Predicting vulnerability for brain tumor: data-driven approach utilizing machine learning

Yutika Amelia Effendi[1], Amila Sofiah[1], Niko Azhari Hidayat[2], Awol Seid Ebrie[3], Zainy Hamzah[4]
[1]Robotics and Artificial Intelligence Engineering, Faculty of Advanced Technology and Multidiscipline,
Universitas Airlangga, Surabaya, Indonesia
[2]Industrial Engineering, Faculty of Advanced Technology and Multidiscipline, Universitas Airlangga, Surabaya, Indonesia
[3]Department of Industrial and Data Engineering, College of Engineering, Pukyong National University, Busan, South Korea
[4]Faculty of Medicine and Health, University of Muhammadiyah Jakarta, Jakarta, Indonesia

Article Info

Article history:

Received Jan 26, 2024
Revised Apr 16, 2024
Accepted May 6, 2024

Keywords:

Brain tumor
Early detection
Health risk
Healthcare
Machine learning
Technology
Vulnerability

ABSTRACT

Brain tumors, whether benign or malignant, present a complex and multifaceted challenge in healthcare, affecting individuals across various age groups. Predicting the vulnerability of brain tumors using health risk factors and symptoms is crucial, yet there have been limited research studies, particularly those integrating artificial intelligence (AI) technology. This research explores machine learning models such as support vector machines (SVMs), multi-layer perceptrons (MLPs), and logistic regression (LR) for the early detection of brain tumors. Evaluation metrics, including accuracy, precision, recall, and F1-score, are employed to assess model performance. The results indicate that the SVM outperforms other models, providing a robust foundation for predictive accuracy. To enhance accessibility and usability, the research also integrates these models into a mobile application predictor. The application is beneficial for assisting individuals in early detection by identifying potential risk factors and symptoms that may lead to a brain tumor. In conclusion, the integration of machine learning through a mobile application represents a transformative approach to personalized healthcare. By empowering individuals with cutting-edge technology, this research strives to enhance early detection and decision-making regarding potential brain tumor risks and symptoms, ultimately contributing to improved patient outcomes and quality of life.

Corresponding Author:

Yutika Amelia Effendi
Robotics and Artificial Intelligence Engineering, Faculty of Advanced Technology and Multidiscipline
Universitas Airlangga
Surabaya, Indonesia
Email: yutika.effendi@ftmm.unair.ac.id

1. INTRODUCTION

The development of brain tumors is a multifaceted process influenced by genetic, environmental, and often unknown factors [1]. Various elements, such as genetic predisposition, exposure to environmental toxins, and a history of radiation therapy to the head, contribute to the complex etiology of these tumors [2]. Notably, brain tumors can affect individuals across age groups, from the youngest to adults, with certain genetic conditions and syndromes increasing the risk in children [3].

Brain tumors, whether benign or malignant, pose a significant threat to both neurological and overall health [4]. The complexity of these tumors complicates the accurate prediction of patient vulnerability, impeding timely intervention and personalized treatment strategies. This is where machine learning, a subset of artificial intelligence (AI) technique, emerges as a transformative solution, using algorithms to analyze extensive datasets, unveiling patterns and correlations not easily discernible by human observation [5]–[7]. Predicting the vulnerability of brain tumors using health risk factors and symptoms is crucial, yet there have been limited research studies, particularly those integrating AI technology. Some studies have focused on predicting survival in patients with brain tumors using AI methods, such as those applied to magnetic resonance imaging (MRI) [8], and have made significant use of computed tomography (CT) and X-ray imaging to identify malignant brain tumors [9].

Therefore, the integration of machine learning in predicting vulnerability for brain tumor opens new avenues for early detection and targeted interventions [10]. Unlike traditional diagnostic methods reliant on static data points, machine learning, excels in processing dynamic and evolving information, assimilating diverse datasets for a comprehensive understanding [11], [12]. This approach facilitates the identification of subtle patterns indicative of vulnerability, enabling timely and precise therapeutic interventions.

Recognizing the importance of early testing for brain tumors [13], [14], this project aims to develop a predictive brain tumor model through a mobile application seamlessly integrated with machine learning models. The application, designed for assessing and monitoring individual health conditions such as risk factors and symptoms, serves as a reliable tool for gauging vulnerability likelihood through the model-sign feature. By harnessing machine learning, the application not only offers a convenient means of staying attuned to health status but also contributes to early detection and informed decision-making regarding potential brain tumor risks. This synergy between the mobile application and machine learning models represents a significant stride towards personalized and proactive healthcare, empowering individuals with cutting-edge technology at their fingertips [15], [16].

The machine learning model utilized in this study consists of support vector machine (SVM), multi-layer perceptron (MLP), and logistic regression (LR). Both risk factors and symptoms obtained from various sources leading to brain tumor diagnosis are researched using these models, respectively. To evaluate the performance of these three machine learning models, various metrics, including accuracy, precision, recall, and F1-score, are employed. The experimental results demonstrate that the SVM presents the best results in performance matrix assessment compared to the MLP and LR. Finally, this research finding provides a valuable opportunity to assist individuals in early detection by identifying potential risk factors and symptoms leading to brain tumors using a mobile application anytime and anywhere with ease.

This study meticulously follows a conventional five-part structure. It begins with an informative introduction outlining the research’s scope in section 1. It then proceeds to cover the theoretical basis in section 2, the methodology in section 3, and the results and discussion sections in section 4, culminating in a concise yet comprehensive conclusion encapsulating the key findings presented in the final chapter, section 5.

2. THEORETICAL BASIS

In this section, we will delve into a comprehensive literature review that examines the utilization of AI technology within the realm of brain tumor. The discussion will meticulously highlight key findings and insights gleaned from a plethora of existing studies in this rapidly evolving field. By synthesizing and analyzing the collective knowledge from diverse sources, we aim to provide a comprehensive understanding of the current landscape of AI technology’s applications in brain tumor.

