COMPARATIVE ANALYSIS OF BRAIN TUMOR PREDICTION BY IMPLEMENTING LOGISTIC REGRESSION OVER SUPPORT VECTOR MACHINE

Nishanth Ganesh[1], Dr.Sivaprasad[2*]

Article History: Received: 12.12.2022 **Revised:** 29.01.2023 **Accepted:** 15.03.2023

Abstract

Aim: This research is about the novel brain tumor detection of Brain tumors by implementing Machine learning algorithms like Logistic Regression (LR) and comparing its accuracy with a Support Vector Machine (SVM).
Materials and Methods: Two groups, namely Logistic Regression and Support Vector Machine algorithm used to find the accuracy of Brain tumor prediction with 20 samples each to evaluate this study. The sample size was calculated using G power with pretest power at 80% and alpha value of 0.05. Brain MRI of Normal and Brain tumors were used as data models to train with the LR and SVM algorithms.
Results: The accuracy of LR is 87.5% and the SVM with 75% and there is a statistical significance observed as 0.890.
Conclusion: Logistic Regression algorithm has more accuracy compared to the Support Vector Machine.

Keywords: Novel Brain tumor detection, MRI image, Machine learning, Logistic Regression(LR), Support Vector Machine(SVM), Accuracy.

[1]Research Scholar, Department of Biomedical Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, Tamilnadu, India, Pincode:602105
[2*]Project Guide, Department of Biomedical Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, Tamilnadu, India, Pincode: 602105

1. Introduction

The brain tumor is an abnormal growth of cells in a mass that is both cancerous and non-cancerous which is very hard to treat therefore early detection of brain tumors will be very useful for doctors and physicians to treat the tumor effectively (Hattingen and Pilatus 2015). The most preferred and generally used technique to diagnose Brain tumors is Magnetic Resonance imaging or MRI image, sometimes PET and CT imaging are also used. In the field of medical image processing, machine learning techniques provide an accurate prediction and diagnosis compared to human-assisted manual classification (Kevin Zhou, Rueckert, and Fichtinger 2019). Image segmentation is one of the most crucial parts in detecting the tumor and lesions also improving the sensitivity and specificity of the tumor which is the primary part of medical image processing (Kevin Zhou, Greenspan, and Shen 2017). Due to poor accuracy in manual prediction techniques many of the researchers have implemented Machine learning and deep learning algorithms to predict the brain tumors in large sets of samples of MRI images with better accuracy (Menze and Bakas 2021);(Wu, Shen, and Sabuncu 2016). There are various cycles to distinguish brain tumors in MRI images. In Image processing, feature extraction, and more calculations have been executed for the identification of brain tumors, yet a persuasive and exact method to identify the tumor's exact position and to analyze in a short period is in incredible need. The execution and complexity associated with the medical image segmentation methods are improved by applying feature extraction procedures and mathematical models. From this, we are able to understand that machine learning techniques play a vital role in medical image processing and these algorithms also predict various diseases like Brain tumors, liver cancer, lung tumors, etc. The bloom of machine learning and deep learning techniques in image processing have been developed to provide better segmentation and accuracy in medical imaging (Lu et al. 2017).

