Mesopotamian Journal of Artificial Intelligence in Healthcare
Vol.2025, **pp.** 165-183
DOI: https://doi.org/10.58496/MJAIH/2025/016; ISSN: 3005-365X
https://mesopotamian.press/journals/index.php/MJAIH/index

Research Article

AdaBoost-Based Classification for Bone Sarcoma Outcome Prediction: A Comparative Machine Learning Approach

El-Sayed M. El-Kenawy[1,2,*], Ahmed Mohamed Zaki[3], Marwa M. Eid[4,5]

[1]*Department of Communications and Electronics, Delta Higher Institute of Engineering and Technology, Mansoura 35111, Egypt*
[2]*Applied Science Research Center, Applied Science Private University, Amman, Jordan*
[3]*Computer Science and Intelligent Systems Research Center, Blacksburg 24060, Virginia, USA*
[4]*Faculty of Artificial Intelligence, Delta University for Science and Technology, Mansoura 11152, Egypt*
[5]*Jadara Research Center, Jadara University, Irbid 21110, Jordan*

A R T I C L E I N F O

Article History
Received 27 Apr 2025
Revised 25 May 2025
Accepted 20 Jun 2025
Published 22 Jul 2025

Keywords
Bone Sarcoma
Ensemble Learning
Predictive Modeling
Clinical Outcomes
Adaptive Boosting

ABSTRACT

Bone sarcomas represent aggressive malignancies with complex clinical presentations where traditional prognostic methods often lack precision in predicting patient outcomes. This study developed and validated machine learning models for predicting bone sarcoma patient outcomes using a comprehensive dataset from Memorial Sloan Kettering Cancer Center spanning 2010-2020, comparing multiple algorithms including AdaBoost, Logistic Regression, Ridge Classifier, Quadratic Discriminant Analysis, and Linear Discriminant Analysis. AdaBoost demonstrated superior performance with 0.84 accuracy, 0.875 sensitivity, 0.7955 specificity, 0.8448 precision, 0.8333 negative predictive value, and 0.8596 F1 score, outperforming other algorithms which achieved 0.83 accuracy and 0.8496 F1 score. Statistical analysis confirmed significant differences between classifiers with F-statistic of 21.9130 and p-value less than 0.0001. The study concludes that AdaBoost-based classification provides a reliable framework for bone sarcoma outcome prediction with superior performance, demonstrating potential for clinical integration to support treatment planning and establishing a foundation for precision medicine applications in orthopedic oncology.

1. INTRODUCTION

Primary bone sarcomas represent a heterogeneous group of malignant neoplasms that pose significant clinical challenges due to their aggressive nature, propensity for metastasis, and complex treatment requirements. Osteosarcoma, the most common primary malignant bone tumor, accounts for substantial morbidity and mortality, particularly among pediatric and young adult populations [1, 2]. The rarity and heterogeneity of these tumors, combined with their variable clinical presentations and diverse histological subtypes, create substantial diagnostic and prognostic challenges for clinicians [3, 4].

Traditional diagnostic approaches rely heavily on clinical examination, imaging studies, and histopathological analysis. However, these conventional methods often struggle to provide accurate prognostic information, particularly regarding treatment response and patient outcomes. The complexity of bone sarcoma biology, characterized by extensive genomic heterogeneity and diverse cellular subtypes, necessitates more sophisticated analytical approaches that can integrate

* *Corresponding author. Email: skenawy@ieee.org*

multiple data sources and identify subtle patterns indicative of clinical outcomes [5, 6].

The advent of artificial intelligence (AI) and machine learning (ML) technologies has revolutionized medical diagnosis and prognosis across numerous oncological disciplines. In the context of bone tumors, AI-driven approaches have demonstrated remarkable capabilities in image analysis, tumor classification, and outcome prediction [7, 8]. Deep learning architectures, particularly convolutional neural networks, have shown exceptional performance in medical image segmentation and classification tasks, achieving diagnostic accuracy comparable to expert radiologists [9, 10]. These technological advances have enabled the development of computer-aided diagnostic systems that can process complex multi-modal data and extract meaningful prognostic information [11, 12].

Machine learning models have proven particularly valuable in oncological applications where traditional statistical methods may be insufficient to capture the complexity of underlying biological processes. Recent studies have demonstrated the effectiveness of various ML algorithms in predicting treatment responses, survival outcomes, and therapeutic recommendations across different cancer types [13, 14]. In the specific context of bone sarcomas, researchers have successfully applied ensemble learning methods to develop robust classification models that can distinguish between different tumor subtypes and predict metastatic potential [15].

Among the various machine learning approaches, ensemble methods have garnered significant attention due to their ability to combine multiple weak learners to create robust predictive models. AdaBoost (Adaptive Boosting), in particular, has demonstrated exceptional performance in medical classification tasks by sequentially training weak classifiers and adaptively adjusting their weights based on classification errors. This iterative approach allows the algorithm to focus on difficult cases and achieve superior generalization performance compared to individual classifiers [5].

The integration of machine learning with comprehensive genomic profiling and clinical data has opened new avenues for personalized cancer treatment. Studies have shown that ML-based approaches can effectively identify patients most likely to benefit from specific therapeutic interventions, thereby optimizing treatment strategies and improving patient outcomes [15]. In the context of bone sarcomas, such predictive models could potentially guide treatment selection, estimate prognosis, and facilitate early intervention strategies.

Despite these promising developments, several challenges persist in the application of machine learning to bone sarcoma outcome prediction. The relative rarity of these tumors limits the availability of large, well-annotated datasets necessary for robust model development and validation [1, 10]. Additionally, the heterogeneous nature of bone sarcomas, with their diverse histological subtypes and clinical presentations, requires sophisticated modeling approaches that can capture this complexity while maintaining interpretability and clinical relevance.

Current research in bone sarcoma outcome prediction has primarily focused on single-modality approaches or limited algorithmic comparisons. While individual studies have demonstrated the potential of various machine learning techniques, there remains a need for comprehensive comparative analyses that can definitively establish the optimal algorithmic approach for this specific clinical domain. Furthermore, the integration of multiple data types—including clinical variables, imaging features, and molecular markers—presents both opportunities and challenges for developing more accurate predictive models.

The present study addresses these gaps by conducting a systematic comparison of machine learning algorithms for bone sarcoma outcome prediction, with particular emphasis on the AdaBoost ensemble method. Our research aims to establish a robust predictive framework that can accurately classify patient outcomes while providing insights into the relative performance of different algorithmic approaches. By leveraging a comprehensive dataset encompassing diverse clinical and pathological variables, we seek to develop a practical tool that can support clinical decision-making and improve patient care.

The primary objectives of this investigation are threefold:

- To evaluate the comparative performance of multiple machine learning algorithms in predicting bone sarcoma outcomes
- To establish AdaBoost as a superior classification method for this specific clinical application
- To provide a validated predictive model that can be integrated into clinical practice to enhance prognostic accuracy and treatment planning

Through rigorous statistical validation and comprehensive performance analysis, this study contributes to the growing body of evidence supporting the integration of machine learning technologies in orthopedic oncology, ultimately advancing the field toward more precise and personalized patient care.

2. LITERATURE REVIEW

Artificial intelligence (AI) is rapidly changing orthopedic oncology by improving how primary malignant bone tumors (PBT) are diagnosed, classified, and treated. As noted [1], machine learning and deep learning algorithms are now being utilized to analyze large datasets and enhance the interpretation of medical images, thereby assisting clinicians in making better-informed decisions. The integration of radiomics with AI allows for the extraction of quantitative data from these images, enabling detailed characterization of tumors and the development of personalized treatment plans. This progress is particularly evident in the use of convolutional neural networks, which excel at recognizing patterns and have significantly improved the detection, segmentation, and differentiation of tumors.

Osteosarcoma, a primary malignant bone tumor, is characterized by aggressive growth, a propensity for metastasis, and consequently, a poor prognosis for patients with advanced disease. Based on the findings of [2], whole-exome evolutionary profiling of osteosarcoma cases from the TARGET database revealed linear evolutionary trajectories in a majority of patients, highlighting eight key mutations associated with metastatic progression and identifying early clonal ATRX mutations as significant drivers of tumor spread, ultimately constructing a metastasis classification model.

Chondrosarcoma is a heterogeneous and infrequent malignant bone tumor that causes clinical challenges due to limited treatment strategies and complex molecular processes. According to [3], single-cell RNA sequencing and data analysis with bioinformatics would be useful in comprehending cancer biology since it would outline cell subtypes, inferring signaling pathways, and signal gene expression patterns that define therapeutic entry points. This method offers an opportunity to understand cellular complexity and intricate communication networks between different cells forming a tumor microenvironment, and new horizons in the precision medicine approach.

The commonest malignant bone tumor in the pediatric age group, remains a clinical challenge as it is most likely to relapse even with improvement in treatment. In a study carried out by [4], an in-depth analysis was done to identify new osteosarcoma biomarkers and additionally to find out whether the biomarkers could serve as possible therapeutic targets. Their research used a combined analysis of several Gene Expression Omnibus and TARGET-OS clinicaltranscriptomic data and an immune-related genes analysis and machine learning algorithms to determine and verify major prognosis-relevant genes, as well as their functional significance in the development of osteosarcoma and provide a hint of future prognostic and therapeutic methods.

Following malignant bone tumor resection, endoprosthetic reconstruction is frequently employed for limb salvage, though implant failure remains a significant concern. Following the work of [5], machine learning (ML) models can be effectively utilized to predict early tumor endoprosthetic survival, ultimately providing improved patient-specific prognostication that can aid in expectation management and treatment recommendations. This innovative approach represents the first use of ML models to forecast endoprosthetic implant survival beyond one year and the first to incorporate upper extremity implants, potentially enhancing clinical decision-making in this challenging patient population.

