



OPEN Phenotypic classification of opium poppy genotypes (*Papaver somniferum* L.) based on morpho-phenological traits

Yasin Özgen, Nilgun Bayraktar & Ugur Ozkan✉

Accurate classification of opium poppy (*Papaver somniferum* L.) genotypes based on phenotypic variation is essential for efficient breeding programs targeting specific traits such as flower and seed color. Opium poppy populations exhibit considerable variation in their morpho-phenological traits, particularly in flower and seed color, which is the focus of the present study. Descriptive plant measurements used to classify flower and seed color included germination time, flowering time, maturity time, plant height, number of capsules per plant, capsule length, capsule width, stigma rays, seed/capsule percentage, capsule yield per plant, seed yield per plant, capsule yield, seed yield, 1000-seed weight, morphine content, noscapine content, morphine yield, and noscapine yield. The opium poppy dataset ($nrow=200$) was divided into a training subset (70%) ($nrow=140$) and a test subset (30%) ($nrow=60$). ML algorithms, including Naïve Bayes, Support Vector Machine (SVM), weighted k-nearest neighbors (k-NN), Learning Vector Quantization (LVQ), and two decision-tree-based classification algorithms, bagging CART and Random Forest, were used for classification. According to the classification of flower and seed colors, the classification accuracies of Naïve Bayes, SVM, Weighted k-NN, and LVQ were found to be 95.00%, 91.67%, 78.33%, and 60.00% for flower color, and 78.33%, 91.67%, 50.00%, and 50.00%, respectively. Naïve Bayes and SVM classification yielded more accurate results for FC and SC, respectively.

Keywords Alkaloid, Breeding, Morphine, Noscapine, Seed

Opium poppy (*Papaver somniferum* L.) is the most economically important species of the *Papaver* genus and has been used for medicinal and culinary purposes throughout human history¹. Opium poppy is a herbaceous and dicotyledonous species originating from the Mediterranean coastal region². Opium poppy is primarily cultivated for the alkaloids found in its capsules and seeds. As a narcotic plant, it is grown under the supervision of the United Nations, particularly in countries where its cultivation is legal (Türkiye, India, Hungary, Spain, France, the Czech Republic, China, and Australia). Among these countries, Türkiye has the largest area dedicated to poppy opium cultivation and employs traditional agricultural practices^{3,4}.

Alkaloids can be found at different levels in other plant organs, except the seeds, and plants synthesize more than 80 different alkaloids. Alkaloids typically accumulate at their highest levels during the green or dry capsule phases. Alkaloid extraction methods vary by country; dry capsules are used in Türkiye, a green plant in France, and opium gum (obtained through capsule lancing in India). The most important alkaloids include morphine, noscapine, tebaine, and codeine. The alkaloids listed are key chemical components in medicine and pharmaceuticals. Seeds, another essential economic outcome, can be used directly or turned into flour for cakes and pastries with nutritional value rich in healthy fixed oils ($\leq 50\%$), proteins ($\leq 13\%$), and essential minerals (Ca, Fe, Mg, Mn, P, and Zn)^{5,6}. In poppy opium plants, flowers usually bloom at the tips of the stems and branches, with an average of 3 to 5. These flowers are of various colors, including white, blue, red, and shade. Interestingly, there is a genetic connection between flower and seed colors. For example, plants with white petals tend to produce white or pale seeds, such as yellow^{5,7}.

The use of powerful machine learning (ML) algorithms has gained importance, owing to recent developments in computer technology, rather than classical statistical methods when statistical assumptions are violated⁸. By combining big data and high-performance computing, ML enables a better analysis and understanding

Faculty of Agriculture, Department of Field Crops, Ankara University, Ankara, Turkey. ✉email: ugurozkan@ankara.edu.tr

of complex agricultural processes. It helps define and uncover patterns, optimize operations, and improve decision-making in data-intensive farming environments⁹. Modern agriculture is facing several major challenges, including climate change and soil degradation¹⁰. With such enormous pressures on agriculture, the optimization and acceleration of agricultural practices and techniques are urgently needed. The importance of ML in agricultural studies, especially in regression and classification problems, has increased annually owing to its effectiveness. ML can be applied to various steps of sustainable agriculture, such as crop yield prediction, soil and crop type classification, and classification of crops based on the growth phase^{11,12}. Classification plays a fundamental role across all scientific disciplines in diverse applications and serves as a foundational tool for organizing knowledge. It enables the identification and differentiation of diverse entities, phenomena, concepts, and provides a structured framework¹³. Classification models include training and test subsets. While the training subset was used to build the classification model, the test subset was used to predict the class label for a new sample¹⁴. In a previous study on classification problems, Rahman, et al¹³., Bhat, et al¹⁵., and Senapaty, et al¹⁶. noted the soil classification, soil suitability for crop selection, and crop recommendation using ML algorithms. Koklu and Ozkan¹⁷ and Taner, et al¹⁸. focused on the multi-class classification of dry bean and apple varieties, respectively.

Flower color and seed color are genetically linked in opium poppy plants. Accordingly, the seed phenotype can potentially be predicted during the flowering stage through remote sensing and image processing techniques. In breeding programs, this approach may facilitate the selection of other traits associated with flower color in the genotype populations with diverse floral pigmentation. Furthermore, the early prediction, detection, and elimination of undesired genotype individuals in production fields can contribute to minimizing economic losses^{7,19}. The current study focuses on color classification using ML algorithms, which were first documented for flower and seed colors in opium poppy plants. The classification of flowers and seed color provides valuable insights into their production and plays a crucial role in seed quality assessment. Traditionally, these tasks are performed by specialists through the visual inspection of individual samples, which is a labor-intensive and time-consuming process²⁰. The current study aimed to classify flower and seed color for opium-poppy plants using data from 23 lines and two standard plants of opium poppy plants in 2017 and 2018. Popular ML algorithms were employed, including Naïve Bayes, Support Vector Machine (SVM), weighted k-nearest neighbors (k-NN), Learning Vector Quantization (LVQ), and two decision-tree-based classification algorithms: bagging Classification and Regression Tree (CART) and Random Forest.

Material and method

Study area and materials

The experimental field was located within the Application and Research Area of Ankara University in Türkiye (39°57'44" N latitude, 32°51'35" E longitude at an elevation of 938 m) (Fig. 1). In the current study, lines from the opium poppy collection of the Department of Field Crops, advanced to the F₃ stage through morphological observations, preliminary yield, and alkaloid analyses as a result of breeding studies, were used. Among the promising lines identified through alkaloid analyses conducted in 2016, 23 lines and two standard varieties were selected and included in the current study (Table 1).

The standard cultivars (ST 1, ST 2) used in the experimental field was locally registered genotypes certified by the Turkish Grain Board (TMO). Their widespread cultivation potential, high adaptability to environmental

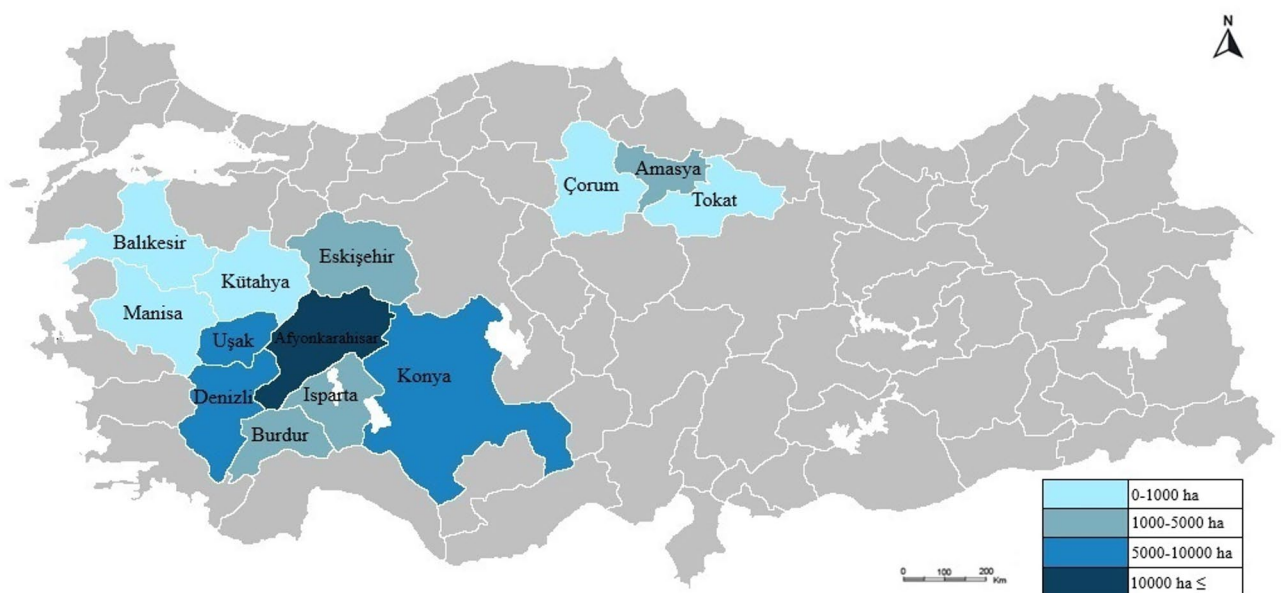


Fig. 1. Legal opium poppy cultivation cities in Türkiye (Microsoft Paint in Windows 10, <https://www.microsoft.com/windows>).

