

Article

Predicting Tumor Dynamics Post-Staged GKRS: Machine Learning Models in Brain Metastases Prognosis

Ana-Maria Trofin ¹, Călin Gh. Buzea ^{2,3}, Răzvan Buga ^{1,2,*}, Maricel Agop ⁴, Lăcrămioara Ochiuz ¹ ,
Dragos Teodor Iancu ^{1,5} and Lucian Eva ^{2,6}

- ¹ University of Medicine and Pharmacy “Grigore T. Popa” Iași, 700115 Iași, Romania; ana-maria.trofin@umfiasi.ro (A.-M.T.); lacramioara.ochiuz@umfiasi.ro (L.O.); dt_iancu@yahoo.com (D.T.I.)
² Clinical Emergency Hospital “Prof. Dr. Nicolae Oblu” Iași, 700309 Iași, Romania; calinb2003@yahoo.com (C.G.B.); elucian73@yahoo.com (L.E.)
³ National Institute of Research and Development for Technical Physics, IFT Iași, 700050 Iași, Romania
⁴ Physics Department, Technical University “Gheorghe Asachi” Iași, 700050 Iași, Romania; m.agop@yahoo.com
⁵ Regional Institute of Oncology, 700483 Iași, Romania
⁶ University Apollonia, 700511 Iași, Romania
* Correspondence: bugarazvan@yahoo.com

Abstract: This study assesses the predictive performance of six machine learning models and a 1D Convolutional Neural Network (CNN) in forecasting tumor dynamics within three months following Gamma Knife radiosurgery (GKRS) in 77 brain metastasis (BM) patients. The analysis meticulously evaluates each model before and after hyperparameter tuning, utilizing accuracy, AUC, and other metrics derived from confusion matrices. The CNN model showcased notable performance with an accuracy of 98% and an AUC of 0.97, effectively complementing the broader model analysis. Initial findings highlighted that XGBoost significantly outperformed other models with an accuracy of 0.95 and an AUC of 0.95 before tuning. Post-tuning, the Support Vector Machine (SVM) demonstrated the most substantial improvement, achieving an accuracy of 0.98 and an AUC of 0.98. Conversely, XGBoost showed a decline in performance after tuning, indicating potential overfitting. The study also explores feature importance across models, noting that features like “control at one year”, “age of the patient”, and “beam-on time for volume V1 treated” were consistently influential across various models, albeit their impacts were interpreted differently depending on the model’s underlying mechanics. This comprehensive evaluation not only underscores the importance of model selection and hyperparameter tuning but also highlights the practical implications in medical diagnostic scenarios, where the accuracy of positive predictions can be crucial. Our research explores the effects of staged Gamma Knife radiosurgery (GKRS) on larger tumors, revealing no significant outcome differences across protocols. It uniquely considers the impact of beam-on time and fraction intervals on treatment efficacy. However, the investigation is limited by a small patient cohort and data from a single institution, suggesting the need for future multicenter research.

Keywords: gamma knife radiosurgery (GKRS); brain metastasis; tumor dynamics forecasting; machine learning models; feature importance



Citation: Trofin, A.-M.; Buzea, C.G.; Buga, R.; Agop, M.; Ochiuz, L.; Iancu, D.T.; Eva, L. Predicting Tumor Dynamics Post-Staged GKRS: Machine Learning Models in Brain Metastases Prognosis. *Diagnostics* **2024**, *14*, 1268. <https://doi.org/10.3390/diagnostics14121268>

Academic Editor: Dechang Chen

Received: 3 May 2024

Revised: 3 June 2024

Accepted: 12 June 2024

Published: 15 June 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Brain metastases (BMs), often referred to as secondary brain cancers, pose significant treatment dilemmas and urgently need strategies to lessen the impact on those diagnosed [1]. There’s an increasing incidence of BMs, likely influenced by improvements in conventional therapies like surgery, radiation, and chemotherapy, which have led to longer survival rates for patients. Without any medical interventions, people with BMs usually face a median survival time of approximately 2 months after their diagnosis, especially when the disease affects the central nervous system [2]. The advent of targeted treatment methods, such as Gamma Knife radiosurgery (GKRS), is becoming a preferred approach

for BM management due to its ability to better target tumors locally with fewer side effects compared to traditional whole brain radiation therapy, and its effectiveness is on par with that of surgical removal [3–6]. However, GKRS is limited in treating larger BMs (exceeding 3 cm in diameter or 10 cc in volume) because of the potential for radiation-induced harm, like radiation toxicity [7–10]. The Gamma Knife ICON, using mask fixation for patient positioning, represents an advanced solution for hypo-fractionated treatment of large BMs. Despite these advancements, the median survival rate for patients with BMs hovers around one year [11,12].

Lung cancer stands as a major global cancer threat, categorized mainly into non-small cell lung cancer (NSCLC) and small cell lung cancer. NSCLC is further divided into adenocarcinoma, squamous cell carcinoma, and large cell carcinoma, as classified by the World Health Organization [13,14]. Between 16% and 60% of non-small cell lung cancer (NSCLC) patients develop brain metastases (BM), irrespective of whether their disease is driven by oncogenes or lacks actionable mutations [15,16]. Addressing BMs in NSCLC requires an integrated approach to achieve timely control of the disease. Unfortunately, the effectiveness of standard treatments such as surgery, radiotherapy, and corticosteroid palliation is limited, and these treatments are often associated with severe neurotoxic effects that can delay or impair systemic therapy, leading to high mortality rates [17–20]. The introduction of advanced neuro-imaging techniques and more effective systemic therapies has extended the lives of NSCLC patients, thereby increasing the overall incidence of BMs during their lifetime [21]. Treatment options like stereotactic radiosurgery (SRS), either as a standalone treatment or following surgery and whole-brain radiotherapy (WBRT) for specific cases is considered [22], although they carry the risk of significant neurocognitive side effects. The use of GKRS for NSCLC patients with BMs is debated due to their limited OS. Nevertheless, some policies, like those implemented in Korea, permit the use of GKRS for BM treatment every three months.

Beyond lung cancer, other cancers such as breast, colorectal, prostate, ovarian, renal cancers, and melanoma also metastasize to the brain, significantly impacting patient outcomes and treatment approaches. Breast cancer, for example, has a notable propensity to spread to the brain, particularly in patients with HER2-positive or triple-negative subtypes, challenging clinicians to tailor treatments for both systemic disease and brain involvement [23]. Colorectal cancer, though less commonly associated with BM, poses a significant risk when metastasis occurs, necessitating a multidisciplinary approach to manage both the primary disease and brain metastases [24]. Prostate cancer rarely metastasizes to the brain, but when it does, it signifies a late-stage disease and poor prognosis, highlighting the need for innovative therapeutic strategies [25,26]. Ovarian and renal cancers, similarly, can lead to brain metastases, with renal cell carcinoma being more prone to spread to the brain, requiring careful consideration of treatment options to address this aggressive disease behavior [27,28]. Brain metastases (BM) occur in almost 50% of patients with cutaneous melanoma (CM) and are the third most common metastatic site [29].

Machine learning (ML), as a branch of artificial intelligence, is instrumental in crafting models that learn from data autonomously, without direct programming. Among these, tree-based ML algorithms are particularly appreciated for their straightforwardness and clarity, making the decision paths visual and interpretable. Decision trees stand out by mapping decisions through nodes and labels, ensuring not only high accuracy in classification but also clear presentation of information, which is crucial in healthcare decision-making [30,31]. However, challenges such as overfitting can arise from small data sets. Techniques like random forests and boosted decision trees are therefore employed to overcome these hurdles, offering improved prediction accuracy through a detailed examination of the interconnections among variables in the dataset [32].

In cases where the data consist of clinical features, deep learning (DL) can be highly beneficial, particularly in uncovering intricate relationships and interactions among various clinical parameters that may not be apparent through traditional modeling techniques. Additionally, deep learning models, such as Convolutional Neural Networks, could potentially

integrate diverse types of clinical data, including sequential patient records, to forecast tumor dynamics more accurately, offering a more holistic approach to understanding and predicting patient outcomes post-GKRS [33].

This is why, in our investigation of patients experiencing brain metastases from various primary cancers, we used, in addition to the usual ML algorithms, such as Logistic Regression, Support Vector Machines (SVM), and K nearest neighbors (KNN), tree-based modeling techniques—including Decision Tree, Random Forest, and Boosted decision tree classifiers XGBoost—to anticipate the dynamics of tumors within 3 months following GKRS treatment. Furthermore, we also integrated advanced deep learning models, such as Convolutional Neural Networks (CNNs), to see if we can enhance our predictive accuracy and gain deeper insights into the complex patterns of tumor progression. This approach allowed us to pinpoint essential factors and feature permutations that are crucial in estimating the prognosis for these patients.

2. Research Methodology

2.1. Overview of Research Design and Participant Details

We outline the framework of our study, detailing the procedures followed and the characteristics of the individuals who participated in our research.

The study protocol retrospectively reviewed the medical records of patients treated with GKRS with mask-based fixation of BMs from various primary cancers, between July 2022 and March 2024 at “Prof. Dr. Nicolae Oblu” Emergency Clinic Hospital—Iasi. All experiments were carried out in accordance with relevant guidelines and regulations. The study used only pre-existing medical data; therefore, patient consent was not required, and since it was retrospective, there was no need for approval from the Ethics Committee of Clinical Emergency Hospital “Prof. Dr. Nicolae Oblu” Iasi.

A total of 77 patients (45 males and 32 females; age range 39 to 85 years old; median age, 64 years old) who were previously diagnosed with BMs were enrolled in this study. General characteristics including age, sex, C1yr—tumor volume at one year control, MTS extra cranial—existence of extra cranial metastases, receiving pretreatment, deceased before 1 year, Karnofsky performance scale (KPS) score [34], number of lesions, beam-on time over the number of isocenters for each of the 3 volumes treated, total tumor volume, and tumor dynamics (progression or regression within 3 months following GKRS treatment), are summarized in Table 1. The study design is shown in Figure 1.

Table 1. Patient demographics for brain metastases from various primary cancers.

Characteristics	Value
Number of patients	77
Age (yr) Median (range)	64 (39–85)
Sex	
Male (%)	45 (58.44%)
Female (%)	32 (41.56%)
C1yr—control over one year (cm ³) Median(range)	17 missing data 0.9 (0–30)
Patients with extra cranial MTS	6 missing data 54
Receiving pre-treatment, systemic treatment	68
Deceased before 1 year	1 missing data 25

Table 1. Cont.

