Artvin Çoruh Üniversitesi
Orman Fakültesi Dergisi
ISSN:2146-1880, e-ISSN: 2146-698X
Yıl: 2026, Cilt: 27, Sayı: 1, Sayfa:66-80

Artvin Coruh University
Journal of Forestry Faculty
ISSN:2146-1880, e-ISSN: 2146-698X
Year: 2026, Vol: 27, Issue: 1, Pages:66-80

ofd.artvin.edu.tr

Classification of Strength Properties of Commercially Important Wood Types Grown in the United States by Machine Learning

Amerika Birleşik Devletleri'nde Yetişen Ticari Öneme Sahip Odun Türlerinin Dayanım Özelliklerinin Makine Öğrenmesi ile Sınıflandırılması

Kenan KILIÇ*[1]
[1]Yozgat Bozok University, Yozgat Vocational School, Department of Design, Interior Design Program, Yozgat

Eser bilgisi/*Article info*
Araştırma makalesi / *Research article*
DOI: 10.17474/artvinofd.1717771
*Sorumlu yazar/*Corresponding author*
Kenan KILIÇ
e-mail: kenan.kilic@bozok.edu.tr
Gönderilme tarihi /*Submission date*
11.06.2025
Kabul tarihi /*Acceptance date*
25.08.2025
Yayımlanma tarihi/ *Publication date*
24.02.2026

Keywords:
Machine learning
Wood classification
Wood properties
Hardwood
Softwood
Mechanical properties

Anahtar kelimeler:
Makine öğrenme
Odun sınıflandırma
Odun özellikleri
Geniş yapraklı ağaç odunu
İğne yapraklı ağaç odunu
Mekanik özellikler

Abstract
Commercially important wood species grown in the United States are divided into hardwood and softwood based on their mechanical properties. Comparative analysis is conducted by optimising six different machine learning algorithms: SVM, XGBoost, Random Forest, Logistic Regression, KNN, and Decision Tree. Preliminary processes such as completing missing data, coding categorical data, and standardisation are applied to the dataset to make it suitable for machine learning algorithms. Experiments were conducted using the stratified 10-fold cross-validation method. Hyperparameter optimisation was performed with GridSearchCV. The SVM algorithm provides the best accuracy with 96.90%. This model is followed by XGBoost with 95.13% accuracy and an AUC of 0.9891, followed by Random Forest with 94.25% accuracy. Logistic Regression performs with 90.27% accuracy, Decision Tree with 90.71% accuracy, and KNN with 88.05% accuracy. Results show that kernel-based (SVM) and ensemble-based (XGBoost, RF) models provide higher classification performance than linear and instance-based models. These models have the potential to improve wood quality control processes, increase resource efficiency, and contribute to sustainable forestry practices.

Öz
Amerika Birleşik Devletleri'nde yetiştirilmekte olan ticari öneme sahip ağaç türlerinin mekanik özelliklerine dayanarak geniş yapraklı ağaç (hardwood) ve iğne yapraklı ağaç (softwood) odunlarının ayrımı yapılmaktadır. Altı farklı makine öğrenmesi algoritması; SVM, XGBoost, Random Forest, Logistic Regression, KNN ve Decision Tree optimize edilerek karşılaştırmalı analiz gerçekleştirilmektedir. Veri seti üzerinde eksik verilerin tamamlanması, kategorik verilerin kodlanması ve standartlaştırma gibi ön işlemler uygulanarak veriler makine öğrenme algoritmaları için uygun hale getirilmektedir. Araştırmada, Stratified 10-Fold Cross Validation yöntemi kullanılarak deneyler gerçekleştirilmiştir. Hiperparametre optimizasyonu GridSearchCV ile gerçekleştirilmiştir. Doğruluk açısından en iyi sonucu %96.90 ile SVM algoritması vermektedir. Bu modeli %95.13 doğruluk ve 0.9891 AUC değeri ile XGBoost, ardından %94.25 doğrulukla Random Forest takip etmiştir. Logistic Regression %90.27, Decision Tree %90.71 ve KNN %88.05 daha düşük doğrulukla performans göstermektedir. Sonuçlar, kernel tabanlı (SVM) ve topluluk tabanlı (XGBoost, RF) modellerin doğrusal ve örnek tabanlı modellere göre daha yüksek sınıflandırma başarımı sunduğunu göstermektedir. Bu modeller, odun kalite kontrol süreçlerini iyileştirme, kaynak verimliliğini artırma ve sürdürülebilir ormancılık uygulamalarına katkı sağlama potansiyeli taşımaktadır.

INTRODUCTION

Wood is one of the oldest construction and production materials in human history. For over five thousand years, from the Bronze Age to the Industrial Revolution, it was widely used in almost every society as a fuel, shelter, and production tool, and continues to be used today (Perlin 2005). The primary reason for this prevalence is that wood is mechanically durable, has a diverse density, and is malleable. Indeed, even ancient civilizations possessed the knowledge to distinguish wood's technical properties, such as hardness, strength, and density, according to its intended use (Bartolucci et al. 2020).

Based on their botanical characteristics, wood species are generally divided into two main categories: hardwoods (angiosperms) and softwoods (gymnosperms). Hardwoods are derived from broad-leaved plants with seeds enclosed in fruit, while softwoods include conifers, conifers, and fast-growing species. This classification is not only a biological distinction but also a fundamental criterion for determining industrial uses. Hardwoods are preferred for furniture and decoration due to their density and durability, while softwoods, with their lightness and ease of workability, are widely used in the construction and packaging sectors (Bozkurt 1971, Osamah 2016).

The historical development of wood in the construction industry is also noteworthy. Despite its ease of access and workability, it was replaced by alternative materials such as stone, steel, and concrete with the Industrial Revolution (Batur 2004, Bilici 2006). However, in recent years, with increasing sustainability concerns, interest in natural materials has revived. Its environmentally advantageous properties have made wood a material worthy of re-evaluation within contemporary construction technologies (Gündüz et al. 2009).

Recently, artificial intelligence and machine learning (ML) techniques have come to the fore for the classification and quality control of wood species based on their mechanical properties. The effects of structural factors such as tree species, growth ring width, and diameter at breast height on wood density are being modelled, thereby improving classification accuracy (Ramananantoandro et al. 2016). New regression models developed to predict wood density yield successful results with lower standard errors (Vieilledent et al. 2018). Furthermore, software such as CarDen provides practical advantages in field studies by automating density measurements in central regions (Jacquin et al. 2019). Studies in mangrove forests have also revealed relationships between salinity levels and density variation among species (Virgulino-Júnior et al. 2020). In addition, different approaches such as puncture resistance measurements (Martínez et al. 2020), experimental density analyses on yellow birch (Jones et al. 2023), and topological convolutional neural networks (Kantavichai and Turnblom 2022) offer significant contributions to the estimation of wood density and mechanical properties. Artificial neural networks (ANN) are attracting attention as a powerful and flexible method for quality estimation (Monteiro et al. 2023).