Since 2013, 27 publications have been published in Science Direct, and 35 papers have been published in IEEE and 21 publications in Springer based on the prediction of brain tumors using the proposed algorithm and the existing approach. In the referred research paper Sachdeva proposed a multi-class brain tumor classification and prediction method based on statistical composition and intensity features. The features were trained by an artificial neural network (ANN) model. In improvement, the accuracy of this method was enhanced from 77% to 91% by reducing the dimensionality of the characteristics vector with principal component analysis (PCA) (Sachdeva et al. 2013). Eman Abdel-Maksoud proposed a method for image segmentation that combines the K-means clustering technique with the Fuzzy C-means algorithm. To give an accurate brain tumor detection, it is followed by thresholding and level set segmentation steps. In terms of least calculation time, the proposed technique can benefit from K-means clustering for image segmentation. Furthermore, it can benefit from the Fuzzy C-means in terms of accuracy. In terms of accuracy, processing time, and performance, the suggested image segmentation approach was evaluated by comparing it to some segmentation algorithms ("Brain Tumor Segmentation Based on a Hybrid Clustering Technique" 2015)). Hunner et al., exhibited the use of thresholding techniques to locate brain tumors and described a tumor recognition comparative research. The acquired data are displayed, demonstrating that the Sobel edge detection operator can efficiently identify tumors and extract tumor boundaries. It determines the tumor's size and stage. For accurately detecting cancers, MRI image pictures seem to be the most useful. The use of digital imaging methods to diagnose a brain tumor in MRI images is crucial in this research (Hunnur, Raut, and Kulkarni 2017). In this paper devkota et al., proposed an approach for brain tumor segmentation is established. It focuses on how to improve tumor segmentation using mathematical morphology and FCM (Fuzzy C-Means) algorithm which improves the computation time, but the propounded result has not been tested up to the estimation stage and precipitates as-detects cancer with 92% and classifier has a precision of 86% ("Image Segmentation for Early Stage Brain Tumor Detection Using Mathematical Morphological Reconstruction" 2018), ("Brain Tumor Segmentation Based on a Hybrid Clustering Technique" 2015).

Our institution is keen on working on latest research trends and has extensive knowledge and research experience which resulted in quality publications (Rinesh et al. 2022; Sundararaman et al. 2022; Mohanavel et al. 2022; Ram et al. 2022; Dinesh Kumar et al. 2022; Vijayalakshmi et al. 2022; Sudhan et al. 2022; Kumar et al. 2022; Sathish et al. 2022; Mahesh et al. 2022; Yaashikaa et al. 2022). Based on the above analysis, inaccurate prediction of brain tumors is the main problem. Early detection of brain tumors have become extremely crucial in terms of patient survival. The computer-aided diagnosis method based on machine learning is better in aspects of accuracy and efficient utilization of picture information. Brain tumor MRI images are utilized in the study to detect and classify tumors. The study proposes a novel brain tumor detection method and improving the accuracy and comparing the accuracy of Logistic Regression with the Support Vector Machine.

2. Materials and Methods

The study is conducted in the Image Processing Lab, Saveetha School of Engineering. For the novel brain tumor detection of brain tumors, the proposed effort incorporates two groups. Logistic Regression had 20 samples in Group 1 and Support Vector Machine had 20 samples in Group 2. Using clinclac.com, a pre-test analysis was created with G power of 80 percent and an alpha value of 0.05. For sample preparation in groups 1 and 2, brain MRI images were used. The software tool utilized was a Jupyter notebook running Anaconda with an Intel Core i5 processor and 4GB of RAM. The system had a Windows operating system, a 64-bit processor, and a 917GB hard drive (Feng et al. 2019).

Brain MRI images are chosen using the proposed method by filtering out low-resolution and irregular images. Selected images are subjected to a preprocessing and segmentation process, which simplifies the image representation and makes data analysis easier. In feature extraction, three feature descriptors are used to extract global information from an image, color, shape, and textures. For novel brain tumor detection, The training dataset is used by logistic regression algorithms in order to predict brain tumors. A Support Vector Machine is used to repeat the technique. The accuracy of both LR and SVM is observed and compared.

The calculations for predicting brain tumors were performed using the below equations (1 to 4). The accuracy of both LR and SVM algorithms is measured by the ratio of correctly predicted observations to the total observations given in equation (1). The recall is the ratio between the positive observations predicted correctly and all the observations in the class yes (2). The ratio of accurately predicted positive observations to total expected positive data is known as precision (3). F1 Score is the weighted average of Accuracy and Recovery (4). A true positive is when the model predicts the positive class properly. A true negative, on the other hand, is a result in which the model accurately predicts the negative class. A false positive occurs when the model predicts the positive class inaccurately. A false negative is an outcome in which the model predicts the negative class incorrectly.

