

Traditional Machine Learning Algorithms for Bone Tumor Treatment Prediction: A Comprehensive Performance Assessment

Helbi Mathew
Department of Informatics
Otto Von Guericke University
Magdeburg, Germany

Abstract— Bone tumor refers to unusual growth of cells in the skeletal system, characterized by varying degrees of aggressiveness and clinical indications. Handling these tumors presents significant challenges due to their complex characteristics and case specific factors. Typical approaches to bone tumor treatment, including surgery, radiation therapy, and chemotherapy depend on factors such as tumor type, position, and stage. Despite advances in technology and medical treatment, achieving individualized, fast opinion, and affordable treatment options remain grueling. This highlights the necessity for innovative strategies. Recently, machine learning algorithms have emerged as important tools in oncology, offering advancements in diagnostic accuracy, prognostic assessment, and treatment decision-making. This paper aims to identify the most effective traditional machine learning algorithm for prognosticating bone tumor treatment. The competence of algorithms such as logistic regression, decision trees, support vector machines, random forests, and k-nearest neighbors in predicting treatment pathways was assessed using clinical datasets. Random Forest demonstrated superior performance among all algorithms, achieving the top notch values for performance criteria. Whereas, Naïve Bayes exhibited comparatively poor performance. Machine learning algorithms offer a rapid as well as case-specific method for bone tumor treatment prediction. Through the development of a prognostic system derived from this study, future endeavors can shape individualized treatment strategies for individuals with bone tumors, potentially leading to further cost effective therapy.

Keywords—Bone tumor, machine learning, performance analysis.

I. INTRODUCTION

Bone tumors represent a significant challenge in oncology, particularly among individuals under the age of 20, where they rank as the third leading reason for cancer related mortality [1]. The division of bone tumors into benign, intermediate, or malignant categories is based on the 5th edition of the World Health Organization's classification system, which was published in 2020 [2]. In recent times there has been a notable increase in the prevalence of both benign and malignant bone cancers. While benign tumors outnumber malignant ones, the latter category is a larger threat.

Each type of bone tumor exhibits distinct biological characteristics: benign tumors generally remain stable, often needing just curettage or periodic monitoring [3]. Whereas to treat intermediate tumors more aggressive treatment strategies are needed. Malignant tumors, which have aggressive natural nature and propensity for distant metastasis, necessitate comprehensive remedial approaches, including surgery, chemotherapy, and radiotherapy [4],[5]. Given the heterogeneous clinical presentations of bone tumors, a precise differential diagnosis is pivotal for arriving at treatment decisions. A great many methods are presently employed in bone tumor diagnosis, with the foremost among them being:

- 1) Imaging Modalities: X-rays, CT scans, MRIs, and PET scans are generally used to visualize and assess bone lesions.
- 2) Biopsy: Tissue biopsy involves using a small bone tissue sample for histological examination, for the definitive diagnosis of tumor type and grade.
- 3) Blood Tests: Certain blood markers like alkaline phosphatase and lactate dehydrogenase may be high in people with bone tumors, facilitating diagnosis and treatment monitoring.
- 4) Histopathological Analysis: Microscopic examination of samples from biopsy or surgical resection enables pathologists to examine cellular features, and growth patterns of bone tumors, enabling precise diagnosis and grouping.
- 5) Bone Scintigraphy: This imaging approach injects a radioactive tracer that accumulates in areas of increased bone metabolism, helping in detecting and assessing bone tumor involvement.

In recent times, there has been a swell of interest in exercising machine learning techniques for bone tumor treatment. Numerous research studies have explored the application of ML algorithms in this area, aiming to increase the accuracy and effectiveness of treatment decision-making. In 2020 [6], the effectiveness of machine learning in differentiating between low-grade and high-grade cartilaginous bone tumors from radiomic parameters from unenhanced MRIs was assessed. The results showed that the AdaboostM1 classifier achieved an accuracy of 75% and corresponding AUCs of 0.78, suggesting that exercising machine learning methods holds significant flair in diagnosing cartilaginous bone tumors of varying degrees and

could be beneficial in preoperative tumor assessment. Su et al. [7] conducted a study to produce a clinical prediction model for pulmonary metastasis in osteosarcoma patients and assess the factors affecting its incidence. They conducted univariate and multivariate logistic regression analyses to identify threat factors and constructed a nomogram grounded on these findings. The nomogram displayed promising predictive ability and clinical significance, offering the potential for more substantiated identification and treatment guidance in osteosarcoma patients to enhance prognosis.