LINES	Flower Color	Seed Color	References
LINE 1	Purple	Blue	Valizadeh, et al ²¹ . <i>Hybridization genotypes</i>
LINE 2	White	White	
LINE 3	Purple	Blue	
LINE 4	White	Pink	
LINE 5	White	Pink	İpek ²² <i>Hybridization genotypes</i>
LINE 6	White	White	
LINE 7	Purple	Pink	
LINE 8	Purple	Blue	
LINE 9	White	White	Department of Field Crops, Seed Gene Bank, Ankara, Türkiye <i>Selection genotypes</i>
LINE 10	Purple	Blue	
LINE 11	Purple	Pink	
LINE 12	Pink	Pink	
LINE 13	Purple	Pink	
LINE 14	Pink	Brown	
LINE 15	Purple	Blue	
LINE 16	Purple	Blue	
LINE 17	Purple	Pink	
LINE 18	Purple	Green	
LINE 19	Purple	Blue	
LINE 20	White	White	Javani ²³ <i>Hybridization genotypes</i>
LINE 21	Purple	Blue	
LINE 22	Purple	Blue	
LINE 23	Purple	Blue	Department of Field Crops, Seed Gene Bank, Ankara, Türkiye, <i>Selection genotypes</i>
ST 1	White	White	
ST 2	White	White	

Table 1. Flower and seed colors of lines and standard varieties and their locations of origin. **ST:** standard plant.

conditions, superior capsule and seed yield, as well as elevated alkaloid content were the primary factors influencing their selection for the study. ST 1 and ST 2 were named as “Ofis 1”, “Ofis 8”, respectively.

Visual records of the flower and seed samples were captured under natural daylight conditions to ensure uniform illumination. The samples were placed on gray A4 paper to minimize background interference. All images were taken from a fixed distance of 50 cm using a Canon EOS 650 digital camera, and standardized camera settings (ISO 400, $f/4$, and 1/250 s) were applied across all samples to ensure consistency. The resulting digital images were analyzed using *ImageJ* software²⁴. Each image was evaluated based on its pixel matrix and the color distribution was manually measured within defined regions of interest (ROIs) on the samples. The RGB values of each pixel served as the primary basis for the analysis, and these values were converted to the CIELAB color space when necessary. Prior to color analysis, all images were normalized using a gray background as a reference. Color measurements were performed using the *ImageJ* interface, with the mean color values calculated from three different points on each sample. To enhance reproducibility, all images were saved at a fixed resolution in the JPEG or TIFF format. No automatic macros or plugins were used in the analysis, and all the procedures were manually conducted.

Climate and soil properties of the study area

The total precipitation (379.9 mm) was below 100 mm in the first year (270.8 mm), whereas in the second year (377.8 mm), it was close to the long-term average. The average temperature during the vegetation period was similar to the long-term average (10.3 °C) in the first year (10.1 °C), but approximately two degrees higher in the second year (12.3 °C). The average relative humidity (63.4%) was lower than the long-term average in both years, 59.0% in the first year and 62.1% in the second year (Table 2).

The soil was sampled from a depth of 30 cm at various points within the experimental plot and had a clay loam texture. It had a pH of 7.45 (approximately neutral) and contained 1.30% organic matter. The nitrogen (N) content was determined to be 0.067%, the potassium (K_2O) content was 254.00 ppm, the phosphorus (P_2O_5) content was 8.80 ppm, and the electrical conductivity was 0.10 dS m^{-1} ²⁵.

Descriptive measurements of opium poppy plants

According to the randomized block design, all opium poppy materials were planted in plots of 5 m^2 , with four replicates, in October. A total of 5 kg da^{-1} nitrogen (N) and 5 kg da^{-1} pure phosphorus (P_2O_5) were applied. The following characteristics were measured: germination time (days), flowering time (days), maturity time (days), plant height (cm), number of capsules per plant (units), capsule length (mm), capsule width (mm), number of stigma rays (units), seed-to-capsule ratio (%), capsule yield per plant (g), seed yield per plant (g), capsule yield

Climate Factors	Months	Oct.	Nov.	Dec.	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Total	Mean
Precipitation (mm)	2016-17	2.90	21.8	47.5	30.4	7.50	46.1	19.8	47.4	47.4	0.00	270.8	----
	2017-18	19.9	33.6	37.4	46.0	29.9	72.7	2.40	87.8	37.7	10.4	377.8	----
	LYM	30.1	33.3	45.5	42.7	36.1	41.3	47.5	51.7	37.5	14.2	379.9	----
Mean Temperature (°C)	2016-17	13.8	6.90	-0.30	-1.30	2.90	8.10	11.0	15.8	20.4	25.6	----	10.3
	2017-18	12.6	7.10	4.70	3.10	6.50	10.0	15.3	18.0	21.4	24.5	----	12.3
	LYM	13.2	7.10	2.50	0.30	1.90	6.10	11.2	15.8	19.8	23.3	----	10.1
Relative Humidity (%)	2016-17	54.6	54.3	74.3	76.8	66.7	60.2	50.6	55.9	58.0	38.1	----	59.0
	2017-18	56.3	69.7	78.3	77.0	73.1	63.3	44.9	59.6	53.4	45.2	----	62.1
	LYM	60.1	69.6	76.7	76.7	71.8	64.2	59.5	57.4	52.7	45.4	----	63.4

Table 2. Meteorological data for 2016–2017, 2017–2018, and the long years mean (LYM).

Traits	Min	Max	Mean	SE	N
Germination time (days)	7.00	26.00	16.46	0.31	200
Flowering time (days)	185.00	219.00	203.16	0.59	200
Maturity time (days)	241.00	287.00	264.46	0.93	200
Plant height (cm)	65.00	130.00	98.31	0.85	200
Capsules (number plant ⁻¹)	1.70	9.30	4.85	0.12	200
Capsule length (mm)	41.00	65.00	52.19	0.34	200
Capsule width (mm)	28.00	58.00	39.66	0.37	200
Stigma rays (number)	6.00	17.00	10.68	0.16	200
Seed-to-capsule ratio (%)	0.30	2.00	1.02	0.02	200
Capsule yield (g plant ⁻¹)	1.20	6.00	3.28	0.07	200
Seed yield (g plant ⁻¹)	1.00	7.00	3.43	0.08	200
Capsule yield (kg da ⁻¹)	60.00	184.00	110.30	1.58	200
Seed yield (kg da ⁻¹)	59.00	188.00	112.50	1.72	200
1000-seed weight (g)	0.20	0.72	0.49	0.01	200
Morphine content (%)	0.34	2.07	0.88	0.02	200
Noscapine content (%)	0.00	1.01	0.26	0.01	200
Morphine yield (g da ⁻¹)	0.44	2.94	0.96	0.02	200
Noscapine yield (g da ⁻¹)	0.00	1.00	0.28	0.01	200

Table 3. Descriptive measurements. **Min:** Minimum, **Max:** Maximum, **SE:** Standard error.

per plant (g plant⁻¹), capsule yield per decare (kg da⁻¹), seed yield per plant (g plant⁻¹), seed yield per decare (kg da⁻¹), 1000-seed weight (g), morphine content (%), noscapine content (%), morphine yield (g da⁻¹), and noscapine yield (g da⁻¹). The descriptive plant measurement statistics are presented in Table 3.

Statistical approach

Experimental design and data pre-processing

A balanced randomized complete block design (RCBD) scheme was conducted using four replicates. The treatments consisted of twenty five opium poppy lines ($n=25$) and multiple years ($n=2$). Shapiro-Wilk and Levene's tests were used to check the normal distribution of traits and the equality of the error variances and residual normality, respectively.