Characteristics	Value
KPS score	9 missing data (11.69%)
100	26 (33.77%)
90	8 (10.39%)
80	22 (28.57%)
70	12 (15.58%)
The number of lesions	
Median (range)	2 (1–30)
1–3	52
4–6	12
7–10	8
>10	5
Beam-on time on V1 (min/cm ³)	
Median (range)	0.82 (0.47–2.33)
Beam-on time on V2 (min/cm ³)	1 missing data
Median (range)	0.83 (0.60–3.00)
Beam-on time on V3 (min/cm ³)	2 missing data
Median (range)	0.83 (0.46–4.00)
Total tumor volume (# of patients with):	
<5 cm ³	34
≤10 cm ³	13
>10 cm ³	30
Tumor dynamics (# of patients with):	
- Progression	6
- Regression	71

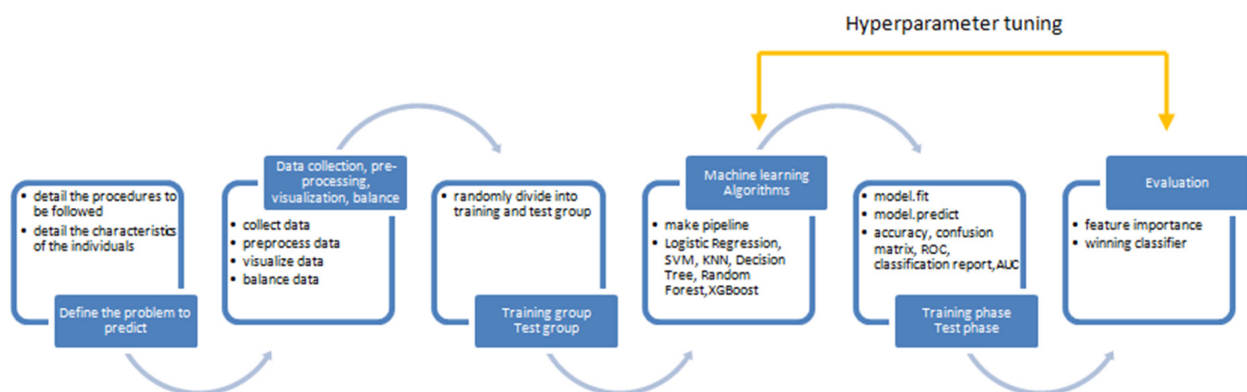


Figure 1. Schematic of the machine learning deep learning study design.

2.2. Strategy for Gamma Knife Radiosurgery Implementation

We discuss, in what follows, the systematic approach adopted for administering Gamma Knife radiosurgery (GKRS), focusing on the meticulous planning and execution phases essential for the treatment. All patients underwent GKRS using the Leksell Gamma Knife ICON (Elekta AB, Stockholm, Sweden).

All MRI examinations were performed on a 1.5 Tesla whole-body scanner (GE 174 SIGMA EXPLORER) that was equipped with the standard 16-channel head coil. The MRI 175 study protocol consisted of the following:

1. The conventional anatomical MRI (cMRI) protocol for clinical routine diagnosis of brain tumors, including, among others, an axial fluid-attenuated inversion recovery (FLAIR) sequence as well as a high-resolution contrast-enhanced T1-weighted (CE T1w) sequence.

2. The advanced MRI (advMRI) protocol for clinical routine diagnosis of brain tumors was extended by axial diffusion-weighted imaging (DWI; b values 0 and 1000 s/mm²) sequence and a gradient echo dynamic susceptibility contrast (GE-DSC) perfusion MRI sequence, which was performed using 60 dynamic measurements during administration of 0.1 mmol/kg-bodyweight gadoterate meglumine.

All magnetic resonance images were registered with Leksell Gamma Plan (LGP, Version 11.3.2, TMR algorithm), and any images with motion artifacts were excluded. The tumor volumes were calculated by LGP without margin. Generally, the prescription of a total dose of 30 Gy, delivered in 3 stages of GRKS, was selected based on the linear quadratic model [35,36] and the work of Higuchi et al. from 2009 [37]. The GKRS planning was determined through a consensus between the neurosurgeon, radiation oncologist, and expert medical physicist.

2.3. Labeling of Medical Data

In the process of labeling medical data, we extracted relevant features from the broad details available in the electronic medical record (EMR) system. One of the prevalent hurdles in machine learning (ML) applications is the issue of incomplete data within the EMR. To tackle this, we applied two principal strategies: imputation and exclusion of data.

Imputation is the method of filling in missing information with plausible values, which proves particularly beneficial when the missing data is scarce. In our research, we opted to replace missing entries in categorical data with the value found in the previous row, whereas for numerical attributes, we used the median value of the dataset. Additionally, for the transformation of categorical variables, label encoding techniques were used [38].

Unbalanced datasets pose challenges in ML when one class vastly outweighs others, which is the case for our study, where only 6 patients in 77 (7.8%) showed signs of progression of lesions. Approaches to address this issue include resampling (oversampling/undersampling), class weighting, cost-sensitive learning, ensemble methods, and data augmentation. The appropriate method depends on the problem and dataset, necessitating evaluation of all classes for accurate results [39].

2.4. Data Manipulation Techniques

Our preference for the Python programming language in this study was driven by its simplicity in handling data operations, coupled with its access to a broad spectrum of freely available libraries [40]. The project leveraged the capabilities of open-source Python 3.0 libraries, including NumPy for numerical data manipulation, pandas for data structures and analysis, Matplotlib and seaborn for data visualization, and TensorFlow, Keras, and scikit-learn for machine learning and data mining tasks.

Logistic Regression is effective in handling binary outcomes, such as the presence or absence of tumor growth. It can utilize various patient and tumor characteristics to predict the probability of specific outcomes, aiding in the risk stratification and management of patients. Support Vector Machine excels in classifying complex patient data into distinct categories, such as predicting the likelihood of tumor recurrence. Its ability to handle high-dimensional data makes it invaluable in analyzing the myriad factors influencing brain metastases, from genetic markers to treatment responses. K Nearest Neighbors offers a straightforward yet powerful method for prognosis by comparing a patient's data against those of similar patients. This similarity-based approach is especially beneficial in medical cases where individual patient characteristics significantly influence the disease trajectory, allowing for personalized prediction of tumor dynamics. Together, Decision Trees and Random Forest algorithms harness the intricate data landscape of BMs to deliver nuanced and individualized prognostic insights. They enable a structured analysis of the factors driving tumor behavior, facilitating targeted and evidence-based treatment strategies that can lead to better patient management and outcomes [41,42].

XGBoost is a machine learning algorithm that belongs to the ensemble learning category, specifically the gradient boosting framework. It uses decision trees as base learners

and employs regularization techniques to enhance model generalization. Known for its computational efficiency, feature importance analysis, and handling of missing values, XGBoost is widely used for tasks such as regression, classification, and ranking; therefore, it was a key component of our methodology [43]. To complement the traditional machine learning models, we also incorporated deep learning (DL) techniques using Convolutional Neural Networks (CNNs). These networks can model complex relationships and interactions between the various clinical, demographic, and biological features of our dataset. By employing layers of interconnected neurons, CNNs are capable of capturing intricate patterns and nonlinearities in the data, which enhances our predictive accuracy and provides deeper insights into the factors influencing tumor dynamics. This integration of deep learning into our methodology significantly strengthens our analytical capabilities, enabling more effective risk stratification and personalized treatment strategies. The dataset was randomly divided into training group ($n = 54$) and test group ($n = 23$). We typically adhere to the conventional 80:20 training-to-test set ratio. However, due to the limited volume of data at our disposal, we opted to modify this ratio (70:30) to ensure a more robust evaluation of the test set performance. This adjustment allows for a more comprehensive assessment under constrained data conditions. The input variables were general characteristics. Hyperparameter tuning was conducted using the GridSearchCV function from scikit-learn, targeting the optimization of model parameters. The dependent variable under investigation was tumor dynamics, i.e., progression or regression within 3 months following GKRS treatment, categorized into “Regression” or “Progression”.

To assess the performance of the ML algorithms and DL models, we computed metrics such as accuracy, sensitivity, and specificity, along with the receiver operating characteristic (ROC) curve and the area under the curve (AUC).

Subsequent to model training, we identify and validate important features and permutation features, with permutation feature importance serving as a versatile technique for evaluating the significance of features across various fitted models, provided the data are structured in tabular form [44].

3. Results

3.1. Overview and Analysis of Data

The cohort study included 77 patients with BM, focusing on their age at the time of GKRS treatment. The KPS scores were evaluated by physicians. Pathologic diagnoses, statuses of receiving chemotherapy, and pretreatment records were obtained from the EMR. The number of lesions, tumor volume, the number of fractions, and prescription doses were documented in the LGP. Detailed data descriptions are summarized in Table 2.

Let us interpret the information in the bar charts from Figure 2a. Each chart shows the distribution of patients with tumor regression (labeled ‘0’) versus tumor progression (labeled ‘1’) across different variables within the dataset.

These charts collectively offer valuable insights into factors associated with tumor progression and survival rates in patients. For instance, gender differences and the presence of external metastases are prominently associated with the progression of the tumor. Pre-treatment status is less clear-cut and could be influenced by many factors that the chart does not specify. Finally, the strong correlation between tumor progression and one-year mortality underscores the seriousness of tumor progression as an indicator of patient prognosis. It is important to note that these are correlative relationships, and causation should not be inferred without further, controlled study.

Table 2. Summary of features categorized for machine learning algorithms.

Feature	Feature Number	Description	Categorical/Numeric Data for Machine Learning Algorithm	Type
Age	0	Age at time of treated GKRS	No categorical	Discrete
Sex	1	Biological sex	0 = Female; 1 = Male Label encoding	Numeric
C1yr—control over one year	2	Volume of lesion measured at the 1 year control	No categorical	Numeric
Patients with extra cranial MTS	3	Patients having detected with extracranial metastases	0 = No; 1 = Yes Label encoding	Numeric
Receiving pre-treatment	4	Before GKRS, treated by surgery or radiotherapy, or performed chemotherapy	0 = No pretreatment; 1 = Pretreatment Label encoding	Numeric
Deceased before 1 year	5	Patients who passed away before 1 year after receiving GKRS	0 = Alive; 1 = Deceased Label encoding	Numeric
KPS score	6	KPS score runs from 0 to 100. Three physicians allow to evaluate the patient ability to receive GKS for BM.	No categorical	Numeric
The number of lesions	7	This divided the 4 groups based on the number of lesions.	No categorical	Discrete
Beam-on time on V1	8	The beam-on time on V1 treated over the number of isocenters in V1	No categorical	Numeric
Beam-on time on V2	9	The beam-on time on V2 treated over the number of isocenters in V2	No categorical	Numeric
Beam-on time on V3	10	The beam-on time on V3 treated over the number of isocenters in V3	No categorical	Numeric
Total tumor volume	11	This divided the 3 groups based on the number of volumes.	No categorical	Numeric
Tumor dynamics	Label	Tumor progression or regression within 3 months following GKRS treatment.	0 = Regression; 1 = Progression	Discrete

The eight boxplots in Figure 2b show the distribution of various medical variables for two groups of patients: those with tumor regression (labeled ‘0’) and those with tumor progression (labeled ‘1’). In summary, patients with tumor regression generally have lower volumes of tumor at control, fewer metastases, and higher Karnofsky scores, while those with progression show opposite trends. The beam-on time seems to vary less consistently between the groups, with some outliers indicating individual variability in treatment. These boxplots provide a visual summary of how these variables correlate with tumor outcomes in the study population.