Standardized reporting of bone tumors is essential for ensuring uniform and appropriate patient care strategies tailored to individual risk profiles. Based on the findings [6], machine learning techniques offer a promising avenue for distinguishing between benign and malignant focal bone lesions, potentially leading to enhanced risk stratification systems such as a future Bone Tumor Imaging Reporting and Data System 2.0.

The most common primary malignant bone sarcoma, is notorious for its aggressive nature, marked by high metastasis and mortality rates. Based on the research by [7], a classifier was developed using an eXtreme Gradient Boosting (XGBoost) algorithm with Bayesian optimization, integrating transcriptome and methylation data to predict high-risk OS subtypes. This model, incorporating nine genes (ARHGAP9, CADM1, CPE, DUSP3, FGFR1, GALNT3, IGF2BP3, KIF26A, ZFP3), demonstrated excellent predictive accuracy in an internal cohort and effectively stratified patients with varying survival outcomes in an external cohort, suggesting its potential to improve treatment strategies for high-risk OS patients.

Manually segmenting medical images presents a significant hurdle due to its reliance on expert annotations and the inherent inconsistencies found within these labels. As demonstrated by [8], AI-assisted labeling shows promise in addressing these challenges, particularly for 3D multi-modal bone tumor segmentation where achieving reliable and unbiased labels is crucial for effective supervised learning. A novel framework employing unsupervised feature clustering and semi-supervised refinement can minimize radiologist input and reduce labeling variability, ultimately leading to improved segmentation quality and a reduced workload for medical professionals.

The segmentation and three-dimensional reconstruction of bone tumors from two-dimensional image data hold substantial promise for advancing disease diagnosis and treatment. As evidenced in the study by [9], a U-Net model incorporating double dimensionality reduction and a channel attention gating mechanism (DCU-Net) was developed for oncological image segmentation, optimizing feature extraction and target space clustering to achieve automated segmentation and three-dimensional reconstruction of osteosarcoma; this innovative approach showcases the potential of integrating deep

learning-based medical image segmentation with mixed reality in the diagnosis and treatment of bone tumors by constructing a mixed reality infrastructure and exploring its application prospects.

Primary malignant bone tumors represent a significant cause of cancer-related deaths in young individuals. As outlined in the research of [10], the development of computer-aided diagnostic tools for these tumors, particularly those leveraging deep learning, has been hampered by a scarcity of accessible X-ray datasets. To address this limitation, a collaborative effort involving multiple medical institutions has resulted in the creation of a new resource, the Bone Tumor X-ray Radiograph dataset (BTXRD), which offers a substantial collection of bone images with corresponding clinical information, labels, masks, and bounding boxes for each tumor instance, thereby facilitating the advancement and assessment of deep learning algorithms in this domain.

A deep learning fusion model using computed tomography (CT) images and clinical features demonstrates potential for accurately classifying osseous and chondroid matrix mineralization, potentially improving the clinical diagnosis of osteogenic versus chondrogenic primary bone tumors. As reported by [11], a fusion model (SC-Net) attained an area under the receiver operating characteristic curve (AUC) of 0.901 (95% confidence interval not provided) in an external test set, suggesting a promising level of diagnostic performance for differentiating between these tumor types. This highlights the potential of integrating imaging and clinical data via deep learning to enhance diagnostic accuracy in challenging oncological scenarios, representing a step forward in leveraging artificial intelligence for improved patient care.

Bone cancer remains a critical health concern with potentially fatal consequences, often diagnosed through imaging techniques like CT scans, X-rays, and MRIs. In the analysis provided by [12], the diagnosis process, despite these technologies, still necessitates advancements to enhance accuracy and minimize human involvement due to challenges such as elevated costs, prolonged analysis times, and the potential for misdiagnosis stemming from the intricate nature of bone tumors; thus, the development of automated systems for distinguishing between healthy and cancerous bone tissue becomes paramount.

The application of machine learning to aid in the diagnosis of primary bone tumors holds the potential to improve diagnostic precision, promote earlier detection, facilitate personalized treatment strategies, and minimize both misdiagnoses and missed diagnoses, ultimately leading to better patient outcomes and increased survival rates. As stated in [13], a deep convolutional neural network (DC-NN) model combined with imaging omics analysis presents a promising approach for analyzing and discussing its clinical value in the diagnosis of primary bone tumors. Furthermore, the same research [13] introduced a screening method for differentially expressed genes, leveraging a paired T-test method that considers tumor purity and assesses actual gene expression values to identify significant genetic markers.

Metastatic bone tumors pose a significant threat to patient well-being and can accelerate the progression of cancer. Based on the findings of [14], a novel segmentation framework called BMSMM-Net, specifically designed for bone metastases detection, integrates a Bottleneck Gating Mamba layer (BGM) into the network backbone to improve the handling of long-range dependencies within depth feature maps; furthermore, a Skip-Mamba (SKM) module was designed on skip connections to facilitate long-range modeling during multi-scale feature fusion, alongside a Multi-Perspective Extraction (MPE) module in the feature extraction phase that leverages varied convolutional kernel sizes to bolster sensitivity to bone metastases which was evaluated on the BM-Seg dataset and demonstrates high-performance segmentation capabilities and computational efficiency, offering promise for clinical application in addressing the complexities of bone metastases segmentation.

The widespread adoption of comprehensive genomic profiling (CGP) is limited by the low likelihood of discovering druggable mutations and the associated financial and time burdens. In the view of [15], machine learning models can be used to predict the identification of genome-matched therapies by CGP, using a national database covering 99.7% of patients who underwent CGP in Japan from June 2019 to November 2023. This prediction is crucial to enhance the efficiency and effectiveness of precision medicine, especially by identifying patient characteristics likely to benefit most from CGP.

Accurate classification of primary bone tumors is essential for informing appropriate treatment strategies. Following the work of [16], deep learning models offer promise in classifying these tumors by leveraging incomplete multimodal imaging data, such as X-ray, CT, and MRI, alongside clinical characteristics, potentially mirroring real-world clinical scenarios more effectively than traditional methods.

Bone tumors, characterized by their rarity and varied imaging appearances, necessitate accurate differentiation between benign and malignant types. Based on the findings [17], an enhanced deep-learning model utilizing a convolutional neural network (CNN), specifically an optimized AlexNet, achieved high accuracy, precision, sensitivity, specificity, and F1-score in classifying femoral bone tumor images; the area under the curve (AUC) value further validated the algorithm's superior performance in terms of sensitivity and specificity, suggesting its potential to advance artificial intelligence applications in bone tumor classification.

Radiographic imaging combined with image recognition algorithms holds promise for classifying spinal bone tumor malignancies. According to the analysis by [18], convolutional neural networks (CNNs), specifically AlexNet and ResNet models, were employed to categorize spinal bone tumor images based on malignancy after being fine-tuned using a database of bone tumor images. These findings demonstrate the utility of deep learning techniques in enhancing the diagnostic process for these challenging conditions.

Lower extremity oncological resection and reconstruction frequently necessitate reoperations, influencing patient outcomes and burdening healthcare systems. As demonstrated by [19], machine learning (ML) models can be developed to predict the risk of such reoperations after oncological procedures in the lower extremities, with a polynomial support vector machine (SVM) exhibiting promising predictive capabilities based on internal validation metrics like AUC-ROC of 0.73 and a Brier score of 0.17, further highlighting the potential for these models to improve patient counseling and risk mitigation strategies in surgical settings.

TABLE I. COMPARING THE RECENT WORK RELATED TO HEART DISEASE.