Data preprocessing is a crucial step in preparing data for the application of ML algorithms. Not all attributes in a dataset contribute equally to the analytical goals; therefore, careful selection and transformation of the relevant traits are essential for ensuring accurate and meaningful results. Given that real-world data are often noisy, incomplete, or inconsistent, various preprocessing approaches, such as data cleaning, normalization, transformation, and reduction, are employed to enhance data quality and facilitate effective analysis²⁶. In the current study, the dataset was shuffled to increase randomness and change the current ranking and systematic orders. Subsequently, a seed was created in the R environment because of the same reproductive ability as the shuffled data. The opium poppy dataset ($nrow = 200$) was divided into a training subset (70%, $nrow = 140$) and a test subset (30%, $nrow = 60$). In the modeling stage, the current study used the classification of opium poppy flower color (FC) ($n = 3$) and seed color (SC) ($n = 5$). Flower color contained 60% purple, 32% white, and 8% pink flower color data, while seed color contained 40% blue, 28% pink, 24% white, 4% green, and 4% brown seed color data. To address the challenge of class imbalance, we employed multiple random seeds and performed a manual consistency check. This process verified that all color classes, including the minority green and brown seed classes, were present in both the training and test subsets, thereby ensuring the models were

exposed to the full range of phenotypic variation. However, it is acknowledged that the small absolute number of minority class samples (e.g., $n = 8$ each for green and brown seeds) inherently limits the statistical power for their reliable prediction. Principal Component Analysis (PCA) was performed using JMP v.17 to explore multivariate variability of morpho-phenological traits²⁷.

Development of modelling

Classification models were used to determine the association between studied traits and specific classes. Numerous algorithms facilitate this process, and each potentially produces different results depending on the characteristics of the dataset¹⁷. Therefore, selecting the most suitable classifier for the given data ensures accurate and reliable decision making. In the current study, the six most common machine learning classification techniques were used: Naïve Bayes, Support Vector Machine (SVM), weighted k-nearest neighbors (k-NN), Learning Vector Quantization (LVQ), and two decision-tree-based classification algorithms: bagging Classification and Regression Tree (CART) and Random Forest.

Machine learning algorithms and hyper-tuning

Naïve Bayes is a probabilistic classification algorithm that operates under a strong independence assumption, positing that each trait contributes independently to the probability of a given class²⁸. This model estimates the likelihood of class membership by computing the conditional probability of each class given the presence of specific traits and combining these with the prior probability of each class²⁹. The traits were considered as $f_1, f_2, f_3, \dots, f_i$, where O is the opium-poppy plant and C is the class. Equation (1) shows that $P(O|C)$ is the posterior probability of C for data O , $P(C)$ is the prior probability, and $P(w_i|C)$ is the product of likelihoods in the Naïve Bayes model¹⁶ (Eq. 1). The R package “naivebayes” was used to create the Naïve Bayes model in the study³⁰. To prevent zero possibilities, “laplace” was determined as 1 for the Naïve Bayes models of FC and SC.

$$P(O|C) \propto P(C) \prod_{i=1}^n P(w_i|C) \quad (1)$$

SVMs are a class of kernel-based algorithms that are known for their powerful computational capabilities in solving classification and regression tasks. By projecting data into higher-dimensional spaces, SVMs identify optimal hyperplanes that maximize the separation between different classes, making them particularly effective in complex nonlinear problems³¹. In the multiclass SVM with classes in the current study, hyperplanes are constructed to separate each pair of classes. The one-vs-rest (OvR) strategy was employed to classify a new sample. In this approach, a distinct hyperplane is formulated for each class, distinguishing it from the combined set of all the other classes. In the formulas below, w is the direction of the plane, b is the threshold value, and x is the trait vector. Also, y_i represents the hyperplane used to separate the data points of the classes. The decision function $w \cdot x_i + b$ calculates the distance between x_i and the hyperplanes (Eq. 2). The R package “e1071” was used to create the SVM model in the study³². The radial basis function (RBF) was determined as the kernel function for the SVM models of the FC and SC in the current study. The “cost” was also identified as 8 and 32 for the best tuned SVM model for FC and SC, respectively.

$$\begin{aligned} f(x) &= 0 \\ f(x) &= w \cdot x + b \\ f(x) &> 0 \\ f(x) &> 0 \\ y_i(w \cdot x_i + b) \end{aligned} \quad (2)$$

The weighted k-NN algorithm classifies based on a given “ k ” value, considering the class of the nearest neighbor. In this algorithm, a vector is classified using the known vectors of the class¹⁷. During the training process, a new data point is selected, and the Euclidean distance to its nearest neighboring data points is computed. The Euclidean distance formula is shown in Eq. 3. Represented coordinate values of the 1st point are x_1, y_1 , and x_2, y_2 for the 2nd point. In addition, the training subset data are x_1 , and y_1 , and the test subset data are x_2, y_2 . The weighted k-NN algorithm can be refined by weighting each of the k neighbors based on their distance to the query point x_q , assigning higher weights w_i to closer neighbors¹³ (Eq. 4). Equation 5 defines the k-NN weighting function. If x_q exactly matches a training point x_i , the output is set to $F(x_q) = f(x_i)$ to avoid division by zero. Using all training instances with weighting makes the algorithm global but increases computational cost¹³.

$$\text{Euclidean distance} = \sqrt{[(x_2 - x_1)^2 + (y_2 - y_1)^2]} \quad (3)$$

$$F(x_q) = \frac{\sum_{i=1}^k w_i f(x_i)}{\sum_{i=1}^k w_i} \quad (4)$$

$$w_i = \frac{1}{d(x_q, x_i)^2} \quad (5)$$

The R package “*kkn*” was used to create a weighted k-NN model in the study³³. The “*train.kkn*” function was used to adjust the best hyper-parameters, with “*k* = 6” and “*k* = 3”, “*optimal*” and “*triangular*” kernels were used for the weighted k-NN models of FC and SC.

LVQ operates by partitioning the input space into distinct decision regions, each represented by a codebook vector. For each class, an input vector is assigned to the codebook vector closest to it. The algorithm iteratively adjusts these regions and their corresponding codebook centers until the decision boundaries and centers stabilize³⁴. The Euclidean distance was calculated using the same formula for the weighted k-NN in the current study (Eq. 3). The R package “*rsample*” was used to create a weighted LVQ model in the study³⁵. The LVQ model has “*codebook size*” as 100, “*niter*” as 1000, and “*alpha*” as 0.30, which are the same values for FC and SC.

Decision trees are small, non-intersecting algorithms that use rules to discriminate or partition traits with similar response variables. Bagging CART was applied to the dataset as a decision tree algorithm in the current study. Bagging is a meta-algorithm designed to increase the robustness and accuracy of classification models in machine learning and to prevent overfitting by reducing variance³⁶. This algorithm is generally used for CART (classification and regression Tree). The “*bag*” function of the R package “*caret*” was used for the tree bagging CART model of FC and SC. Before creating the CART model, the training subset was cross-validated with ten k-folds (k-fold = 10) using the “*trainControl*” function, which is a common configuration standard for decision trees. Random Forest is a popular ensemble technique based on combining or ensembling various decision tree models to address the problem of overfitting³⁷. The R packages of “*mlr*” and “*randomForest*” were used for hyper-tuning and creating the decision-tree model in the study, in the same order^{38,39}. The tuning parameters were *ntree*, *mtry*, and *node size* for both FC and SC, which were adjusted to 15, 4, and 12 and 18, 5, and 19, respectively. The decision tree algorithm and heat maps, which have a higher accuracy rate, are shown (Fig. 3).

Classification performance metrics

In the classification process, model training, evaluation, and optimization, R Studio’s “*caret*” package is used for a systematic approach⁴⁰. To evaluate the models’ performance, plot confusion matrices and performance metrics were obtained for all classification models with the help of the “*metr*” R package⁴¹. A plot confusion matrix is a tabular representation commonly employed to assess the performance of a classification model against a dataset with known classes or labels. It comprises critical parameters: true positives (*tp*), false positives (*fp*), true negatives (*tn*), and false negatives (*fn*), which provide a comprehensive evaluation of the model metrics¹⁵. This matrix enables precise insights into the model’s predictive capabilities and error rates. The performance metrics of the classification models are listed, presented, and explained in detail in Table 4. In Kappa formula (Table 4), *P_o* represents the relative observed agreement among raters, while *P_e* represents the hypothetical probability of the match occurring by chance.