2.1. Brain tumor

Brain tumors, whether benign or malignant, are abnormal growths of cells within the brain [17]. These tumors can arise from various cell types and locations in the brain, leading to a wide range of symptoms [18]. Common signs of a brain tumor may include persistent headaches, seizures, changes in vision, difficulty balancing, and cognitive impairments [19]. The exact cause of brain tumors remains unclear, with factors such as genetics, exposure to radiation, and certain genetic conditions potentially contributing to their development. Given the complexity of the brain and the critical functions it performs, the presence of a brain tumor requires careful evaluation and diagnosis to determine the most appropriate course of treatment.

Clinical detection of brain tumors involves a combination of medical history assessment, physical examination, and advanced imaging techniques. Neurological examinations assess the patient’s coordination, reflexes, and sensory functions, providing valuable insights into potential brain abnormalities. Medical professionals may also utilize imaging studies such as MRI and CT scans to visualize the structure of the brain and identify any abnormalities, including the presence of tumors. Additionally, advanced imaging technologies allow for detailed analysis of blood flow and metabolic activity in the brain, aiding in the characterization of tumors and their potential impact on surrounding tissues [20].

While imaging studies provide valuable information, a definitive diagnosis often requires a biopsy [13]. A biopsy involves removing a small tissue sample from the suspected tumor for laboratory analysis. This procedure helps determine the type of cells present, whether the tumor is benign or malignant, and guides treatment planning. In some cases, cerebrospinal fluid analysis or molecular testing may be employed to gather further insights into the nature of the tumor [20]. Collaborative efforts between neurologists, neurosurgeons, and oncologists play a crucial role in the clinical detection and confirmation of brain tumors. Early and accurate diagnosis is pivotal for initiating timely and appropriate treatment strategies, thereby improving the prognosis and quality of life for individuals affected by brain tumors [10], [14].

2.2. Brain tumor research methods

Numerous studies have integrated AI methodologies in various aspects of brain tumor research, including detection, screening, imaging, and disease prediction, employing techniques such as deep learning and machine intelligence [6], [21]–[24]. Investigations have explored the utilization of AI in grading Gliomas. Gutta *et al.* [25] conducted a comparative study evaluating features learned from a convolutional neural network (CNN) against standard radiomic features for predicting glioma grades. Sun *et al.* [26] assessed the efficacy of a LR model based on radiomics in predicting glioma grades. Additionally, Zhuge *et al.* [27] explored the automatic differentiation of low-grade glioma (LGG) and high-grade glioma (HGG) on conventional MRI images, employing CNNs for their study.

Concurrently, machine learning was applied to distinguish between different types of brain tumors. Shrot *et al.* [28] achieved successful differentiation of brain tumors by utilizing both basic and advanced MRI-based radiomics, involving 141 patients. The results demonstrated that the classification incorporated morphologic MRI, perfusion MRI, and diffusion-tensor imaging (DTI) metrics, employing feature subset selection via SVMs, with a binary SVM classification accuracy ranging from 81.6% to 97.0%. Dong *et al.* [29] successfully differentiated between pilocytic astrocytoma (PA) and glioblastoma multiforme (GBM) using MRI quantitative radiomic features through a decision tree model, based on a cohort of 66 patients. The results indicated a training set accuracy of 87%, sensitivity of 90%, and specificity of 83%, while the validation set exhibited an accuracy of 86%, sensitivity of 80%, and specificity of 91%. Additionally, Chakrabarty *et al.* [30] employed a CNN to discern between various tumor types (HGG, LGG, metastases, meningioma, pituitary adenoma, acoustic neuroma, and healthy tissue), achieving commendable accuracy, precision, and recall.

Furthermore, researchers have explored the capacity of AI technology to predict isocitrate dehydrogenase (IDH) mutation, the combined loss of the short arm chromosome 1 (1p19q) codeletion status, and Methylguanine-DNA methyltransferase (MGMT) promoter methylation status among glioma patients [17]. Despite the significant volume of research utilizing AI technology in brain tumor studies, there remains a scarcity of studies specifically addressing vulnerability through the integration of health risk factors and symptoms. This research aims to fill this gap by assisting in the determination of an individual’s vulnerability to brain tumors, leveraging the capabilities of machine learning.

3. METHOD

In this section, we will provide a comprehensive overview, covering aspects such as the study area, dataset employed, data preprocessing techniques applied, and the machine learning models utilized. This detailed exploration will offer valuable insights into the methodologies used for predicting vulnerability for brain tumor. By thoroughly exploring these aspects, we aim to provide a nuanced understanding of the methodologies employed and their implications for the accuracy and reliability of the predictive models.

3.1. Dataset and data preprocessing

The data pipeline for this research comprises several key steps: gathering the data, cleaning it, conducting exploratory data analysis (EDA), designing the model, and performing training and validation. The study was conducted in Surabaya, East Java, Indonesia, utilizing datasets collected from a Surabaya hospital spanning the period from 2012 to 2018 for training and testing the models. The dataset used for the ‘model-sign’ includes risk factors and symptoms of individuals diagnosed with brain tumors, obtained from various clinical journals, brain tumor books, and discussions with several neurosurgeons. A detailed breakdown is provided in Table 1, which showcases 9 risk factors and 11 symptoms relevant to individual vulnerability to brain tumors. For R9 (hormonal birth control), it applies only to female patients.

In our endeavor to create high-quality training datasets, a meticulous approach to data preprocessing is essential. This involves several key steps, including the removal of attributes (columns) that are largely empty and contribute little insight to the model. Additionally, the dataset attributes will be encoded into binary coding to ensure efficient representation for the learning algorithms. Addressing missing information in vital columns is a priority, as is the normalization of datasets to mitigate biases during the learning process. The overarching aim of these data preprocessing steps is to procure a dataset that is not only comprehensive but also free from noise, thereby establishing a robust foundation for effective model training.