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

$$\text{Recall} = \frac{TP}{TP+FN} \quad (2)$$

$$\text{Precision} = \frac{TP}{FP+TP} (3)$$

$$\text{F1-Score} = 2 \times \frac{Precision \times Recall}{Precision+Recall} \quad (4)$$

Statistical Analysis

The IBM SPSS software was used to do the statistical analysis. For data analysis using both LR and SVM algorithms. Twenty samples were taken for each of the proposed methods, and the predicted accuracy was recorded for analysis in the SPSS tool's MS Excel sheet. The independent variables are numerous locations with varying tumor impacts on the brain, while the dependent variables are accuracy (Stehlik-Barry and Babinec 2017).

3. Results

Table 1 denotes the accuracy obtained from a novel Brain tumor detected 20 samples of two groups. LR is group 1 which is the proposed method has an accuracy of 87.5 % and group 2 is the SVM algorithm which has an accuracy of 75 %. From the table 1 comparison, it is analyzed that LR has better accuracy than SVM.

Table 2 denotes group statistics of novel brain tumor detected samples obtained from statistical analysis by the IBM spss tool for LR and SVM algorithms. The mean value of LR is 86.3 % and SVM is 74.14 %. The standard deviation and standard mean error of LR is better than SVM.

Table 3 denotes the independent sample test obtained from statistical analysis using algorithms like LR and SVM with the IBM spss tool. The significance value obtained from the proposed method is 0.890.

Figure 1 denotes the comparison difference between the accuracy obtained from LR and SVM algorithms with error bar values. From fig. 1. It is known that LR has better accuracy than SVM.

Figure 2 denotes the workflow of novel brain tumor detection of brain tumors by using Logistic Regression and Support Vector Machine. In fig.2 the five main processes involved in the importing libraries, data acquisition, data preprocessing, image segmentation and feature extraction.

4. Discussion

In this study Logistic Regression has obtained higher accuracy than Support Vector Machines in the novel brain tumor detection of brain tumor with a significant value of 0.890. A study done by lina chato had done research using machine learning and deep learning techniques to predict the overall survival of brain tumor patients using MRI images. The volumetric analysis and the location features were extracted from the provided data set. The slice images were used to extract the statistical and intensity features. The texture features were used to train machine learning algorithms and to produce the best prediction model. The percentage accuracy

was attained 46% with 10 fold accuracy. The overall accuracy for linear obtained was 68.8% as a result of accuracy (Chato and Latifi 2017).

In research done by alain jungo had done a study using deep learning versus classical regression of brain tumor patient survival prediction. The results have been achieved on the BraTS 2018 training dataset by 100 satisfied CV runs which are reported as mean standard variation and on another with 33 held out (HO) samples. It consists of a baseline of logistic regression with a single feature of age and the accuracy results in 69.7% (Suter et al. 2019). In another study done by Tonmoy Hossain has used the Convolution neural network and has detected the brain tumor, there were 2 distinct models were used and a comparison of other models was done. There were many filtering techniques used in the process to obtain satisfactory results. The logistic regression was used in the filtering techniques and we have obtained 87% as a result of Accuracy (Hossain et al. 2019).

This paper also limits with only a few datasets and data varies for different strains of the tumor and the sample size can be increased for better accuracy. In the future, the proposed work will be improved by having less computational time compared to the existing algorithm, and also the computer-aided diagnostics will also be well developed to classify or predict brain tumors. The Techniques for brain tumor detection and studying tumors in clinical imaging have already shown promising potential, and this trend will definitely continue in the future.

5. Conclusion

Based on the result and tabulations, the Logistic Regression has an accuracy of 87.5% and the Support Vector Machine has an accuracy of 75%. Therefore we conclude that Logistic Regression provides more accuracy in predicting the brain tumor than the Support Vector Machine. The novel brain tumor detection rate of SVM can be augmented by combining other advanced algorithms.

Declarations

Conflict of interest

The authors declare no potential conflict of interest

Authors Contribution

Author NG was involved in data collection, Data analysis, Manuscript writing, Author SP was involved in Conceptualization, Data validation, and Critical review of the manuscript

Acknowledgment

The authors would like to thank Saveetha School of Engineering, Saveetha Institute of Medical and Technical Science (formerly known as Saveetha University) management for providing all necessary facilities to complete the project successfully.