Results

Classification performance metrics of opium poppy lines according to flower color and seed color

Machine Learning (ML) is a powerful and adaptable methodology widely utilized across various disciplines to address complex problems, perform predictive modeling, extract patterns, and facilitate data-driven analysis. In the current study, the training subset served as the resource for model learning, enabling the optimization of model parameters, whereas the test subset was utilized to rigorously evaluate the model’s classification performance. The results of the models developed in the current study are summarized in Table 5, which provides a comprehensive overview of the current study’s findings.

Metrics such as accuracy, kappa, precision, recall, specificity, and F1-score are used in the process of determining an algorithm’s overall performance. According to the classification of FC and SC, the classification accuracies of Naïve Bayes, SVM, Weighted k-NN, and LVQ were found to be 95.00%, 91.67%, 78.33%, and

Metric	Formula
Accuracy	$\frac{\sum_{i=1}^l \frac{tp_i + tn_i}{tp_i + fn_i + fp_i + tn_i}}{l}$
Kappa	$K = \frac{P_o - P_e}{1 - P_e}$
Error rate	$\frac{\sum_{i=1}^l \frac{fp_i + fn_i}{tp_i + fn_i + fp_i}}{l}$
Precision	$\sum_{i=1}^l \frac{tp_i}{tp_i + fp_i}$
Recall	$\sum_{i=1}^l \frac{tp_i}{tp_i + fn_i}$
Specificity	$\frac{tn}{fp + tn} = 1 - fp$
F1-score	$\frac{2 * P_M * r_M}{P_M + r_M} = \frac{2 * \sum_{i=1}^l \frac{tp_i}{tp_i + fp_i} * \sum_{i=1}^l \frac{tp_i}{tp_i + fn_i}}{\sum_{i=1}^l \frac{tp_i}{tp_i + fp_i} + \sum_{i=1}^l \frac{tp_i}{tp_i + fn_i}}$

Table 4. Calculation and explanations of multi-class classification performance metrics.

Performance metrics	Naïve Bayes	SVM	Weighted k-NN	LVQ	Bagging CART	Random Forest
Flower color (FC)						
Accuracy	95.00	91.67	78.33	60.00	60.00	66.67
Kappa	89.47	83.93	60.33	20.79	14.69	23.03
Error rate	5.00	8.33	21.66	40.00	40.00	33.33
Precision	80.64	93.92	70.09	-	52.48	73.40
Recall	75.92	73.09	76.42	41.03	44.88	48.30
Specificity	97.03	94.76	87.27	73.20	71.04	73.12
F1-score	78.21	82.21	73.12	-	48.38	58.27
Seed color (SC)						
Accuracy	78.33	91.67	50.00	50.00	45.00	40.00
Kappa	68.05	87.70	25.28	26.56	21.12	8.32
Error rate	21.66	8.33	50.00	50.00	55.00	60.00
Precision	-	94.62	50.39	-	39.65	-
Recall	60.75	83.64	39.82	31.48	47.87	23.45
Specificity	93.77	97.47	84.99	85.28	84.14	81.71
F1-score	-	88.79	44.48	-	66.01	52.58

Table 5. Performance metrics obtained for different ML algorithms using the test subset. Note: “-” indicates undefined metric due to absence of predicted instances for a class. **SVM**: Support vector machine, **k-NN**: K-nearest neighbors, **LVQ**: Learning vector quantization, **CART**: Classification and regression tree.

60.00% for FC, and 78.33%, 91.67%, 50.00%, and 50.00% for SC, respectively. SVM and Naïve Bayes yielded more accurate results for FC and SC compared to the other ML algorithms. While the SVM model achieved highest F1-scores of 82.21 and 88.79%, and the performance varied drastically across algorithms. Tree-based models (Bagging CART, Random Forest) exhibited critically low F1-scores of 48.38–58.27% for FC; 52.58–66.01% for SC, indicating fundamental limitations in handling phenotypic complexity. The very low frequency of minority classes inherently limited the statistical power for their reliable prediction and evaluation, which is a recognized challenge in imbalanced classification tasks. The occurrence of undefined precision and F1-scores (NaN) for some algorithms (Table 5) is a direct result of models failing to correctly predict these extremely rare classes in the test set, not their complete absence from it.

Bagging CART and Random Forest were used as decision-tree-based ML algorithms in the current study. The classification performance metrics of the Random Forest were greater than those of bagging CART in the FC. Similar to Naïve Bayes, weighted k-NN, and LVQ, FC is more accurate than SC in Random Forest. The accuracy and error rates of Random Forest and bagging CART for FC were 66.67% and 33.33% and 60.00% and 40.00%, respectively. Precision and recall of these were 73.40% and 48.30% and 52.48% and 44.88%, respectively. According to the SC classification, bagging CART and Random Forest exhibited extremely low performance metrics. The classification accuracies of bagging CART and Random Forest were 45.00% and 40.00%.

Multivariate trait variability by principal component analysis (PCA) and trait importance
PCA was used to visualize inherent variability in morpho-phenological traits in the current study (Fig. 2). It was also identify traits driving separation between flower/seed color groups.

The first two principal components (PCs) collectively explained 43.00% of the total variance (PC1:27.50%, PC2:15.50%), revealing distinct patterns in trait associations and group separation. PC1 (27.50% variance) was strongly driven by yield-related traits, including capsule yield per decare (highest positive loading), seed yield per plant, and morphine yield per decare. Traits with negative loadings on PC1 included germination time and maturity time. PC2 (15.50% variance) was dominated by the developmental and morphological traits (Fig. 2). Plant height (strong positive loading), maturity time, flowering time, capsule width, and stigma rays showed moderate positive contributions.

As mentioned for group separation by FC, purple-flowered genotypes clustered in the positive quadrant of PC1 and were associated with high capsule/seed yield. White-flowered genotypes occupied negative PC1 and were correlated with longer germination time. Pink-flowered genotypes displayed intermediate positioning, overlapping both groups, but showed a slight separation along PC2 (driven by plant height). For SC, the blue-seeded genotypes strongly aligned with high PC1 values (yield traits). White-seeded genotypes were grouped in the negative PC1/PC2 quadrant and were linked to delayed phenology (germination/maturity time). Brown and green-seeded genotypes showed minimal separation, indicating shared trait profiles. PCA confirmed that biological distinctness drives ML efficacy when the models are appropriately designed. Well-separated groups, such as purple/white flowers, enabled high accuracy in probabilistic/kernel models, which was 95.00% for Naïve Bayes and 91.67% for SVM.

Table 6 shows key traits of FC and SC classification, and also trait importance was analyzed for the top-performing models (Naïve Bayes for FC; SVM for SC). For FC classification, noscapine yield and capsule yield per decare were the strongest predictors, explaining Naïve Bayes’ 95% accuracy. Purple-flowered genotypes created distinct probabilistic thresholds leveraged by Naïve Bayes. For SC classification, SVM relied heavily on