Table 1. Dataset information

Risk factors	Symptoms
R1. Genetics	S1. Headache
R2. Job/work	S2. Seizures
R3. Trauma	S3. Vomit
R4. Allergy	S4. Loss of consciousness
R5. CNS infection	S5. Neurocognitive
R6. Smoking	S6. Paresis
R7. Alcohol	S7. Sensory disorders
R8. Radiation	S8. Focal seizures
R9. Hormonal birth control	S9. Aphasia
	S10. Visual impairment
	S11. Lower cranial nerve

Additionally, we divided the data into training and testing sets at an 85% to 15% ratio. Following the cleaning process, the model-sign data consists of 1,477 rows and 20 columns, which were then divided into training and testing datasets. The data preprocessing results are shown in Table 2. Based on Table 2, we have identified ten cases featuring individual health risk factors and symptoms. For instance, individual 1 in row 1 exhibits symptoms S1 (headache), S3 (vomiting), S6 (paresis), S10 (visual impairment), and S11 (lower cranial nerve), and similar patterns are observed across all individual cases in the datasets.

Table 2. Data preprocessing result

Risk factors									Symptoms										
R1	R2	R3	R4	R5	R6	R7	R8	R9	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11
0	0	0	0	0	0	0	0	0	1	0	1	0	0	1	0	0	0	1	1
0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0
0	0	0	0	0	0	0	0	0	1	0	1	1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0
0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	1	0	1	0	0	1	0	0	0	1	1
0	0	0	0	0	0	0	0	1	1	1	0	0	0	1	0	0	0	0	0

3.2. Machine learning model

As mentioned in the previous section, this study focuses on assessing an individual’s health conditions to determine the likelihood of vulnerability to brain tumors, termed as the ‘model sign.’ In the context of machine learning, ‘model sign’ could refer to the output or prediction generated by a machine learning model [31]. For instance, in a binary classification problem, the model’s sign might signify whether the predicted class is positive (1) or negative (0).

To ascertain an individual’s vulnerability for the model sign, we investigated various machine learning classification models. The tested models include the LR model, an MLP model with 5 layers (1 input, 3 hidden, and 1 output) incorporating dropout layers and early stopping, as well as the SVM model. We chose these classification models based on their diversity and strengths in capturing different aspects of data patterns.

LR stands as a fundamental model traditionally leveraged for predicting continuous outcomes [32]. Its inclusion in this context serves the purpose of exploring its simplicity and assessing baseline performance within a classification framework [33]. Moving on to the MLP, this model, characterized by a specific architecture, operates as an artificial neural network adept at learning intricate relationships within data [34]. The incorporation of dropout layers prevents overfitting, while early stopping enhances efficiency by halting training when performance plateaus, making it an apt choice for capturing nuanced patterns [35]. In contrast, the SVM, known for its prowess in handling high-dimensional data and nonlinear relationships, is selected to investigate its potential in delineating a clear boundary between vulnerable and non-vulnerable categories [36]. The deliberate choice of these models reflects a strategic approach to problem evaluation, considering diverse perspectives on factors such as linearity, non-linearity, and the ability to capture complex patterns [37]. Each model contributes unique strengths, and their collective utilization enables a comprehensive understanding of the data, potentially leading to improved prediction accuracy.

Once we acquire three distinct models through the classification process, the subsequent step involves evaluating their performance. This assessment entails the use of various metrics, including accuracy, precision, recall, and F1-score [38], [39]. Accuracy, calculated as the ratio of correctly predicted instances (true positives and true negatives) to the total instances, provides a broad overview of overall correctness [38]. Precision, defined as the ratio of true positives to the sum of true positives and false positives, offers insights into the accuracy of positive predictions [38]. Recall, calculated as the ratio of true positives to the sum of true positives and false negatives, emphasizes the model's ability to correctly identify positive instances [39]. Finally, the F1-score, a combination of recall and precision, serves as a comprehensive metric balancing the trade-off between false positives and false negatives [39].

Figure 1 illustrates the machine learning model approach utilized in this study. The process of building the machine learning model for this research involves selecting machine learning algorithms (LR, MLP, and SVM), defining the hyperparameters, training the models, and evaluating their performance. Subsequently, the models are utilized for predictions, encompassing deploying the model, making predictions, and monitoring its performance.

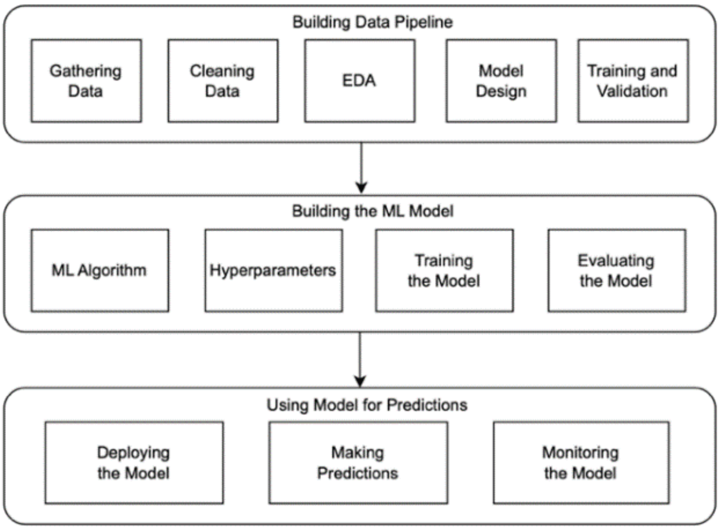


Figure 1. Machine learning model deployment

4. RESULTS AND DISCUSSION

In this section, we provide a thorough examination of the research findings, delving into the nuances of the results obtained. Moreover, the discussion extends beyond mere presentation, offering a robust analysis of the implications and significance of the outcomes in the context of the broader research landscape. By engaging in a comprehensive discourse, this section aims to provide a deeper understanding of the research's contributions and potential avenues for future exploration.

