.Original Image

Discussion

A comparative evaluation of the nine algorithms mentioned above has been conducted using the following performance metrics:

- **Accuracy:** The proportion of correctly classified instances.

$$\text{Accuracy} = (TP+TN)/(TP+FP+FN+TN) \quad (1)$$

- **Precision:** The proportion of true positives among all positive predictions.

$$\text{Precision} = TP/(TP+FP) \quad (2)$$

- **Recall:** The proportion of true positives among all actual positive instances.

$$\text{Recall} = TP/(TP+FN) \quad (3)$$

- **F1-Score:** The harmonic mean of precision and recall.

$$\text{F1-score} = (2*\text{Precision}*\text{Recall})/(\text{Precision} + \text{Recall}) \quad (4)$$

Where, the following attributes have been used in the measurement:

- True Positive (TP)
- True Negative (TN)
- False Positive (FP)
- False Negative (FN)

Table 1 presents a comparison of the performance metrics of 6 machine learning algorithms with above four attributes. The algorithms evaluated include SVM, Logistic Regression, K- Nearest Neighbors, NB, Decision Tree and Random Forest.

Table1. Attributes for the calculation of performance metrics of the six ML algorithms

Algorithm	TP	TN	FP	FN
SVM	95	219	24	31
Logistic Regression	92	220	23	34
KNN	102	210	33	24
NB	79	205	38	47
DT	104	215	28	22
Random Forest	120	215	25	6

The performance of different machine learning algorithms can be evaluated using various metrics, including accuracy, precision, recall, and F1-score have been depicted in

Table 2 Comparison of the different ML algorithms based on accuracy, recall, precision and F1-score

Received: 2025-08-08
Accepted: 2025-08-25
Published Online: 2025-09-20
DOI: 10.70454/JRIST.2025.10101

Journal of Recent Innovation in Science and Technology
E-ISSN: 3117-3926

model	accuracy	recall	precision	F1-score
SVM	0.851	0.754	0.798	0.775
LR	0.846	0.730	0.800	0.763
KNN	0.845	0.804	0.756	0.779
NB	0.769	0.627	0.675	0.650
DT	0.865	0.825	0.788	0.806
RF	0.908	0.952	0.811	0.876

By examining the table, we can compare the performance of these models and identify the strengths and weaknesses of each. For instance:
- Random Forest(RF) has the highest accuracy (0.908) and recall (0.952).
- Support Vector Machine (SVM) has a relatively low recall (0.754) but higher precision (0.798).
- NB has the lowest accuracy (0.769) and recall (0.627).

The results presented in Table 2 reveal that the Gradient Boosting classifier has achieved the highest accuracy among all the implemented machine learning models. With an accuracy of 0.924, Gradient Boosting outperformed the other models, followed closely by Random Forest, which achieved an accuracy of 0.908.

ROC curves of the different algorithms for performance comparison

The ROC curves corresponding to SVM, Logistic Regression, KNN, NB, DT classifier, Random Forest classifier have been shown in below figure.

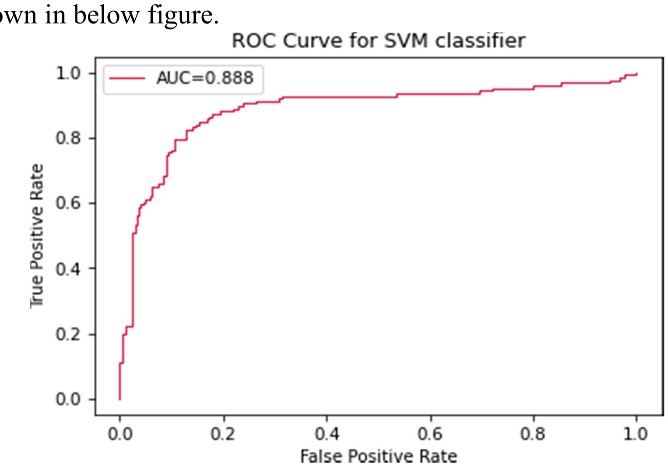

Fig.3.ROC curve of SVM classifier

Received: 2025-08-08
Accepted: 2025-08-25
Published Online: 2025-09-20
DOI: 10.70454/JRIST.2025.10101