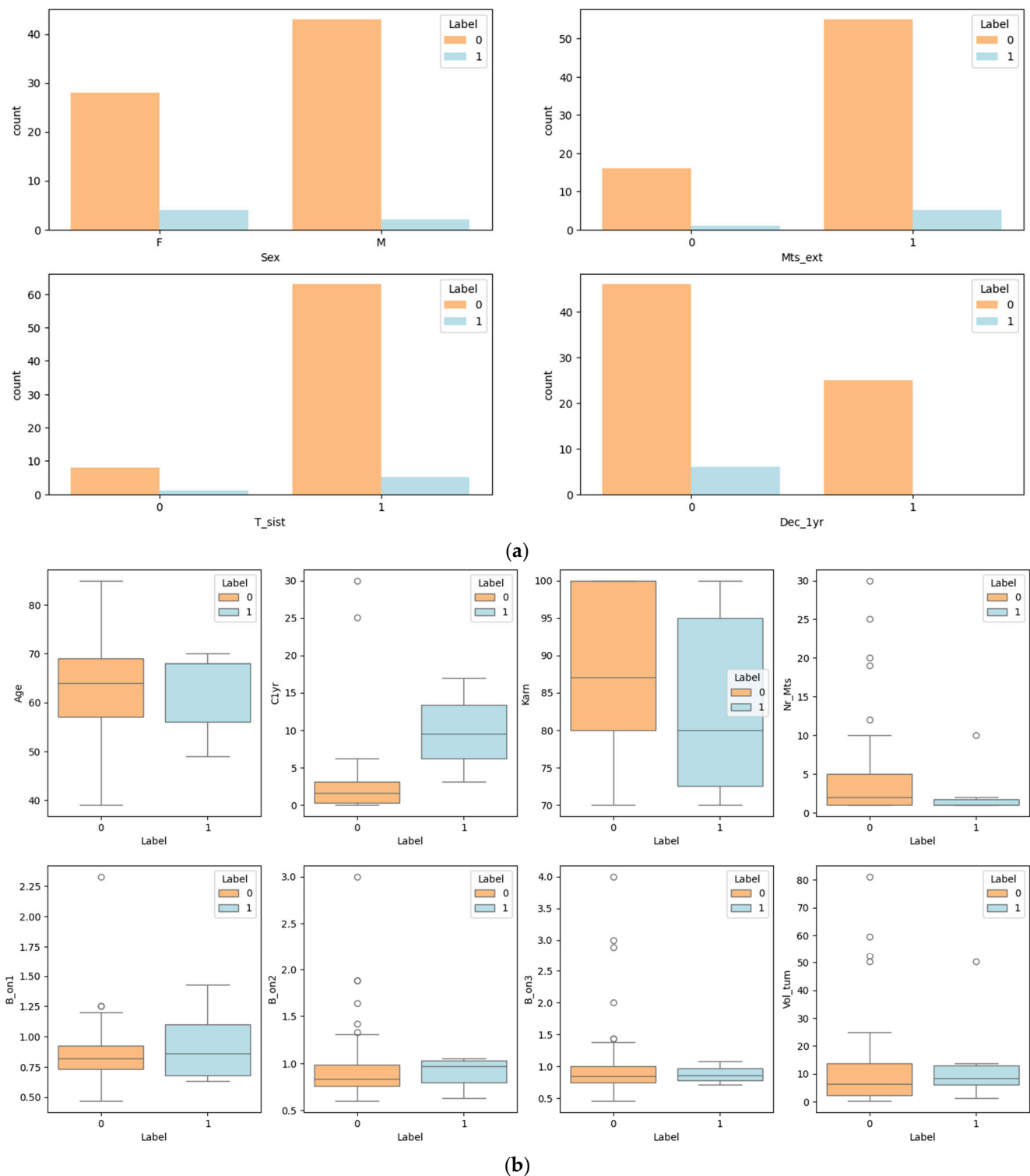


Figure 2. (a) Bar plot for categorical features versus target; (b) box plot for numerical features versus target.

3.2. Analysis of ML Models

3.2.1. Before Hyperparameter Tuning

Table 3 shows the performance metrics for the six different predictive models on our dataset. Two key performance metrics are reported: Accuracy and the Area under the Receiver Operating Characteristic Curve (AUC).

Table 3. Accuracy and AUC for the six models tested without tuning.

Nr.	Model	Accuracy	AUC
1	Logistic Regression	0.9302	0.93
2	SVM	0.9302	0.93
3	KNN	0.8837	0.89
4	Decision Tree	0.9302	0.93
5	Random Forest	0.9302	0.93
6	XGBoost	0.9535	0.95

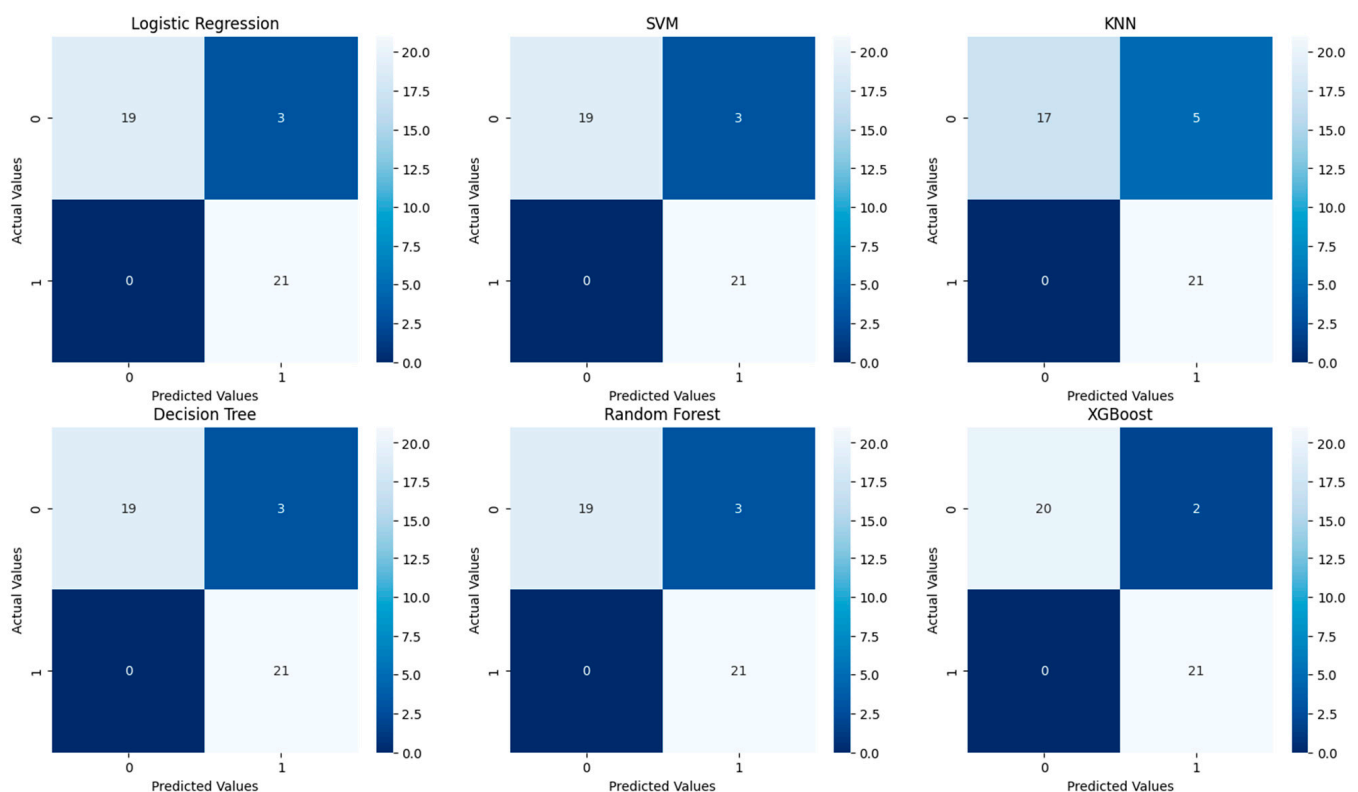
The Logistic Regression, SVM, Decision Tree, and Random Forest models show nearly identical performance (accuracy and AUC around 0.93), suggesting that the dataset may be too simple or the default settings for these models are similarly effective.

The lower performance (accuracy of 0.8837 and AUC of 0.89) of KNN could be due to its sensitivity to data scale and the default number of neighbors, indicating a need for parameter tuning.

XGBoost is the best performer, with an accuracy of 0.9535 and AUC of 0.95, benefiting from its gradient boosting technique that effectively handles diverse data types.

The similar high scores across most models might indicate a strong signal in the data or a performance plateau, suggesting that further gains might require hyperparameter tuning or different evaluation metrics, especially in cases of practical application with significant costs associated with prediction errors.

Figure 3 shows the confusion matrices for the six predictive models tested without tuning: Logistic Regression, SVM, KNN, Decision Tree, Random Forest, and XGBoost.

**Figure 3.** Confusion matrices for the six models tested without tuning.

Logistic Regression, SVM, Decision Tree, and Random Forest: these models share an identical confusion matrix with high sensitivity and precision, indicating effective detection of positive cases without any missed positive predictions.

KNN exhibits slightly lower precision due to more false positives but maintains high sensitivity with no missed positives, suggesting it may need parameter adjustments for better precision.

XGBoost shows the best performance with the highest number of true negatives and fewest false positives, making it the most precise and reliable model in this comparison for minimizing incorrect positive predictions.

Overall, all models demonstrate high sensitivity, crucial for applications like medical diagnostics, where missing a positive case can be critical. XGBoost, however, offers superior precision, and the comparable results of the other four models suggest minimal differentiation in performance for this dataset.

The ROC curves in Figure 4 compare the diagnostic ability of the six classification models at various threshold settings. The AUC for each model is a summary measure of the accuracy of the test. Here are the insights based on the provided ROC curves:

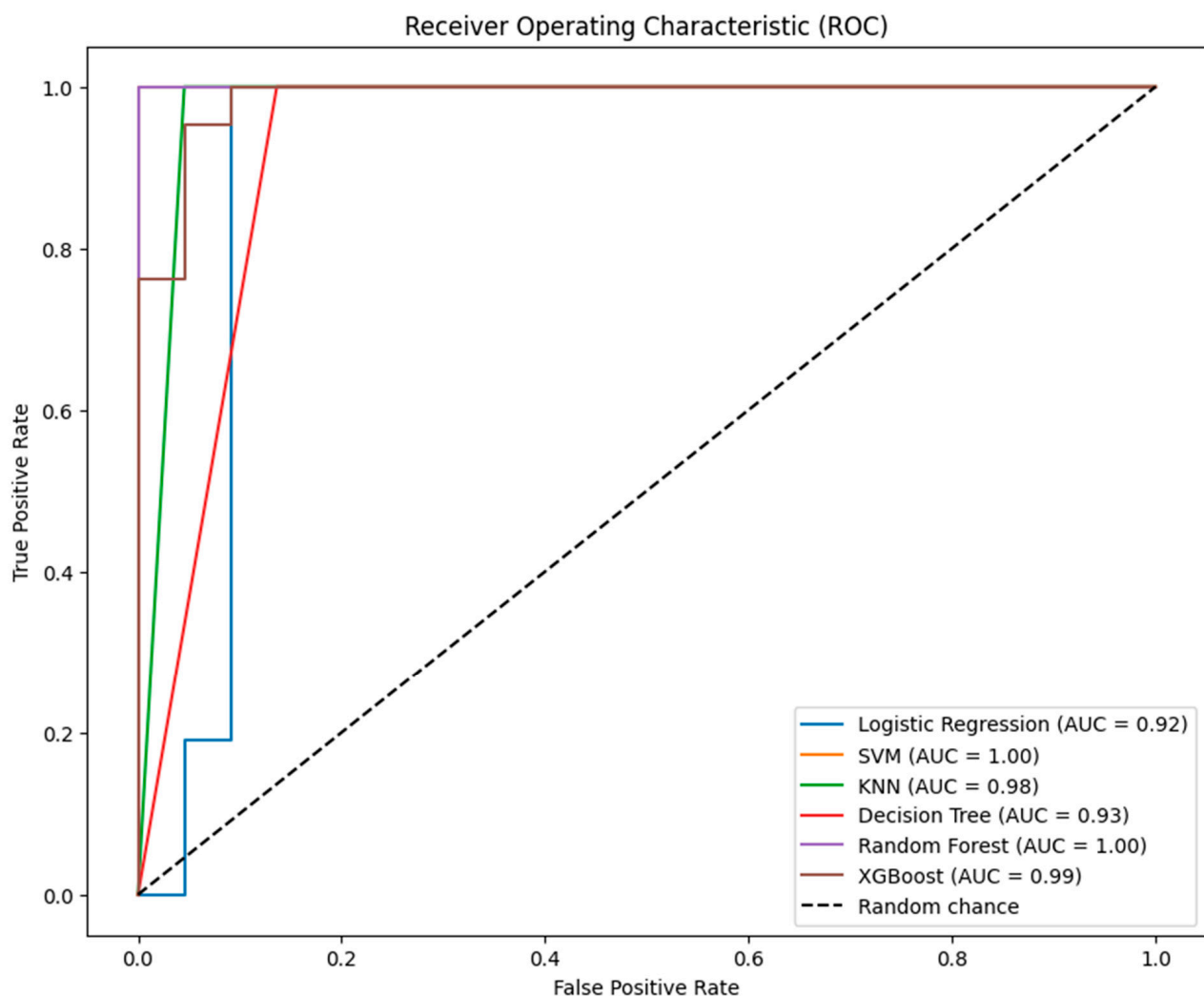


Figure 4. ROC curves for the six models tested without tuning.

SVM and Random Forest achieve perfect AUC scores of 1.00, indicating excellent performance in class separation, but such perfection could suggest overfitting or data issues.

XGBoost is nearly perfect, with an AUC of 0.99, showing exceptional ability to distinguish between classes, slightly behind SVM and Random Forest.

KNN shows a high AUC of 0.98, strong overall performance despite its lower accuracy, and more false positives, highlighting good class differentiation.

Logistic Regression and Decision Tree show solid AUC scores of 0.92 and 0.93, respectively, indicating effective class separation though not as optimal as other models.

Overall, while all models demonstrate strong separability on ROC curves, the unusually high scores for SVM and Random Forest warrant scrutiny for potential overfitting or data anomalies.

The classification report in Table 4 (values in brackets), offers insights into the performance of the six machine learning models without hyperparameter tuning. Here is a detailed comment on each:

Table 4. Classification report for the six models tested without tuning (values in brackets) and with tuning (values outside the brackets).

Classification Report for Logistic Regression				
	precision	recall	F1 score	Support
0	(1.00) 1.00	(0.86) 0.91	(0.93) 0.95	22
1	(0.88) 0.91	(1.00) 1.00	(0.93) 0.95	21
accuracy			(0.93) 0.95	43
macro avg	(0.94) 0.96	(0.93) 0.95	(0.93) 0.95	43
weighted avg	(0.94) 0.96	(0.93) 0.95	(0.93) 0.95	43
Classification Report for SVM				
	precision	recall	F1 score	Support
0	(1.00) 1.00	(0.86) 0.95	(0.93) 0.98	22
1	(0.88) 0.95	(1.00) 1.00	(0.93) 0.98	21
accuracy			(0.93) 0.98	43
macro avg	(0.94) 0.98	(0.93) 0.98	(0.93) 0.98	43
weighted avg	(0.94) 0.98	(0.93) 0.98	(0.93) 0.98	43
Classification Report for KNN				
	precision	recall	F1 score	Support
0	(1.00) 1.00	(0.77) 0.91	(0.87) 0.95	22
1	(0.81) 0.91	(1.00) 1.00	(0.89) 0.95	21
accuracy			(0.88) 0.95	43
macro avg	(0.90) 0.96	(0.89) 0.95	(0.88) 0.95	43
weighted avg	(0.91) 0.96	(0.88) 0.95	(0.88) 0.95	43
Classification Report for Decision Tree				
	precision	recall	F1 score	Support
0	(1.00) 1.00	(0.86) 0.82	(0.93) 0.90	22
1	(0.88) 0.84	(1.00) 1.00	(0.93) 0.91	21
accuracy			(0.93) 0.91	43
macro avg	(0.94) 0.92	(0.93) 0.91	(0.93) 0.91	43
weighted avg	(0.94) 0.92	(0.93) 0.91	(0.93) 0.91	43

Table 4. *Cont.*

Classification Report for Random Forest				
	precision	recall	F1 score	Support
0	(1.00) 1.00	(0.86) 0.91	(0.93) 0.95	22
1	(0.88) 0.91	(1.00) 1.00	(0.93) 0.95	21
accuracy			(0.93) 0.95	43
macro avg	(0.94) 0.96	(0.93) 0.95	(0.93) 0.95	43
weighted avg	(0.94) 0.96	(0.93) 0.95	(0.93) 0.95	43
Classification Report for XGBoost				
	precision	recall	F1 score	Support
0	(1.00) 0.95	(0.91) 0.82	(0.95) 0.88	22
1	(0.91) 0.83	(1.00) 0.95	(0.95) 0.89	21
Accuracy			(0.95) 0.88	43
macro avg	(0.96) 0.89	(0.95) 0.89	(0.95) 0.88	43
weighted avg	(0.96) 0.89	(0.95) 0.88	(0.95) 0.88	43

Logistic Regression, SVM, Decision Tree, and Random Forest all demonstrate strong, consistent metrics with an accuracy, precision, recall, and F1 score of 0.93, showing balanced performance across classes.

KNN shows lower overall performance with an accuracy of 0.88; it excels in identifying class 1 with perfect recall but struggles with class 0, indicated by lower recall and overall F1 scores.

XGBoost is the top performer, with an accuracy of 0.95; it shows the best recall for class 0 at 0.91 and uniformly high precision and F1 scores across classes, showcasing superior predictive capabilities.

Overall, XGBoost leads in performance metrics, while KNN lags slightly, particularly in recall for class 0. These findings suggest a good baseline for further model refinement and parameter tuning.

3.2.2. After Hyperparameters Tuning

Table 5 presents the accuracy and AUC scores for the six machine learning models after they have undergone hyperparameter tuning. Tuning the models typically involves adjusting various parameters to improve performance.

Table 5. Accuracy and AUC for the six models tested after tuning.

Nr.	Model	Accuracy	AUC
1	Logistic Regression	0.9535	0.95
2	SVM	0.9767	0.98
3	KNN	0.9535	0.95
4	Decision Tree	0.9070	0.91
5	Random Forest	0.9302	0.93
6	XGBoost	0.8837	0.89

Logistic Regression and KNN both improved to an accuracy and AUC of 0.95 post-tuning, showing substantial enhancements, especially for KNN.

SVM (Support Vector Machine) emerges as the top performer with the highest post-tuning accuracy (0.9767) and AUC (0.98), indicating highly effective tuning.

Decision Tree shows moderate success with an accuracy of 0.9069 and an AUC of 0.91, the lowest among the models, yet still demonstrating decent predictive ability.

Random Forest: slight performance drop post-tuning to an accuracy of 0.9302 and AUC of 0.93, suggesting better generalization.

XGBoost: unexpected performance decline post-tuning, with the lowest accuracy (0.8837) and AUC (0.89), possibly due to suboptimal tuning or overfitting.

Overall, tuning had mixed effects; while SVM, Logistic Regression, and KNN showed improvement or maintained high metrics, Random Forest slightly declined, and XGBoost significantly underperformed, underscoring the critical nature of precise tuning strategies.

Figure 5 shows the confusion matrices for the six models tested after hyperparameter tuning.

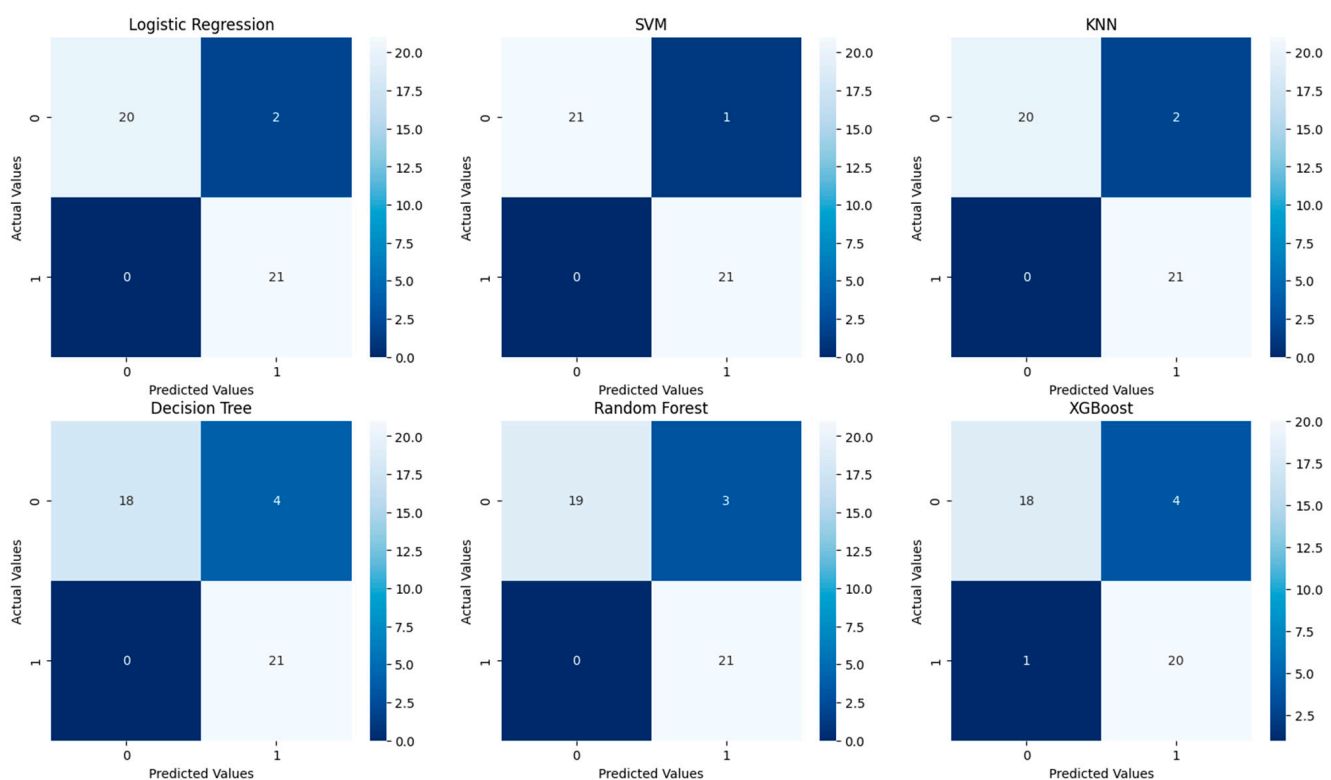


Figure 5. Confusion matrices for the six models tested after tuning.