4.1. Results

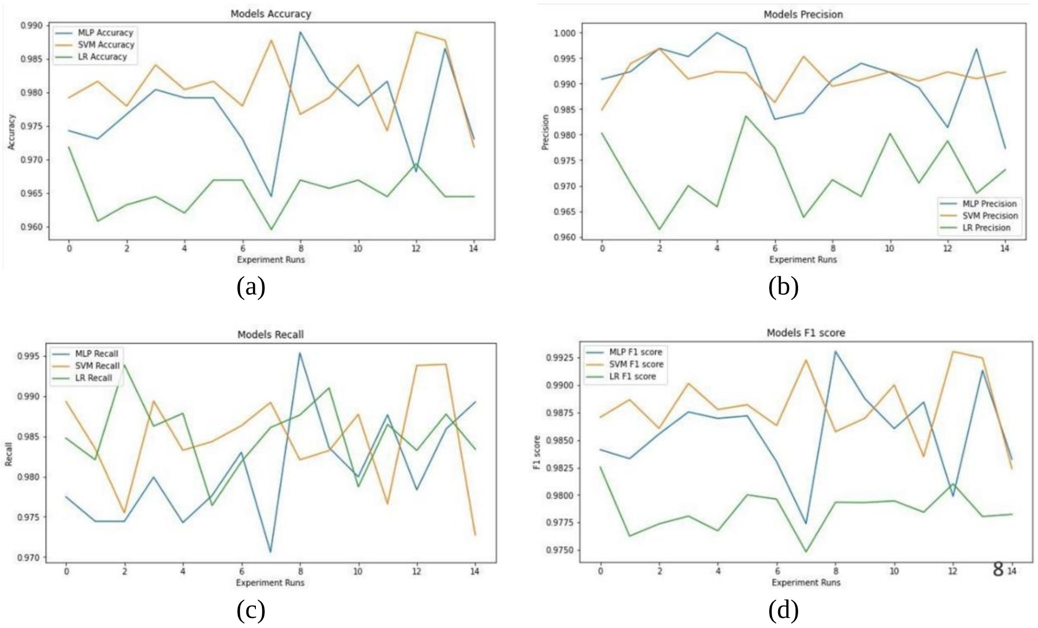
The experimental design scheme is comprehensively outlined in Table 3. The focal point of the training process lies in the model sign, which encompasses 20 input variables derived from the risk factors and symptoms. This setting is then executed by each machine learning algorithm: LR, SVM, and MLP. The model's output serves as individual classification, categorizing individuals as vulnerable or not vulnerable to brain tumors based on the defined parameters. This approach ensures a targeted examination of the factors contributing to vulnerability, thereby enhancing the precision of the predictive model.

Table 3. Experimental design scheme

Model to train	Model sign			
Training and input variables	R1. Genetics	R6. Smoking	S2. Seizures	S7. Sensory disorders
	R2. Job/Work	R7. Alcohol	S3. Vomit	S8. Focal seizures
	R3. Trauma	R8. Radiation	S4. Loss of consciousness	S9. Aphasia
	R4. Allergy	R9. Hormonal birth control	S5. Neurocognitive	S10. Visual impairment
	R5. CNS infection	S1. Headache	S6. Paresis	S11. Lower cranial nerve
Machine learning model	-LR model			
	-MLP model			
	-SVM model			
Output of model	Classification as either vulnerable or not vulnerable to brain tumors			

After obtaining the results of each model, we evaluate the accuracy, precision, recall, and F1-score, which are presented in graphical representation as shown in Figure 2. Analyzing these graphical representations allows for a visual comparison of the trends in accuracy, precision, recall, and F1-score among the SVM, MLP, and LR models, enabling a comprehensive understanding of how each model performs over the course of the experiments. Based on Figure 2(a), the x-axis, denoting the experiment runs, provides a chronological or sequential view of the conducted experiments. On the y-axis, accuracy values indicate how well each model performed in terms of correctly classifying instances. Each line on the graph represents the accuracy achieved by a specific model (SVM, MLP, and LR) in a particular experiment run. Observing trends over multiple runs allows us to assess the consistency and stability of each model's performance. Higher points on the y-axis indicate greater accuracy, and variations across experiment runs can reveal insights into the robustness and reliability of the models. The same principles apply to Figure 2(b), explaining models' precision, Figure 2(c) depicting models' recall, and Figure 2(d) illustrating models' F1-score.

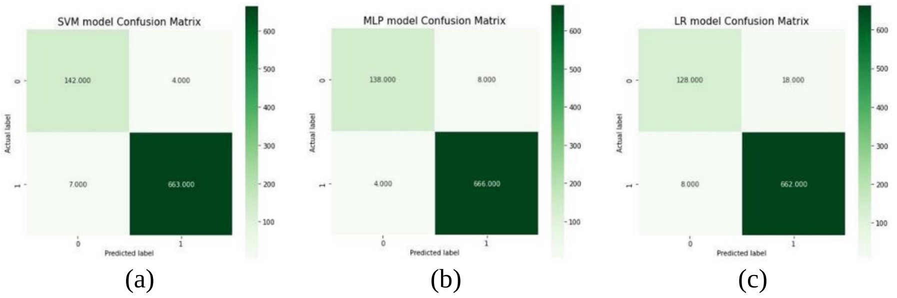
Moreover, in-depth analysis, Figure 3 explains the model's confusion matrix by breaking down the predicted and actual labels to assess the model's classification performance. A confusion matrix is presented in table form, visually representing the performance of a classification algorithm by comparing predicted and actual labels. It typically consists of four components, as explained in section 3.2: true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). Figure 3(a) elucidates the SVM model confusion matrix, Figure 3(b) illustrates the MLP model confusion matrix, and Figure 3(c) presents the LR model confusion matrix. In the end, these matrices enable a thorough evaluation of each model's performance in classifying instances, shedding light on its strengths and potential areas for improvement.