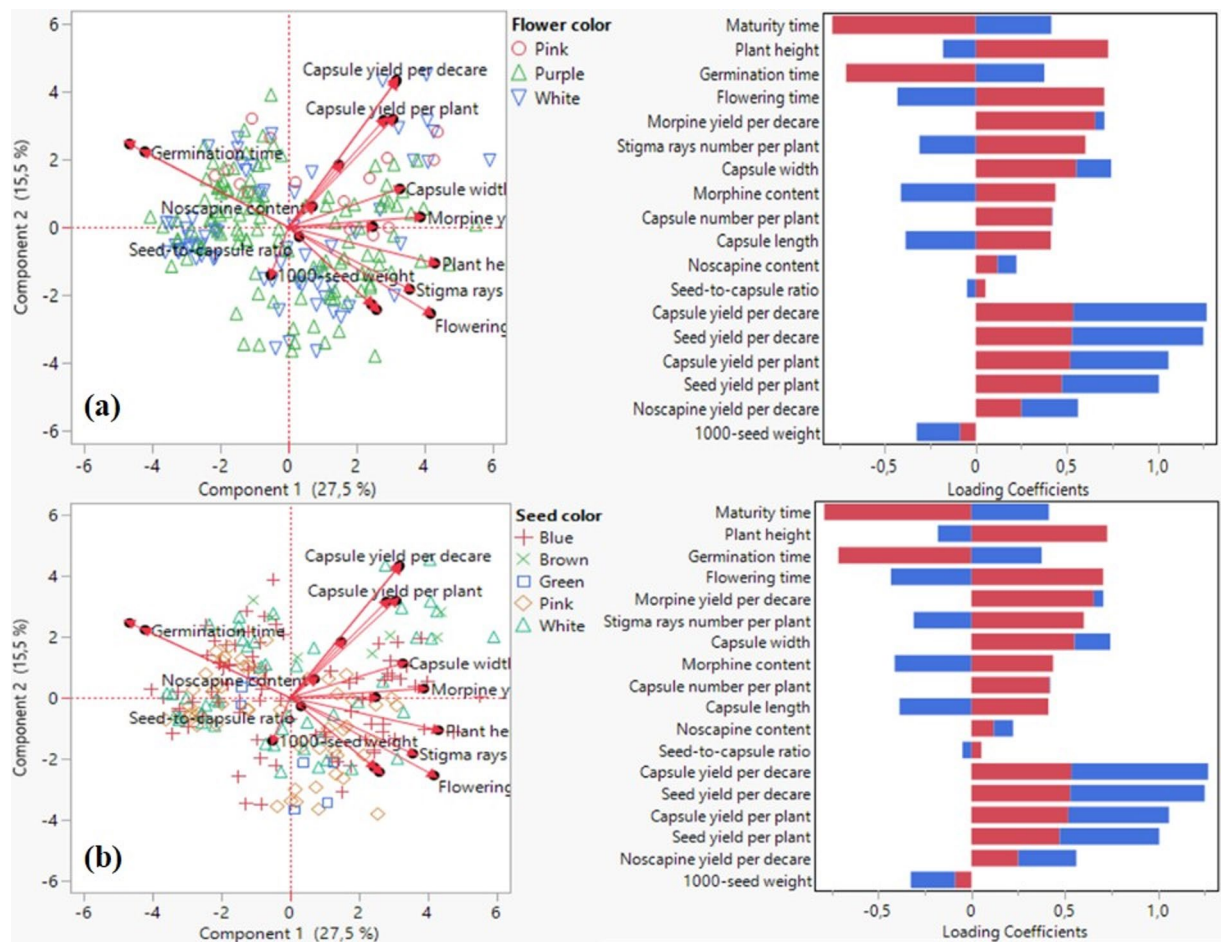


Fig. 2. Biplot analysis opium poppy plants based on flower (a) and seed color (b), and their loading coefficients (Red: PC1, Blue: PC2).

Trait	Model	Importance (%)
Flower color (FC)		
Noscapine yield	Naïve Bayes	28.30
Capsule yield	Naïve Bayes	25.10
Morphine content	Naïve Bayes	18.70
Germination times	Naïve Bayes	12.90
Capsule number per plant	Naïve Bayes	9.80
Seed color (SC)		
Seed yield per plant	SVM	24.50
1000-seed weight	SVM	21.80
Morphine content	SVM	18.20
Capsule width	SVM	15.60
Plant height	SVM	11.30

Table 6. Top five trait importance of predictive traits for FC and SC classification.

seed yield per plant and 1000-seed weight (Table 6). This aligns with the PCA results (Fig. 2a), where these traits dominated PC1 separation. Notably, morphine content, which are though biochemically significant, ranked only fifth in trait importance, suggesting morphological traits drive visual color classification more than alkaloid profiles.

Random forest decision tree and confusion matrices of the machine learning algorithms

In the classification process of FC using the Random Forest model, when the NY was ≤ 0.35 , the capsule number (CN) contributed to group differentiation. If the CN is ≤ 1.73 and above ≥ 1.73 , the FC is classified as white and purple. Another block that helped differentiate the groups was CY, where the NY was > 0.35 . When the CY was below 133 and > 133 , the FC was classified as purple or pink. In the classification process of SC using Random Forest, when CY is ≤ 135 and $CN > 1.97$, SC is classified as pink, whereas all others are classified as blue. If CY was > 135 and NY was greater or smaller than 0.34, SC was classified as brown or white. The integration of the decision tree and heat map offers a comprehensive view of the data and highlights the key traits of the decision tree, as shown in Fig. 3. NY, CN, and CY were the main traits in the two heat maps for FC and SC, respectively. The color range from light green to yellow (> 0.5) indicated a positive relationship between the traits. For FC, when NY was > 0.35 , NY and CN had a positive correlation between these two traits; however, when NY was ≤ 0.35 , the relationship was entirely opposite.

The confusion matrices of FC and SC are presented to visualize and summarize the performance of the ML algorithms in the current study (Figs. 4 and 5). Naïve Bayes and SVM had 3 and 6 misclassifications, while LVQ and bagging CART had 24 misclassifications in 60 samples in FC (Fig. 4). For the other classification factors in the current study, SVM had a minimum of five misclassifications, whereas Random Forest had a maximum of 36 misclassifications out of 60 samples (Fig. 5). The results indicated that FC was the most accurately classified factor in the current study.

Discussion

The performance metrics of the ML algorithms, as shown in Fig. 6, were evaluated using accuracy and kappa as key indicators of efficacy. Additionally, the classification effectiveness of FC and SC was analyzed, as illustrated in Fig. 6. SVM emerged as the most robust classifier for both FC and SC (91.67% accuracy), attributable to its kernel-based handling of nonlinear trait interactions and resilience to class imbalance, critical given SC's five uneven categories of SC. Its one-vs-rest (OvR) strategy optimally separates classes even with imbalanced data (e.g., rare brown/green seeds), thereby minimizing misclassification. This contrasts with distance-based algorithms (k-NN, LVQ), which were affected from trait scaling and phenotypic overlapping in PCA. Naïve Bayes outperformed only for FC (95.00%), where PCA confirmed trait independence (Fig. 2a). Consistent with the findings of the current study, SVM and Naïve Bayes are not sensitive to dataset composition (balanced or imbalanced), whereas k-NN and Random Forest are highly sensitive⁴². Precision and recall are preferable tradeoffs for sensitivity and specificity if there is an imbalanced dataset, as in the current study. Precision and recall were highest in SVM, at 93.92% and 73.09% for FC and 94.62% and 83.64% for SC, respectively. Brownlee⁴³ and Rumora, et al⁴⁴. stated that SVM employs spectrally weighted kernels to reduce redundant information, prioritizes the selection of optimal traits based on their relative significance, and enhances classification performance. Zhou, et al⁴⁵. also mentioned that the SVM demonstrated robustness and exhibited consistent accuracy for imbalanced training set samples, similar to FC and SC, which were the classification factors in the current study. Taner, et al¹⁸. reported 98.17% accuracy in SVM algorithm, in which 15 misclassifications of 872 predictions in the classification of multiclass (ten class) apple varieties based on color. In addition, Koklu and Ozkan¹⁷ achieved 93.44% accuracy classifying 7 dry bean varieties using morphological features and SVM algorithm.

PCA further clarified the ML classification performance. The dominance of yield components (capsule/seed yield) and alkaloid metrics (noscapsine yield) as primary drivers of variability (PC1) aligns with Random Forest's reliance on these traits for node splitting (Fig. 3). Notably, the clear separation of purple-flowered genotypes in PCA is characterized by high yield and low noscapsine, and also explains Naïve Bayes' exceptional accuracy (95.00%) for FC. Conversely, brown/green seeds exhibited minimal PCA separation (Fig. 2b), correlating with high misclassification rates and underscoring the biological complexity underlying seed pigmentation. These synergies confirm that ML algorithms capture physiologically meaningful patterns, in which classification efficacy directly reflects measurable phenotypic divergence in the germplasm. However, the unexpectedly low accuracy of the tree-based models (Random Forest: 40% SC; Bagging CART: 45% SC) contrasts with their typical robustness. This is attributed to the severe class imbalance in seed color (SC). In addition, the trait independence assumption was violated for RF randomization owing to strong trait correlations of yield-alkaloid relationships in PCA. A decision tree of Random Forest, along with a heat map, provided an approximate distribution of sample proportions per leaf and represents the model's confidence in classifying samples within each in the current study⁴⁶ (Fig. 3). The two main blocks of noscapsine (NY) and capsule yield (per decare) (CY) in the heat map illustrate their importance as determining factors in classifying FC and SC, respectively.

