Analysis of machine learning algorithms in brain tumour prediction

M Gupta¹, S K Sharma¹, R Saxena¹, S Arora²

¹School of Information and Communication Technology, Gautam Buddha University, Greater Noida 201312, India
²Department of Computer Science, NIET, Greater Noida 201312, India

manaliguptasingh@gmail.com

Abstract. The tumour is fundamentally an excessive development of dangerous cells in any part of the body, while a tumour in a brain is an unreasonable development of cancerous cells in the brain. Brain tumour can be either benign or malignant. The benign brain tumour has structural consistency and does not include active (cancer) cells, but the malignant brain tumour has no structure consistency and includes active cells. The primary concern is to segment, detect, and extract the infected tumour area from magnetic resonance images (MRI) which are being performed by radiologists or medical experts, and their accuracy is totally dependent on their experience only. Thus, it becomes very essential to overcome these limitations by the use of artificial intelligence. The current paper uses various machine learning algorithms as well as their features to design a structure to predict brain tumour at an early phase by using different classifiers and comparing their respective accuracy parameters.

1. Introduction
In today’s world, IT and e-health care provides the best interface or acts as a bridge between the clinical experts and the patients to provide the improved health care. Machine learning likely covers all the fields of medicine starting from drug detection to the final phase of decision making. The process of practice of medicine has been completely changed by the new era of machine learning. Presently, a radiologist is restricted by his pace, exhaustion and lack of knowledge or practice analyses all the medical images which are an important part of a patient. To train a trained radiologist it takes immense effort, time and huge financial cost. Delayed or incorrect diagnosis cause harm to the patient. According to World Health Organization and American Brain Cancer Association, the general grading scheme uses a scale from grade one to grade four to classify types of tumours i.e., low grade and high-grade tumour depending on their behaviour. Low grade tumour or benign tumours are slow growing and fall under grade one (I) and grade two (II) whereas high grade tumour or malignant tumours are fast growing and fall under grade three (III) and grade four (IV) and can be termed as cancerous growths. If benign tumour is left untreated or not treated on time it changes into a malignant tumour. A person can be affected by brain tumour at any age and its effect on the individual’s body can differ from person to person. A patient who gets diagnosed with grade two tumours has to undergo serial examination scans either by magnetic resonance imaging (MRI) or by computed tomography (CT) in a timely manner. Under the surgical process only low grade or benign tumours can be cured however radiotherapy, chemotherapy or a combination of both are the treatments of malignant or the high-grade tumours.
2. Related Work
Over the years, a good amount of research work has been accomplished to forecast the brain tumour using machine learning technique. Authors proposed distinctive dimensionality reduction and cross-validation method by implementing Naive Bayes (NB) [1], Gaussian Process Classification (GPC) [2], Support Vector Machine (SVM) [3], Artificial Neural Network (ANN) [4], Ada Boost (AB) [5], Logistic Regression (LR) [6], Decision Tree (DT), and Random Forest (RF) [7]. Broad tests were likewise completed for dismissing the outliers and filling missing qualities by processing the mean and the median for improving the accuracy and performance of Machine Learning model. Various methods have been developed to classify the tumours [8] during the last few years in which binary thresholding, Principal Component Analysis (PCA) [9, 10] and Support Vector Machine (SVM) classifier [11] methods are commonly used as segmentation, feature extraction, and classification methods, respectively [12]. Convolution Neural Network was used for classification [13], segmentation [14], and so on especially in image recognition and classification. It is observed from the above literature, that various approaches have been proposed for classification. In all the techniques that are proposed, the Support Vector Machine methodology can be called as best for classification. To classify the brain tumour from MRI correctly as cancerous or non-cancerous with better precision and accuracy is the aim of this research paper. The rest of the paper is ordered as follows: Section 3 deals with the exposition of the Materials and methods used in the study, and Section 4 has been devoted to presenting the results of the experimental assessment of the calculations. Finally, Section 5 is the conclusion and the future scope.