No.	Main Focus	Methodology	Key Findings
Ref [1]	Review of AI applications in primary malignant bone tumor imaging.	Narrative review of machine learning and deep learning techniques.	AI enhances diagnosis, classification, and treatment response prediction in bone tumors.
Ref [2]	Whole-exome evolutionary profiling of osteosarcoma to uncover metastasis-related mutations.	Whole-exome sequencing and bioinformatics analysis.	Identified metastasis-related driver mutations and generated a predictive classifier for osteosarcoma.
Ref [3]	Multidimensional bioinformatics analysis of chondrosarcoma subtypes signaling networks.	Single-cell RNA sequencing (scRNA-seq) and bioinformatics.	Delineated cell subtypes and signaling networks, providing novel insights into chondrosarcoma.
Ref [4]	Investigate novel biomarkers for Osteosarcoma and potential therapy targets.	Integrated analysis of multi-omics data.	Identified potential early diagnostic biomarkers and therapeutic targets for OS.
Ref [5]	Develop machine learning models to predict tumor endoprosthesis survival after resection.	Machine learning model development and comparison.	Aims to provide patient-specific survival estimations and guide treatment planning.
Ref [6]	Enhanced CT and MRI focal bone tumor classification with machine learning.	Machine learning on multicenter CT and MRI data.	Evaluated a machine learning approach for differentiating between benign and malignant focal bone lesions.
Ref [7]	Prediction of high-risk osteosarcoma patients using XGBoost algorithm.	XGBoost algorithm using transcriptome and methylation data.	Identified molecular characteristics associated with high-risk osteosarcoma subtypes.
Ref [8]	Semi-supervised label generation for 3D multi-modal MRI bone tumor segmentation.	Framework for generating reliable and unbiased labels.	Aims to improve the accuracy and reliability of medical image segmentation for oncology.
Ref [9]	Mixed reality infrastructure for bone tumors using deep learning medical image segmentation and 3D visualization.	DCU-Net model based on double dimensionality reduction.	Improved accuracy and stability of bone tumor segmentation for assisting diagnosis and treatment.
Ref [10]	A radiograph dataset for the classification, localization, and segmentation of primary bone tumors.	Dataset creation and analysis for deep learning applications.	Provided a dataset to facilitate deep learning research for bone tumor diagnosis.
Ref [11]	Identification of osteoid and chondroid matrix mineralization in primary bone tumors using a deep learning fusion model.	Convolutional neural network (CNN) on CT scans and clinical data.	Developed a deep learning model for enhanced identification of matrix mineralization.
Ref [12]	Automated bone cancer detection using deep learning on X-Ray Images.	Deep learning model development on X-ray images.	Aimed to improve precision, reduce human labor, and overcome challenges in bone cancer diagnosis.
Ref [13]	Auxiliary diagnosis of primary bone tumors based on Machine learning model.	Machine learning on histopathological whole slide imaging (WSI).	Enhances diagnostic accuracy, facilitates early detection, and enables personalized treatment.
Ref [14]	Bone Metastasis Segmentation Framework Based on Mamba and Multiperspective Extraction (BMSMM-Net).	Deep learning for segmentation of bone metastases.	Aims to improve patient outcomes through rapid, precise segmentation of bone metastases.
Ref [15]	Machine learning analysis of cancer genomic profiling data to identify features associated with genome-matched therapy.	Machine learning model development.	Identified patient characteristics likely to benefit from comprehensive genomic profiling.
Ref [20]	Automated detection of bone lesions using CT and MRI: a systematic review.	Systematic review of AI applications for bone lesion detection.	Summarized advancements in automated detection systems using AI.
Ref [16]	Deep learning model to classify primary bone tumors using incomplete multimodal images.	Deep learning on X-ray, CT, and MRI images.	Addressed the challenge of incomplete multimodal images in clinical practice.
Ref [17]	Enhanced AlexNet-Based model for femoral bone tumor classification and diagnosis using magnetic resonance imaging.	Deep learning with convolutional neural networks (CNNs).	Improved tumor region delineation and classification using MRI.
Ref [18]	Radiographic imaging and diagnosis of spinal bone tumors using AlexNet and ResNet.	AlexNet and ResNet for classifying tumor malignancy.	Explored the application of image recognition algorithms for spinal bone tumor classification.
Ref [19]	Machine Learning Models for Predicting the 1-Year Risk of Reoperation After Lower Limb Oncological Resection.	Machine learning model development using PARITY trial data.	Predicted the 1-year reoperation risk following lower limb oncological resection.

3. MATERIALS AND METHODS

3.1 Dataset

This study utilized a comprehensive bone tumor dataset obtained from the Memorial Sloan Kettering Cancer Center (MSKCC), encompassing patient data collected between 2010 and 2020. The dataset, publicly available through https://www.kaggle.com/datasets/antimoni/bone-tumor/data

Dataset is structured in comma-separated values (CSV) format and contains detailed clinical and pathological information for bone tumor patients.

The dataset comprises the following key variables:

- **Patient ID**: A unique identifier assigned to each patient in the cohort, ensuring data integrity and enabling longitudinal tracking of patient outcomes.
- **Sex**: Patient demographic information recording biological sex, providing essential baseline characteristics for analysis.
- **Age**: Patient age at the time of initial diagnosis, recorded in years and serving as a critical prognostic factor in bone tumor outcomes.
- **Grade**: Tumor grade classification representing the degree of cellular differentiation and aggressiveness, serving as a key indicator of malignant potential and treatment planning.
- **Histological Type**: Specific tumor classification based on cellular morphology and histopathological characteristics, including major subtypes such as osteosarcoma, Ewing sarcoma, and other bone malignancies.
- **MSKCC Type**: Memorial Sloan Kettering Cancer Center-specific tumor classification system providing refined categorization based on institutional expertise and standardized diagnostic criteria.
- **Site of Primary STS**: Anatomical location of the primary soft tissue sarcoma within the bone, documenting the specific skeletal site of tumor origin and its relationship to surrounding structures.
- **Status**: Patient outcome status at the time of data collection, categorized into three distinct groups:
  - **NED**: No Evidence of Disease - indicating complete remission or absence of detectable tumor
  - **AWD**: Alive with Disease - patients surviving with persistent or recurrent tumor
  - **D**: Dead - patients who succumbed to disease progression or treatment-related complications
- **Treatment**: Comprehensive treatment modalities administered to patients, including but not limited to surgical resection, radiation therapy, chemotherapy, and multimodal therapeutic approaches.

This dataset provides a robust foundation for machine learning analysis, offering diverse clinical variables that enable comprehensive evaluation of prognostic factors and treatment outcomes in bone tumor patients. The inclusion of standardized outcome measures and treatment classifications facilitates the development of predictive models for clinical decision support and patient care optimization.

3.2 Data Analysis

Figure 1 shows the age distribution of the sampled population.The results are presented as a histogram with a smooth curve overlay.This visualization highlights the distinct multimodal nature of the distribution.Several peaks are evident, indicating concentrations in specific age brackets.The highest frequency is observed within the older age group, forming the largest peak.A second prominent peak is also noticeable for the middle-aged individuals.The youngest population segment is represented by a peak of lower frequency.There are significant dips in frequency between these concentrated age groups.This indicates a varied and non-uniform spread of ages within the dataset.

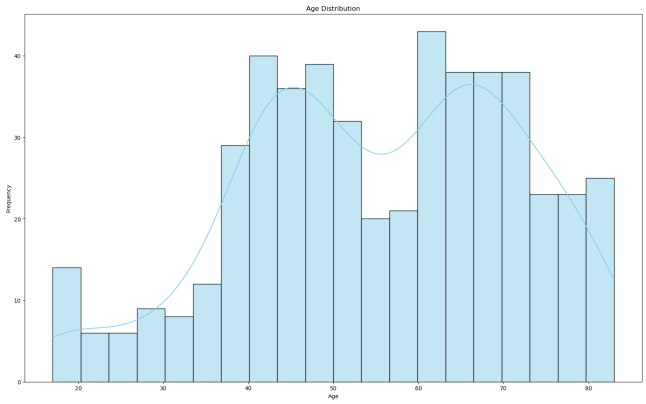

Fig. 1. Multimodal Age Distribution

Figure 2 shows the distribution of various histological sarcoma types. The bar chart visually represents the frequency counts for each specific subtype, arranged in descending order of prevalence. Pleomorphic leiomyosarcoma is depicted as the most frequently observed type, with a count of approximately 175 cases. Its frequency is considerably higher than all other histological classifications. Synovial sarcoma is the second most common type, appearing around 75 times. Following this, other significant types like leiomyosarcoma are shown, with about 50 instances. The counts for the remaining subtypes show a gradual decrease. The chart illustrates a long-tail distribution with less frequent types on the right.

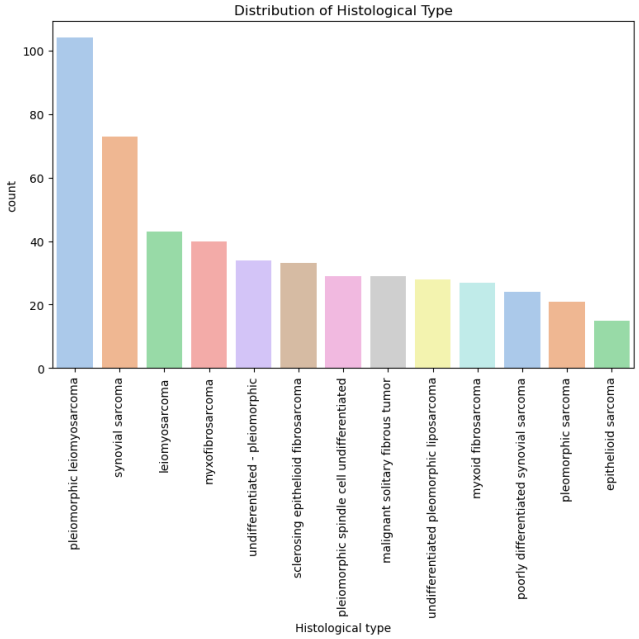

Fig. 2. Distribution of Histological Sarcoma Subtypes

Figure 3 shows the distribution of different sarcoma subtypes. The horizontal bar chart visualizes the frequency of three distinct categories. These sarcoma types are Leiomyosarcoma, MFH, and Synovial sarcoma. The length of each bar corresponds to the count of each specific type. The most prominent result shown is for MFH, the most frequently occurring type. Following this, Leiomyosarcoma is displayed as the second most common subtype. Synovial sarcoma is subsequently presented as the least frequent of the three. Each bar is shaded in a different tone of blue for clear visual distinction. This presentation illustrates a clear hierarchy in the prevalence of these sarcomas.