Today, machine learning-based wood classification studies are moving beyond basic physical properties such as density to comprehensive models that also include multidimensional mechanical parameters such as modulus of elasticity (MOE), flexural strength (MOR), and microscopic structures. For example, the Multi-Input Residual Neural Network (MIRN) model developed by Ma and colleagues combined density data with microscopic image analysis, providing lower error rates in estimating MOE and MOR compared to traditional CNN and ResNet-50-based models (Ma et al. 2025). Another study reported that deep learning models performed comparably to human experts in classification based on microscopic images (Nieradzik et al. 2024). There are also studies where powerful classical machine learning algorithms, such as XGBoost, classify wood species based on mechanical properties with high accuracy. A study conducted in Turkey demonstrated that automated machine learning (AutoML) systems were integrated into classification processes in forestry and yielded successful results in decision support systems (Eker et al. 2023). This current literature clearly demonstrates that both deep learning and advanced classical machine learning approaches provide reliable, accurate, and applicable solutions for classification based on the mechanical properties of wood.

In light of this, a review of the existing literature reveals that comprehensive studies focusing on multi-model comparisons based on mechanical properties, particularly across a wide range of commercially important hardwood and softwood species in the United States, are quite limited. Most studies focus solely on a single algorithm and fail to provide large-scale analyses of model performance and classification stability. Therefore, by comparatively evaluating six different machine learning algorithms, including Random Forest, XGBoost, Logistic Regression, SVM, KNN, and Decision Tree, this study contributes to decision support processes for commercial wood classification and offers a holistic approach to improving machine learning-based classification accuracy. In this respect, the study aims to fill a significant gap in the forest products industry and sustainable natural resource management.

MATERIAL AND METHODS

Data Set

A dataset containing the physical and mechanical properties of various tree species growing in the United States was used. The dataset, obtained from the Kaggle platform, contains a total of 226 samples and 16 features. Each sample represents the characteristics of a specific tree species.

The dataset used was generated based on information from the "Wood Handbook," a technical report prepared by Kretschmann and published by the Forest Products Laboratory of the U.S. Department of Agriculture (USDA). This resource is a comprehensive reference detailing the mechanical properties of wood materials used in engineering. It contributes to a more accurate understanding of wood materials used in the building and construction sectors (Kretschmann 2010).

The dataset includes categorical variables such as the common name of the species, genus, species, scientific name and classification hardwood or softwood, moisture content (Green or 12%), as well as numerical properties such as density, bending strength, modulus of elasticity, work done at maximum load, impact strength, parallel and perpendicular compressive strengths, shear and tensile strength, and surface hardness.

In terms of class distribution, the data set included 132 hardwood (58.4%) and 94 softwood (41.6%) samples. Half of the samples were measured at green moisture content, while the other half were measured at 12% moisture. Missing data were identified in some variables during the missing data analysis; for example, there was a 25.66% missing data rate for the "tension perpendicular to grain" variable, a 12.83% missing data rate for the "side hardness" variable, and an 11.94% missing data rate for the "impact bending" variable. Missing data rates for other numerical variables generally ranged from 1% to 3%. There were no missing observations for categorical variables. Statistical methods using mean or median values were preferred for processing missing data.

When the statistical distribution of numerical variables was examined, the average specific gravity was calculated as 0.488, the average modulus of rupture as 9712 psi, and the average modulus of elasticity as 1.39 million psi. These values provide an important basis for understanding the differences in strength between species. The data set contains the strength characteristics of some commercially important wood species growing in the United States. Each attribute and description of the data set is given in Table 1.

Table 1. Characteristics of some commercial woods grown in the USA

Features	Description
Common species name	Common name of the wood species (e.g., "Sugar Maple")
Genus	Type of wood species (e.g., "Acer")
Species	Scientific name of the wood species (e.g., "saccharum")
Classification	Classification of the wood species ("Hardwood" or "Softwood")
Moisture content	Moisture content of the wood material ("Green" or "12%)
Specific gravity	Specific gravity of the wood material
Modulus of rupture	Modulus of rupture, the amount of stress that a beam can withstand under a load from the centre before it breaks
Modulus of elasticity	Modulus of elasticity, the resistance of the material to deformation
Work to the maximum load	The amount of work that the wood material does before it deforms under maximum load
Impact bending	Impact resistance, measured at the height of a blow with a 0.71 kg (50 lb) hammer
Compression parallel to the grain	Compressive strength parallel to the grain direction, maximum crushing strength
Compression perpendicular_to_grain	Compressive strength perpendicular to the grain direction, proportional limit fibre stress
Shear parallel to the grain	Shear strength parallel to the grain direction, resistance to shear
Tension perpendicular to the grain	Tensile strength perpendicular to the grain direction, maximum tensile strength
Side hardness	Lateral stiffness, measured when the load is applied perpendicular to the grain direction, Janka Hardness Test