The confusion matrix provided a detailed breakdown of the classification performance for each FC and SC, highlighting both the accuracy and misclassification rates with the same aim as Koklu and Ozkan¹⁷ and Taner, et al¹⁸. Each row and column of the matrix corresponds to the predicted and observed classes, respectively. The diagonal elements of the matrix indicate the number of correctly classified instances for each class, whereas the off-diagonal elements represent instances misclassified into other classes. The confusion matrix was generated by all ML algorithms using morpho-phenological traits such as plant height, capsule length and width, capsule yield, seed yield, morphine content and yield, and noscapsine content and yield. The confusion matrix also revealed that the frequency of all major diagonal elements for FC was higher than that for SC. On the other hand, undefined precision/F1-scores (Table 5) occurred when classifiers failed to predict minority seed-color classes (e.g., brown/green), resulting in zero true positives. This reflects inherent challenges in classifying under-represented phenotypes and highlights the need for balanced datasets or oversampling techniques in future work.

The dependent variables, FC and SC, in the current study had three and five classes, respectively. As the number of classes of the dependent variable increased, the accuracy decreased and the error rate increased, except for the SVM. According to the consistent explanation of Del Moral, et al⁴⁷. and in line with the current

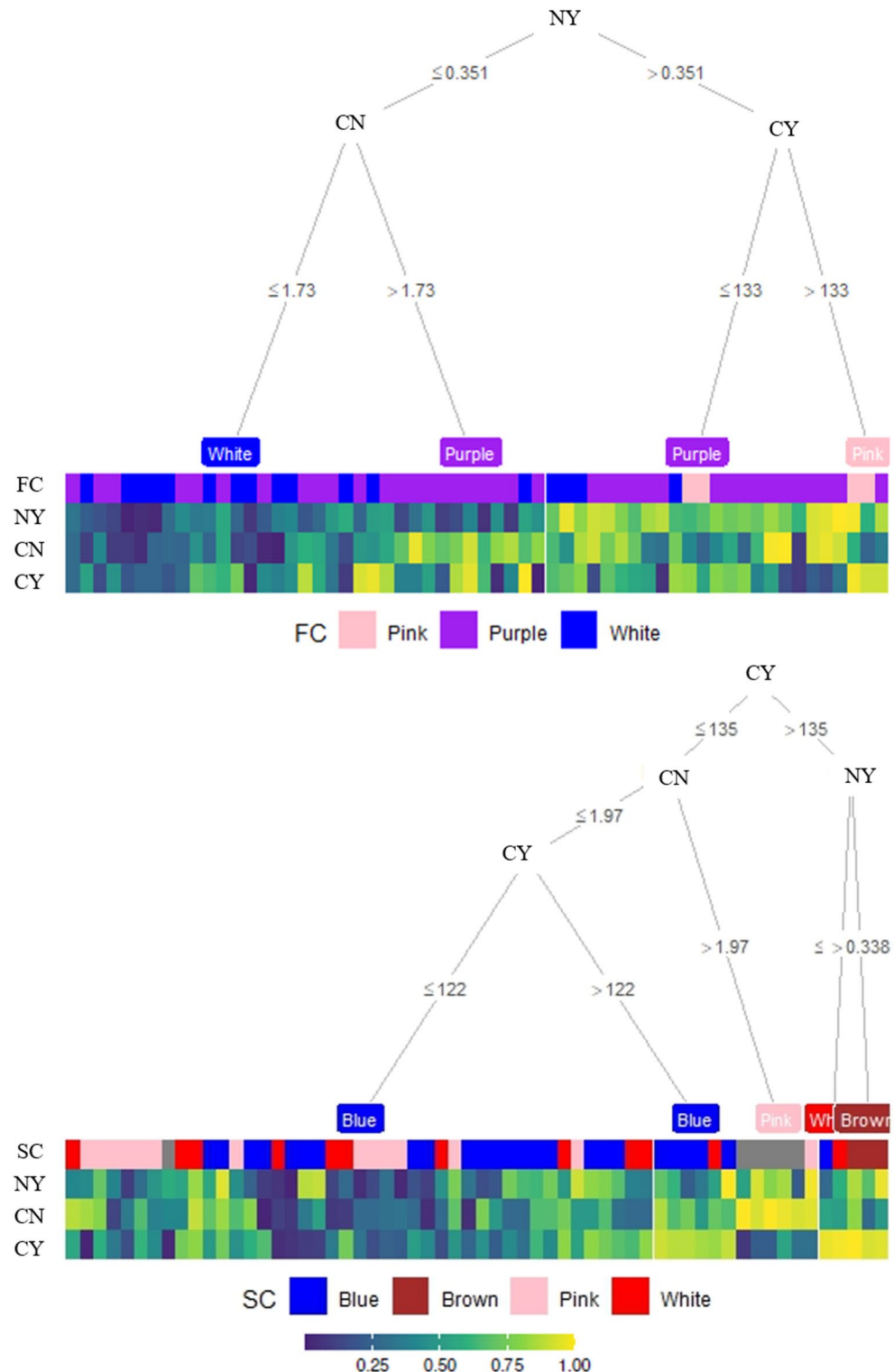


Fig. 3. Random forest decision trees and trait correlation heat maps according to FC and SC. note: CN: Capsule number per plant, CY: Capsule yield per decare, NY: Noscapine yield per decare.

study, as the number of classes increases in classification problems, accurately assigning a new instance to the correct class becomes more challenging than when fewer classes are present. Treating the FC classification (three classes) as a multiclass problem resulted in greater performance compared with the SC classification (five classes) in the current study. In addition, Del Moral, et al⁴⁷, noted that when learning algorithms are required to solve two or more decision boundaries or classification tasks for a single dependent factor, the classification performance tends to decrease. SVM achieved a comparable accuracy for SC and FC of 91.67%. For all other

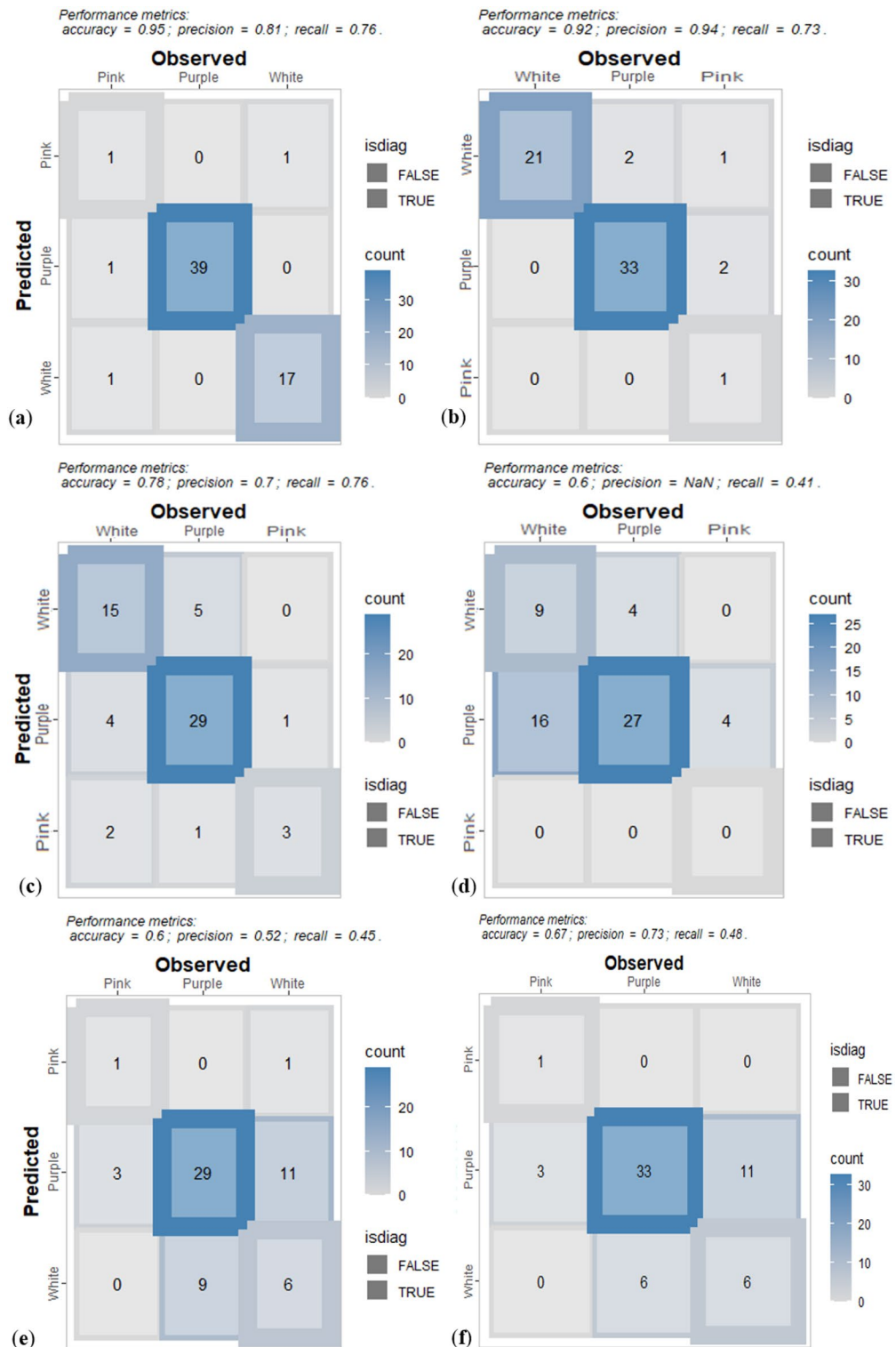


Fig. 4. Confusion matrices of FC. note: (a) Naïve Bayes, (b) SVM, (c) Weighted k-NN, (d) LVQ, (e) Bagging CART, (f) Random Forest.