(a) (b)

(c) (d)

Figure 2. The result of; (a) models accuracy, (b) models precision, (c) models recall, and (d) models F1-score



(a) (b) (c)

Figure 3. The results of; (a) SVM model confusion matrix, (b) MLP model confusion matrix, and (c) LR model confusion matrix

Concluding the model evaluation, Table 4 succinctly encapsulates the mean values for various metrics across each machine learning model. The results highlight that the SVM exhibits superior performance metrics, surpassing both the MLP and LR models. With accuracy, precision, recall, and F1-score values of 0.9809, 0.9914, 0.9847, and 0.9881, respectively, the SVM emerges as the most robust model among the evaluated options. These metrics provide a comprehensive overview of the models' capabilities, offering valuable insights into their comparative strengths and effectiveness in classifying instances.

Table 4. Mean values for model metrics

Metrics	MLP	SVM	LR
Accuracy	0.9772	0.9809	0.9652
Precision	0.9908	0.9914	0.9722
Recall	0.9808	0.9847	0.9852
F1	0.9857	0.9881	0.9786

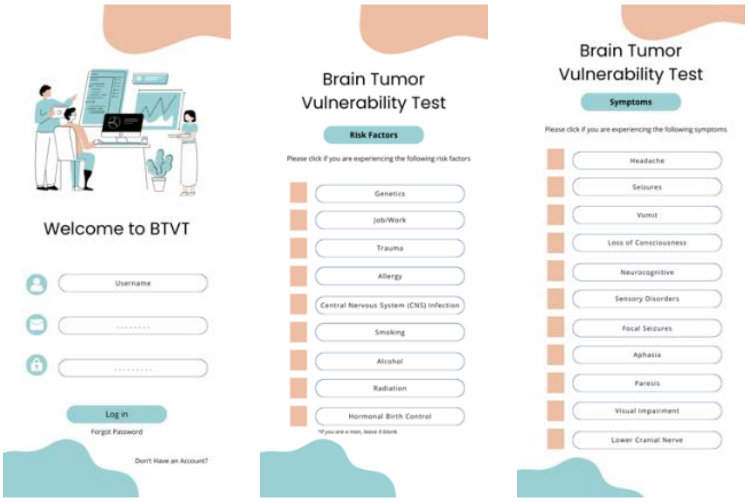
4.2. Discussion

Predicting in brain tumor research is a complex task that may involve non-linear relationships within the data and a significant number of features. Based on our experimental results, SVM is well-suited for creating a clear boundary between vulnerable and non-vulnerable instances. The superiority of the SVM over the MLP and LR models could be attributed to several factors, including handling non-linearity in data, robustness to complex patterns, optimal separation of classes, effective handling of high-dimensional data, robustness to outliers, hyperparameter tuning, and dataset characteristics [40], [41]. However, the choice of the best model is often task-specific and depends on the unique characteristics of the dataset. While SVM performed well in this context which is to predict vulnerability of brain tumor reseach utilizing the health risk factors and symptoms, its effectiveness may vary in different scenarios. Model selection should consider the specific requirements and nature of the predictive task.

The vulnerability in brain tumor research studied and predicted in this research is derived from observations made in various clinical journals [1]–[4], [18]–[20], brain tumor books [42]–[44], and discussions with several neurosurgeons. Various risk factors have been associated with the development of brain tumors. While the exact cause of most brain tumors is often unknown, certain factors may increase the risk, including genetics, job/work, trauma, allergy, central nervous system (CNS) infection, smoking, alcohol, radiation and hormonal birth control. It is important to note that these factors may vary depending on the type and location of the brain tumor. Additionally, many people diagnosed with brain tumors do not have any known risk factors, emphasizing the complexity and multifactorial nature of these conditions. Regular health check-ups and consultations with healthcare professionals are crucial for monitoring and managing potential risks.

In line with risk factors, individuals experiencing symptoms related to brain tumors may present various complaints to their doctors. It is crucial to note that these symptoms can vary depending on the type, size, and location of the tumor. Common complaints and symptoms related to brain tumors include headache, seizures, vomit, loss of consciousness, neurocognitive, paresis, sensory disorders, focal seizures, aphasia, visual impairment and lower cranial nerve. It is important to emphasize that these symptoms can be caused by various medical conditions, and the presence of one or more of these complaints does not necessarily indicate a brain tumor. However, if individuals experience persistent, worsening, or concerning symptoms, it is crucial to seek medical attention promptly for a thorough evaluation and diagnosis. Early detection and intervention are essential for improving outcomes in cases of brain tumors.

In summary, this research aims to assist individuals in early detection by identifying potential risk factors and symptoms that may lead to a brain tumor. Our machine learning model provides individual classification, categorizing individuals as either vulnerable or not vulnerable to brain tumors based on 9 risk factors and 11 symptoms. There is no research specifically addressing the vulnerability of brain tumors through the integration of health risk factors and symptoms with AI technology. Previous related research have focused on forecasting patient survival rates in cases of brain tumors by employing AI techniques, such as those used in MRI [8], and extensively utilized CT and X-ray imaging for the detection of malignant brain tumors [9], as well as brain tumor segmentation based on deep learning [45], [46]. Therefore, this research will be beneficial for individuals to recognize their health status and serves as a foundational study in predicting the vulnerability of brain tumors. For a precise diagnosis, individuals using our mobile application predictor are encouraged to visit and consult with a neurosurgeon and undergo various tests. Concluding this discussion, we also showcase the initial user interface (UI) design of the mobile application predictor in Figure 4. Figure 4(a) depicts the home page of the mobile application, Figure 4(b) displays the page where users input risk factors, and Figure 4(c) exhibits the page showing symptoms before checking the vulnerability results.



Please click if you are experiencing the following risk factors

*If you are a man, leave it blank.