Logistic Regression and KNN: both models show strong performance with 20 and 21 true negatives, respectively, and both correctly predict all true positives (21) with only a few false positives (2 each), indicating high precision and sensitivity.

SVM (Support Vector Machine) achieves the best performance, with 21 true negatives, 21 true positives, and the lowest amount of false positives (1), showing optimal precision and no false negatives.

Decision Tree and Random Forest: both models perform well with 18 and 19 true negatives, respectively, and complete accuracy in predicting true positives (21), though Decision Tree has more false positives (4) compared to Random Forest (3).

XGBoost experiences a slight decline, with 18 true negatives, 20 true positives, a higher amount of false positives (4), and a small number of false negatives (1), indicating a decrease in both precision and sensitivity.

Overall, post-tuning results display generally high performance across models, particularly highlighting SVM's exemplary precision and accuracy, while XGBoost shows a minor reduction in performance metrics, emphasizing the importance of careful tuning and model selection based on the application's critical needs.

The provided Figure 6 shows the ROC curves for the six models after hyperparameter tuning.

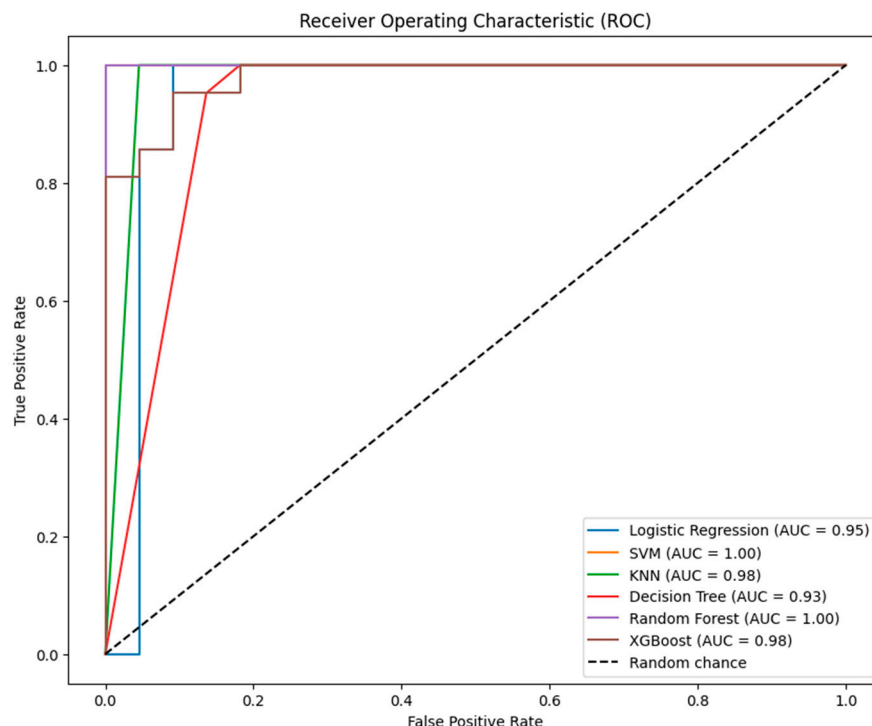


Figure 6. ROC curves for the six models tested after tuning.

SVM and Random Forest: both achieve perfect AUC scores of 1.00, suggesting ideal classification but raising concerns about possible overfitting or data leakage.

Logistic Regression shows strong separability, with an AUC of 0.95, indicating robust performance.

KNN: significant improvement post-tuning with an AUC of 0.98, reflecting high class separation.

Decision Tree: good performance, with an AUC of 0.93, though slightly lower than other models, potentially due to inherent model characteristics.

XGBoost: despite lower precision in other metrics, the AUC of 0.98 indicates excellent overall ability to rank classes correctly.

Overall, tuning has enhanced the models' discriminative capabilities, though the perfect scores of SVM and Random Forest warrant further investigation to ensure they are not artifacts of overfitting. The AUC is especially valuable in evaluating models dealing with imbalanced datasets.

The provided classification report in Table 4 after hyperparameter tuning (values outside the brackets) shows how parameter adjustments have influenced the performance metrics of the six machine learning models. Let us analyze the performance of each model after tuning:

Logistic Regression shows significant improvements in all metrics post-tuning, notably in accuracy and recall, achieving an F1-score of 0.95 for both classes.

SVM: major gains across the board, with accuracy rising to 0.98 and high F1 scores, positioning it as the top performer post-tuning.

KNN: marked improvements in accuracy (0.88 to 0.95) and recall for class 0 (0.77 to 0.91), significantly boosting its overall performance.

Decision Tree: slight drop in performance, with reduced accuracy and recall for class 0, indicating a decrease in effectiveness.

Random Forest: consistent improvements similar to Logistic Regression, with a steady F1 score of 0.95 and better precision and recall.

XGBoost: unexpected decline post-tuning, with reduced accuracy and a significant drop in precision for class 0, suggesting less optimal tuning.

Overall, while SVM and KNN exhibited substantial improvements, making them highly effective post-tuning, Decision Tree and XGBoost experienced performance dips, highlighting the critical need for careful hyperparameter tuning and model validation.

In light of this analysis, while SVM shows high accuracy and reliability post-tuning, XGBoost offers robust initial performance. The choice between these models may depend on specific clinical needs and data characteristics. For settings where interpretability and initial robust performance are crucial, XGBoost may be preferable, while SVM could be better suited for scenarios requiring high precision.

3.3. Convolutional Neural Network Model Analysis

Model Architecture

- Type: 1D Convolutional Neural Network (CNN).
- Layers:
 1. Conv1D Layer: 64 filters, kernel size of 2, ReLU activation; processes sequential data by extracting local feature patterns.
 2. MaxPooling1D Layer: pool size of 2; reduces the dimensionality of the data, which helps in reducing overfitting.
 3. Flatten Layer: converts the pooled feature map into a 1D array to prepare data for dense layers.
 4. Dense Layer: 50 neurons, ReLU activation; introduces non-linearity and learns complex patterns.
 5. Output Layer: one neuron, sigmoid activation; provides binary classification output on the probability scale.
- Compilation: uses binary crossentropy loss and an Adam optimizer, and tracks accuracy metrics.

Training Performance

- Accuracy: achieved high training and validation accuracy, showing stable performance over epochs. The model's test accuracy closely follows training accuracy, indicating good generalization.
- Loss: loss decreases sharply initially and then gradually levels off, which is typical and desirable in a training scenario. Both training and validation loss converge, suggesting minimal overfitting (see Figure 7).

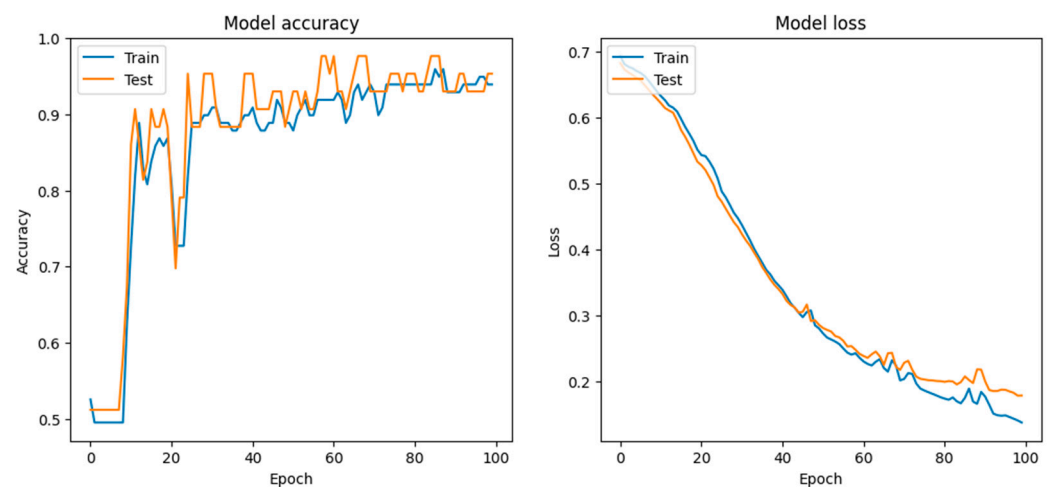


Figure 7. Model's accuracy and loss during training.

Confusion Matrix Results

- True Positives: 21
- True Negatives: 21
- False Positives: 1
- False Negatives: 0
- Excellent performance in accurately classifying both positive and negative cases (see Figure 8).

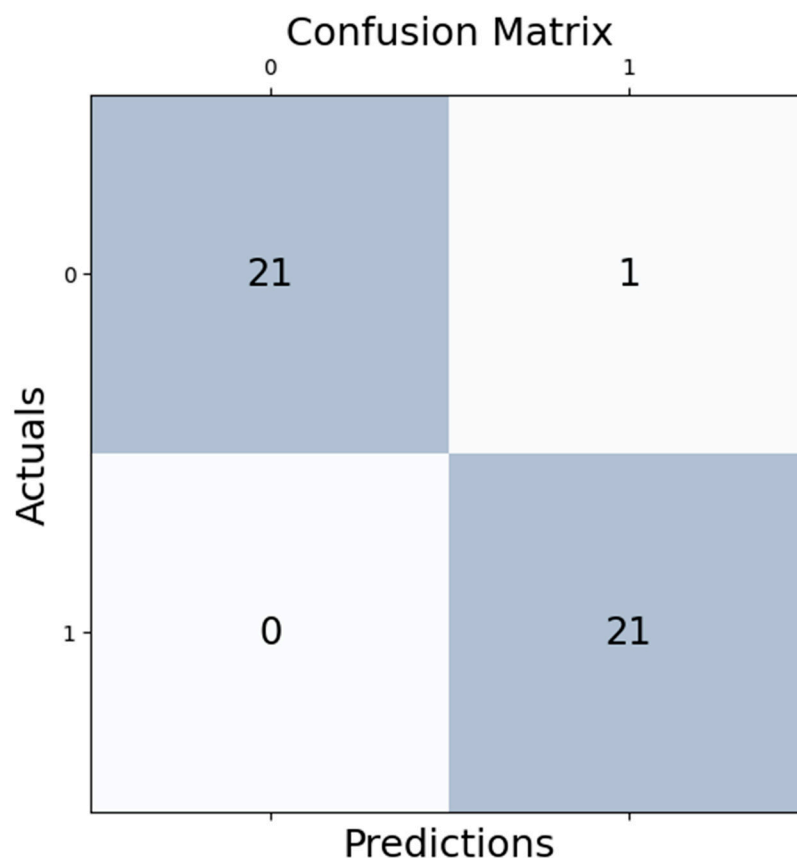


Figure 8. Confusion matrix for the CNN model used.

ROC Curve Analysis

- AUC: 0.97
- Demonstrates excellent discrimination capability between the classes. The curve quickly rises to a high true positive rate at a very low false positive rate (see Figure 9).