Funding

We thank the subsequent institution for providing financial support that encouraged us to complete the research

1. M2 Santhe, Bengaluru
2. Saveetha University
3. Saveetha Institute of Medical and Technical Science
4. Saveetha School of Engineering

6. REFERENCES

"Brain Tumor Segmentation Based on a Hybrid Clustering Technique." 2015. *Egyptian Informatics Journal* 16 (1): 71–81.

Chato, Lina, and Shahram Latifi. 2017. "Machine Learning and Deep Learning Techniques to Predict Overall Survival of Brain Tumor Patients Using MRI Images." *2017 IEEE 17th International Conference on Bioinformatics and Bioengineering (BIBE)*. https://doi.org/10.1109/bibe.2017.00-86.

Dinesh Kumar, M., V. Godvin Sharmila, Gopalakrishnan Kumar, Jeong-Hoon Park, Siham Yousuf Al-Qaradawi, and J. Rajesh Banu. 2022. "Surfactant Induced Microwave Disintegration for Enhanced Biohydrogen Production from Macroalgae Biomass: Thermodynamics and Energetics." *Bioresource Technology* 350 (April): 126904.

Feng, Jin-Zhou, Yu Wang, Jin Peng, Ming-Wei Sun, Jun Zeng, and Hua Jiang. 2019. "Comparison between Logistic Regression and Machine Learning Algorithms on Survival Prediction of Traumatic Brain Injuries." *Journal of Critical Care* 54 (December): 110–16.

Hattingen, Elke, and Ulrich Pilatus. 2015. *Brain Tumor Imaging*. Springer.

Hossain, Tonmoy, Fairuz Shadmani Shishir, Mohsena Ashraf, M. D. Abdullah Al Nasim, and Faisal Muhammad Shah. 2019. "Brain Tumor Detection Using Convolutional Neural Network." *2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT)*. https://doi.org/10.1109/icasert.2019.8934561.

Hunnur, Shrutika Santosh, Akshata Raut, and Swati Kulkarni. 2017. "Implementation of Image Processing for Detection of Brain Tumors." In *2017 International Conference on Computing Methodologies and Communication (ICCMC)*. IEEE. https://doi.org/10.1109/iccmc.2017.8282559.

"Image Segmentation for Early Stage Brain Tumor Detection Using Mathematical Morphological Reconstruction." 2018. *Procedia Computer Science* 125 (January): 115–23.

Kevin Zhou, S., Hayit Greenspan, and Dinggang Shen. 2017. *Deep Learning for Medical Image Analysis*. Academic Press.

Kevin Zhou, S., Daniel Rueckert, and Gabor Fichtinger. 2019. *Handbook of Medical Image Computing and Computer Assisted Intervention*. Academic Press.

Kumar, J. Aravind, J. Aravind Kumar, S. Sathish, T. Krithiga, T. R. Praveenkumar, S. Lokesh, D. Prabu, A. Annam Renita, P. Prakash, and M. Rajasimman. 2022. "A Comprehensive Review on Bio-Hydrogen Production from Brewery Industrial Wastewater and Its Treatment Methodologies." *Fuel*. https://doi.org/10.1016/j.fuel.2022.123594.

Lu, Le, Yefeng Zheng, Gustavo Carneiro, and Lin Yang. 2017. *Deep Learning and Convolutional Neural Networks for Medical Image Computing: Precision Medicine, High Performance and Large-Scale Datasets*. Springer.

Mahesh, Narayanan, Srinivasan Balakumar, Uthaman Danya, Shanmugasundaram Shyamalagowri, Palanisamy Suresh Babu, Jeyaseelan Aravind, Murugesan Kamaraj, and Muthusamy Govarthanan. 2022. "A Review on Mitigation of Emerging Contaminants in an Aqueous Environment Using Microbial Bio-Machines as Sustainable Tools: Progress and Limitations." *Journal of Water Process Engineering*. https://doi.org/10.1016/j.jwpe.2022.102712.

Menze, Bjoern, and Spyridon Bakas. 2021. *Multimodal Brain Tumor Segmentation and Beyond*. Frontiers Media SA.