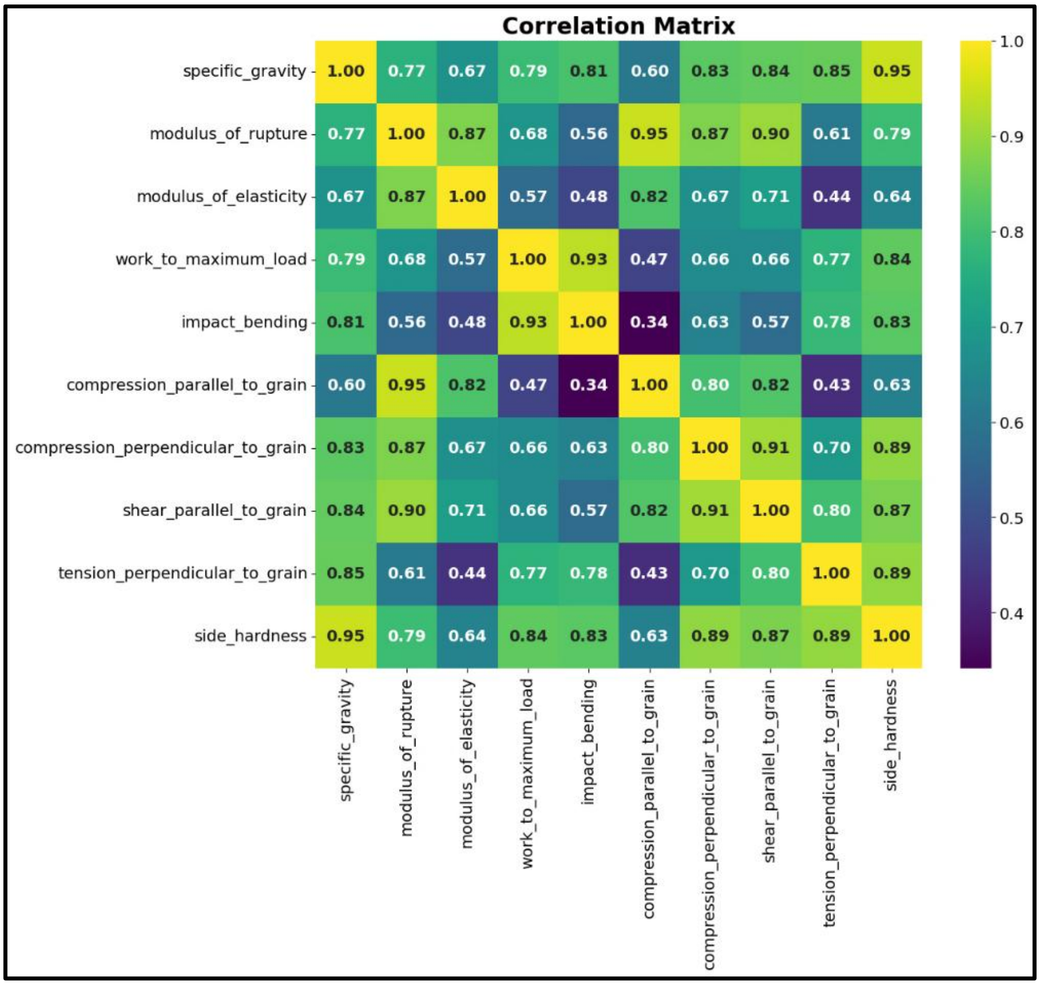

Figure 1. Correlation matrix of the features found in the data set

Data Preprocessing

The dataset used in this study underwent various preprocessing steps to improve the performance of the classification algorithms. First, missing values in the dataset were examined in detail. Missing data in numerical variables were completed using the arithmetic mean, while missing data in categorical variables were completed using the most frequently occurring value (mode). This ensured that data integrity was maintained and the number of samples was not reduced.

The target variable in the dataset is the "classification" column, which classifies tree species as hardwood or softwood. This variable was digitized using the Label Encoding method to facilitate processing by machine learning algorithms; Hardwood was coded as 0 and Softwood as 1. Categorical attributes among the independent variables (e.g., common species name, genus, species, scientific name, moisture content) were digitized using the Ordinal Encoding method. The target variable was not included in this transformation. All numerical features measured at different scales were standardised using StandardScaler. This standardisation process eliminated scale differences between variables during model training, enabling the algorithm to learn more balancedly and reliably. The primary goal of standardisation is to bring the mean of each variable to zero and its standard deviation to one. After the pre-processing steps were completed, the modelling process was started, and it was observed that these steps had a positive effect on the classification performance.

Machine Learning-Based Classification Algorithms

Random Forest

It is an ensemble learning method that combines decision trees. The advantages of this method include its resistance to overfitting, its ability to work well with incomplete data, and its high accuracy. It also has a strong generalisation capacity on datasets containing a large number of features. However, the model is difficult to interpret, and computational costs may increase in large datasets. In addition, there are high memory and processing power requirements (Breiman 2001).

Logistic Regression

Logistic regression is a simple and fast algorithm used for binary classification. The outputs of the model can be easily interpreted, but its accuracy may be low in nonlinear relationships and complex data sets. In addition, since it is based only on linear relationships, its generalisation ability is limited (Hosmer et al. 2013).

Extreme Gradient Boosting Classifier (XGB)

XGB is a fast and efficient classifier developed by optimising the Gradient Boosting algorithm. It exhibits high performance, especially in large data sets and complex problems. It offers advantages such as fast training time and low memory usage. However, it may encounter difficulties such as the risk of overfitting and complex parameter settings (Chen and Guestrin 2016).

Support Vector Machine

SVMs stand out as powerful and effective classification algorithms among supervised learning methods. The version used for classification, the Support Vector Classifier (SVC), attempts to find optimal boundaries that provide maximum margins between different classes. While fundamentally based on the principle of linear separability, it can also yield successful results on nonlinear data structures with the help of kernel functions. SVC demonstrates particularly high performance in high-dimensional datasets and situations where class separation is difficult. The algorithm can be flexibly customised with different kernel types, such as linear, polynomial, radial basis function (RBF), and sigmoid (Cortes and Vapnik 1995).

K Neighbours Classifier (k-NN)

The K-nearest neighbour (KNN) algorithm is an easy-to-understand and simple-to-implement classification method. The training process is fast because the basic calculations are performed in the testing phase. It performs well on small datasets. However, processing time increases as the dataset grows, and accuracy may decrease on data with high-dimensional features due to the "dimensionality problem" (Cover and Hart 1967).

Decision Tree

Decision trees offer models that are easy to interpret and visualise. They can work effectively with both numerical and categorical data types. Their rapid training process and simple structure are their key advantages. However, their tendency to overfit and the model's sensitivity to small changes in the dataset are their major drawbacks (Quinlan 1986).

Hyperparameter Optimisation

Hyperparameter optimisation was applied to improve model performance. The optimal parameters for each classification algorithm were determined using the GridSearchCV method. Grid Search allows selecting the settings that yield the best results by trying all combinations of predefined hyperparameter values. Accuracy was used as the performance metric.

The hyperparameters and their ranges used in the optimisation process are as follows:

- In the Random Forest algorithm, the number of trees (n estimators), maximum depth (max depth), minimum number of samples per leaf (min samples leaf), and maximum number of features (max features) were set.
- In XGBoost, the learning rate, tree depth (max depth), subsample, colsample bytree, and gamma values were optimised.
- For the Support Vector Machine (SVM), searches were conducted on the C parameter, kernel type, and gamma.
- In the Logistic Regression model: The regularisation coefficient (C), penalty function, solver algorithm, and maximum iteration number (max iter) were determined.
- In the K-Nearest Neighbour (KNN) algorithm, the number of neighbours (n neighbours), the weighting method (weights), and the distance metric were examined.
- In the Decision Tree model, the splitting criterion, maximum depth (max depth), minimum number of samples for splitting (min samples split), and minimum number of samples per leaf (min samples leaf) were set.

The optimisation process was performed using a 10-layer (k=10) k-fold cross-validation method that preserves the class distribution in the dataset. This tested the generalizability of the model across different data subsets and reduced the risk of overfitting.

Performance comparisons were made using the best hyperparameters determined by Grid Search for each model. This systematic and orderly approach ensured a fair evaluation of the algorithms' performance. The hyperparameter values of the classification algorithms used in the study are presented in Table 2.