models, SC classification was markedly worse; for example, the Random Forest dropped from 66.67% (FC) to 40.00% (SC). This highlights SC's greater complexity due to class imbalance and overlapping phenotypes of SC's. In the current study, a key challenge was the severe class imbalance, with the 'green' and 'brown' seed colors representing only 4% of the samples. Simple random train-test splitting risks under-representing these minority classes. Post-split verification confirmed their presence, but small sample sizes compromised the models' ability to predict rare phenotypes, as reflected in the undefined precision and F1-scores (Table 5). The

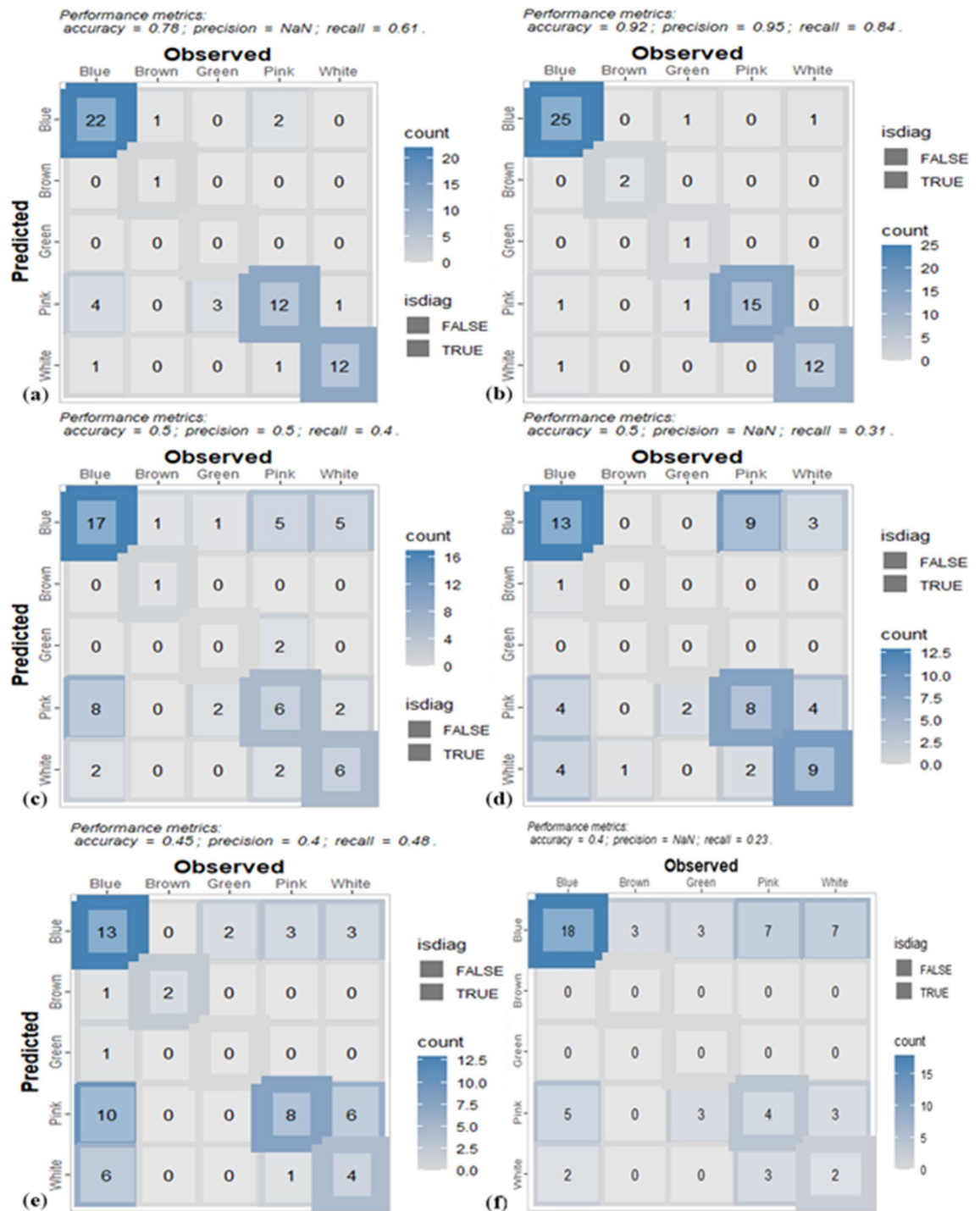


Fig. 5. Confusion matrices of SC. note: (a) Naïve Bayes, (b) SVM, (c) Weighted k-NN, (d) LVQ, (e) Bagging CART, (f) Random Forest.

models performed well for the majority classes, such as blue, white, and pink seeds. Future work should employ synthetic minority oversampling (SMOTE) or cost-sensitive learning to address this imbalance and improve the classification of all phenotypes.

Seeds of the opium poppy are one of the plant's two primary economically valuable products, show considerable variation across countries depending on consumer preferences and dietary habits. Among the morphologically diverse seed color types, white and blue seeds are the most commercially exploited and are regarded as having the highest economic value¹⁹. SC is an additional breeding goal for dietary utilization; opium-poppy seeds have a higher appreciation in the market if they are dark blue⁷. In addition to blue cultivars, specific varieties have been selectively bred for their white seed coat coloration, which is considered advantageous as a potential substitute