In their study, Pereira et al. [8] aimed to develop a machine learning model using CT radiomic features to prognosticate metastasis occurrence after an osteosarcoma diagnosis. They determined that the Random Forest algorithm performed best, achieving 73% accuracy and an AUC of 0.79. It found a noninvasive method for predicting pulmonary metastasis risk in osteosarcoma patients. Altogether, these researches contribute to the growing body of work using machine learning to predict treatment outcomes in bone tumor patients, potentially advancing personalized medicine and clinical management strategies.

Machine learning techniques offer a promising approach for acquiring knowledge from treatment history datasets of bone tumor patients. By erecting predictive models, ML algorithms can help guide treatment strategies. Nevertheless, it is tough to choose the best algorithm for this task, from the numerous available algorithms. This study utilizes six prominent ML algorithms—SVM, NB, KNN, DT, Logistic Regression, and Random Forest—to dissect clinical data and offer treatment recommendations. This approach aims to utilize the predictive power of machine learning to improve decision making in bone tumor management.

II. TREATMENT STRATEGIES FOR BONE TUMOR

To deal with bone tumors an extensive strategy that combines various disciplines to achieve the best possible treatment results is needed. The choice of treatment strategy is grounded on many factors, like the type, location of tumor, size, and stage of the tumor, as well as the patient's overall health and preferences. A detailed description of the major treatment options is given below:

A. Chemotherapy

Chemotherapy involves using potent drugs to eradicate or slow cancer cell growth. Administered orally or intravenously, these drugs travel all through the body to target cancer cells.

Research, including trials from the early 1980s [9], compared outcomes of adjuvant chemotherapy with surgery alone, leading to the development of neo-adjuvant treatment protocols. Today, contemporary regimens often involve multiagent neo-adjuvant protocols, significantly enhancing patient survival rates.

B. Surgery

The objective of surgical intervention is to remove tumors while minimizing damage to surrounding tissues. Techniques depend on tumor size, location, and aggressiveness, from simple excision to complex procedures like limb-salvage

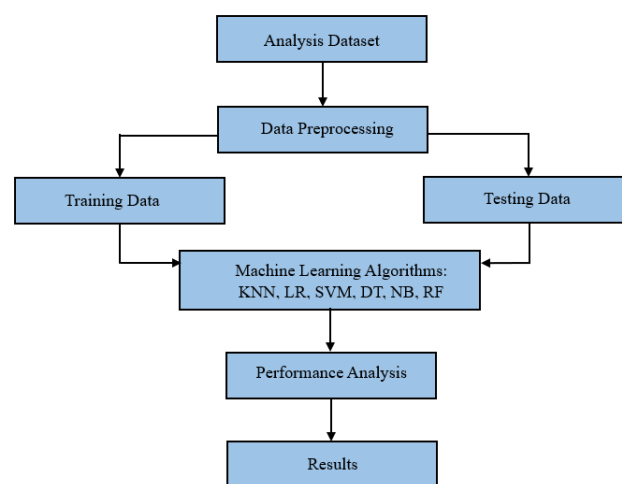


Fig. 1. Methodology Overview

surgery or amputation. Limb-salvage surgery preserves limb function and appearance. Advancements like computer assisted navigation and 3D printing have improved surgical precision. Even after amputation, high-grade osteosarcoma patients often require chemotherapy. For low-grade tumors, surgical excision alone may be sufficient, eliminating the need for chemotherapy.

C. Surgery

Radiotherapy uses high-energy radiation beams to target and destroy cancer cells in bone tissue, playing a crucial role in managing bone tumors by relieving pain and controlling growth. Studies by Chow et al. [10] compared different radiotherapy approaches, showing comparable pain relief and quality of life outcomes. These trials demonstrate the effectiveness of radiotherapy in providing symptomatic relief and improving patient comfort, highlighting its significance in palliative care and local disease control for bone tumor patients.