3. Materials and methods

3.1 Dataset Description
The accuracy and effectiveness of various classifiers were evaluated over the BRATS MICCAI 2015 dataset having 3762 sets of T1 subjective tumour type images and are of three kinds namely Meningioma, Glioma and pituitary, of which 1683 were having tumour and 2079 were of non-tumour. Image acquired using MRI is affected by noise and unwanted artefacts which needs to be removed before the image is processed to determine tumour. The preferred tool in pre-processing technique is median filter which filters noise and unusual observation by substituting the value of a pixel by the median of the intensity values in its immediate neighbourhood [10]. In comparison to other filters, the median filter preserves edges in an image as well as it do not blur the image. Various machine learning techniques have been implemented using Weka (3.9.4). The complete description of the dataset is shown in Table 1.
Table 1. Description of BRATS MICCAI brain tumor dataset

| S. No | Attribute     | Mean     | Standard Deviation | Min Value | Max Value |
|-------|---------------|----------|--------------------|-----------|-----------|
| 1     | Mean          | 9.489    | 5.728              | 0.079     | 33.24     |
| 2     | Variance      | 711.101  | 467.467            | 3.146     | 2910.582  |
| 3     | Standard Deviation | 25.182  | 8.774              | 1.774     | 53.95     |
| 4     | Entropy       | 0.074    | 0.07               | 0.001     | 0.395     |
| 5     | Skewness      | 4.103    | 2.561              | 1.886     | 36.931    |
| 6     | Kurtosis      | 24.389   | 56.435             | 3.942     | 1371.64   |
| 7     | Contrast      | 127.961  | 109.5              | 3.195     | 3382.574  |
| 8     | Energy        | 0.205    | 0.129              | 0.025     | 0.59      |
| 9     | ASM           | 0.059    | 0.058              | 0.001     | 0.348     |
| 10    | Homogeneity   | 0.479    | 0.128              | 0.105     | 0.811     |
| 11    | Dissimilarity | 4.698    | 1.85               | 0.681     | 27.828    |
| 12    | Correlation   | 0.956    | 0.026              | 0.549     | 0.99      |
| 13    | Coarseness    | 0        | 0                  | 0         | 0         |
| 14    | Class         | Tested Positive: Brain Tumor | Tested Negative: Non-Brain Tumor |

3.2 Methodology

The present study has been conducted using the machine learning algorithms such as Naïve Bayes, Support Vector Machine, Neural Networks, K nearest neighbor, Statistical Classifier (J48) and Ensemble Classifiers using voting technique. The classifier algorithms with their respective confusion matrix are discussed below.

3.2.1 Naïve Bayes

For binary (two-class) and multi-class classification problems Naïve Bayes classification algorithm is used. This method is simple and easy to understand particularly when binary or categorical input values are used. In this classifier each feature identifies individually without depending on the other features for example if on the basis of colour, taste and shape a fruit is identified then orange, sweet tart and round fruit will be identified as orange. It is generally used for huge data sets, and it is comparatively the easiest and the fastest machine learning algorithm. In most cases good results have been obtained when it is used. The confusion matrix is used to explain the Naïve Bayes classifier’s performance in Table 2 shown below.

Naïve Bayes classification method is based on conditional probability and makes use of Bayes rule which signifies that by the prior probability \( P(H) \) when hypothesis \( H \) is true, probability of the evidence \( P(E) \) and likelihood probability \( P(E|H) \), we can calculate the conditional probability of the hypothesis when evidence is there \( P(H|E) \). Below is the formula for calculating the conditional probability.

\[
P(H|E) = \frac{(P(E|H) \times P(H))}{P(E)}
\]

Table 2. Confusion Matrix of Naïve Bayes algorithm

|                | Brain tumor | Non-Brain Tumor |
|----------------|-------------|-----------------|
| Brain tumor    | 1583        | 100             |
| Non-Brain Tumor| 17          | 2062            |

Classification has classified the patients into:

- 1583(True Positives): patients with brain tumor that were correctly classified.
- 2062 (True Negatives): Patients without brain tumor and were correctly classified.
- 100 (False Positives): Patients that did not have brain tumor, but algorithm misclassified by saying that they have brain tumor.
- 17 (False Negative): Patients who have brain tumor, but algorithm classified wrongly by saying that they do not have brain tumor.