Mohanavel, Vinayagam, K. Ravi Kumar, T. Sathish, Palanivel Velmurugan, Alagar Karthick, M. Ravichandran, Saleh Alfarraj, Hesham S. Almoallim, Shanmugam Sureshkumar, and J. Isaac JoshuaRamesh Lalvani. 2022. "Investigation on Inorganic Salts K2TiF6 and KBF4 to Develop Nanoparticles Based TiB2 Reinforcement Aluminium Composites." *Bioinorganic Chemistry and Applications* 2022 (January): 8559402.

Ram, G. Dinesh, G. Dinesh Ram, S. Praveen Kumar, T. Yuvaraj, Thanikanti Sudhakar Babu, and Karthik Balasubramanian. 2022. "Simulation and Investigation of MEMS Bilayer Solar Energy Harvester for Smart Wireless Sensor Applications." *Sustainable Energy Technologies and Assessments*. https://doi.org/10.1016/j.seta.2022.102102.

Rinesh, S., K. Maheswari, B. Arthi, P. Sherubha, A. Vijay, S. Sridhar, T. Rajendran, and Yosef Asrat Waji. 2022. "Investigations on Brain Tumor Classification Using Hybrid Machine Learning Algorithms." *Journal of Healthcare Engineering* 2022 (February): 2761847.

Sachdeva, Jainy, Vinod Kumar, Indra Gupta, Niranjan Khandelwal, and Chirag Kamal Ahuja. 2013. "Segmentation, Feature Extraction, and Multiclass Brain Tumor Classification." *Journal of Digital Imaging* 26 (6): 1141–50.

Sathish, T., V. Mohanavel, M. Arunkumar, K. Rajan, Manzoore Elahi M. Soudagar, M. A. Mujtaba, Saleh H. Salmen, Sami Al Obaid, H. Fayaz, and S. Sivakumar. 2022. "Utilization of Azadirachta Indica Biodiesel, Ethanol and Diesel Blends for Diesel Engine Applications with Engine Emission Profile." *Fuel*. https://doi.org/10.1016/j.fuel.2022.123798.

Stehlik-Barry, Kenneth, and Anthony J. Babinec. 2017. *Data Analysis with IBM SPSS Statistics*. Packt Publishing Ltd.

Sudhan, M. B., M. Sinthuja, S. Pravinth Raja, J. Amutharaj, G. Charlyn Pushpa Latha, S. Sheeba Rachel, T. Anitha, T. Rajendran, and Yosef Asrat Waji. 2022. "Segmentation and Classification of Glaucoma Using U-Net with Deep Learning Model." *Journal of Healthcare Engineering* 2022 (February): 1601354.

Sundararaman, Sathish, J. Aravind Kumar, Prabu Deivasigamani, and Yuvarajan Devarajan. 2022. "Emerging Pharma Residue Contaminants: Occurrence, Monitoring, Risk and Fate Assessment – A Challenge to Water Resource Management." *Science of The Total Environment*. https://doi.org/10.1016/j.scitotenv.2022.153897.

Suter, Yannick, Alain Jungo, Michael Rebsamen, Urspeter Knecht, Evelyn Herrmann, Roland Wiest, and Mauricio Reyes. 2019. "Deep Learning Versus Classical Regression for Brain Tumor Patient Survival Prediction." *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*. https://doi.org/10.1007/978-3-030-11726-9_38.

Vijayalakshmi, V. J., Prakash Arumugam, A. Ananthi Christy, and R. Brindha. 2022. "Simultaneous Allocation of EV Charging Stations and Renewable Energy Sources: An Elite RERNN-m2MPA Approach." *International Journal of Energy Research*. https://doi.org/10.1002/er.7780.

Wu, Guorong, Dinggang Shen, and Mert Sabuncu. 2016. *Machine Learning and Medical Imaging*. Academic Press.

Yaashikaa, P. R., P. Senthil Kumar, S. Jeevanantham, and R. Saravanan. 2022. "A

Review on Bioremediation Approach for Heavy Metal Detoxification and Accumulation in Plants." *Environmental Pollution* 301 (May): 119035.