Journal of Recent Innovation in Science and Technology
E-ISSN: 3117-3926

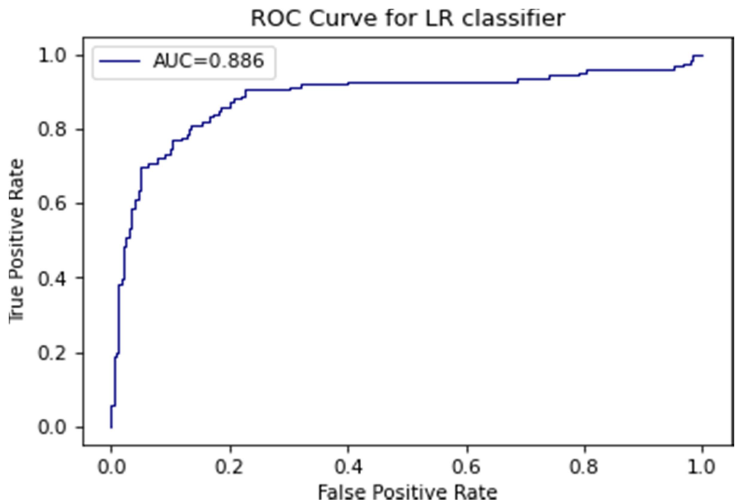

Fig.4. ROC curve of Logistic Regression classifier

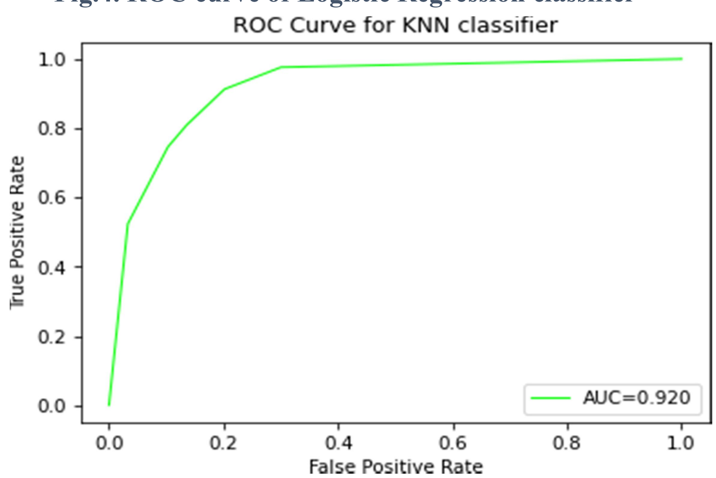

Fig.5. ROC curve of KNN classifier

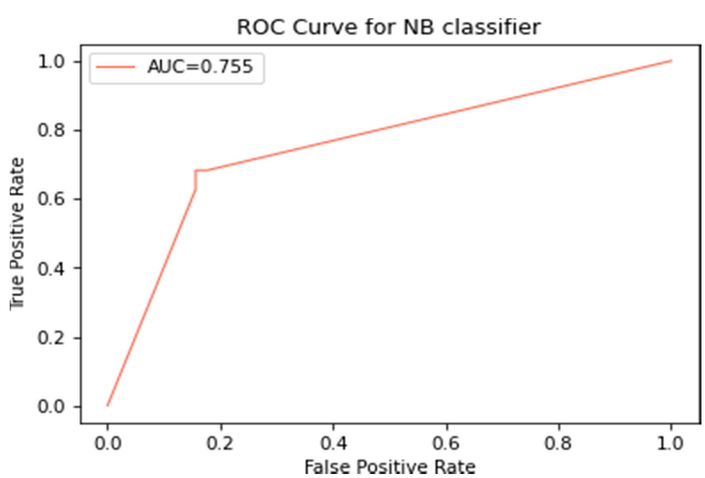

Fig.6. ROC curve of NB classifier

Received: 2025-08-08
Accepted: 2025-08-25
Published Online: 2025-09-20
DOI: 10.70454/JRIST.2025.10101