III. METHODOLOGY

To achieve the goal of this research, a methodology involving several key steps has been employed. Firstly, a comprehensive dataset comprising relevant patient records relevant to bone tumors was gathered, ensuring it contained necessary attributes for analysis. Next, various machine learning classification techniques were utilized to process the dataset and extract valuable insights. Finally, the predictive capabilities of these techniques were thoroughly assessed, evaluating their performance in guiding bone tumor treatment strategies.

An overview of the approach is provided in Fig. 1. Relevant patient data is collected and preprocessed to address class imbalances and remove irrelevant attributes. Various machine learning techniques are then applied, and their performance is evaluated on the test dataset to determine the most suitable classifier for treatment prediction.

A. Datasets and Attributes

The dataset for this study was obtained from Kaggle. The data in the dataset was collected from 500 patients at the Memorial Sloan Kettering Cancer Center between 2010 and 2020.

TABLE I. DATASET DESCRIPTION

S. No.	Attribute	Description
1	Patient ID	Object
2	Sex	Object
3	Age	Numeric
4	Grade	Object
5	Histological Type	Object
6	MSKCC type	Object
7	Site of primary STS	Object
8	Status (NED, AWD, D)	Object
9	Treatment	Object

The dataset comprises crucial patient and tumor details essential for the analysis. It includes demographic factors like Sex and Age at diagnosis, providing insights into potential age or gender-related tumor patterns. Tumor characteristics such as Grade, Histological type, and MSKCC type delineate aggressiveness and specific histopathological profiles. Disease status (NED, AWD, D) is documented, with the Status attribute, and Treatment attribute outlines administered interventions (surgery, radiation, chemotherapy). These features are summarized in Table I for reference.

B. Machine Learning Techniques

Classification algorithms are essential components of machine learning, proficient at categorizing data into distinct classes based on their attributes or features. They play a critical role across many domains, including healthcare, where they aid in tasks such as disease diagnosis, prognosis, and treatment prediction. In the realm of bone tumor management, these

algorithms can prove invaluable in predicting optimal treatment strategies tailored to individual patient traits and tumor profiles. The six prominent classification algorithms analyzed here are discussed below:

1) Random Forest: It is an ensemble learning method renowned for its versatility in constructing multiple decision trees during training and subsequently outputting the mode of the classes for classification tasks. Its robustness and efficiency in handling high-dimensional data with noise have made it a popular choice in various domains. Moreover, its capability to effectively manage missing data has been demonstrated in a study conducted by Shah et al. [11], further highlighting its reliability and applicability across diverse datasets and research contexts.

2) Logistic Regression: It is a widely used foundational method that models the probability of a binary outcome. It operates as a linear model by predicting the probability of occurrence of an event through fitting data to a logistic function. A notable instance of its application in healthcare is its integration with deep learning techniques to forecast cardiovascular risk factors from retinal fundus photographs, as demonstrated in a study by Poplin et al. [12]. This underscores the algorithm's utility and adaptability in addressing complex medical challenges through predictive modeling.

3) K-Nearest Neighbors: KNN is a simple and effective classification algorithm that classifies data points based on the majority class among their k nearest neighbors. It is a non-parametric method that does not make strong assumptions about the underlying data distribution. It is one of the most popular and simplest classification and regression classifiers used in machine learning today. KNN has been evaluated and enhanced to address imbalanced datasets, highlighting its versatility and adaptability [13].

4) Support Vector Machines: SVM is a powerful supervised learning model used for classification, outlier identification and regression analysis. SVM aims to find the hyperplane that best separates data points into different classes, maximizing the margin between classes. SVMs are best utilized for binary classification problems. Cortes and Vapnik [14] introduced SVMs in 1995, laying the foundation for their widespread application in various fields, including healthcare.

5) Naïve Bayes: Naive Bayes classifiers are particularly favored for their simplicity and efficiency in handling large datasets with numerous predictors. In their seminal work, Domingos and Pazzani [15] not only analyzed the theoretical underpinnings of Naive Bayes classifiers but also showcased their effectiveness in practical settings. Their research provided valuable insights into the algorithm's capabilities and contributed significantly to its widespread adoption in diverse domains.