Please click if you are experiencing the following symptoms

(a) (b) (c)

Figure 4. UI design of; (a) home page, (b) risk factors page, and (c) symptoms page

5. CONCLUSION

Maintaining good health is paramount for overall well-being, and vigilance in detecting and addressing factors such as brain tumors plays a crucial role in ensuring a healthy life. This study showcases the integration of machine learning in predicting vulnerability for brain tumor, marking a groundbreaking development in healthcare. By harnessing the computational prowess of AI technology, individuals can now more easily assess whether they are vulnerable or not vulnerable to brain tumors. We demonstrated the application of machine learning models, including LR, MLP, and SVM, on real dataset. According to our experimental results, the SVM outperformed other models in the evaluation metrics for the classification problem. Consequently, this model is employed to provide feedback to an individual based on the given health risk factors and symptoms. With promising results, the trained models are now ready for practical application, leveraging more abundant and recent data. The methodologies implemented in this study have the potential to offer valuable insights for individuals and healthcare centers. Navigating the intersection of technology and medicine, the synergy between machine learning and neuro-oncology holds immense promise for shaping the future of brain tumor care.

ACKNOWLEDGEMENTS

This research was supported by the Faculty of Advanced Technology and Multidiscipline, Universitas Airlangga, Indonesia, under contract number 112/UN3.1.17/PT/2022. The authors also extend their gratitude to the Digital Health, Computational Learning, and Intelligent Systems Research Group for their invaluable assistance during the training and testing phases.

REFERENCES

[1] K. M. Reilly, "Brain tumor susceptibility: the role of genetic factors and uses of mouse models to unravel risk," *Brain Pathology*, vol. 19, no. 1, pp. 121–131, Jan. 2009, doi: 10.1111/j.1750-3639.2008.00236.x.

[2] A. Vienne-Jumeau, C. Tafani, and D. Ricard, "Environmental risk factors of primary brain tumors: a review," *Revue Neurologique*, vol. 175, no. 10, pp. 664–678, Dec. 2019, doi: 10.1016/j.neurol.2019.08.004.

[3] Q. T. Ostrom *et al.*, "Risk factors for childhood and adult primary brain tumors," *Neuro-Oncology*, vol. 21, no. 11, pp. 1357–1375, Nov. 2019, doi: 10.1093/neuonc/noz123.

[4] L. Nayak and F. M. Iwamoto, "Primary brain tumors in the elderly," *Current Neurology and Neuroscience Reports*, vol. 10, no. 4, pp. 252–258, Jul. 2010, doi: 10.1007/s11910-010-0110-x.

[5] D. B. Olawade, O. J. Wada, A. C. David-Olawade, E. Kunonga, O. Abaire, and J. Ling, "Using artificial intelligence to improve public health: a narrative review," *Frontiers in Public Health*, vol. 11, Oct. 2023, doi: 10.3389/fpubh.2023.1196397.

[6] M. Illimoottil and D. Ginat, "Recent advances in deep learning and medical imaging for head and neck cancer treatment: MRI, CT, and PET Scans," *Cancers*, vol. 15, no. 13, p. 3267, Jun. 2023, doi: 10.3390/cancers15133267.

[7] A. B. Abdusalomov, M. Mukhiddinov, and T. K. Whangbo, "Brain tumor detection based on deep learning approaches and magnetic resonance imaging," *Cancers*, vol. 15, no. 16, p. 4172, Aug. 2023, doi: 10.3390/cancers15164172.

[8] C. di Noia *et al.*, "Predicting survival in patients with brain tumors: current state-of-the-art of ai methods applied to MRI," *Diagnostics*, vol. 12, no. 9, p. 2125, Sep. 2022, doi: 10.3390/diagnostics12092125.

[9] M. M. Khan, A. S. Omee, T. Tazin, F. A. Almalki, M. Aljohani, and H. Algethami, "A novel approach to predict brain cancerous tumor using transfer learning," *Computational and Mathematical Methods in Medicine*, vol. 2022, pp. 1–9, Jun. 2022, doi: 10.1155/2022/2702328.

[10] M. Aggarwal, A. K. Tiwari, M. P. Sarathi, and A. Bijalwan, "An early detection and segmentation of brain tumor using deep neural network," *BMC Medical Informatics and Decision Making*, vol. 23, no. 1, p. 78, Apr. 2023, doi: 10.1186/s12911-023-02174-8.

[11] A. Hosny, C. Parmar, J. Quackenbush, L. H. Schwartz, and H. J. W. L. Aerts, "Artificial intelligence in radiology," *Nature Reviews Cancer*, vol. 18, no. 8, pp. 500–510, Aug. 2018, doi: 10.1038/s41568-018-0016-5.

[12] I. H. Sarker, "Deep learning: a comprehensive overview on techniques, taxonomy, applications and research directions," *SN Computer Science*, vol. 2, no. 6, p. 420, Nov. 2021, doi: 10.1007/s42979-021-00815-1.

[13] P. M. Brennan *et al.*, "Early diagnosis of brain tumours using a novel spectroscopic liquid biopsy," *Brain Communications*, vol. 3, no. 2, Apr. 2021, doi: 10.1093/braincomms/fcab056.

[14] D. Walker, W. Hamilton, F. M. Walter, and C. Watts, "Strategies to accelerate diagnosis of primary brain tumors at the primary-secondary care interface in children and adults," *CNS oncology*, vol. 2, no. 5, pp. 447–462, Sep. 2013, doi: 10.2217/cns.13.36.

[15] K. B. Johnson *et al.*, "Precision medicine, AI, and the future of personalized health care," *Clinical and Translational Science*, vol. 14, no. 1, pp. 86–93, Jan. 2021, doi: 10.1111/cts.12884.

[16] S. A. Alowais *et al.*, "Revolutionizing healthcare: the role of artificial intelligence in clinical practice," *BMC Medical Education*, vol. 23, no. 1, p. 689, Sep. 2023, doi: 10.1186/s12909-023-04698-z.

[17] A. S. Peddinti, S. Maloji, and K. Manepalli, "Evolution in diagnosis and detection of brain tumor – review," *Journal of Physics: Conference Series*, vol. 2115, no. 1, p. 012039, Nov. 2021, doi: 10.1088/1742-6596/2115/1/012039.