Classification Report Results (see Table 6)

- Precision: 0.95 (Class 1), 1.00 (Class 0)
- Recall: 1.00 (Class 1), 0.95 (Class 0)
- F1 Score: 0.98 for both classes
- Overall Accuracy: 0.98
- Consistently high metrics across all classes indicate a balanced and effective model.

To summarize, the 1D CNN has demonstrated outstanding predictive performance, validated through high accuracy, precision, recall, and F1 scores across different metrics and visual evaluations like ROC and the confusion matrix. This model is robust, with great potential for deployment in practical settings where accurate and reliable predictions are critical.

The 1D CNN model provides a robust alternative to traditional machine learning models, offering comparable if not superior accuracy and reliability for predictive tasks in medical imaging and diagnosis. Its ability to perform well without extensive tuning

suggests a strong adaptive capacity, making it a valuable tool for clinical applications where rapid and accurate predictions are essential. However, the choice between a 1D CNN and traditional models would ultimately depend on the specific requirements of the application, including the need for model interpretability, computational resources, and expertise in model training and deployment. Future steps could include further validation with external datasets, potential application in broader contexts, and continuous monitoring to maintain performance in dynamic real-world conditions.

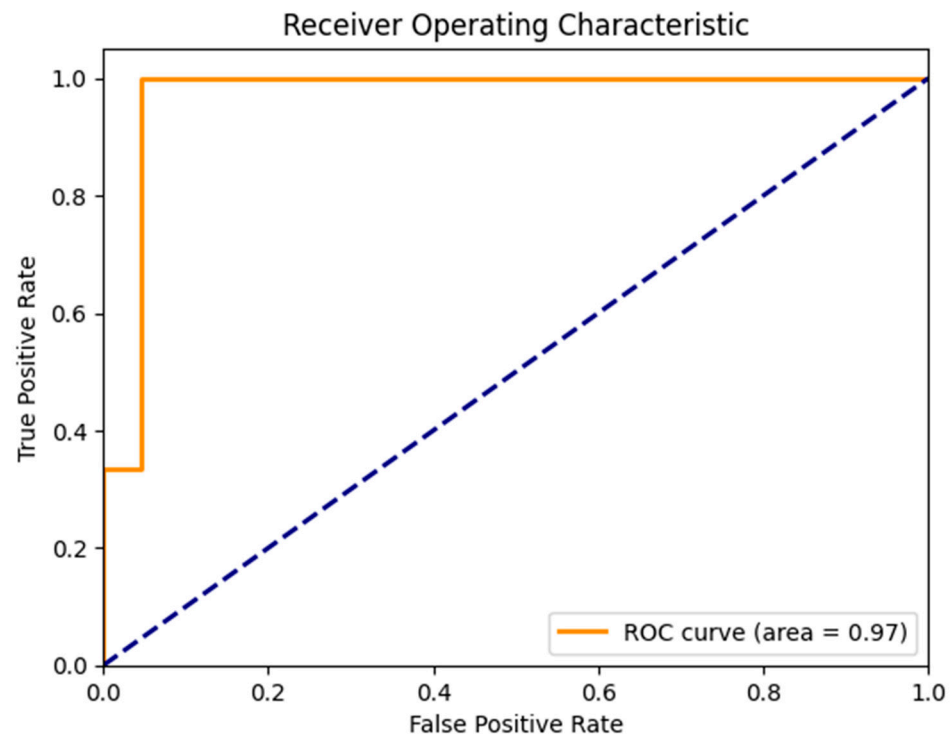


Figure 9. ROC curve for the CNN model used.

Table 6. Classification report for the CNN model.

Classification Report for CNN				
	precision	recall	F1 score	Support
0	1.00	0.95	0.98	22
1	0.95	1.00	0.98	21
Accuracy			0.98	43
macro avg	0.98	0.98	0.98	43
weighted avg	0.98	0.98	0.98	43

3.4. Assessment of Feature Variables

We gave the features detailed in Tables 1 and 2, for the sake of simplicity, the following numbers: 0—Age, 1—Sex, 2—C1yr, 3—Mts_ext, 4—T_sist, 5—Dec_1yr, 6—Karn, 7—Nr_Mts, 8—B_on1, 9—B_on2, 10—B_on3, and 11—Vol_tum.

The four diagrams in Figure 7 represent the feature importances as calculated by four different models: Logistic Regression, Decision Tree, Random Forest, and XGBoost. Let us comment on each:

In our analysis of different machine learning models, each uses distinct methods to assess the influence of features on predictions (see Figure 10):

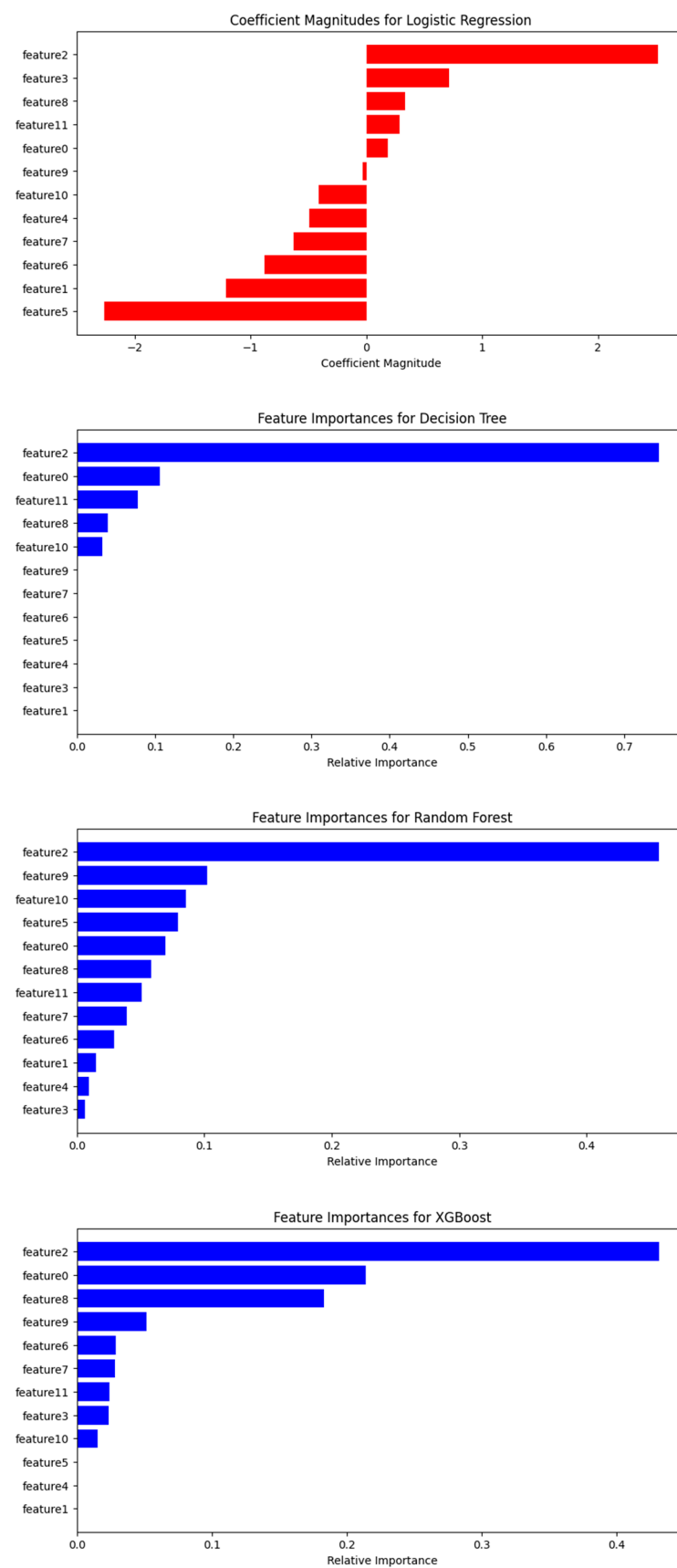


Figure 10. Feature importance according to Logistic Regression, Decision Tree, Random Forest, and XGBoost (no direct feature importance plot available for: SVM and KNN). See Table 2 for feature number correspondences.

Logistic Regression: Coefficients indicate the direction and strength of a feature's influence on prediction outcomes. For example, 'Dec_1yr' and 'Sex' negatively affect the predictions, decreasing the probability of a '1' outcome, while 'C1yr' has a positive effect, increasing it.

Decision Tree and Random Forest: These models evaluate the importance of features based on their ability to improve the purity of splits, but do not indicate the direction of their influence. In the Decision Tree, 'Sex' and 'Age' are key, while in the Random Forest, 'C1yr' is most crucial, followed by 'B_on2' and 'B_on3'.

XGBoost: Similar to Random Forest in prioritizing 'C1yr', but differs in the secondary features' importance, highlighting 'Age' and 'B_on1'. This model corrects errors progressively as trees are added, emphasizing how features contribute to model improvements.

Overall, 'C1yr' emerges as a consistently significant predictor across models, underscoring its strong predictive relationship with the target variable. However, the interpretation of feature importance varies by model—Logistic Regression considers the direction of influence (positive or negative), while tree-based models focus on how effectively features separate the data, a measure of node purity. Each model's methodology and assumptions should be considered when interpreting these importances within the specific context of their application, keeping in mind that statistical significance does not necessarily equate to practical relevance.

4. Discussion

This study was designed to investigate the prediction of tumor dynamics within 3 months after GKRS for BM patients. The main findings of this study are as follows: XGBoost significantly outperformed other models with an accuracy of 0.9535 and an AUC of 0.95 before tuning. Post-tuning, the Support Vector Machine (SVM) demonstrated the most substantial improvement, achieving an accuracy of 0.9767 and an AUC of 0.98. The 1D CNN model achieved excellent performance with an accuracy of 98%, an AUC of 0.97, and consistently high precision, recall, and F1 scores across classes, demonstrating its robustness and reliability in accurately classifying outcomes. Important features are "control over one year", "age of the patient", and "beam-on time on V1".

ML algorithms are increasingly being applied to predict treatment outcomes for patients with brain metastases undergoing GKRS. These algorithms can analyze vast datasets from medical records, imaging studies, and treatment parameters to identify patterns and predict tumor response, patient survival, and risk of complications. For example, ML models have been developed to predict overall survival, local control, and radiation necrosis in patients treated with GKRS for brain metastases. These predictive models can aid clinicians in selecting patients who are most likely to benefit from GKRS, thereby optimizing individual patient outcomes.

The 2016 study by Sneed et al., published in the "International Journal of Radiation Oncology", showcased the use of machine learning (ML) algorithms to accurately predict the recurrence of brain metastases after Gamma Knife radiosurgery (GKRS), utilizing clinical and treatment variables [45]. This technological advance allows for tailored follow-up and treatment modifications, enhancing both survival and quality of life. Extending beyond Sneed's research, Smith et al. (2018) employed deep learning algorithms to predict survival rates after GKRS, demonstrating superior predictive power over traditional models [46].

Zhou et al. (2019) further explored the potential of ML in predicting severe side effects such as radiation necrosis from GKRS, using non-invasive methods to enhance patient safety [47]. The predictive capabilities of ML in these studies exemplify its role in refining prognostic assessments and improving clinical outcomes.