6) Decision Trees: It is a versatile classification algorithm that recursively partitions the feature space into regions, assigning labels to instances based on the majority class within

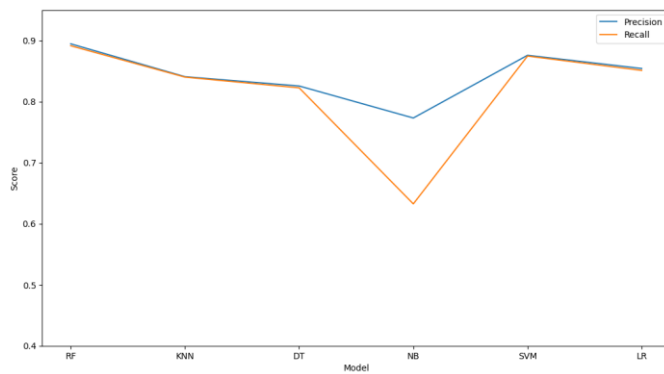
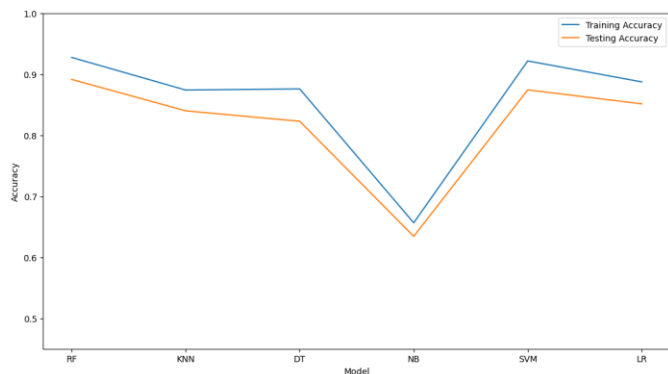


Fig. 2. Comparative performance of the analyzed machine learning algorithms: Left - Training and Testing accuracies. Right - Precision and Recall.

each region. Decision Trees are easily interpretable and are suitable for both numerical and categorical data. Quinlan [16] introduced decision tree algorithms like ID3 and C4.5, paving the way for their widespread adoption in various domains, including healthcare.

IV. EXPERIMENTAL RESULTS AND ANALYSIS

In this study on bone tumor treatment prediction, N-fold cross-validation technique has been employed, where N has been set to 5, and used to evaluate the performance of classification algorithms. This approach involves dividing the dataset into 5 folds, ensuring each fold is representative of the overall dataset. In every iteration, four folds are used for training and the remaining one fold is used for testing. This is repeated 5 times, using one fold serving as the testing set each time. The performance metrics, including precision, recall, F-measure, and accuracy, were computed for each iteration, and the final prediction accuracies were averaged across all 5 datasets to obtain comprehensive results.

A. Performance Measures

Performance measures in machine learning are numeric standards employed to evaluate the efficiency and accuracy of predictive models. These measures yield valuable insights into how well a model generalizes to unseen data and aids in comparative performance analysis of different algorithms.

TABLE II. ALGORITHM ACCURACY

S. No.	Classification Algorithms	Training Accuracy
1	Random Forest	89.14%
2	KNN	84%
3	Decision Tree	82.28%
4	SVM	87.43%
5	Logistic Regression	85.14%
6	Naïve Bayes	63.43%

Common performance measures for classification tasks include accuracy, precision, recall, AUC-ROC, F-measure. These performance measures play a crucial role in evaluating and fine-tuning machine learning models to ensure their fruitfulness in practical applications. This section defines the

evaluation metrics that are used for the experiments in this study. The metrics are derived from the counts of True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN).