Journal of Recent Innovation in Science and Technology
E-ISSN: 3117-3926

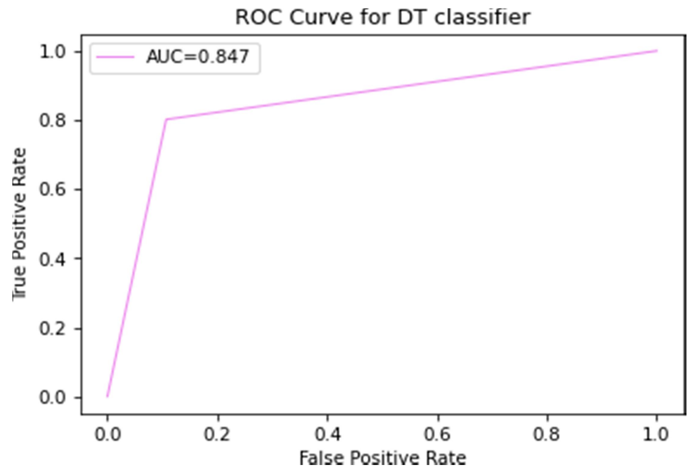

Fig.7.ROCcurveofDTclassifier

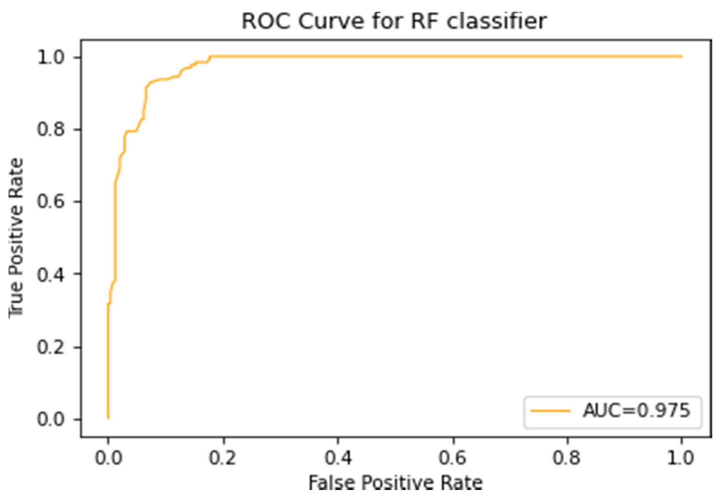

Fig.8. ROC curve of Random Forest classifier

Conclusion and Scope of Future Work

This paper presents a detailed overview of how ML algorithms can be used for medical image processing. ML has improved and paved the way for efficient diagnosis, recognition and prediction in numerous domains of healthcare, brain tumor detection and classification being one of them. Six ML algorithms have been used to predict whether a patient has a brain tumor or not based on a dataset comprising of brain MRI images.

The ML algorithms that have been used are Support Vector Machine (SVM), Logistic Regression, K-Nearest Neighbor (KNN), Naïve Bayes (NB), Decision Tree (DT) classifier and Random Forest classifier. A performance comparison of the different ML algorithmshas beenconducted basedon a few performance metrics such as accuracy, recall, precision, F1- Score and ROC curves. After the evaluation of the test scores, it has been concluded that random forest technique is the best classifier among all the other ML classifiers that have been used.

In future, one of the most important improvements that can be made is adjusting the architecture so that it can be used during brain surgery, for classifying and accurately locating the tumor.

Received: 2025-08-08
Accepted: 2025-08-25
Published Online: 2025-09-20
DOI: 10.70454/JRIST.2025.10101

Journal of Recent Innovation in Science and Technology
E-ISSN: 3117-3926

References

[1] DeAngelis, Lisa M. "Brain tumors." *New England journal of medicine* 344.2 (2001): 114-123.
[2] Herholz, Karl, et al. "Brain tumors." *Seminars in nuclear medicine*. Vol. 42. No. 6. WB Saunders, 2012.
[3] Amin, Javaria, et al. "Brain tumor detection using statistical and machine learning method." *Computer methods and programs in biomedicine* 177 (2019): 69-79.
[4] Amin, Javaria, et al. "Brain tumor detection and classification using machine learning: a comprehensive survey." *Complex & intelligent systems* 8.4 (2022): 3161-3183.
[5] Painuli, Deepak, and Suyash Bhardwaj. "Recent advancement in cancer diagnosis using machine learning and deep learning techniques: A comprehensive review." *Computers in Biology and Medicine* 146 (2022): 105580.
[6] Hatami, Toktam, et al. "A machine learning approach to brain tumors segmentation using adaptive random forest algorithm." *2019 5th Conference on Knowledge Based Engineering and Innovation (KBEI)*. IEEE, 2019.
[7] Solanki, Shubhangi, et al. "Brain tumor detection and classification using intelligence techniques: an overview." *IEEE Access* 11 (2023): 12870-12886.
[8] Singh, Yogesh, Pradeep Kumar Bhatia, and Omprakash Sangwan. "A review of studies on machine learning techniques." *International Journal of Computer Science and Security* 1.1 (2007): 70-84.
[9] Singh, Soni, K. R. Ramkumar, and Ashima Kukkar. "Machine learning techniques and implementation of different ML algorithms." *2021 2nd Global Conference for Advancement in Technology (GCAT)*. IEEE, 2021.
[10] Kotsiantis, Sotiris B., Ioannis D. Zaharakis, and Panayiotis E. Pintelas. "Machine learning: a review of classification and combining techniques." *Artificial Intelligence Review* 26 (2006): 159-190.