[18] F. Hanif, K. Muzaffar, K. Perveen, S. M. Malhi, and S. U. Simjee, "Glioblastoma multiforme: a review of its epidemiology and pathogenesis through clinical presentation and treatment," *Asian Pacific Journal of Cancer Prevention*, vol. 18, no. 1, pp. 3–9, 2017, doi: 10.22034/APJCP.2017.18.1.3.

[19] F. Ghandour, A. Squassina, R. Karaky, M. Diab-Assaf, P. Fadda, and C. Pisanu, "Presenting psychiatric and neurological symptoms and signs of brain tumors before diagnosis: A systematic review," *Brain Sciences*, vol. 11, no. 3, pp. 1–20, Feb. 2021, doi: 10.3390/brainsci11030301.

[20] J. H. Rees, "Diagnosis and treatment in neuro-oncology: an oncological perspective," *British Journal of Radiology*, vol. 84, no. SPEC. ISSUE 2, pp. S82–S89, Dec. 2011, doi: 10.1259/bjr/18061999.

[21] U. Raghavendra *et al.*, "Brain tumor detection and screening using artificial intelligence techniques: current trends and future perspectives," *Computers in Biology and Medicine*, vol. 163, p. 107063, Sep. 2023, doi: 10.1016/j.compbiomed.2023.107063.

[22] M. Afridi, A. Jain, M. Aboian, and S. Payabvash, "Brain tumor imaging: applications of artificial intelligence," *Seminars in Ultrasound, CT and MRI*, vol. 43, no. 2, pp. 153–169, Apr. 2022, doi: 10.1053/j.sult.2022.02.005.

[23] D. C. Yadav and S. Pal, "Measure the superior functionality of machine intelligence in brain tumor disease prediction," in *Artificial Intelligence-Based Brain-Computer Interface*, Elsevier, 2022, pp. 353–368. doi: 10.1016/B978-0-323-91197-9.00005-9.

[24] J. Huang, N. A. Shlobin, S. K. Lam, and M. DeCuypere, "Artificial intelligence applications in pediatric brain tumor imaging: a systematic review," *World Neurosurgery*, vol. 157, pp. 99–105, Jan. 2022, doi: 10.1016/j.wneu.2021.10.068.

[25] S. Gutta, J. Acharya, M. S. Shiroishi, D. Hwang, and K. S. Nayak, "Improved glioma grading using deep convolutional neural networks," *American Journal of Neuroradiology*, vol. 42, no. 2, pp. 233–239, Feb. 2021, doi: 10.3174/ajnr.A6882.

[26] X. Sun *et al.*, "A logistic regression model for prediction of glioma grading based on radiomics," *Journal of Central South University (Medical Sciences)*, vol. 46, no. 4, pp. 385–392, 2021, doi: 10.11817/j.issn.1672-7347.2021.200074.

[27] Y. Zhuge *et al.*, "Automated glioma grading on conventional MRI images using deep convolutional neural networks," *Medical Physics*, vol. 47, no. 7, pp. 3044–3053, Jul. 2020, doi: 10.1002/mp.14168.

[28] S. Shrot, M. Salhov, N. Dvorski, E. Konen, A. Averbuch, and C. Hoffmann, "Application of MR morphologic, diffusion tensor, and perfusion imaging in the classification of brain tumors using machine learning scheme," *Neuroradiology*, vol. 61, no. 7, pp. 757–765, Jul. 2019, doi: 10.1007/s00234-019-02195-z.

[29] F. Dong *et al.*, "Differentiation between pilocytic astrocytoma and glioblastoma: a decision tree model using contrast-enhanced magnetic resonance imaging-derived quantitative radiomic features," *European Radiology*, vol. 29, no. 8, pp. 3968–3975, Aug. 2019, doi: 10.1007/s00330-018-5706-6.

[30] S. Chakrabarty, A. Sotiras, M. Milchenko, P. Lamontagne, M. Hileman, and D. Marcus, "MRI-based identification and classification of major intracranial tumor types by using a 3D convolutional neural network: a retrospective multi-institutional analysis," *Radiology: Artificial Intelligence*, vol. 3, no. 5, p. e200301, Sep. 2021, doi: 10.1148/ryai.2021200301.

[31] N. Farsad, N. Shlezinger, A. J. Goldsmith, and Y. C. Eldar, "Data-driven symbol detection via model-based machine learning," in *IEEE Workshop on Statistical Signal Processing Proceedings*, IEEE, Jul. 2021, pp. 571–575. doi: 10.1109/SSP49050.2021.9513859.

[32] V. V. P. Wibowo, Z. Rustam, A. R. Laeli, and A. A. Said, "Logistic regression and logistic regression-genetic algorithm for classification of liver cancer data," in *2021 International Conference on Decision Aid Sciences and Application, DASA 2021*, IEEE, Dec. 2021, pp. 244–248. doi: 10.1109/DASA53625.2021.9682242.

[33] M. Martinez-Garcia, I. Inza, and J. A. Lozano, "Learning a logistic regression with the help of unknown features at prediction stage," in *Proceedings - 2023 IEEE Conference on Artificial Intelligence, CAI 2023*, IEEE, Jun. 2023, pp. 298–299. doi: 10.1109/CAI54212.2023.00133.

[34] D. C. Shubhangi and A. K. Pratibha, "Asthma, Alzheimer's and dementia disease detection based on voice recognition using multi-layer perceptron algorithm," in *Proceedings of the 2021 IEEE International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems, ICSES 2021*, IEEE, Sep. 2021, pp. 1–7. doi: 10.1109/ICSES52305.2021.9633923.

[35] A. O. Ibrahim, S. M. Shamsuddin, A. Y. Saleh, A. Abdelmaboud, and A. Ali, "Intelligent multi-objective classifier for breast cancer diagnosis based on multilayer perceptron neural network and differential evolution," in *Proceedings - 2015 International Conference on Computing, Control, Networking, Electronics and Embedded Systems Engineering, ICCNEEE 2015*, IEEE, Sep. 2016, pp. 422–427. doi: 10.1109/ICCNEEE.2015.7381405.