Table 2. Hyperparameters and values of classification algorithms

Algorithms	Hyperparameters and Value Ranges
Random Forest	n estimators: 100, 200, 300; max depth: None, 10, 20, 30; min samples split: 2, 4, 6; min samples leaf: 1, 2, 4; max features: sqrt, log2
Logistic Regression	C: [0.001, 0.01, 0.1, 1, 10, 100, 1000]; penalty: ['l2', 'none']; solver: ['lbfgs', 'saga', 'newton-cg']; max iter: [500, 1000, 2000]
XGBoost	n estimators: [100, 200, 300]; max depth: [3, 5, 7, 10]; learning rate: [0.01, 0.1, 0.2]; subsample: [0.7, 0.8, 1.0]; colsample bytree: [0.7, 0.8, 1.0]; gamma: [0, 0.1, 0.2]
SVM	C: [0.1, 1, 10, 100]; kernel: ['linear', 'rbf', 'poly']; gamma: ['scale', 'auto']
KNN	n neighbours: [3, 5, 7, 9, 11, 15]; weights: ['uniform', 'distance']; metric: ['euclidean', 'manhattan', 'minkowski']
Decision Tree	criterion: ['gini', 'entropy']; max depth: [None, 5, 10, 20, 30]; min samples split: [2, 5, 10]; min samples leaf: [1, 2, 4]

Experimental Setup

Various machine learning algorithms, such as Random Forest, Logistic Regression, XGBoost, SVM, KNN, and Decision Tree, were used to classify wood species. To ensure reliable and accurate evaluation of the models, the Stratified 10-Fold Cross Validation method, which preserves the class distribution in the dataset, was chosen. This method divides the dataset into 10 equal parts, each part serving as the test set, while the remaining parts are used for training. During the data preprocessing process, missing values in numerical columns were filled with the mean value of the relevant column, and missing categorical variables were filled with the most frequent value (mode).

Categorical variables were converted to numerical format using OrdinalEncoder, and the target variable was encoded with LabelEncoder to make it usable by the model. The numerical features used in the model were standardised using StandardScaler because they were measured at different scales. Thus, all variables were scaled to have a mean of zero and a standard deviation of one. The optimal hyperparameters for each model were determined using the GridSearchCV method, and 10-fold cross-validation was used in this process. After selecting the optimal parameters, the models were retrained, and their performance was examined in detail using the Stratified 10-Fold Cross Validation method.

The experiments were conducted in a cloud environment equipped with an Intel Xeon 2.00 GHz processor provided by the Kaggle platform. This environment has sufficient computational capacity for data processing and machine learning applications.

Using the GridSearchCV method applied in the research, the optimal hyperparameters for each classification algorithm were determined, and these parameters were optimised using 10-fold Stratified K-Fold cross-validation. The best parameter sets are presented in Table 3.

Table 3. Best hyperparameters obtained with GridSearchCV

Algorithms	Best Hyperparameter Values
Random Forest	n estimators=100, max depth=20, min samples split=5, min samples leaf=1
Logistic Regression	C=10, penalty='l2', solver='lbfgs', max_iter=1000
XGBoost	n estimators=200, max depth=5, learning rate=0.1, subsample=1.0, colsample bytree=1.0, gamma=0
SVM	C=10, kernel='rbf', gamma='scale'
KNN	n neighbors=7, weights='distance', metric='minkowski'
Decision Tree	criterion='gini', max depth=10, min samples split=5, min samples leaf=1

Evaluation Metrics

Accuracy

Chicco and Jurman (2020) stated that the accuracy metric is the ratio of correctly predicted data points to the total data points.

Precision

Powers (2020) stated that the precision score is the ratio of true positives to the total number of all samples predicted as positive.

Recall

As defined by Powers (2020), the sensitivity score is the ratio of true positives to the total of true positive samples.

F1 Score

According to Chicco and Jurman (2020), the F1 score is a metric that calculates the harmonic mean of precision and sensitivity. In all these metrics, the worst value is 0 and the best value is 1.

ROC-AUC

Receiver Operating Characteristic - Area Under the Curve is a widely used performance metric to evaluate the discriminatory power of classification models. The ROC curve visualises the relationship between the True Positive Rate (Sensitivity) and the False Positive Rate at different threshold values (Adegun and Viriri 2020).

The evaluation metrics used in the research are given in Table 4.

Table 4. Evaluation metrics and formulas

Metrics	Formuls
Accuracy	(TP + TN) / (TP + TN + FP + FN)
Precision	TP / (TP + FP)
Recall	TP / (TP + FN)
F1-Score	2 × (Precision×Recall) / (Precision+Recall)
AUC	$\int_0^1$ TPR(FPR) dFPR

RESULTS AND DISCUSSION

Six different machine learning algorithms—Random Forest, Logistic Regression, XGBoost, SVM, KNN, and Decision Tree—were used to classify hardwood and softwood using the strength properties of commercially important trees grown in the United States. The performance of each algorithm was analysed using precision, recall, F1-score, accuracy, and ROC-AUC metrics. The study presents the average classification report, average complexity matrices, and average ROC-AUC values obtained from 10 k-fold cross-validation. This significantly enhances the success of the Random Forest classification algorithm in capturing the varying behaviours of the data in each section and in facilitating real-world problems.

The average classification report of the Random Forest classification algorithm is presented in Table 5.

Table 5. Random Forest average classification report

	Precision	Recall	F1-Score	Support
Hardwood	0.9612	0.9394	0.9502	132
Softwood	0.9175	0.9468	0.9319	94
Accuracy			**0.9425**	226
Macro Avg	0.9394	0.9436	0.9491	226
Weighted Avg	0.9431	0.9425	0.9426	226

According to Table 5, the overall accuracy rate of the model is 94.25%. For the Hardwood class, precision was 96.12%, recall was 93.94%, and F1-score was 95.02%. For the Softwood class, precision was 91.75%, recall was 94.68%, and F1-score was 93.19%. When the model's average performance across classes (macro average) was examined, precision was 93.94%, recall was 94.36%, and F1-score was 94.91%. The weighted averages were found to be 94.31%, the recall was 94.25%, and the F1-score was 94.26%, respectively. These data demonstrate that the Random Forest algorithm successfully distinguished between both classes and demonstrated balanced performance without any significant discrimination between classes. The average classification performance of the Logistic Regression algorithm is summarised in Table 6.

Table 6. Logistic Regression average classification report

	Precision	Recall	F1-Score	Support
Hardwood	0.9167	0.9167	0.9167	132
Softwood	0.8830	0.8830	0.8830	94
Accuracy			**0.9027**	226
Macro Avg	0.8998	0.8998	0.8998	226
Weighted Avg	0.9027	0.9027	0.9027	226

An examination of Table 6 shows that the overall model accuracy rate was calculated as 90.27%. This result reflects the generalisation power of linear-based classification algorithms such as Logistic Regression. For the hardwood class, the precision, recall, and F1-score values were all at 91.67%, demonstrating that the model consistently recognised this class.