Tables and Figures

Table 1. Percentage of accuracy acquired between Logistic Regression and Support vector machine of 20 samples each with a mean accuracy of 87.5 for Logistic Regression and 75 for Support vector machine.

Samples (MRI)	LR	SVM
1	87.19	73.64
2	86.79	74.36
3	86.15	74.55
4	85.85	74.52
5	85.57	73.44
6	85.22	74.53
7	86.91	74.96
8	85.89	74.6
9	86.14	73.14
10	85.45	74.41
11	86.48	74.44
12	87.5	73.72
13	87.39	74.85
14	85.64	73.83
15	85.62	73.34
16	86.39	74.51
17	86.29	73.17
18	86.82	73.57
19	86.8	75
20	86.42	74.29

Table 2. Statistical analysis of Logistic regression and Support vector machine algorithms with mean accuracy, Standard deviation, and Standard Error Mean. It is observed that the Logistic regression algorithm performed better than the Support vector machine algorithm.

Group Statistics					
	Samples	**N**	**Mean**	**Std. Deviation**	**Std. Error Mean**

Accuracy	LR	20	86.3255	.66226	.14809
	SVM	20	74.1435	.60180	.13457

Table 3. Statistical analysis of LR and SVM algorithms with independent sample tests. It is observed with a statistically significant difference of 0.890

Independent Samples Test										
		Levene's Test for Equality of Variances		**t-test for Equality of Means**						
		F	**Sig**	**t**	**df**	**Sig. (2-tailed)**	**Mean Difference**	**Std. Error Difference**	**95% Confidence Interval of the Difference**	
									Lower	**Upper**
Accuracy	**Equal variances assumed**	.01	.89	60.88	38	.000	12.18	.20	11.77	12.58
	Equal variances not assumed			60.88	37.65	.000	12.18	.20	11.77	12.58

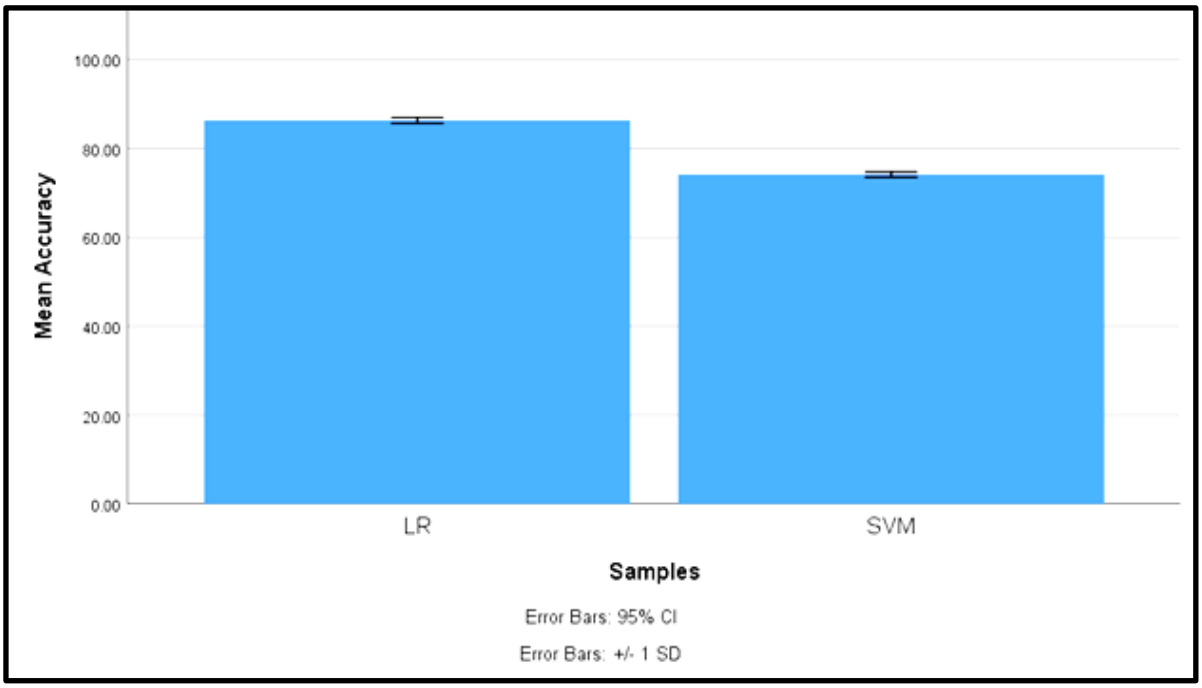

Fig. 1. Bar Chart Comparison between LR and SVM algorithm. The mean accuracy value of the LR is better than the SVM algorithm. X-Axis: LRvs SVM and Y-Axis: Mean Accuracy, SD ± 1.