[36] N. Mohan and V. Jain, "Performance analysis of support vector machine in diabetes prediction," in *Proceedings of the 4th International Conference on Electronics, Communication and Aerospace Technology, ICECA 2020*, IEEE, Nov. 2020, pp. 1–3, doi: 10.1109/ICECA49313.2020.9297411.

[37] F. Maulidina, Z. Rustam, and J. Pandelaki, “Lung cancer classification using support vector machine and hybrid particle swarm optimization-genetic algorithm,” in *2021 International Conference on Decision Aid Sciences and Application, DASA 2021*, IEEE, Dec. 2021, pp. 751–755. doi: 10.1109/DASA53625.2021.9682259.

[38] L. E. Pomme, R. Bourqui, R. Giot, and D. Auber, “Relative confusion matrix: efficient comparison of decision models,” in *Proceedings of the International Conference on Information Visualisation*, IEEE, Jul. 2022, pp. 98–103. doi: 10.1109/IV56949.2022.00025.

[39] J. L. García-Balboa, M. V. Alba-Fernández, F. J. Ariza-López, and J. Rodríguez-Avi, “Homogeneity test for confusion matrices: A method and an example,” in *International Geoscience and Remote Sensing Symposium (IGARSS)*, IEEE, Jul. 2018, pp. 1203–1205. doi: 10.1109/IGARSS.2018.8517924.

[40] B. J. B. Nair, S. Yadhukrishnan, and A. Manish, “A comparative study on document images classification using logistic regression and multiple linear regressions,” in *Proceedings of the 2023 2nd International Conference on Augmented Intelligence and Sustainable Systems, ICAISS 2023*, IEEE, Aug. 2023, pp. 1096–1104. doi: 10.1109/ICAISS58487.2023.10250671.

[41] C. Saini, K. D. Mahato, C. Azad, and U. Kumar, “Breast cancer prediction using different machine learning algorithms: a comparative study,” in *2023 International Conference on Artificial Intelligence and Applications, ICAIA 2023 and Alliance Technology Conference, ATCON-1 2023 - Proceeding*, IEEE, Apr. 2023, pp. 1–6. doi: 10.1109/ICAIA57370.2023.10169729.

[42] G. Seano, *Brain tumors*, vol. 158. in Neuromethods, vol. 158. New York, NY: Springer US, 2021. doi: 10.1007/978-1-0716-0856-2.

[43] L. E. Abrey and W. P. Mason, *Fast facts: brain tumors*. S. Karger AG, 2011. doi: 10.1159/isbn.978-1-905832-89-7.

[44] N. A. Mohile and A. A. Thomas, *Brain tumors: a pocket guide*. Cham: Springer International Publishing, 2023. doi: 10.1007/978-3-031-41413-8.

[45] J. Shedbalkar and K. Prabhushetty, “Deep transfer learning model for brain tumor segmentation and classification using UNet and chopped VGGNet,” *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*, vol. 33, no. 3, pp. 1405–1415, Mar. 2024, doi: 10.11591/ijeecs.v33.i3.pp1405-1415.

[46] B. Kuntiyellannagari, B. Dwarakanath, and P. V. P. Reddy, “Hybrid model for brain tumor detection using convolution neural networks,” *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*, vol. 33, no. 3, pp. 1775–1781, Mar. 2024, doi: 10.11591/ijeecs.v33.i3.pp1775-1781.

BIOGRAPHIES OF AUTHORS

Yutika Amelia Effendi obtained her B.Comp.Sc. (Bachelor of Computer Science) and M.Comp.Sc. (Master of Computer Science) degrees in Informatics from Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia. Presently, she holds a position as a lecturer at the robotics and artificial intelligence engineering study program within the Department of Engineering, Faculty of Advanced Technology and Multidiscipline, Universitas Airlangga. Her research interests encompass process mining, artificial intelligence, industrial data analytics, knowledge engineering, and business process management. She can be contacted at email: yutika.effendi@ftmm.unair.ac.id.

Amila Sofiah graduated from master’s program in Electrical Engineering at Bandung Institute of Technology, Indonesia. She is currently a staff member in the robotics and artificial intelligence engineering within the Faculty of Advanced Technology and Multidiscipline, Airlangga University. Her research interests include biomedical signal processing, instrumentation and control, and robotics. She can be contacted at email: amila.sofiah@ftmm.unair.ac.id.

Niko Azhari Hidayat earned his doctoral degree from Airlangga University and is a medical doctor specializing in Thoracic and Cardiovascular Surgery. Currently, he is a staff member in the Industrial Engineering Study Program within the Faculty of Advanced Technology and Multidiscipline, Airlangga University. His professional interests include technology, focusing on Digital Health Technology, Vascular Telehealth, HRD Health Edutechnology, and Medicosociopreneurship. He can be contacted at email: niko-a-h@fk.unair.ac.id.

Awol Seid Ebrie is currently studying for a Ph.D. degree in Industrial Data Science and Engineering (IDSE) in the Department of Industrial and Data Engineering at Pukyong National University, Busan, South Korea. He earned his B.Sc. degree in Statistics in 2008 from the University of Gondar and completed his M.Sc. degree in Biostatistics in 2012 at Jimma University, Ethiopia. He can be contacted at email: awolseid@pukyong.ac.kr.

Zainy Hamzah is a neurosurgeon at Jakarta Cempaka Putih Islamic Hospital, Omni Hospital Pekayon, and Mayapada Hospital Jakarta Selatan. In addition to his medical practice, he serves as a teaching staff in neurosurgery and heads the Department of Surgery at the Faculty of Medicine and Health, University of Muhammadiyah Jakarta, Indonesia. Zainy is also an active member of the Indonesian Neurosurgical Society (INS), serving as one of the core administrators. He can be contacted at email: zainy.hamzah@umj.ac.id.