For the softwood class, the precision, recall, and F1-score values were 88.30%. The balanced results obtained in both classes indicate that there was no significant performance difference between classes and that the model consistently discriminated.

In terms of macro averages, the precision, recall, and F1-score values were each determined as 89.98%. This demonstrates that the average performance between classes is quite high and balanced. The weighted average was 90.27%, confirming the overall performance of the model when considering the distribution of classes in the dataset. The Logistic Regression algorithm's successful results at this level demonstrate that the classes in the dataset are linearly separable and that the model effectively achieves this separation. The average classification performance of the XGBoost classification algorithm is presented in Table 7.

Table 7. XGBoost average classification report

	Precision	Recall	F1-Score	Support
Hardwood	0.9549	0.9621	0.9585	132
Softwood	0.9462	0.9362	0.9412	94
Accuracy			**0.9513**	226
Macro Avg	0.9506	0.9491	0.9498	226
Weighted Avg	0.9513	0.9513	0.9513	226

An examination of Table 7 reveals that the XGBoost algorithm performed quite well on the test data. The model's overall accuracy was calculated as 95.13%, demonstrating high classification success. For the hardwood class, the 95.49% precision and 96.21% recall values achieved an F1-score of 95.85%. Similarly, for the softwood class, the 94.62% precision, 93.62% recall, and 94.12% F1-score were achieved. These findings demonstrate that the XGBoost model can effectively and balancedly distinguish both classes.

According to the macro-average values, the precision was calculated as 95.06%, the recall was 94.91%, and the F1-score was 94.98%. This demonstrates that there was no imbalance in the model's overall performance, even when each class was treated with equal importance. The weighted average was 95.13%, which is in full agreement with the overall accuracy rate. Although sample sizes vary across classes, the model's predictive power is unaffected and consistently delivers high accuracy. The average classification performance of the SVM algorithm is detailed in Table 8.

Table 8. SVM average classification report

	Precision	Recall	F1-Score	Support
Hardwood	0.9699	0.9773	0.9736	132
Softwood	0.9677	0.9574	0.9626	94
Accuracy			**0.9690**	226
Macro Avg	0.9688	0.9674	0.9681	226
Weighted Avg	0.9690	0.9690	0.9690	226

An examination of Table 8 reveals that the SVM classification algorithm exhibits very high performance on the two-class problem. The overall model accuracy rate is 96.90%, demonstrating that it can clearly distinguish between classes and provides an extremely good fit to the dataset.

Examining the dataset by class, for the Hardwood class, precision is 96.99%, recall is 97.73%, and F1-score is 97.36%. Similarly, for the Softwood class, precision is 96.77%, recall is 95.74%, and F1-score is 96.26%. The high precision and recall values in both classes indicate that the model minimises both false positive and false negative classifications and achieves balanced learning across classes.

The 96.81% F1-score obtained using the macro average indicates that the model performs consistently across both classes, regardless of the sample size. The weighted averages, at 96.90%, closely match the model's overall accuracy, demonstrating that imbalanced data distribution does not negatively impact model performance.

The SVM model's ability to identify nonlinear decision boundaries was instrumental in achieving this high success rate. Classification using the RBF (Radial Basis Function) kernel function, in particular, successfully distinguished complex patterns. Furthermore, optimising hyperparameters using the GridSearchCV method maximised the model's overall performance. The average classification results of the KNN classification algorithm are presented in Table 9.

Table 9. KNN average classification report

	Precision	Recall	F1-Score	Support
Hardwood	0.9565	0.8333	0.8907	132
Softwood	0.8018	0.9468	0.8638	94
Accuracy			**0.8805**	226
Macro Avg	0.8792	0.8901	0.8795	226
Weighted Avg	0.8922	0.8805	0.8814	226

An examination of Table 9 shows that the overall accuracy rate of the K-Nearest Neighbour (KNN) algorithm was calculated as 88.05%. This accuracy level demonstrates that the model achieved a basic level of classification success; however, it performed relatively poorly compared to other models.

When examining the classes separately, although a precision value of 95.65% was achieved for the Hardwood class, the recall rate remained at 83.33%, indicating that the model incorrectly classified some genuine Hardwood samples as Softwood. Conversely, 94.68% recall and 80.18% precision were achieved for the Softwood class. This suggests that the model identified samples in the Softwood class with higher accuracy, but tended to incorrectly classify some Hardwood samples as Softwood.

When macro averages and weighted averages are considered, F1-scores range between 88% and 89%. This indicates that the model experiences inconsistencies in class separation and is particularly sensitive to sample distribution and data structure.

This limited performance of the KNN algorithm can be explained by the reduced effectiveness of distance-based decision-making methods, especially in high-dimensional data structures. This method, which relies on distance calculations between data points, has difficulty distinguishing complex class boundaries, which reduces overall performance. The average classification performance of the Decision Tree classification algorithm is detailed in Table 10.

Table 10. Decision Tree average classification report

	Precision	Recall	F1-Score	Support
Hardwood	0.9051	0.9394	0.9219	132
Softwood	0.9101	0.8617	0.8852	94
Accuracy			**0.9071**	226
Macro Avg	0.9076	0.9005	0.9036	226
Weighted Avg	0.9072	0.9071	0.9067	226

When Table 10 is examined, the overall accuracy rate of the Decision Tree algorithm was calculated as 90.71%. This rate indicates that the model largely completed the classification task. The 90.51% precision and 93.94% recall values obtained for the Hardwood class indicate that this class was recognised by the model with high accuracy. In contrast, the 86.17%

recall value for the Softwood class indicates that the model incorrectly classified some Softwood samples. This indicates that more clearly defining the boundaries between classes could improve model performance.