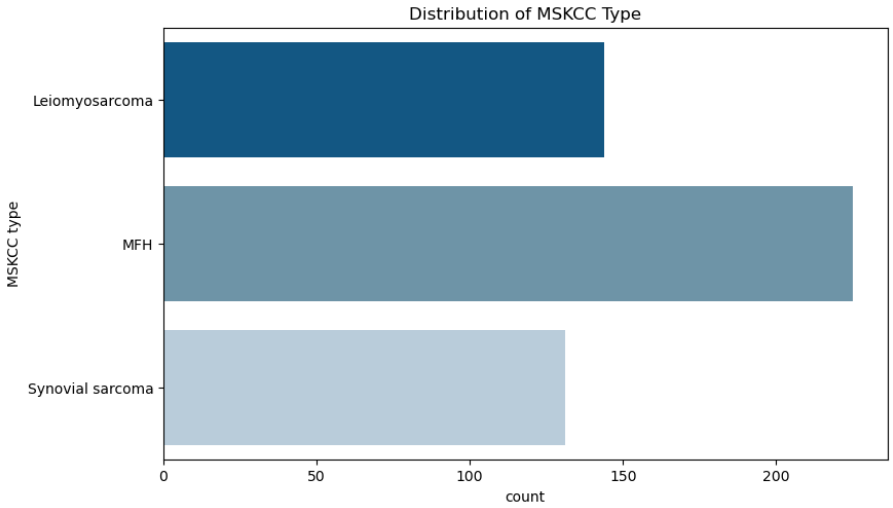

Fig. 3. Frequency of Sarcoma Subtypes

Figure 4 shows the distribution of primary STS across various anatomical locations presented in a bar chart. The data clearly reveals a non-uniform pattern of occurrence. Among the sites, the left thigh is the most prevalent location, exhibiting the highest frequency count in the analysis. The right thigh is the second most common site for primary STS. Ranking third, the right buttock is another significant location, though with a lower count than the thighs. The visualization emphasizes that the thighs are the most common regions.

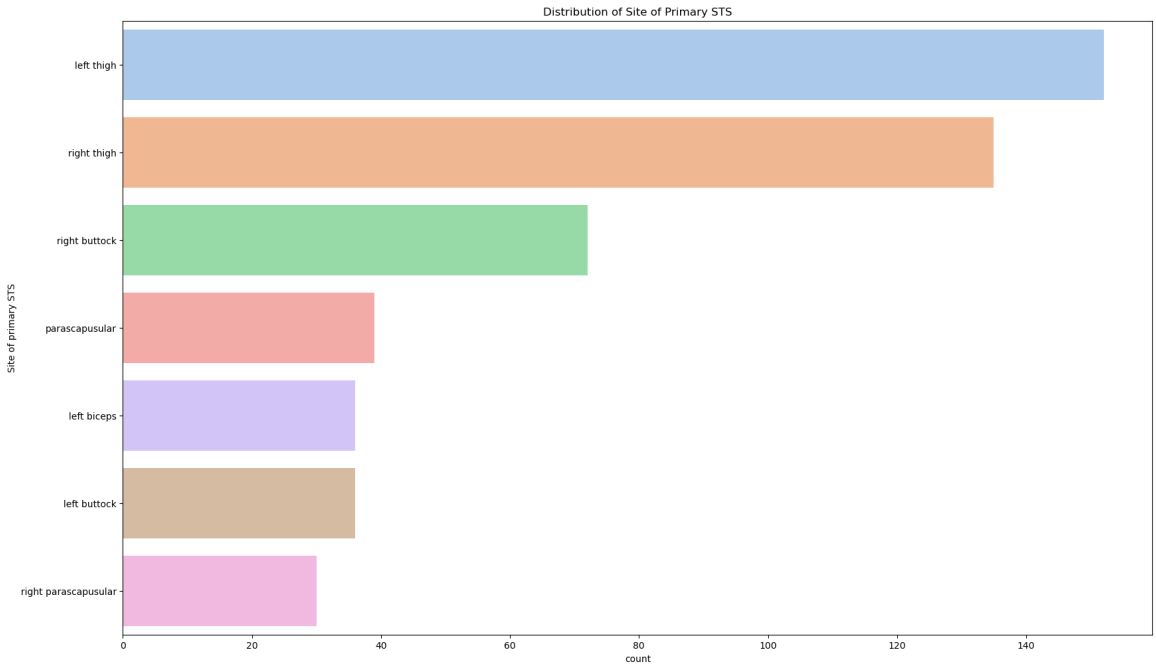

Fig. 4. Anatomical Distribution of Primary Soft Tissue Sarcoma

Figure 5 shows the distribution of patient outcomes across three different MSKCC sarcoma types. The best outcome presented, no evidence of disease (NED), is most prevalent in patients with Leiomyosarcoma. In stark contrast, for patients who are alive with disease, the MFH type is distinctly the most frequent. For the deceased patient status, Synovial sarcoma has the highest count.

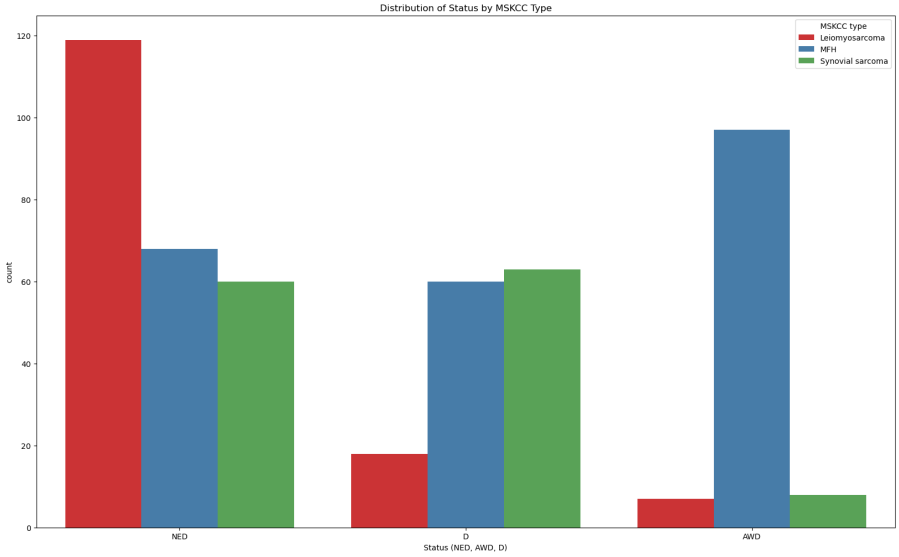

Fig. 5. Patient Outcomes by MSKCC Sarcoma Subtype

Figure 6 shows the distribution of primary treatments categorized by patient outcome status. For patients who have No Evidence of Disease (NED), the most successful treatment is Radiotherapy with Surgery. This combination shows a substantially higher count for this group. For individuals who are Alive with Disease (AWD), the best and most frequent model is the trimodal therapy. This consists of Radiotherapy, Surgery, and Chemotherapy, which is the most common intervention for this status.

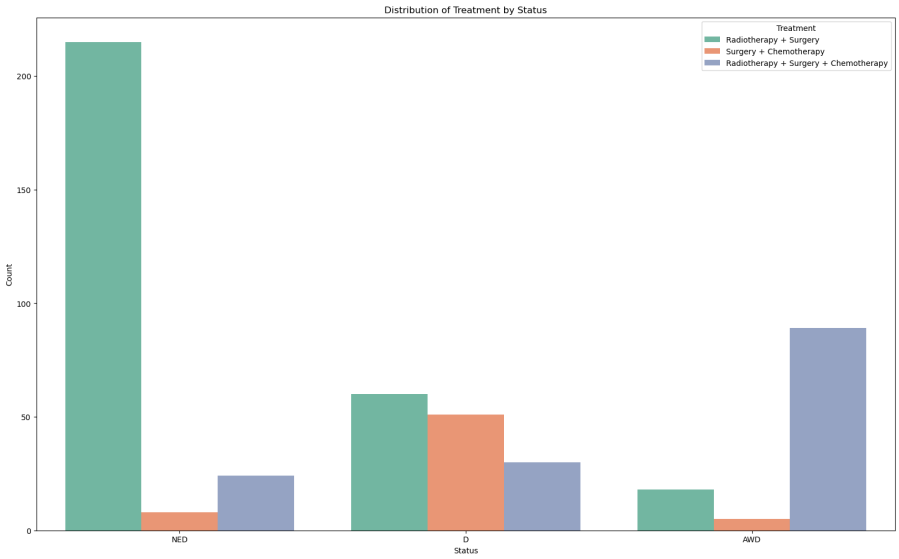

Fig. 6. Comparison of Treatment Modalities by Patient Survival Status

3.3 AdaBoost

AdaBoost (Adaptive Boosting) represents a sophisticated meta-algorithm that constructs a strong classifier through the weighted combination of multiple weak learners, typically decision stumps or shallow decision trees [21]. The algorithm implements a forward stagewise additive modeling approach, iteratively minimizing the exponential loss function while maintaining computational efficiency through greedy optimization strategies [22].

3.3.1. Theoretical Framework and Algorithmic Mechanics

The fundamental theoretical framework of AdaBoost is grounded in the PAC (Probably Approximately Correct) learning theory and margin-based generalization bounds [23]. The algorithm operates under the assumption that weak learners can achieve accuracy slightly better than random guessing, which is then amplified through iterative boosting to achieve strong classification performance [24].

The core algorithmic framework operates through sequential hypothesis generation, where each weak learner $h_t(x)$ is trained on a re-weighted distribution of training examples. This adaptive reweighting mechanism ensures that subsequent iterations focus increasingly on previously misclassified instances, effectively implementing a form of hard example mining [25].

The weight update mechanism follows the exponential reweighting scheme:

$$w_{t+1}(i) = \frac{w_t(i) \exp(-\alpha_t y_i h_t(x_i))}{Z_t} \tag{1}$$

where α_t represents the classifier weight determined by the training error rate, and Z_t serves as the normalization constant ensuring proper probability distribution properties [26].