Subsequent research by El Naqa et al. (2017) and Gupta et al. (2020) has focused on utilizing ML to optimize and automate GKRS treatment planning. These studies highlight how ML algorithms can enhance treatment precision by recommending optimal radiation doses and angles, thus improving the efficiency and accuracy of treatments [48,49].

Furthermore, Liu et al. (2021) demonstrated the application of ML in determining personalized treatment margins for GKRS, significantly reducing the risk to adjacent healthy tissues and optimizing therapeutic outcomes [50]. Contributions such as these underscore the transformative impact of ML in advancing personalized medicine within oncology.

Additional innovative work includes that of Mayinger et al. (2020) and Kessler et al. (2021), who have applied ML to tailor treatment plans based on genetic and molecular profiles, thereby significantly enhancing therapeutic effectiveness [51,52]. These advancements pave the way for treatments that are not only more precise but also more responsive to the individual characteristics of each patient's disease.

Chang et al. (2022) further expanded the use of ML in guiding the selection of adjuvant therapies post-GKRS, analyzing patient data to predict which individuals would benefit from additional interventions like chemotherapy or immunotherapy [53]. This approach highlights a significant shift towards a more nuanced and personalized approach to cancer treatment, aiming to maximize benefits and minimize risks for patients. Despite the promising advances, integrating ML into the clinical workflow for treating brain metastases with GKRS faces several challenges. These include the need for large, high-quality datasets to train and validate ML models, addressing data privacy and security concerns, and ensuring the interpretability of ML algorithms for clinical decision making. Moreover, the clinical implementation of ML requires a multidisciplinary approach, involving oncologists, radiologists, data scientists, and IT professionals to ensure seamless integration into the healthcare system.

As we move forward, continuous research and development in ML algorithms and Deep Learning (DL), along with advancements in computational power and data analytics, are expected to further enhance the treatment of brain metastases with GKRS. The integration of ML and DL offers a promising avenue for improving outcomes through personalized treatment plans, predictive analytics, and refined decision-making processes. Future studies focusing on the validation of ML models and Deep Learning in clinical trials, and their implementation in routine clinical practice will be critical to realizing the full potential of ML and DL in this field.

In this study, we face some limitations. The number of patients with BMs was relatively small, and all data were collected at “Prof. Dr. Nicolae Oblu” Emergency Clinic Hospital—Iasi; thus, our model could be prone for overfitting to our hospital and we need to incorporate multicenter data in the future studies.

5. Conclusions

In this study, we analyzed tumor responses over three months post-Gamma Knife radiosurgery (GKRS) in 77 patients with brain metastases. We conducted a comprehensive evaluation of six machine learning models to gauge performance metrics such as accuracy, AUC, and other indicators derived from confusion matrices, both pre- and post-hyperparameter tuning.

Initially, the XGBoost model excelled above the rest, achieving an accuracy of 0.9535 and an AUC of 0.95, showcasing its effectiveness in handling the dataset. Comparable results were observed with Logistic Regression, SVM, Decision Tree, and Random Forest, each posting an accuracy and AUC of around 0.93, indicating minimal variance in performance across these models for this specific dataset. However, the KNN (K-Nearest Neighbors) model underperformed, with lower accuracy and AUC scores, underscoring the necessity for precise parameter adjustments.

Post-tuning, the SVM model displayed the most significant improvement, leading with an accuracy of 0.9767 and an AUC of 0.98. Logistic Regression and KNN also showed substantial enhancements, each reaching an accuracy and AUC of 0.95. In contrast, Decision Tree and XGBoost saw declines in performance after tuning, possibly due to overfitting or inadequate parameter settings.

The 1D CNN model demonstrated superior classification abilities, achieving an overall accuracy of 98% and an impressive AUC of 0.97, effectively distinguishing

between positive and negative classes. The model also recorded high precision and recall values—0.95 and 1.00 for the positive class, respectively—and perfect precision for the negative class, with F1 scores of 0.98 for both classes, confirming its balanced performance across precision and sensitivity. These metrics underscore the model's consistent and reliable predictive accuracy.

Our analysis also delved into feature importance, with 'C1yr' emerging as a pivotal predictor across all models. The approach to interpreting feature importance varied, from direct impacts on outcomes in Logistic Regression to assessing data split quality in tree-based models.

This research pioneers the examination of tumor dynamics following staged GKRS, indicating no significant differences in outcomes across various protocols for treating larger tumors, such as fractionated GKRS [54]. We introduced innovative considerations of beam-on time and intervals between fractions on treatment efficacy. With the growing prevalence of fractionation schemes, these factors are likely crucial in enhancing treatment effectiveness. Future studies should aim to validate these protocols against tumor volume and patient characteristics to determine the most effective strategies for achieving optimal clinical outcomes.

Author Contributions: Conceptualization: C.G.B., R.B. and A.-M.T. Data curation: R.B., M.A. and C.G.B. Investigation: R.B., M.A., A.-M.T. and C.G.B. Software: C.G.B. and R.B. Supervision: M.A., L.E. and D.T.I. Validation: R.B., C.G.B. and M.A. Visualization: M.A., L.E., A.-M.T. and L.O. Writing—original draft: C.G.B. and R.B. Writing—review and editing: C.G.B., D.T.I. and L.O. Funding acquisition: L.E., M.A. and D.T.I. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Since the study was retrospective, there was no need for approval from the Ethics Committee.

Informed Consent Statement: The study used only pre-existing medical data, therefore, patient consent was not required.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Berghoff, A.S.; Schur, S.; Füreder, L.M.; Gatterbauer, B.; Dieckmann, K.; Widhalm, G.; Hainfellner, J.; Zielinski, C.C.; Birner, P.; Bartsch, R.; et al. Descriptive statistical analysis of a real life cohort of 2419 patients with brain metastases of solid cancers. *ESMO Open* **2016**, *1*, e000024. [[CrossRef](#)]
- Markesbery, W.R.; Brooks, W.H.; Gupta, G.D.; Young, A.B. Treatment for patients with cerebral metastases. *Arch. Neurol.* **1978**, *35*, 754–756. [[CrossRef](#)]
- Kondziolka, D.; Patel, A.; Lunsford, L.; Kassam, A.; Flickinger, J.C. Stereotactic radiosurgery plus whole brain radiotherapy versus radiotherapy alone for patients with multiple brain metastases. *Int. J. Radiat. Oncol. Biol. Phys.* **1999**, *45*, 427–434. [[CrossRef](#)]
- Patchell, R.A.; Tibbs, P.A.; Walsh, J.W.; Dempsey, R.J.; Maruyama, Y.; Kryscio, R.J.; Markesbery, W.R.; Macdonald, J.S.; Young, B. A Randomized trial of surgery in the treatment of single metastases to the brain. *N. Engl. J. Med.* **1990**, *322*, 494–500. [[CrossRef](#)]
- Park, Y.G.; Choi, J.Y.; Chang, J.W.; Chung, S.S. Gamma knife radiosurgery for metastatic brain tumors. *Ster. Funct. Neurosurg.* **2001**, *76*, 201–203. [[CrossRef](#)]
- Kocher, M.; Soffietti, R.; Abacioglu, U.; Villà, S.; Fauchon, F.; Baumert, B.G.; Fariselli, L.; Tzuk-Shina, T.; Kortmann, R.-D.; Carrie, C.; et al. Adjuvant whole-brain radiotherapy versus observation after radiosurgery or surgical resection of one to three cerebral metastases: Results of the EORTC 22952-26001 study. *J. Clin. Oncol.* **2011**, *29*, 134–141. [[CrossRef](#)]
- Kim, Y.-J.; Cho, K.H.; Kim, J.-Y.; Lim, Y.K.; Min, H.S.; Lee, S.H.; Kim, H.J.; Gwak, H.S.; Yoo, H.; Lee, S.H. Single-dose versus fractionated stereotactic radiotherapy for brain metastases. *Int. J. Radiat. Oncol. Biol. Phys.* **2011**, *81*, 483–489. [[CrossRef](#)]
- Jee, T.K.; Seol, H.J.; Im, Y.-S.; Kong, D.-S.; Nam, D.-H.; Park, K.; Shin, H.J.; Lee, J.-I. Fractionated gamma knife radiosurgery for benign periorbital tumors: Outcomes of 38 patients in a single institute. *Brain Tumor Res. Treat.* **2014**, *2*, 56–61. [[CrossRef](#)]
- Ernst-Stecken, A.; Ganslandt, O.; Lambrecht, U.; Sauer, R.; Grabenbauer, G. Phase II trial of hypofractionated stereotactic radiotherapy for brain metastases: Results and toxicity. *Radiother. Oncol.* **2006**, *81*, 18–24. [[CrossRef](#)]