3.2.2 K-Nearest Neighbour: Classifies objects based on feature similarity where features are input variables. Despite being under supervised machine learning KNN does not have any training phase and uses all the data in its training phase in classification method. It does not have any prior knowledge, so it does not assume anything about the data. Because of these properties it is often described as non-parametric and lazy algorithm.

| Table 3. Confusion Matrix of K nearest Neighbour (KNN) algorithm |
|---------------------------------------------------------------|
| Brain tumor | Non-Brain Tumor |
| Brain tumor | 1635 | 48 |
| Non-Brain Tumor | 31 | 2066 |

3.2.3 Support Vector Machine (SVM). It comes under the category of supervised machine learning and analyses the data used for classification analysis. Support vector machine is an illustration of the training data as points in space separated into categories by a clear gap that is as wide as possible. Gap is maintained by drawing a line which is parallel to the hyper plane. There is a boundary between two categories which predicts where new data/examples will belong to. It is effective in high dimensional spaces. Kernels are provided for the decision functions.

| Table 4. Confusion Matrix of SVM algorithm |
|-------------------------------------------|
| Brain tumor | Non-Brain Tumor |
| Brain tumor | 1611 | 72 |
| Non-Brain Tumor | 10 | 2069 |

3.2.4 Neural Network: Neural network’s functioning is like the human brain’s neural network and is used in unsupervised machine learning. It consists of nodes (neurons) which are interconnected in different layers where each node is responsible for some computation. It converts an input value into some output value and can generate large different values of output set without needing to redesign the output criteria as it is adaptable to varying input values. In neural network information is passed from input layer through the hidden layers to the output layer i.e., forwarding in one direction is a feed forward neural network. There are no loops in this network.

| Table 5. Confusion Matrix of Decision Tree algorithm |
|-----------------------------------------------------|
| Brain tumor | Non-Brain tumor |
| Brain tumor | 1655 | 28 |
| Non-Brain tumor | 17 | 2062 |

3.2.5 Decision Tree: It is a classification modelling tool which represents the nodes in a flowchart/tree like graph structure. In this algorithm the data set is broken down into smaller subsets and it is recursive in nature. This process is cyclic for each sub tree which is being rooted at the new nodes. A
tree with the leaf nodes and the decision nodes is the result of a tree. In a tree, each node represents the test case for some attribute, each edge signifies the answers to test case and leaf nodes basically provides the actual output or can be defined as class label. Both categorical data and numerical data can be handled.

| Table 6: Confusion Matrix of Support Vector Machine algorithm |
|---------------------------------------------------------------|
| Brain tumor | Non-Brain tumor |
| Brain tumor   | 1643 | 40 |
| Non-Brain tumor | 22  | 2057 |

3.2.6 **Ensemble method**: It is a technique under machine learning algorithms that combines various base models in order to produce one optimal predictive model. The goal of any problem in machine learning is to find a model which can produce the best or the accurate solution. Comparatively making one model and hoping the model to make the best or the most accurate predictor, ensemble methods can be used which take myriad of models and then the average of those models is done to calculate one final model. In this support vector machine and decision tree algorithms are being used.

| Table 7. Confusion Matrix of Ensemble method Algorithm |
|-------------------------------------------------------|
| Brain tumor | Non-Brain tumor |
| Brain tumor   | 1610 | 73 |
| Non-Brain tumor | 6  | 2073 |

4. Results and discussion
We have used Naïve Bayes, K Nearest neighbour, Support Vector Machine, Decision Trees, Neural network, and Ensemble methods in this research work. We have performed experiments using 10-folds cross validation technique. Parameters such as precision, recall, accuracy, f-measure, receiver operating curve (ROC) and area under the curve (AUC) are used for the classification. Results are discussed in Table 8.

| Table 8. Classification Results |
|---------------------------------|
| Sno  | Classifier          | Precision | Recall | F-Measure | ROC Area | Accuracy |
|------|--------------------|-----------|--------|-----------|----------|----------|
| 1    | Naïve Bayes        | 0.989     | 0.941  | 0.964     | 0.992    | 96.89    |
| 2    | Support vector Machine | 0.994  | 0.957  | 0.975     | 0.976    | 97.61    |
| 3    | K nearest Neighbor | 0.992     | 0.971  | 0.982     | 0.982    | 98.38    |
| 4    | Neural Network     | 0.99      | 0.983  | 0.987     | 0.997    | 98.80    |
| 5    | Decision Trees(j48) | 0.987  | 0.976  | 0.981     | 0.985    | 98.35    |
| 6    | Ensemble (Decision Tree+ SVM) | 0.996  | 0.957  | 0.976     | 0.99     | 98.9     |

**Precision**: is used to approximate the throughput of classification. It is defined as the fraction of related instances among all the possible retrieved instances. For the various machine learning algorithms comparative approach for precision is depicted in Figure 1.
Recall: Basically, it defines the actual positives that were identified correctly. For the various machine learning algorithms comparative approach for precision is depicted in Figure 2.