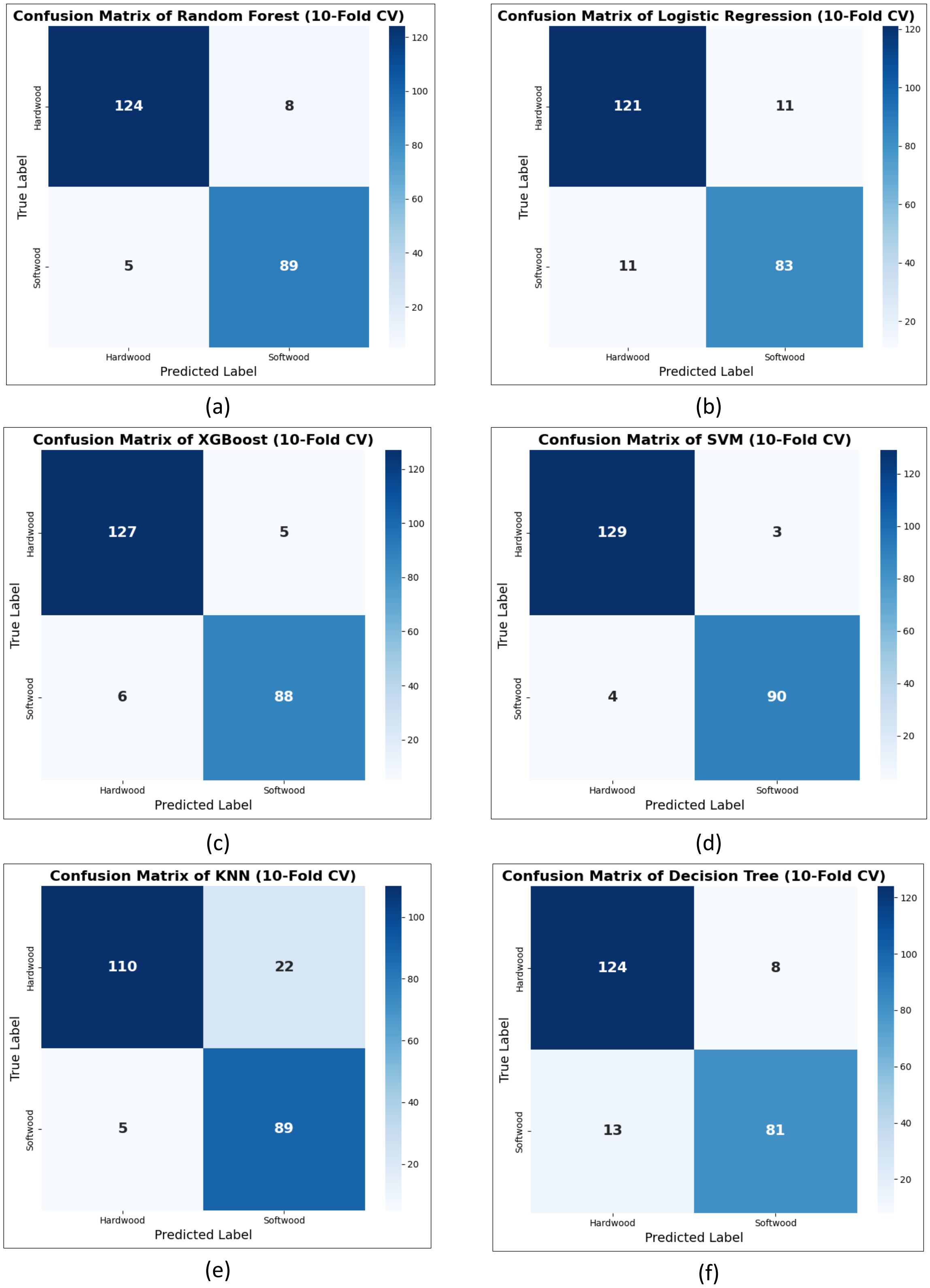

Figure 2. Confusion matrices of the classification algorithms obtained within the scope of the research: (a) Random Forest, (b) Logistic Regression, (c) XGBoost, (d) SVM, (e) KNN, (f) Decision Tree

The macro average F1-score was determined as 90.36%, and the weighted average F1-score was determined as 90.67%. These results indicate that the model generally exhibits balanced performance for both classes. However, improving prediction performance in the Softwood class could further increase the model's overall performance.

The complexity matrices of all classification algorithms obtained within the scope of the research are visually presented in Figure 2. These images constitute an important reference for visually evaluating in which classes the models are more successful or less successful.

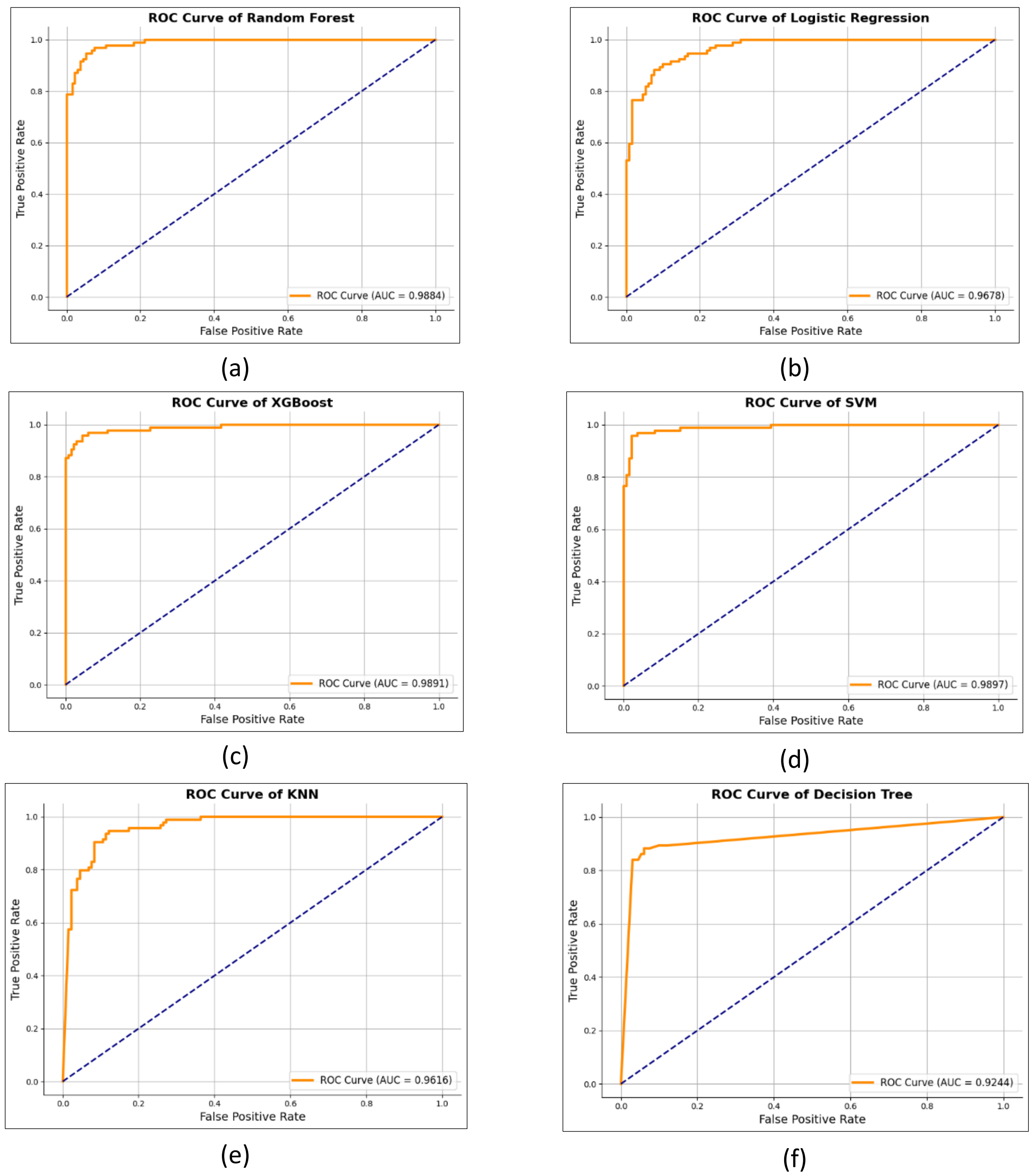

Figure 3. ROC-AUC value matrices of the classification algorithms obtained within the scope of the research: (a) Random Forest, (b) Logistic Regression, (c) XGBoost, (d) SVM, (e) KNN, (f) Decision Tree