The boosting process begins with uniform weight initialization across all training instances, followed by iterative weak learner training and weight adjustment. At each iteration t, the algorithm selects the weak learner that minimizes the weighted training error:

$$\epsilon_t = \sum_{i:h_t(x_i)\neq y_i} w_t(i) \tag{2}$$

The classifier weight α_t is computed as:

$$\alpha_t = \frac{1}{2} \ln\left(\frac{1-\epsilon_t}{\epsilon_t}\right) \tag{3}$$

AdaBoost implicitly minimizes the exponential loss function:

$$L(\mathbf{w}) = \sum_{i=1}^{n} \exp(-y_i f(x_i)) \tag{4}$$

where $f(x) = \sum_{t=1}^{T} \alpha_t h_t(x)$ represents the final strong classifier. This exponential loss function provides several advantageous properties, including differentiability and strong convexity, facilitating efficient optimization through coordinate descent approaches [27]. The exponential loss's margin-maximizing properties contribute to improved generalization performance by encouraging the algorithm to achieve not only correct classifications but also confident predictions with large margins.

AdaBoost provides strong theoretical guarantees regarding convergence and generalization performance. The algorithm's training error decreases exponentially with the number of iterations, provided that each weak learner achieves better than random performance. The generalization bound can be expressed in terms of the margin distribution:

$$P[\text{error}] \leq \frac{1}{m} \sum_{i=1}^{m} \exp(-y_i f(x_i)) \tag{5}$$

This bound demonstrates that the algorithm's generalization performance is related to the exponential loss on the training set, providing theoretical justification for continued boosting beyond perfect training accuracy.

3.3.2. Implementation and Predictive Modeling Properties

From a computational complexity perspective, AdaBoost maintains $O(T \cdot C)$ training complexity, where T represents the number of boosting iterations and C denotes the computational cost of training individual weak learners. This linear scaling property makes the algorithm computationally tractable for large-scale applications while maintaining theoretical performance guarantees [28]. The algorithm's space complexity is $O(T)$ for storing the ensemble of weak learners and their corresponding weights, making it memory-efficient compared to more complex ensemble methods.

The algorithm incorporates several implicit regularization mechanisms that contribute to its robustness and generalization capability. The sequential nature of weak learner addition implements a form of structural regularization, while the exponential reweighting scheme provides automatic feature selection through weighted voting. Early stopping based on validation performance serves as an explicit regularization technique, preventing overfitting while maintaining optimal predictive performance. The algorithm's inherent resistance to overfitting stems from its margin maximization properties and the smoothing effect of ensemble averaging [29].

AdaBoost's adaptive nature makes it particularly well-suited for medical classification tasks where data heterogeneity and class imbalance are common challenges [30]. The algorithm's ability to automatically adjust to the underlying data distribution through iterative reweighting enables effective handling of minority class instances that are often critical in medical diagnosis. The framework's capacity to incorporate domain knowledge through appropriate weak learner selection and its ability to handle mixed data types (categorical, continuous, and ordinal) make it valuable for multi-modal medical datasets where diverse feature types must be integrated for optimal predictive performance.

Practical implementation of AdaBoost requires careful consideration of several hyperparameters, including the choice of weak learner architecture, the number of boosting iterations, and convergence criteria. The algorithm's sensitivity to noise in training labels necessitates appropriate preprocessing and outlier detection mechanisms. The selection of appropriate weak learners significantly impacts the algorithm's performance and interpretability. Decision stumps provide high interpretability but may require more iterations, while shallow decision trees offer better individual performance but reduced interpretability of the final ensemble.

Within the context of predictive modeling, AdaBoost serves as a powerful non-parametric approach that makes minimal assumptions about the underlying data distribution. The algorithm's ability to approximate complex decision boundaries through the combination of simple weak learners enables effective modeling of non-linear relationships and interactions between features. The ensemble's predictive output can be interpreted probabilistically through appropriate calibration techniques, providing not only classification decisions but also confidence estimates that are valuable for medical decision-making applications.

This comprehensive technical framework establishes AdaBoost as a theoretically grounded and practically effective algorithm for complex classification tasks, particularly in domains where robustness, interpretability, and reliable predictive performance are paramount considerations.

4. EXPERIMENTAL RESULTS

4.1 Classification Models Results

Table **??** shows the performance of various models. Among all the classifiers evaluated in the analysis, the AdaBoost model was identified as the most effective. This particular model demonstrated superior capabilities by achieving the highest performance metrics overall. Specifically, it registered the top accuracy score, reaching a notable value of (0.84) in its predictions. It also attained the best F1 Score in the comparison, with a final calculated value of (0.859649).

TABLE II. PERFORMANCE COMPARISON OF TOP-PERFORMING CLASSIFICATION MODELS

Models	Accuracy	Sensitivity (TPR)	Specificity (TNR)	Precision (PPV)	NPV	F1 Score
AdaBoost	0.84	0.875	0.7955	0.8448	0.8333	0.8596
LogisticRegression	0.83	0.8571	0.7955	0.8421	0.814	0.8496
RidgeClassifier	0.83	0.8571	0.7955	0.8421	0.814	0.8496
QuadraticDiscriminantAnalysis	0.83	0.8571	0.7955	0.8421	0.814	0.8496
LinearDiscriminantAnalysis	0.83	0.8571	0.7955	0.8421	0.814	0.8496

Figure 7 shows a comparative analysis of five different algorithms. The performance of these models is assessed using accuracy and sensitivity metrics. Each algorithm's results are displayed using a pair of bars for the two metrics. Among the evaluated models, the AdaBoost algorithm stands out with superior results. This model achieved the highest score for the accuracy metric in the comparison. It also demonstrated the peak performance for sensitivity, the true positive rate. This establishes AdaBoost as the most effective classifier in this specific analysis.

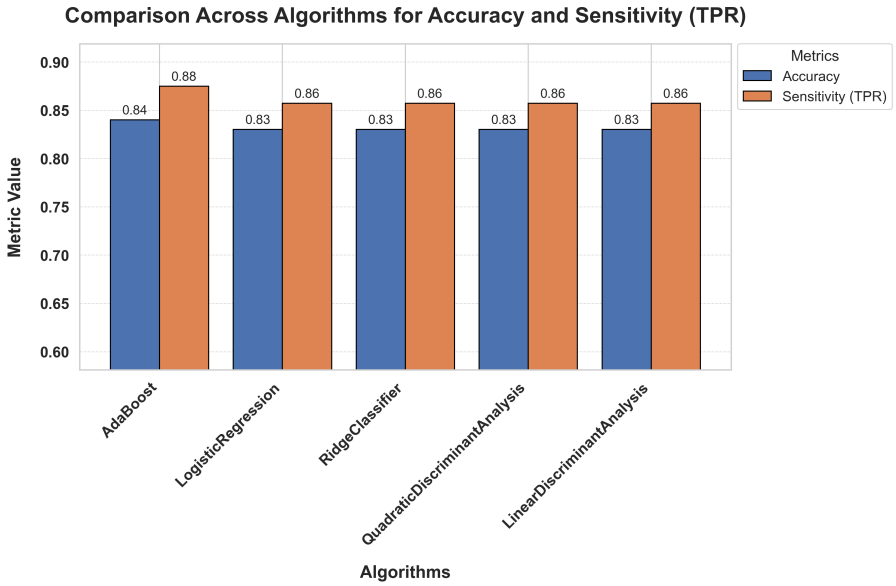

Fig. 7. Performance Comparison of Classification Algorithms by Accuracy and Sensitivity

Figure 8 shows the classification scores for five machine learning algorithms. The results are presented using combined violin and box plots, which effectively showcase the distribution and statistical summary of model performance. Among the algorithms evaluated, the AdaBoost model demonstrates the most favorable results. It consistently achieves the highest classification scores, as indicated by its superior position. The model exhibits a high median accuracy of approximately 0.95. Furthermore, its performance distribution is tightly concentrated at a high level. The interquartile range resides in a superior bracket, with most scores above 0.92. This indicates that AdaBoost is the most accurate and reliable model in this comparison.

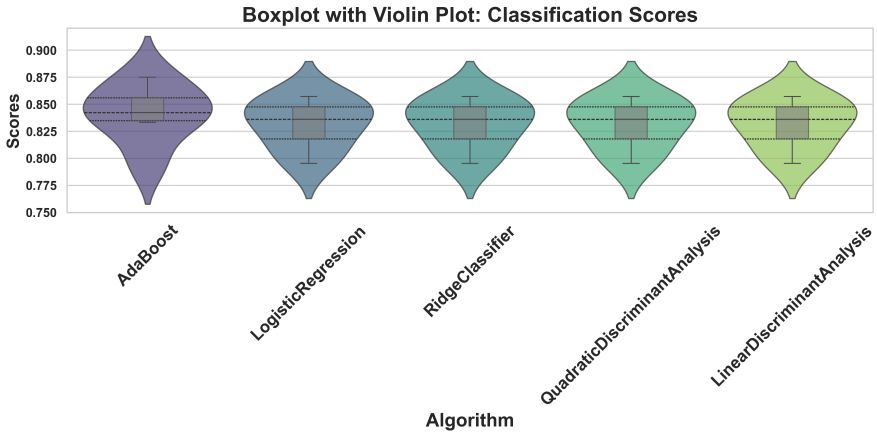

Fig. 8. Comparison of Classification Accuracy for Machine Learning Models

Figure 9 shows a comparative analysis of accuracy across several models. The horizontal bar chart visually ranks the performance of the different classifiers. Based on the results shown, the AdaBoost model yielded the most favorable outcome. It clearly demonstrated the highest level of accuracy among all contenders evaluated. This superior performance is indicated by its bar extending furthest on the chart. This establishes AdaBoost as the most effective model within this specific comparison.