F measure: helps in providing a single score that balances both the concerns of precision and recall in a number. For the various machine learning algorithms comparative approach for precision is depicted in Figure 3.

ROC Curve: A higher X axis value signifies an upper number of false positives than True Negatives and higher Y axis indicates a higher number of True Positives than False Negatives. For the various machine learning algorithms comparative approach for precision is depicted in figure 4.
Figure 4. ROC Comparison

Accuracy: It shows that if the accuracy is higher or more, then the performance of the model in distinguishing the positive and negative class is better. For the various machine learning algorithms comparative approach for precision is depicted in Figure 5.

Figure 5. Accuracy Comparison

5. Conclusion and future work
In this paper, we have attempted to recognize the best machine learning classification method to anticipate the brain tumour at a beginning phase. During our examination work, we have considered six order procedures and assessed them on different performance measures. Examinations were carried on Brain Tumour dataset and tool used was Weka. Experimental results establish that Ensemble classifier technique gave the best accuracy to predict brain tumour. Future work will include data augmentation and convolutional neural network model to classify the tumour which will additionally upgrade accuracy.

References
[1]. D. Garrett, D. A. Peterson, C. W. Anderson and M. H. Thaut, "Comparison of linear, nonlinear, and feature selection methods for EEG signal classification," in IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 11, no. 2, pp. 141-144.
[2]. G. Webb, J. Boughton and Z. Wang, "Not So Naive Bayes: Aggregating One-Dependence Estimators", Machine Learning, vol. 58, no. 1, pp. 5-24.
[3]. S. Brahim-Belhouari and A. Bermak, “Gaussian process for non-stationary time series prediction,” Comput. Statist. Data Anal., vol. 47, no. 4, pp. 705-712.
[4]. C. Cortes and V. Vapnik, "Support vector networks Mach. Learn", vol. 20, no. 3, pp. 273-297.
[5]. A. Reinhardt, T. Hubbard, Using neural networks for prediction of the sub cellular location of proteins, *Nucleic Acids Research*, vol. 26, Issue 9, 1 May 1998, pp. 2230–2236.

[6]. Kégl B (2013) the return of AdaBoost.MH: multi-class Hamming trees. *arXiv*:1312.6086

[7]. A. Saffari, C. Leistner, J. Santner, M. Godec and H. Bischof, "On-line Random Forests," 2009 IEEE 12th International Conference on Computer Vision Workshops, ICCV Workshops, Kyoto, 2009, pp. 1393-1400.

[8]. H. Mohsen, E. El-Dahshan, E. El-Horbaty, A. Salem, *Brain tumor type classification based on support vector machine in magnetic resonance images*, Annals Of Dunarea De Jos University Of Galati, Mathematics, Physics, Theoretical mechanics, Fascicle II, Year IX (XL).

[9]. Karl Pearson F.R.S. (1901) LIII. On lines and planes of closest fit to systems of points in space, *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 2:11, 559-572, DOI: 10.1080/14786440109462720

[10]. Kaya, I.E., Pehlivanl, A.A., Sekizkarde, E.G. and Ibrikci, T. (2017) PCA Based Clustering for Brain Tumour Segmentation of T1w MRI Images. *Computer Methods and Programs in Biomedicine*, vol. 140, pp.19-28.

[11]. Cortes, C., Vapnik, V. Support-vector networks. *Mach Learn* 20, 273–297 (1995).

[12]. N. Abdullah, Lee Wee Chuen, U. K. Ngah and Khairul Azman Ahmad, "Improvement of MRI brain classification using principal component analysis," 2011 *IEEE International Conference on Control System, Computing and Engineering*, Penang, 2011, pp. 557-561, doi: 10.1109/ICCSCE.2011.6190588.

[13]. Abiwinanda N., Hanif M., Hesaputra S.T., Handayani A., Mengko T.R. (2019) Brain Tumour Classification Using Convolutional Neural Network. In: Lhotska L., Sukupova L., Lacković I., Ibott G. (eds) *World Congress on Medical Physics and Biomedical Engineering* 2018. IFMBE Proceedings, vol 68/1. Springer, Singapore

[14]. S. Pereira, A. Pinto, V. Alves and C. A. Silva, "Brain Tumour Segmentation Using Convolutional Neural Networks in MRI Images," in *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1240-1251.