The classification performance of different machine learning algorithms for hardwood and softwood species was evaluated using confusion matrices obtained using a 10-fold cross-validation method. The SVM (Support Vector Machines) model achieved the highest accuracy rate, misclassifying only 7 instances and successfully distinguishing between both classes. The XGBoost and Random Forest algorithms also demonstrated high performance with 11 and 13 incorrect predictions, respectively. In contrast, the K-Nearest Neighbour (KNN) algorithm showed the lowest accuracy with 27 incorrect classifications. The Logistic Regression and Decision Tree models were particularly notable for their higher error rates, particularly in the softwood class. The overall results reveal that the classes in the dataset are not easily separated linearly, and therefore, kernel-based approaches and ensemble methods yield more effective results compared to linear or instance-based models. The ROC-AUC values of the classification algorithms developed in the research are visually presented in Figure 3.

When comparing the ROC curves and AUC (Area Under Curve) values of different classification algorithms, the model with the highest AUC score was XGBoost (AUC=0.9891). This model discriminated between classes most effectively. XGBoost was followed by Random Forest (AUC=0.9884) and SVM (AUC=0.9867), respectively. These three algorithms demonstrated strong performance both in terms of area under the ROC curve and classification accuracy. The KNN model yielded an acceptable AUC value of 0.9616, but it appears to have failed to adequately distinguish between classes, particularly due to errors observed in the confusion matrix. The Logistic Regression algorithm demonstrated moderate success with AUC=0.9678. Conversely, the Decision Tree model had the lowest discrimination power with an AUC value of 0.9242. These results demonstrate that kernel-based and ensemble methods, which are successful in capturing nonlinear structures, outperform traditional linear models.

In terms of accuracy and overall classification performance, the most successful model was SVM. SVM achieved 96.90% accuracy, and was notable for its F1-score of 97.36% in the Hardwood class and 96.26% in the Softwood class. The XGBoost model came in second with 95.13% accuracy and balanced F1-scores, while Random Forest achieved similar success with 94.25% accuracy. However, it fell slightly behind due to its low recall, particularly in the Softwood class.

The simpler algorithms, Decision Tree and Logistic Regression, achieved accuracies of 90.71% and 90.27%, respectively. The success of Logistic Regression indicates the presence of partial linear separability in the dataset. KNN, on the other hand, underperformed the other algorithms with an accuracy of 88.05%. The low sensitivity value, particularly in the Hardwood class, revealed that this class was difficult to recognise.

In general, the results indicate that models capable of learning complex and nonlinear relationships (especially SVM and XGBoost) yielded more successful results, given the structural characteristics of the dataset. Tree-based models such as Random Forest and Decision Tree were also effective, but their success depended more on the model's depth and hyperparameter settings.

These findings are consistent with some studies in the literature. For example, Ramananantoandro et al. (2016) reported achieving over 85% accuracy with variables such as density, ring spacing, and diameter in their study on the Shorea platyclados species. Vieilledent et al. (2018) succeeded in reducing the RMSE to 60–65 kg/m³ in their model for density estimation in tropical wood species. Kantavichai and Turnblom (2022) reported high accuracy in estimating wood modulus of elasticity (MOE) and flexural strength (MOR) using the XGBoost algorithm, with R^2=0.93.

Similarly, deep learning-based hybrid models also offer high accuracy. For example, Ma et al. (2025) achieved remarkable results of $R^2 = 0.95$ for MOE and $R^2 = 0.85$ for MOR with their Multi-Input Residual Neural Network (MIRN) model. Nieradzik et al. (2024), working with microscopic images, achieved similar accuracy rates (94%) to human experts. In our study, achieving results close to these levels without using deep networks demonstrates that classical methods are still valid and powerful alternatives.

Finally, in a study conducted by Eker et al. (2023) using data obtained from Turkish forests, an accuracy rate of 91.2% was achieved using AutoML systems. These results are important because they show that the models in our study can achieve similar levels of success with automatic parameter optimisations.

CONCLUSIONS AND FUTURE WORK

Summary of Findings

Hardwood and softwood classification was performed based on the mechanical properties of commercially important wood species grown in the United States. Results obtained using six different machine learning algorithms were analysed based on their accuracy performance: XGBoost, SVM, Random Forest, Logistic Regression, KNN, and Decision Tree.

The results show that the SVM algorithm achieved the highest accuracy rate of 97.9%. The SVM model, which made only seven incorrect classifications in total, achieved high precision and balance in both classes. SVM was followed by the XGBoost algorithm with 97.4% accuracy and AUC = 0.9891, which also demonstrated strong and balanced performance.

The Random Forest algorithm achieved high accuracy with 96.6% accuracy and AUC = 0.9884, but performed slightly lower than the other two models.

The Logistic Regression algorithm performed moderately with 93.3% accuracy and experienced instability, particularly at class boundaries. The KNN algorithm performed least well with 86.6% accuracy, resulting in the highest overall misclassification.

Decision Tree, with 88.3% accuracy and AUC=0.9242, was insufficient in distinguishing between classes, producing high error rates, particularly in the Softwood class.

Kernel-based (SVM) and ensemble methods (XGBoost, RF) offered higher accuracy and lower error rates compared to classical and sample-based models (Logistic Regression, KNN, Decision Tree).

Applications for the Wood Industry: The high accuracy rates obtained have the potential for direct application in various areas of the wood industry. SVM and XGBoost algorithms, in particular, can provide a robust infrastructure for real-time classification systems that can be integrated into production lines. Such systems can automate material identification processes, reducing faulty production rates and enabling more efficient use of natural resources. Additionally, automated classification systems based on mechanical properties can reduce human-induced errors in quality control processes and increase standardisation in production. This is also important for sustainable forestry practices, as accurate classification allows for more effective implementation of tree species-specific usage scenarios.

Recommendations for Future Work: In future studies, it is recommended that classification models be tested with larger datasets collected from different regions and covering a wider range of tree species. This allows for a more comprehensive assessment of the models' generalisation capabilities.