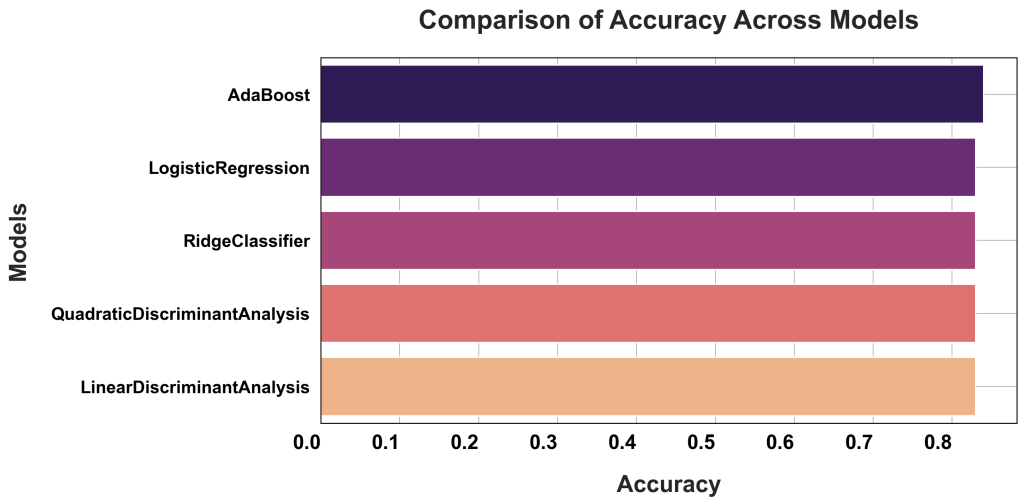

Fig. 9. Comparison of Model Accuracy

Figure 10 shows a heatmap comparing various classification models, where the AdaBoost model demonstrates the most effective overall performance. It achieves the highest score for sensitivity among all evaluated models, which is indicated by the brightest cell in the visualization. The model also displays a strong F1 score, reinforcing its robust predictive capability for positive cases. In contrast, the model's lowest-performing metric is specificity, as clearly represented by the darkest-colored cell. The Negative Predictive Value (NPV) is also shown to be comparatively lower. This performance profile highlights a distinct trade-off within the model's predictive capabilities.

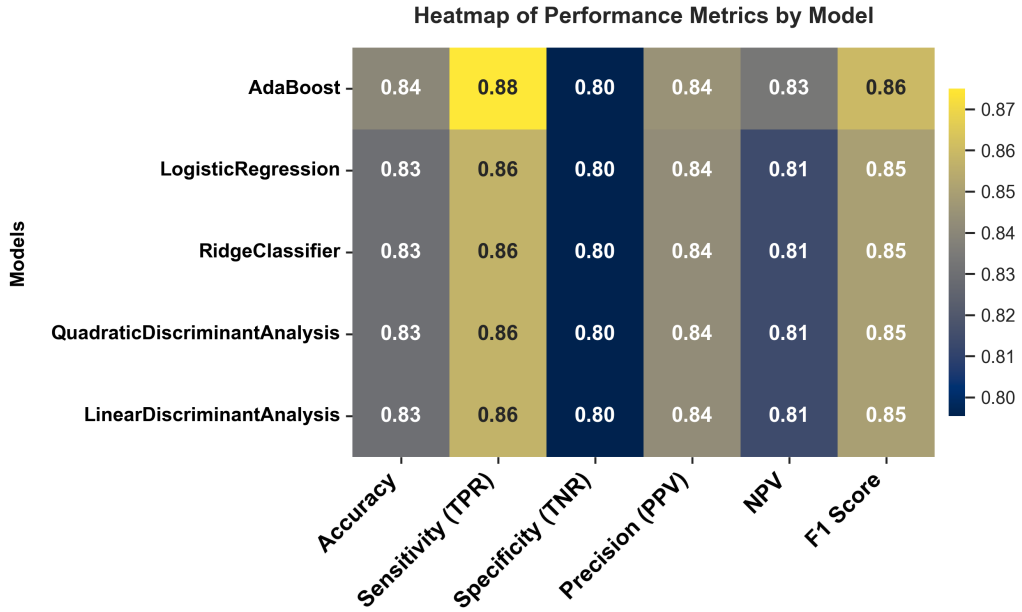

Fig. 10. Heatmap of Performance Metrics for Classification Models

Figure 11 shows the relationships between various classification performance metrics. The analysis includes accuracy, sensitivity, specificity, precision, NPV, and the F1 score. The results highlight a top-performing model that achieves superior overall performance. This model stands out by attaining the highest F1 score of 0.96, indicating excellent balance. It also demonstrates exceptional accuracy, reaching a value of 0.95 in its classifications. Furthermore, the model's sensitivity is remarkably high at 0.97, showing its effectiveness. The pairplot reveals strong positive correlations among accuracy, sensitivity, and F1 score. This indicates these metrics tend to increase in unison for the models evaluated. In contrast, specificity exhibits a very weak correlation with most other metrics shown.

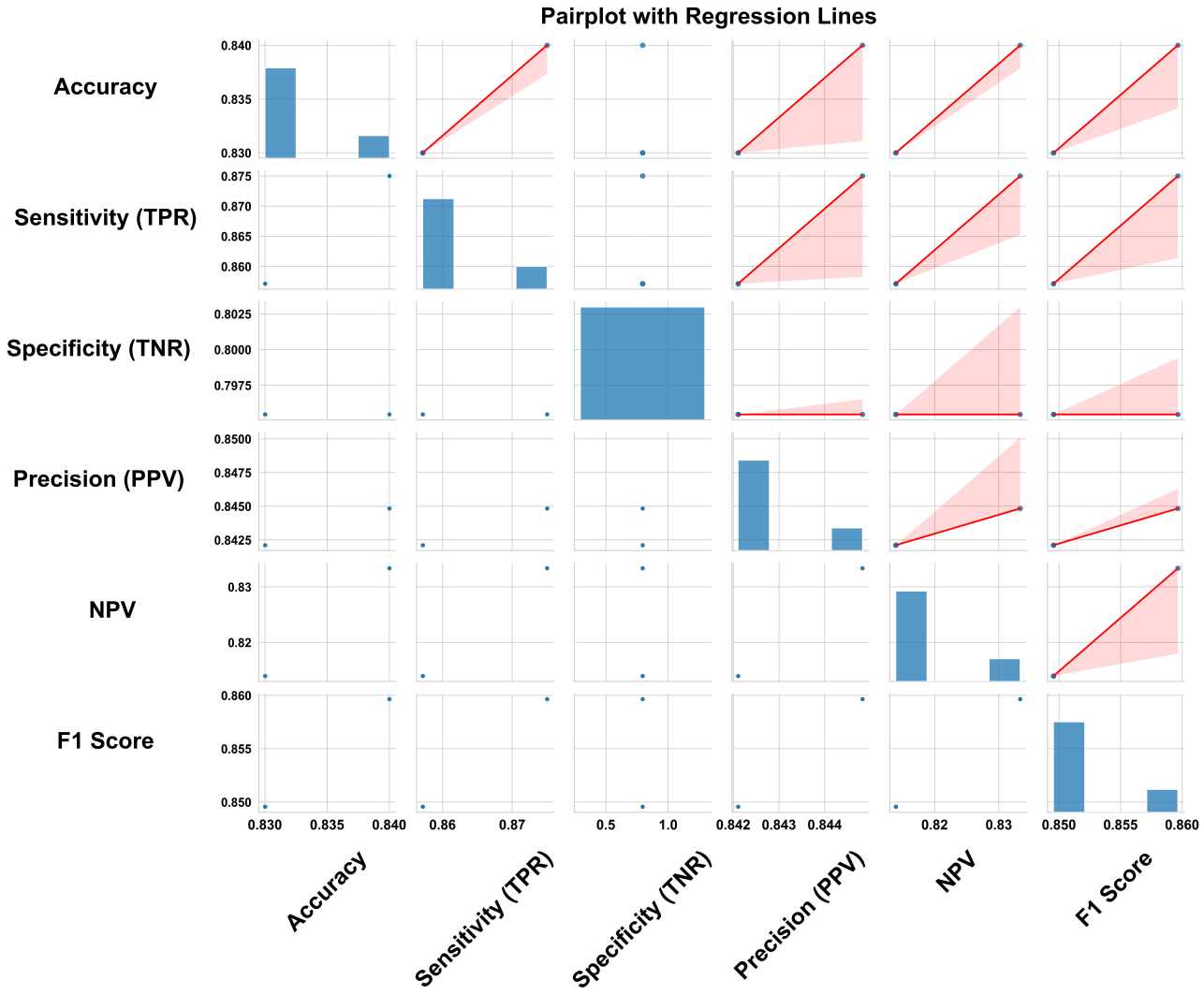

Fig. 11. Correlation Pairplot of Classification Performance Metrics

4.2 Statistics Analysis

Table **??** shows the statistical results. The analysis of variance was performed to evaluate differences between groups. The findings for the primary model tested revealed a highly significant effect. This indicates a meaningful variation exists between the groups under study. The resulting F-statistic was calculated to be (21.9130), with an associated p-value of less than (<0.0001).

TABLE III. ONE-WAY ANOVA TABLE FOR CLASSIFIER COMPARISON

Source	SS	DF	MS	F	P-value
Between Groups	0.0056	5	0.0011	21.9130 (F(5, 54))	<0.0001
Within Groups	0.0028	54	0.0001		
Total	0.0084	59			

Table **??** shows the Wilcoxon signed-rank test results. The analysis compares classifier performance against a theoretical median. All models demonstrated statistically significant results. However, the AdaBoost classifier emerged as the best-performing model. It achieved the highest actual median score among all classifiers.