1) *Precision*: It measures the proportion of correctly predicted positive cases among all predicted positive cases.

$$Precision = \frac{TP}{TP+FP} \tag{1}$$

2) *Recall*: It measures the proportion of correctly predicted positive cases among all actual positive cases.

$$Recall = \frac{TP}{TP+FN} \tag{2}$$

3) *Accuracy*: It measures the proportion of correctly predicted cases among all cases.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{3}$$

4) *F-measure*: It is the harmonic mean of precision and recall, providing a balance between them.

$$F - measure = \frac{2 * Recall * Precision}{Precision + Recall} \tag{4}$$

TABLE III. PRECISION AND RECALL OF CLASSIFICATION ALGORITHMS

S. No.	Classification Algorithms	Precision	Recall
1	Random Forest	0.894	0.891
2	KNN	0.841	0.84
3	Decision Tree	0.825	0.822
4	SVM	0.876	0.874
5	Logistic Regression	0.854	0.851
6	Naïve Bayes	0.773	0.623

B. Comparison Results

In Fig. 2, the training and testing accuracy for predicting bone tumor treatment, of the six algorithms investigated in this study are compared. Random Forest demonstrates slightly superior

performance. Table II shows that Random Forest and SVM achieve high testing accuracies of 89.14% and 87.43%, respectively. Conversely Naive Bayes has the least accuracy of 63.43%. This emphasizes the substantial gap in predictive accuracy between Random Forest and SVM compared to Naive Bayes.

Fig. 2, also illustrates the comparison of precision and recall values achieved by the different algorithms, with Random Forest once again demonstrating superior performance. Table III shows that Random Forest achieves impressive precision and recall scores of 0.894 and 0.891, respectively, while SVM demonstrates values of 0.876 and 0.874 for precision and recall, respectively. Naive Bayes stands out as the least effective algorithm, achieving a precision score of 0.773 and a recall score of 0.623.

These findings further reinforce the superiority of Random Forest and SVM over other algorithms in accurately predicting bone tumor treatment outcomes. The consistent performance of these two methods underscores their reliability and effectiveness in clinical decision-making for bone tumor management.

The bar chart presented in Fig. 3 compares the F1-score attained by various classification algorithms. Table IV shows that Random Forest exhibits remarkably high score of 0.891, which is better than those of all other algorithms.

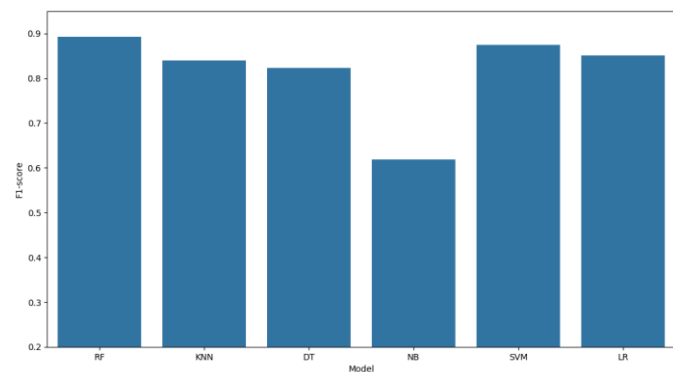


Fig. 3. Comparison of F1-score of ML algorithms.

RESULTS

Following thorough assessment based on precision, recall, F-measure, and accuracy metrics, the best suited machine learning approach for predicting bone tumor treatment has been identified. The evaluation pipeline demonstrates that with only a few demographic details and information on tumor type, Random Forest can effectively predict the ideal treatment pathway for the patient with efficiency. Through testing on a dataset specific to bone tumor cases, Random Forest has emerged as the leading method, exhibiting superior performance compared to other techniques. Referring to the findings depicted in Fig. 2 and 3, it is evident that an impressive accuracy of 89.14% in predicting bone tumor treatment using the provided medical dataset was achieved by Random Forest.

CONCLUSION

In this study, an analysis was conducted focusing on the prediction of bone tumor treatment, with various medical features associated with this condition being considered through the application of machine learning methodologies. Real healthcare datasets were utilized to extract valuable insights aimed at enhancing the prognosis of patients with bone tumors. To achieve effective prediction of bone tumor treatment pathways, experimentation with six broadly recognized machine learning algorithms was undertaken. The findings revealed the best existing classification algorithm for this purpose, which is Random Forest. The results of these experiments offer to equip healthcare professionals with the tools to make informed and prompt clinical judgments, ultimately aiding in the preservation of human life. Building upon the predictive system developed based on this research, future initiatives can revolutionize the landscape of personalized treatment strategies for patients. By leveraging the insights gleaned from the evaluation pipeline, healthcare professionals can tailor treatment plans to the specific needs of each patient, optimizing outcomes and minimizing potential side effects. By accurately predicting the most effective course of action, healthcare providers can streamline treatment protocols, potentially reducing time delays and unnecessary procedures. Ultimately, this could lead to improved patient outcomes, better quality of life, and a more sustainable healthcare system.