Furthermore, the performance of the algorithms can be further improved through the use of feature engineering, dimensionality reduction, hyperparameter optimisation, and automated construction techniques such as AutoML. Future studies on hybrid deep learning models that integrate mechanical properties with microscopic or macroscopic image data will also be beneficial. Such approaches have the potential to provide more flexible and highly accurate solutions compared to classical methods, especially for complex classification problems.

REFERENCES

Adegun AA, Viriri S (2020) FCN-based DenseNet framework for automated detection and classification of skin lesions in dermoscopy images. *IEEE Access,* 8:150377–150396. https://doi.org/10.1109/ACCESS.2020.3016651

Bartolucci B, De Rosa A, Bertolin C, Berto F, Penta F, Siani AM (2020) Mechanical properties of the most common European woods: a literature review. *Frattura ed Integrità Strutturale*, 14(54):249–274. https://doi.org/10.3221/IGF-ESIS.54.19

Batur A (2004) Gelişmiş ahşap yapım sistemleri ve Türkiye koşulları yönünden değerlendirilmesi. Gebze Yüksek Teknoloji Enstitüsü, Mühendislik ve Fen Bilimleri Enstitüsü Yüksek Lisans Tezi, Gebze.

Bilici S (2006) Ahşap konut üretim sistemleri; Almanya örneği. Selçuk Üniversitesi Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi, Konya.

Bozkurt AY (1971) Önemli bazı ağaç türleri odunlarının tanımı, teknolojik özellikleri ve kullanılış yerleri. *İstanbul Üniversitesi Orman Fakültesi Yayınları,* Yayın No: 177, İstanbul, 99 s.

Breiman L (2001) Random forests. *Machine Learning*, 45(1):5–32. https://doi.org/10.1023/A:1010933404324

Chen T, Guestrin C (2016) XGBoost: A Scalable Tree Boosting System. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, New York, pp 785–794. https://doi.org/10.1145/2939672.2939785

Chicco D, Jurman G (2020) Machine learning can predict survival of patients with heart failure from serum creatinine and ejection fraction alone. *BMC Medical Informatics and Decision Making,* 20(1):16. https://doi.org/10.1186/s12911-020-1023-5

Cortes C, Vapnik V (1995) Support-vector networks. *Machine Learning,* 20(3):273–297. https://doi.org/10.1007/BF00994018

Cover T, Hart P (1967) Nearest neighbor pattern classification. *IEEE Transactions on Information Theory,* 13(1):21–27. https://doi.org/10.1109/TIT.1967.1053964

Eker R, Alkiş KC, Uçar Z, Aydın A (2023) Ormancılıkta makine öğrenmesi kullanımı. *Turkish Journal of Forestry*, 24(2):150–177. https://doi.org/10.18182/tjf.1282768

Gündüz G, Yıldırım N, Göksu Ş, Onat SM (2009) Ak dut ağacının anatomik, kimyasal, fiziksel ve mekanik özellikleri. *Düzce Üniversitesi Ormancılık Dergisi,* 5(1):131–149. https://izlik.org/JA92ZE78CK

Hosmer DW, Lemeshow S, Sturdivant RX (2013) Applied Logistic Regression. 3rd Edn. John Wiley & Sons, Hoboken.

Jacquin P, Mothe F, Longuetaud F, Billard A, Kerfriden B, Leban JM (2019) CarDen: a software for fast measurement of wood density on increment cores by CT scanning. *Computers and Electronics in Agriculture,* 156:606–617. https://doi.org/10.1016/j.compag.2018.12.008

Jones G, Liziniewicz M, Lindeberg J, Adamopoulos S (2023) Non-destructive evaluation of downy and silver birch wood quality and stem features from a progeny trial in Southern Sweden. *Forests*, 14(10):2031. https://doi.org/10.3390/f14102031

Kantavichai R, Turnblom EC (2022) Identifying non-thrive trees and predicting wood density from resistograph using temporal convolution network. *Forest Science and Technology*, 18(4):144–149. https://doi.org/10.1080/21580103.2022.2115561

Kretschmann D (2010) Mechanical Properties of Wood. In: Wood Handbook: Wood as an Engineering Material. U.S. Department of Agriculture, Forest Service, Forest Products Laboratory, Madison, pp 5.1–5.46 (General Technical Report FPL-GTR-190).

Ma J, Kuang Z, Fang Y, Huang J (2025) A multi-input residual network for non-destructive prediction of wood mechanical properties. *Forests,* 16(2):355. https://doi.org/10.3390/f16020355

Martínez RD, Balmori JA, Llana DF, Bobadilla I (2020) Wood density determination by drilling chips extraction in ten softwood and hardwood species. *Forests,* 11(4):383. https://doi.org/10.3390/f11040383

Monteiro TC, Araújo Júnior CA, dos Santos JH, Cardoso Silva T, Magalhães do Nascimento T, Ferraresso Conti Junior JL, Pereira da Rocha M (2023) Artificial intelligence to growth stresses predicting in Eucalyptus clones using dendrometric variables and wood density. *Maderas Ciencia y Tecnología*, 25:1–15. http://dx.doi.org/10.4067/s0718-221x2023000100430

Nieradzik L, Sieburg-Rockel J, Helmling S, Keuper J, Weibel T, Olbrich A, Stephani H (2024) Automating wood species detection and classification in microscopic images of fibrous materials with deep learning. *Microscopy and Microanalysis*, 30(3):508–520. https://doi.org/10.1093/mam/ozae038

Usamah MG (2016) Comparison study between hardwood and softwood. *Journal of Babylon University Pure and Applied Sciences,* 23(3):563–564. https://search.emarefa.net/detail/BIM-933238

Perlin J (2005) A Forest Journey: The Story of Wood and Civilization. The Countryman Press, Woodstock.

Powers DMW (2020) Evaluation: from precision, recall and F-measure to ROC, informedness, markedness & correlation. *Journal of Machine Learning Technologies,* 2(1):2229-3981. https://doi.org/10.9735/2229-3981

Quinlan JR (1986) Induction of decision trees. *Machine Learning,* 1(1):81–106. https://doi.org/10.1007/BF00116251

Ramananantoandro T, Ramanakoto MF, Rajoelison GL, Randriamboavonjy JC, Rafidimanantsoa HP (2016) Influence of tree species, tree diameter and soil types on wood density and its radial variation in a mid-altitude rainforest in Madagascar. *Annals of Forest Science,* 73(4):1113–1124. https://doi.org/10.1007/s13595-016-0576-z

Vieilledent G, Fischer FJ, Chave J, Guibal D, Langbour P, Gérard J (2018) The Cirad Wood Density Database. Zenodo. https://zenodo.org/records/1095454, Erişim Tarihi: 08.07.2025.

Virgulino-Júnior PC, Gardunho DC, Silva DN, Fernandes ME (2020) Wood density in mangrove forests on the Brazilian Amazon coast. *Trees,* 34(1):51–60. https://doi.org/10.1007/s00468-019-01896-5