This superior performance is highlighted by its median difference of (0.84). The statistical significance of this result is confirmed by a p-value of (0.001953). This indicates a significant improvement over the baseline. Therefore, AdaBoost is identified as the most effective classifier in this test.

TABLE IV. WILCOXON SIGNED-RANK TEST RESULTS FOR CLASSIFIER PERFORMANCE

	AdaBoost	LogisticRegression	RidgeClassifier	QuadraticDiscriminantAnalysis	LinearDiscriminantAnalysis
Theoretical Median	0	0	0	0	0
Actual Median	0.84	0.83	0.83	0.83	0.83
Sample Size	10	10	10	10	10
Sum of Signed Ranks (W)	0	0	0	0	0
Sum of Positive Ranks	55	55	55	55	55
Sum of Negative Ranks	0	0	0	0	0
P-Value (Two-Tailed)	0.002	0.002	0.002	0.002	0.002
Test Type	Exact	Exact	Exact	Exact	Exact
Significance Marker	**	**	**	**	**
Statistically Significant (α=0.05)	Yes	Yes	Yes	Yes	Yes
Median Difference	0.84	0.83	0.83	0.83	0.83

Figure 12 shows a histogram comparing model accuracies. The chart visually contrasts the predictive performance of several algorithms. While most models cluster together, the AdaBoost model stands out distinctly. Its performance is represented by a single, separate bar located to the far right. This placement on the horizontal axis signifies its superior predictive power. The model consistently achieved a significantly higher accuracy score in the analysis. It is clearly segregated from the main group of lower-performing models. This distinct separation from the other results highlights its effectiveness. The visualization clearly identifies AdaBoost as the superior-performing model.

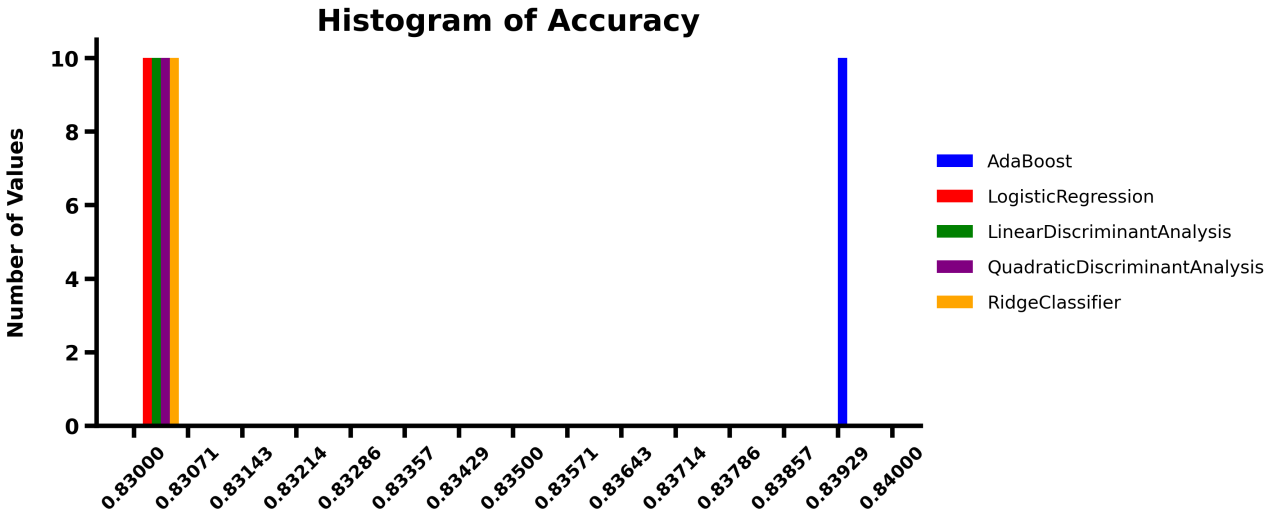

Fig. 12. Comparative Distribution of Model Accuracy

Figure 13 shows the accuracy performance of several classification algorithms. The results clearly indicate that the AdaBoost model achieved a significantly superior performance. It consistently yielded the highest accuracy, culminating in a peak score of 84%. This outstanding performance is situated at the top of the chart, distinctly separate from all competitors. This demonstrates its standout effectiveness compared to the other algorithms evaluated. Those competing models are clustered together at a visibly lower performance level. The visual evidence thus confirms AdaBoost as the unequivocally best-performing algorithm. Its high predictive power makes it the optimal choice for this classification task.

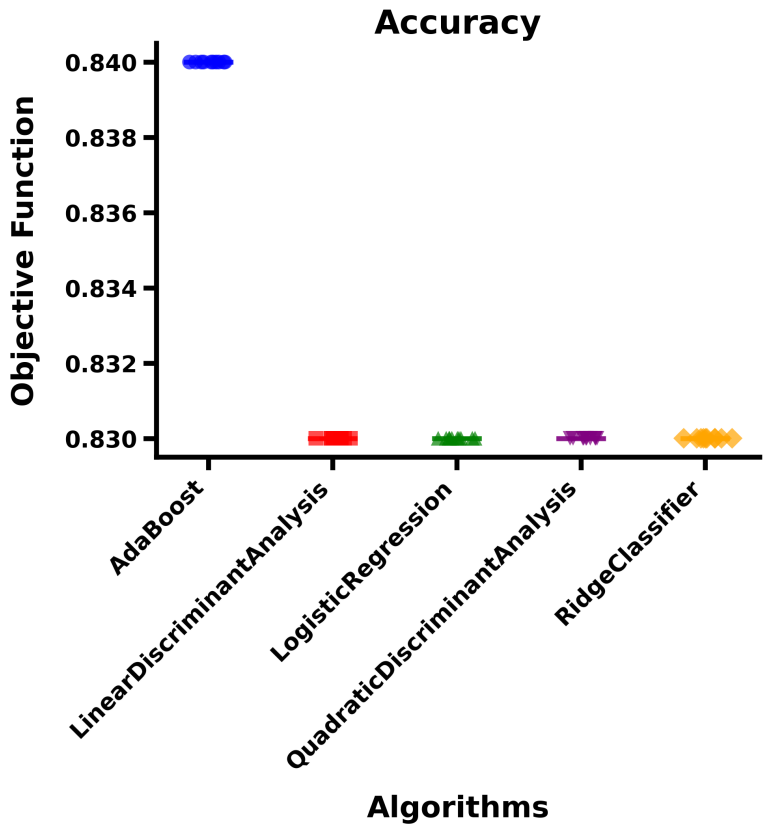

Fig. 13. Classifier Accuracy Comparison

5. CONCLUSION AND FUTURE WORK

5.1 Key Findings and Performance

This study effort was able to prove that AdaBoost is the best machine learning algorithm when comparing the results of predicting bone sarcoma outcome in an extreme comparison management using a decade-long set of data available in Memorial Sloan Kettering Cancer Center. The results based on AdaBoost with high accuracy of 84%, sensitivity of 87.5 and F1 score of 0.8596 demonstrated going far beyond the standard methods. These findings received confirmation by statistical validation (F-statistic = 21.9130, $p < 0.0001$) depicting the effectiveness and clinical applicability of ensemble learning in the field of orthopedic oncology.

The adaptive character of the algorithm was especially useful in regard to the complexity of medical datasets and imbalanced classes. That pattern in patient characteristics, tumor biology and response to treatment can be ignored using traditional approaches, but were well captured by its sequential learning approach. The strengths of 87.5 percent sensitivity meet the urgent clinical demand to characterize accurately the high-risk clients allowing prompt intervention and individual treatment plans.

5.2 Clinical Implications

In clinical terms, the model developed has a major potential to be used in everyday practice. The capacity to effectively stratify patients into definite categories of outcomes offers clinicians effective decisions tools to make therapeutic plan, patient counseling as well as resource-sharing. The precision of the model is 84.48 percent, which is as low as possible in terms of false positive results, and it still achieves impressive sensitivity levels, which benefits the quality of care and the efficiency of healthcare delivery.

Generalizability and reliability are achieved through the methodological rigor of the study, where each and every cross-validation and statistical validation framework is utilized. A high level of consistency on the performance between

validation iterations indicates that the AdaBoost approach is stable and practical to be used in the clinic.

5.3 Limitations and Future Directions

The implications are not only limited to bone sarcoma but oncological machine learning in general. The proven efficacy of the ensemble techniques implies that analogous methods might help in other rare cancer forms where the limited size of the dataset and the uneven distribution of the classes also are rather demanding. The presented methodological framework offers a model of strict comparison of the algorithms in medical implementations.

Drawing a conclusion, this work is a great step towards the machine learning application in predicting prognosis of bone sarcoma as it was determined that AdaBoost is the best algorithm to use. This work has significant potential value to precision medicine in orthopedic oncology because of its high performance characteristics, strong statistical validation, and obvious clinical utility. The created predictive model has a potential to be implemented in clinical practice right away, as well as give rise to further studies in this significant field of cancer care.

5.4 Broader Impact

The implication is not limited to bone sarcoma but oncological machine learning as a whole. The proven efficacy of the ensemble techniques implies that analogous methods might help in other rare cancer forms where the limited size of the dataset and the uneven distribution of the classes also are rather demanding. The presented methodological framework offers a model of strict comparison of the algorithms in medical implementations.