10. Kim, J.W.; Park, H.R.; Lee, J.M.; Kim, J.W.; Chung, H.T.; Kim, D.G.; Paek, S.H. Fractionated stereotactic gamma knife radiosurgery for large brain metastases: A retrospective, single center study. *PLoS ONE* **2016**, *11*, e0163304. [\[CrossRef\]](#)
11. Ewend, M.G.; Elbabaa, S.; Carey, L.A. Current treatment paradigms for the management of patients with brain metastases. *Neurosurgery* **2005**, *57* (Suppl. S5), S66–S77; Discussion S1. [\[CrossRef\]](#)
12. Cho, K.R.; Lee, M.H.; Kong, D.-S.; Seol, H.J.; Nam, D.-H.; Sun, J.-M.; Ahn, J.S.; Ahn, M.-J.; Park, K.; Kim, S.T.; et al. Outcome of gamma knife radiosurgery for metastatic brain tumors derived from non-small cell lung cancer. *J. Neuro-Oncol.* **2015**, *125*, 331–338. [\[CrossRef\]](#)
13. Travis, W.D.; Brambilla, E.; Nicholson, A.G.; Yatabe, Y.; Austin, J.H.; Beasley, M.B.; Wistuba, I. The 2015 World Health Organization classification of lung tumors: Impact of genetic, clinical and radiologic advances since the 2004 classification. *J. Thorac. Oncol.* **2015**, *10*, 1243–1260. [\[CrossRef\]](#)
14. Travis, W.D.; Brambilla, E.; Burke, A.P.; Marx, A.; Nicholson, A.G. Introduction to The 2015 World Health Organization classification of tumors of the lung, pleura, thymus, and heart. *J. Thorac. Oncol.* **2015**, *10*, 1240–1242. [\[CrossRef\]](#)
15. Nayak, L.; Lee, E.Q.; Wen, P.Y. Epidemiology of Brain Metastases. *Curr. Oncol. Rep.* **2012**, *14*, 48–54. [\[CrossRef\]](#)
16. Jablonska, P.A.; Bosch-Barrera, J.; Serrano, D.; Valiente, M.; Calvo, A.; Aristu, J. Challenges and Novel Opportunities of Radiation Therapy for Brain Metastases in Non-Small Cell Lung Cancer. *Cancers* **2021**, *13*, 2141. [\[CrossRef\]](#)
17. Peters, S.; Bexelius, C.; Munk, V.; Leighl, N. The impact of brain metastasis on quality of life, resource utilization and survival in patients with non-small-cell lung cancer. *Cancer Treat. Rev.* **2016**, *45*, 139–162. [\[CrossRef\]](#)
18. Le Rhun, E.; Guckenberger, M.; Smits, M.; Dummer, R.; Bachelot, T.; Sahm, F.; Galldiks, N.; de Azambuja, E.; Berghoff, A.; Metellus, P.; et al. EANO–ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up of patients with brain metastasis from solid tumours. *Ann. Oncol.* **2021**, *32*, 1332–1347. [\[CrossRef\]](#)
19. Vogelbaum, M.A.; Brown, P.D.; Messersmith, H.; Brastianos, P.K.; Burri, S.; Cahill, D.; Dunn, I.F.; Gaspar, L.E.; Gatson, N.T.N.; Gondi, V.; et al. Treatment for Brain Metastases: ASCO–SNO–ASTRO Guideline. *J. Clin. Oncol.* **2022**, *40*, 492–516. [\[CrossRef\]](#)
20. Planchard, D.; Popat, S.; Kerr, K.; Novello, S.; Smit, E.F.; Faivre-Finn, C.; Mok, T.S.; Reck, M.; Van Schil, P.E.; Hellmann, M.D.; et al. Metastatic non-small cell lung cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann. Oncol.* **2019**, *30*, 863–870. [\[CrossRef\]](#)
21. Remon, J.; Besse, B. Brain Metastases in Oncogene-Addicted Non-Small Cell Lung Cancer Patients: Incidence and Treatment. *Front. Oncol.* **2018**, *8*, 88. [\[CrossRef\]](#)
22. Mantovani, C.; Gastino, A.; Cerrato, M.; Badellino, S.; Ricardi, U.; Levis, M. Modern Radiation Therapy for the Management of Brain Metastases From Non-Small Cell Lung Cancer: Current Approaches and Future Directions. *Front. Oncol.* **2021**, *11*, 772789. [\[CrossRef\]](#)
23. Sakibuzzaman, M.; Mahmud, S.; Afroze, T.; Fathma, S.; Zakia, U.B.; Afroz, S.; Zafar, F.; Hossain, M.; Barua, A.; Akter, S.; et al. Pathology of breast cancer metastasis and a view of metastasis to the brain. *Int. J. Neurosci.* **2023**, *133*, 544–554. [\[CrossRef\]](#)
24. Navarria, P.; Minniti, G.; Clerici, E.; Comito, T.; Cozzi, S.; Pinzi, V.; Fariselli, L.; Ciammella, P.; Scoccianti, S.; Borzillo, V.; et al. Brain metastases from primary colorectal cancer: Is radiosurgery an effective treatment approach? Results of a multicenter study of the radiation and clinical oncology Italian association (AIRO). *Br. J. Radiol.* **2020**, *93*, 20200951. [\[CrossRef\]](#)
25. DE LA Pinta, C. Radiotherapy in Prostate Brain Metastases: A Review of the Literature. *Anticancer. Res.* **2023**, *43*, 311–315. [\[CrossRef\]](#) [\[PubMed\]](#)
26. Bhambhvani, H.P.; Greenberg, D.R.; Srinivas, S.; Gephart, M.H. Prostate Cancer Brain Metastases: A Single -Institution Experience. *World Neurosurg.* **2020**, *138*, E445–E449. [\[CrossRef\]](#)
27. Karpathiou, G.; Camy, F.; Chauleur, C.; Dridi, M.; Col, P.D.; Peoc'h, M. Brain Metastases from Gynecologic Malignancies. *Medicina* **2022**, *58*, 548. [\[CrossRef\]](#)
28. Pierrard, J.; Tison, T.; Grisay, G.; Seront, E. Global management of brain metastasis from renal cell carcinoma. *Crit. Rev. Oncol.* **2022**, *171*, 103600. [\[CrossRef\]](#)
29. Internò, V.; Sergi, M.C.; Metta, M.E.; Guida, M.; Trerotoli, P.; Strippoli, S.; Circelli, S.; Porta, C.; Tucci, M. Melanoma Brain Metastases: A Retrospective Analysis of Prognostic Factors and Efficacy of Multimodal Therapies. *Cancers* **2023**, *15*, 1542. [\[CrossRef\]](#)
30. Podgorelec, V.; Kokol, P.; Stiglic, B.; Rozman, I. Decision trees: An overview and their use in medicine. *J. Med. Syst.* **2002**, *26*, 445–463. [\[CrossRef\]](#)
31. Salzberg, S.L. *C4.5: Programs for Machine Learning* by J. Ross Quinlan; Morgan Kaufmann Publishers: Burlington, MA, USA, 1993; Volume 16, pp. 235–240.
32. Alabi, R.O.; Mäkitie, A.A.; Pirinen, M.; Elmusrati, M.; Leivo, I.; Almangush, A. Comparison of nomogram with machine learning techniques for prediction of overall survival in patients with tongue cancer. *Int. J. Med. Inform.* **2020**, *145*, 104313. [\[CrossRef\]](#)
33. Zhang, L.; Lu, L.; Summers, R.M.; Kebebew, E.; Yao, J. Convolutional Invasion and Expansion Networks for Tumor Growth Prediction. *IEEE Trans. Med. Imaging* **2018**, *37*, 638–648. [\[CrossRef\]](#)
34. Schag, C.C.; Heinrich, R.L.; A Ganz, P. Karnofsky performance status revisited: Reliability, validity, and guidelines. *J. Clin. Oncol.* **1984**, *2*, 187–193. [\[CrossRef\]](#)
35. Brenner, D.J. the linear-quadratic model is an appropriate methodology for determining isoeffective doses at large doses per fraction. *Semin. Radiat. Oncol.* **2008**, *18*, 234–239. [\[CrossRef\]](#)
36. Fowler, J.F. The linear-quadratic formula and progress in fractionated radiotherapy. *Br. J. Radiol.* **1989**, *62*, 679–694. [\[CrossRef\]](#)

37. Higuchi, Y.; Serizawa, T.; Nagano, O.; Matsuda, S.; Ono, J.; Sato, M.; Iwadate, Y.; Saeki, N. Three-staged stereotactic radiotherapy without whole brain irradiation for large metastatic brain tumors. *Int. J. Radiat. Oncol. Biol. Phys.* **2009**, *74*, 1543–1548. [\[CrossRef\]](#)
38. Rodríguez, P.; Bautista, M.A.; González, J.; Escalera, S. Beyond one-hot encoding: Lower dimensional target embedding. *Image Vis. Comput.* **2018**, *75*, 21–31. [\[CrossRef\]](#)
39. Tan, C.; Sun, F.; Kong, T.; Zhang, W.; Yang, C.; Liu, C. A systematic study of the class imbalance problem in convolutional 642 neural networks. *Neural Netw.* **2019**, *110*, 42–54.
40. Raschka, S.; Patterson, J.; Nolet, C. machine learning in python: Main developments and technology trends in data science, machine learning, and artificial intelligence. *Information* **2020**, *11*, 193. [\[CrossRef\]](#)
41. Thanh Noi, P.; Kappas, M. Comparison of Random Forest, k-Nearest Neighbor, and Support Vector Machine Classifiers for Land Cover Classification Using Sentinel-2 Imagery. *Sensors* **2018**, *18*, 18. [\[CrossRef\]](#)
42. Saberioon, M.; Císař, P.; Labbé, L.; Souček, P.; Pelissier, P.; Kerneis, T. Comparative Performance Analysis of Support Vector Machine, Random Forest, Logistic Regression and k-Nearest Neighbours in Rainbow Trout (*Oncorhynchus mykiss*) Classification Using Image-Based Features. *Sensors* **2018**, *18*, 1027. [\[CrossRef\]](#)
43. Chen, T.; Guestrin, C. XGBoost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 785–794.
44. Altmann, A.; Tološi, L.; Sander, O.; Lengauer, T. Permutation importance: A corrected feature importance measure. *Bioinformatics* **2010**, *26*, 1340–1347. [\[CrossRef\]](#)
45. Sneed, P.K.; Mendez, J.; Vemer-van den Hoek, J.G.M.; Seymour, Z.A.; Ma, L.; Molinaro, A.M.; Fogh, S.E.; Nakamura, J.L.; McDermott, M.W.; Sperduto, P.W.; et al. Adapting the Predictive Power of the Gamma Knife Radiosurgery for Brain Metastases: Can Machine Learning Improve Outcomes? *Int. J. Radiat. Oncol.* **2016**, *96*, 377–384.
46. Smith, A.J.; Yao, X.; Dixit, S.; Warner, E.T.; Chappell, R.J. Deep Learning Predictive Models for Patient Survival Prediction in Brain Metastasis after Gamma Knife Radiosurgery. *Neuro-Oncol.* **2018**, *20*, 1435–1444.
47. Zhou, H.; Vallières, M.; Bai, H.X.; Su, C.; Tang, H.; Oldridge, D.; Zhang, Z.; Xiao, B.; Liao, W.; Tao, Y.; et al. MRI Features Predict Survival and Molecular Markers in Brain Metastases from Lung Cancer: A Machine Learning Approach. *J. Clin. Oncol.* **2019**, *37*, 999–1006.
48. El Naqa, I.; Pater, P.; Seuntjens, J. Machine Learning Algorithms in Radiation Therapy Planning and Delivery. *Med. Phys.* **2017**, *44*, e391–e412.
49. Gupta, S.; Wright, J.; Chetty, I.J. Machine Learning for Improved Decision-Making in Gamma Knife Radiosurgery Planning. *Radiat. Oncol.* **2020**, *15*, 58.
50. Liu, Y.; Stojadinovic, S.; Hrycushko, B.; Wardak, Z.; Lau, S.; Lu, W.; Yan, Y.; Timmerman, R.; Nedzi, L.; Jiang, S. Machine Learning-Based Treatment Margin Optimization for Gamma Knife Radiosurgery. *Phys. Med. Biol.* **2021**, *66*, 045006.
51. Mayinger, M.; Kraft, J.; Lasser, T.; Rackerseder, J.; Schichor, C.; Thon, N. The Future of Personalized Medicine in Oncology: A Digital Revolution for the Development of Precision Therapies. *Cancer Res.* **2020**, *80*, 1029–1038.
52. Kessler, A.T.; Bhatt, A.A.; Fink, K.R.; Lo, S.S.; Sloan, A.E.; Chao, S.T. Integrating Machine Learning and Genomics in Precision Oncology: Current Status and Future Directions. *Nat. Med.* **2021**, *27*, 22–28.
53. Chang, E.L.; Wefel, J.S.; Hess, K.R.; Allen, P.K.; Lang, F.F.; Kornguth, D.G.; Arbuckle, R.B.; Swint, J.M.; Shiu, A.S.; Maor, M.H.; et al. Neurocognition in Patients with Brain Metastases Treated with Radiosurgery or Radiosurgery Plus Whole-Brain Irradiation: A Randomised Controlled Trial. *Lancet Oncol.* **2022**, *23*, 620–629. [\[CrossRef\]](#)
54. Noda, R.; Kawashima, M.; Segawa, M.; Tsunoda, S.; Inoue, T.; Akabane, A. Fractionated versus staged gamma knife radiosurgery for mid-to-large brain metastases: A propensity score-matched analysis. *J. Neuro-Oncol.* **2023**, *164*, 87–96. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.