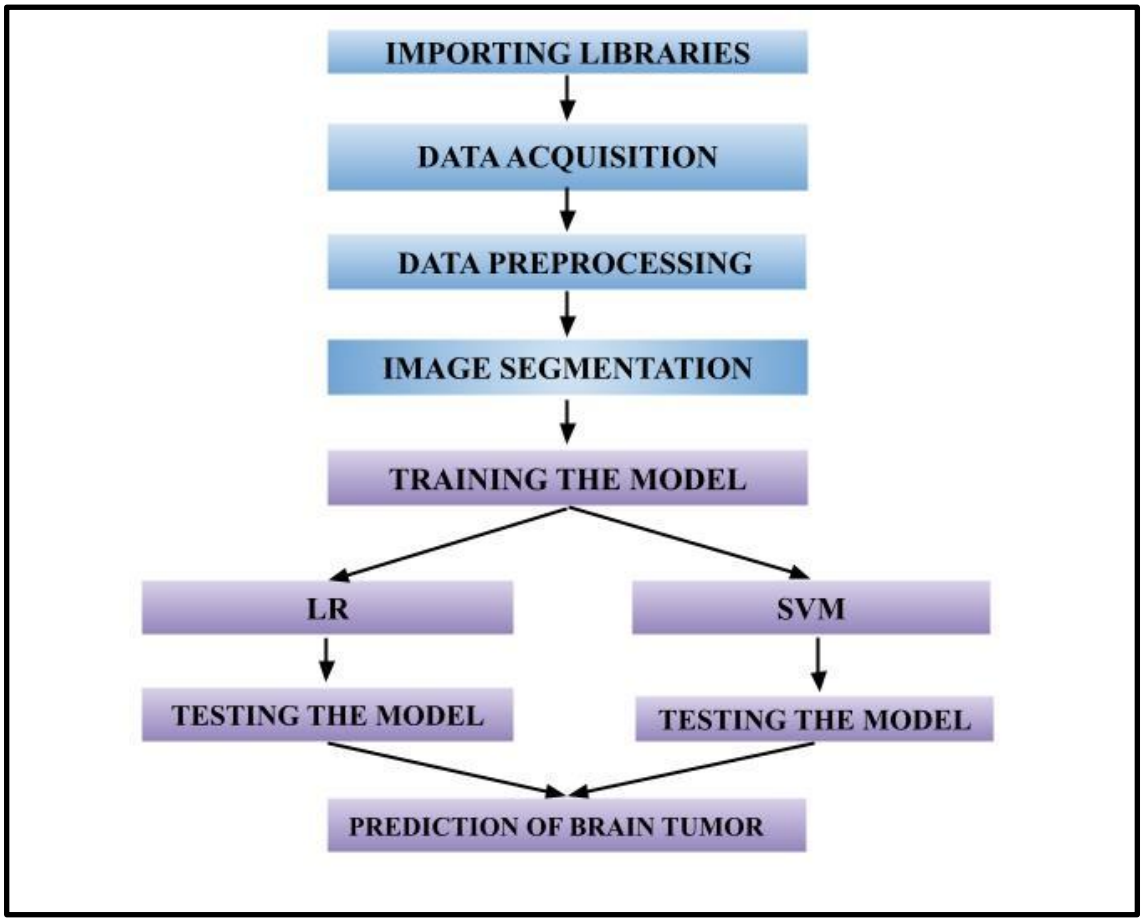

Fig. 2. Proposed workflow for the novel brain tumor detection by using Logistic Regression and Support Vector Machine.