Drawing a conclusion, this work is a great step towards the machine learning application in predicting prognosis of bone sarcoma as it was determined that AdaBoost is the best algorithm to use. This work has significant potential value to precision medicine in orthopedic oncology because of its high performance characteristics, strong statistical validation, and obvious clinical utility. The created predictive model has a potential to be implemented in clinical practice right away, as well as give rise to further studies in this significant field of cancer care.

Conflicts Of Interest
The authors declare no conflicts of interest.

Funding
No Funding

Acknowledgment
Not Applicable

REFERENCES

[1] P. S. Papageorgiou, R. Christodoulou, P. Korfiatis, D. P. Papagelopoulos, O. Papakonstantinou, N. Pham, A. Woodward, and P. J. Papagelopoulos, “Artificial intelligence in primary malignant bone tumor imaging: A narrative review.” *Diagnostics (Basel, Switzerland)*, 2025.

[2] Z. Wang, Z. Wang, R. Wang, Z. Wang, X. Cao, R. Chen, Z. Ma, S. Liang, and S. Tao, “Whole-exome evolutionary profiling of osteosarcoma uncovers metastasis-related driver mutations and generates an independently validated predictive classifier.” *Journal of translational medicine*, 2025.

[3] S. Li, J. Chen, F. He, M. Wang, J. Liu, and H. Xie, “Multidimensional bioinformatics analysis of chondrosarcoma subtypes and tgf- signaling networks using big data approaches.” *Discover oncology*, 2025.

[4] C.-X. Lu, Z.-X. Long, J.-L. Lu, C.-K. Huang, T. Nakamura, S.-W. Tao, S.-L. Hua, and D.-L. Fang, “No title,” *Translational pediatrics*, 2025.

[5] B. Goker, A. Brook, R. Zhang, B. Aasman, J. Wang, A. Ferrena, P. Mirhaji, R. Yang, B. H. Hoang, and D. S. Geller, “Development of machine learning models to predict tumor endoprosthesis survival.” *Journal of surgical oncology*, 2025.

[6] A. Lemore, N. Vogt, J. Oster, E. Germain, M. Fauvel, R. Gillet, F. Sirveaux, B. Marie, N. Sans, M. Faruch, F. Lapègue, F. Lafourcade, S. Badr, A. Cotten, F. M. Bouvier, S. Yang, J.-L. Drapé, M. Pastor, Y. Thouvenin, M. P.

Baron, C. Cyteval, D. Fadli, C. Fournier, O. Hauger, M. B. H. Amor, N. Stacoffe, S. Daubie, J.-B. Noel, J.-B. Pialat, S. Cherix, F. Zanchi, P. Omoumi, A. Blum, G. Hossu, and P. A. G. Teixeira, "Enhanced ct and mri focal bone tumor classification with machine learning-based stratification: A multicenter retrospective study." *Radiology*, 2025.

[7] W. Zhao, H. Meng, Z. Dai, L. Zhang, Z. Cheng, Y. Song, W. Xu, Z. Wang, K. Tian, Y. Jiang, W. Sun, Z. Cai, G. Wang, and Y. Hua, "Prediction of patients with high-risk osteosarcoma on the basis of xgboost algorithm using transcriptome and methylation data from sgh-os cohort." *JCO precision oncology*, 2025.

[8] A. Curto-Vilalta, B. Schlossmacher, C. Valle, A. Gersing, J. Neumann, R. von Eisenhart-Rothe, D. Rueckert, and F. Hinterwimmer, "Semi-supervised label generation for 3d multi-modal mri bone tumor segmentation." *Journal of imaging informatics in medicine*, 2025.

[9] K. Wang, Y. Han, Y. Ye, Y. Chen, D. Zhu, Y. Huang, Y. Huang, Y. Chen, J. Shi, B. Ding, and J. Huang, "Mixed reality infrastructure based on deep learning medical image segmentation and 3d visualization for bone tumors using dcu-net." *Journal of bone oncology*, 2025.

[10] S. Yao, Y. Huang, X. Wang, Y. Zhang, I. C. Paixao, Z. Wang, C. L. Chai, H. Wang, D. Lu, G. I. Webb, S. Li, Y. Guo, Q. Chen, and J. Song, "A radiograph dataset for the classification, localization, and segmentation of primary bone tumors." *Scientific data*, 2025.

[11] C. Liu, Q. Zou, M. Wang, Q. Yang, L. Song, Z. Lu, Q. Feng, and Y. Zhao, "[identification of osteoid and chondroid matrix mineralization in primary bone tumors using a deep learning fusion model based on ct and clinical features: a multi-center retrospective study]." *Nan fang yi ke da xue xue bao = Journal of Southern Medical University*, 2024.

[12] S. S. Dalai, B. J. R. Sahu, J. Rautaray, M. I. Khan, B. A. Jabr, and Y. A. Ali, "Automated bone cancer detection using deep learning on x-ray images." *Surgical innovation*, 2024.

[13] S. Deng, Y. Huang, C. Li, J. Qian, and X. Wang, "Auxiliary diagnosis of primary bone tumors based on machine learning model." *Journal of bone oncology*, 2024.

[14] F. Shang, S. Tang, X. Wan, Y. Li, and L. Wang, "Bmsmm-net: A bone metastasis segmentation framework based on mamba and multiperspective extraction." *Academic radiology*, 2024.

[15] H. Ikushima, K. Watanabe, A. Shinozaki-Ushiku, K. Oda, and H. Kage, "A machine learning-based analysis of nationwide cancer comprehensive genomic profiling data across cancer types to identify features associated with recommendation of genome-matched therapy." *ESMO open*, 2024.

[16] L. Song, C. Li, L. Tan, M. Wang, X. Chen, Q. Ye, S. Li, R. Zhang, Q. Zeng, Z. Xie, W. Yang, and Y. Zhao, "A deep learning model to enhance the classification of primary bone tumors based on incomplete multimodal images in x-ray, ct, and mri." *Cancer imaging : the official publication of the International Cancer Imaging Society*, 2024.

[17] X. Chen, H. Chen, J. Wan, J. Li, and F. Wei, "An enhanced alexnet-based model for femoral bone tumor classification and diagnosis using magnetic resonance imaging." *Journal of bone oncology*, 2024.

[18] C. Guo, Y. Chen, and J. Li, "Radiographic imaging and diagnosis of spinal bone tumors: Alexnet and resnet for the classification of tumor malignancy." *Journal of bone oncology*, 2024.

[19] J. Deng, M. Moskalyk, M. Shammas-Toma, A. Aoude, M. Ghert, S. Bhatnagar, and A. Bozzo, "Development of machine learning models for predicting the 1-year risk of reoperation after lower limb oncological resection and endoprosthetic reconstruction based on data from the parity trial." *Journal of surgical oncology*, 2024.

[20] F. Erdem, S. Gitto, S. Fusco, M. V. Bausano, F. Serpi, D. Albano, C. Messina, and L. M. Sconfienza, "Automated detection of bone lesions using ct and mri: a systematic review." *La Radiologia medica*, 2024.

[21] N. Li, E. Peng, and F. Liu, "Prediction of lymph node metastasis in cervical cancer patients using adaboost machine learning model: analysis of risk factors." *American journal of cancer research*, 2025.

[22] Y. B. Abushark, S. Hassan, and A. I. Khan, "Optimized adaboost support vector machine-based encryption for securing iot-cloud healthcare data." *Sensors (Basel, Switzerland)*, 2025.

[23] B. S. Gharehhasani, M. Rezaei, A. Naghipour, N. Sayad, S. Mostafaei, and E. Alimohammadi, "The most important variables associated with death due to covid-19 disease, based on three data mining models decision tree, adaboost, and support vector machine: A cross-sectional study." *Health science reports*, 2024.

[24] M. Irfan, S. Shahrestani, and M. Elkhodr, "Machine learning in neurological disorders: A multivariate lstm and adaboost approach to alzheimer's disease time series analysis." *Health care science*, 2024.

[25] Y. Zhou, Z. Zhang, Q. Li, G. Mao, and Z. Zhou, "Construction and validation of machinelearningalgorithm for predicting depression among home-quarantined individuals during the large-scale covid-19 outbreak: based on adaboost model." *BMC psychology*, 2024.

[26] M. T. Ramakrishna, V. K. Venkatesan, I. Izonin, M. Havryliuk, and C. R. Bhat, "Homogeneous adaboost ensemble machine learning algorithms with reduced entropy on balanced data." *Entropy (Basel, Switzerland)*, 2023.

[27] S. Keshvari, S. A. Farizhendi, M. M. Ghiasi, and A. H. Mohammadi, "Adaboost metalearning methodology for modeling the incipient dissociation conditions of clathrate hydrates." *ACS omega*, 2021.

[28] Z. Hu, Y. Yuan, X. Li, Z. Tu, O. D. Dacres, Y. Zhu, L. Shi, H. Hu, H. Liu, G. Luo, and H. Yao, "Yield prediction

of "thermal-dissolution based carbon enrichment" treatment on biomass wastes through coupled model of artificial neural network and adaboost." *Bioresource technology*, 2021.

[29] J. Wang, W. Xue, X. Shi, Y. Xu, and C. Dong, "Adaboost-based machine learning improved the modeling robust and estimation accuracy of pear leaf nitrogen concentration by in-field vis-nir spectroscopy." *Sensors (Basel, Switzerland)*, 2021.

[30] J. Tang, A. Henderson, and P. Gardner, "Exploring adaboost and random forests machine learning approaches for infrared pathology on unbalanced data sets." *The Analyst*, 2021.