TABLE IV. F1-SCORE OF CLASSIFICATION ALGORITHMS

S. No.	Classification Algorithms	F1-Score
1	Random Forest	0.891
2	KNN	0.839
3	Decision Tree	0.823
4	SVM	0.874
5	Logistic Regression	0.851
6	Naive Bayes	0.618

REFERENCES

- [1] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics," CA: a cancer journal for clinicians, vol. 68, no. 1, pp. 7-30, January 2018.
- [2] International Agency for Research on Cancer ed, Soft tissue and bone tumours, 5th ed., vol 3, Lyon, France: Int. Agcy. Research on Cancer, 2020.
- [3] H. Fritzsche ,K. D. Schaser, C. Hofbauer, "Benign tumours and tumour-like lesions of the bone : General treatment principles", Orthopade pp. 484-497, Springer, June 2017.
- [4] N. Antropova , B. Q. Huynh, M. L. Giger, "A deep feature fusion methodology for breast cancer diagnosis demonstrated on three imaging modality datasets", Med Phys, pp. 5162-5171, October 2017.
- [5] R. Liu et al., "A deep learning-machine learning fusion approach for the classification of benign, malignant, and intermediate bone tumors," Eur. Radiol., vol. 32, no. 2, pp. 1371-1383, Feb. 2022.
- [6] S. Gitto et al., "MRI radiomics-based machine-learning classification of bone chondrosarcoma," Eur. J Radiology, Elsevier, vol. 128, p. 109043, 2020.
- [7] Z. Su, F. Huang, C. Yin, Y. Yu, and C. Yu, "Clinical model of pulmonary metastasis in patients with osteosarcoma: A new multiple machine learning-based risk prediction," J. Orthop. Surg, vol. 31, no. 2, June 2023.

- [8] H. M. Pereira, M. E. Leite Duarte, I. R. Damasceno, L. A. de Oliveira Moura Santos, and M. H. Nogueira-Barbosa, "Machine learning-based CT radiomics features for the prediction of pulmonary metastasis in osteosarcoma," *Brit. J. Radiol.*, vol. 94, no. 1124, p.20201391, Aug. 2021.
- [9] A. Misaghi, A. Goldin, M. Awad and A. A. Kulidjian, "Osteosarcoma: a comprehensive review," *SICOT-J*, vol. 4, 2018.
- [10] E. Chow et al., "Single versus multiple fractions of repeat radiation for painful bone metastases: a randomised, controlled, non-inferiority trial," *Lancet Oncol.*, vol. 15, no. 2, pp. 164-171, Feb. 2014.
- [11] A. D. Shah, J. W. Bartlett, J. Carpenter, O. Nicholas and H. Hemingway, "Comparison of random forest and parametric imputation models for imputing missing data using MICE: a CALIBER study," *AJE*, vol. 179, no. 6, pp. 764-774, March 2014.
- [12] R. Poplin et al., "Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning," *Nat. Biomed. Eng.*, vol. 2, no. 3, pp. 158-164, March 2018.
- [13] Z. Shi, "Improving k-nearest neighbors algorithm for imbalanced data classification," in *IOP Conf. Ser. Mater. Sci. Eng.*, 2020, p. 012072.
- [14] C. Cortes, and V. Vapnik, "Support-vector networks", *Machine learning*, vol. 20, pp. 273-297, Springer, 1995.
- [15] P. Domingos, and M. Pazzani, "Beyond independence: Conditions for the optimality of the simple bayesian classifier," in *Proc. 13th Intl. Conf. Machine Learning*, 1996, pp. 105-112.
- [16] J. R. Quinlan, "Generating production rules from decision trees," *IJCAI*, vol. 87, pp. 304-307, August 1987.