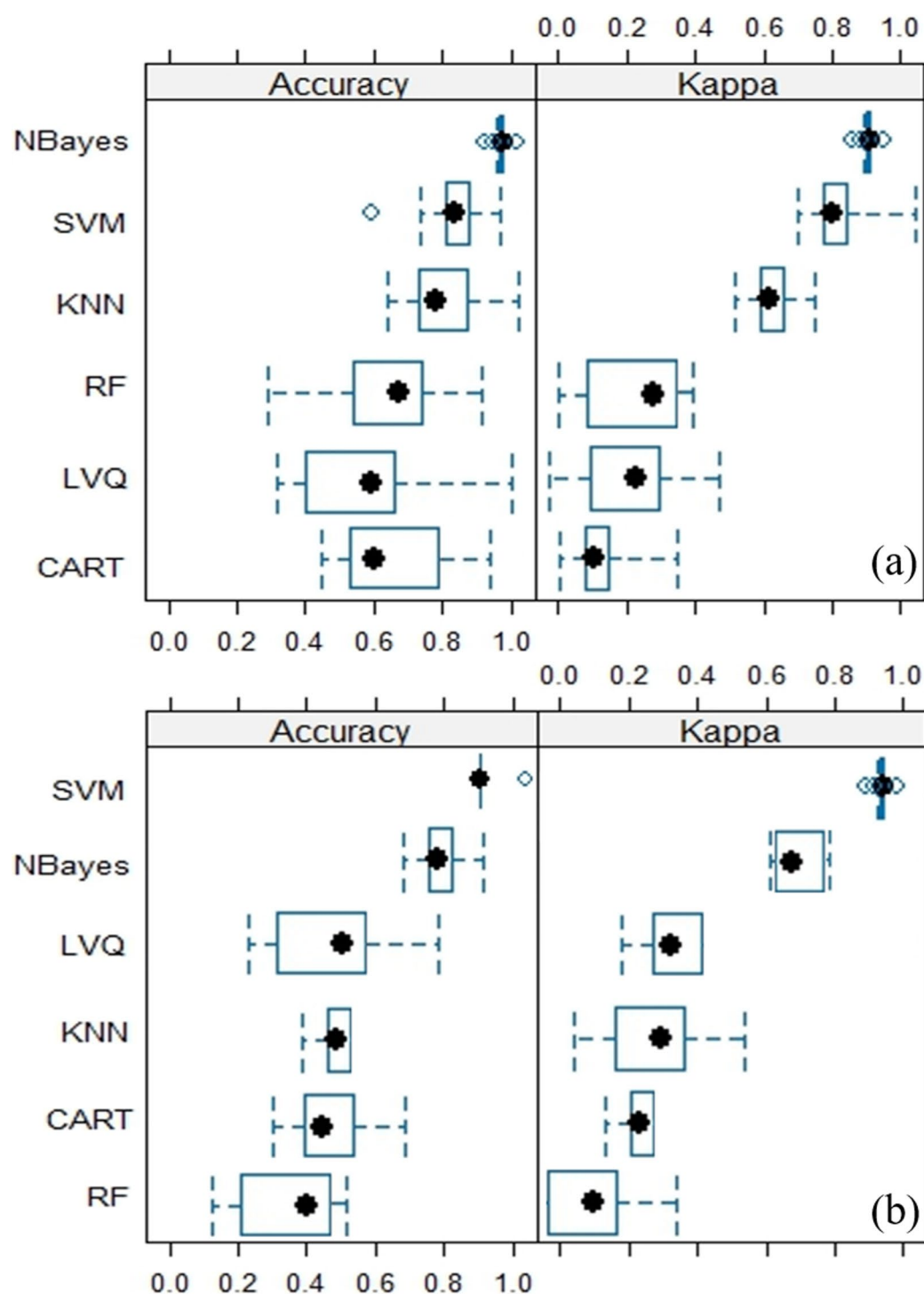


Fig. 6. Comparison of ML algorithms for FC (a) and SC (b). NOTE: CART: Classification and regression tree, KNN: weighted k-nearest neighbor, LVQ: Learning vector quantization, RF: Random forest, SVM: Support vector machine.

for walnuts in bakery applications. Notable examples include *cv.* 'KP-Albakomp'⁴⁸, Bernáth and Németh⁷ noted that *cv.* 'Michalko' and 'Mieszko', which have blue seed color, exhibited a yield potential of 1.00–1.20 t ha⁻¹ and an oil content of 48.00–49.00%. In such cases, classifying the seed color allows the intended use to be determined in advance. Genotypes with irregular flowers or colored petals have ornamental decoration values and may be useful as marker traits for phytochemical characteristics⁷. György, et al.⁴⁹ found that SSR markers were ineffective in distinguishing genotypes based on geographical origin or seed color. However, their analysis revealed some grouping patterns related to variations in the alkaloid content and petal color. This underscores the need for further studies on molecular markers targeting genotypes with diverse alkaloid compositions and

desirable agronomic traits to improve the efficiency of breeding programs. The seed market in Türkiye offers populations characterized by flowers of mixed colors and shapes. Schijfsma, et al⁵⁰, noted that environmental conditions affect seed color differences among opium-poppy varieties during their development.

Conclusion

The current study demonstrated that significant variation exists among opium poppy plants in terms of morpho-phenological characteristics, flower color (FC), and seed color (SC). Using a dataset comprising twenty-three advanced lines and two standard varieties over two growing seasons, multiple ML algorithms were applied to classify FC and SC based on a comprehensive set of plant traits. Naïve Bayes classification yielded more accurate results for FC, with an accuracy of 95.00%. This was followed by the SVM classification at 91.67%. The SVM classification yielded more accurate results, with 91.67% for both SC and FC. Random Forest had an accuracy of 66.67% for FC, whereas bagging CART had an accuracy of 45.00% for SC. PCA revealed that ML classification efficiency was fundamentally linked to the underlying trait architecture; well-separated groups, for example purple vs. white flowers, achieved near-perfect accuracy, whereas overlapping phenotypes (brown seeds) posed greater challenges. This emphasizes the ability of ML to detect biologically coherent patterns beyond superficial color categories. Decision-tree-based ML algorithms had poorer classification performance, especially in classifying SC, compared to FC. The classification of FC and SC enables the intended use to be determined in advance. These findings highlight the effectiveness of ML approaches, particularly probabilistic and kernel-based classifiers, for accurately predicting categorical phenotypic traits in opium poppy plants. These results provide a promising foundation for further integration of ML in plant breeding and germplasm characterization efforts aimed at improving and selecting desirable traits in opium poppy lines. Future research should focus on expanding the dataset to include more genotypes and environmental conditions to enhance the robustness and generalizability of the model. In addition, exploring deep learning techniques or ensemble methods could further improve the classification performance. Integrating genomic or transcriptomic data alongside morpho-phenological traits may also enable more precise phenotype prediction and assist in marker-assisted selection, contributing to more efficient and targeted breeding programs in the opium poppy.

Data availability

The datasets used during the current study are available from the corresponding author on reasonable request.

Received: 5 May 2025; Accepted: 29 January 2026

Published online: 09 February 2026

References

- Díaz-Bárcena, A. & Giraldo, P. Exploring the research evolution of *papaver somniferum* and *cannabis sativa*: A bibliometric comparative analysis. *Ind. Crops Prod.* **203**, 117143 (2023).
- Lal, R. et al. Registration of a high yielding variety CIMAP AJAY morphine rich in concentrate of poppy straw of opium poppy (*Papaver somniferum* L.). *J. Med. Aromat. Plants Sci.* **33**, 204–208 (2011).
- Abhishek, D. & Choudhary, C. K. Genetics of morphine, yield and its candidate characters in opium poppy (*Papaver somniferum* L.). *Electron. J. Plant. Breed.* **1**, 649–655 (2010).
- Kahraman-Yanardağ, Y., Day, S., Bayraktar, N. & Özgen, Y. Evaluation of agronomic and oil characteristics of selected Turkish poppy genotypes under ankara's climate conditions. *Agronomy* **14**, 957 (2024).
- Baser, K. H. C. & Arslan, N. in *Medicinal and Aromatic Plants of the Middle-East* (eds Zohara Yaniv & Nativ Dudai) 305–332 Springer Netherlands, (2014).
- Özcan, M. M. & Atalay, Ç. Determination of seed and oil properties of some poppy (*Papaver somniferum* L.) varieties. *Grasas Y Aceites* **57**, 169–174 (2006).
- Bernáth, J. & Németh, É. (eds) Poppy. *Oil crops*, 449–468 (2009).
- Eyduran, E., Akin, M. & Eyduran, S. Application of multivariate adaptive regression splines through R software. Volume: 1, *Ankara Turkey: Nobel Acad. Publishing*, 256 pg (2019).
- Liakos, K. G., Busato, P., Moshou, D., Pearson, S. & Bochtis, D. Machine learning in agriculture: A review. *Sensors* **18**, 2674 (2018).
- Benos, L. et al. Machine learning in agriculture: A comprehensive updated review. *Sensors* **21**, 3758 (2021).
- Kumar, R., Singh, M., Kumar, P. & Singh, J. in *2015 international conference on smart technologies and management for computing, communication, controls, energy and materials (ICSTM)*. 138–145 (IEEE).
- Sarker, I. H. Machine learning: Algorithms, Real-World applications and research directions. *SN Comput. Sci.* **2**, 160. <https://doi.org/10.1007/s42979-021-00592-x> (2021).
- Rahman, S. A. Z., Mitra, K. C. & Islam, S. M. in *2018 21st international conference of computer and information technology (ICCIT)*. 1–4 (IEEE).
- Tharwat, A. Classification assessment methods. *Appl. Comput. Inf.* **17**, 168–192. <https://doi.org/10.1016/j.aci.2018.08.003> (2020).
- Bhat, S. A., Hussain, I. & Huang, N. F. Soil suitability classification for crop selection in precision agriculture using GBRT-based hybrid DNN surrogate models. *Ecol. Inf.* **75**, 102109 (2023).
- Senapaty, M. K., Ray, A. & Padhy, N. A decision support system for crop recommendation using machine learning classification algorithms. *Agriculture* **14**, 1256 (2024).
- Koklu, M. & Ozkan, I. A. Multiclass classification of dry beans using computer vision and machine learning techniques. *Comput. Electron. Agric.* **174**, 105507 (2020).
- Taner, A. et al. Multiclass Apple varieties classification using machine learning with histogram of oriented gradient and color moments. *Appl. Sci.* **13**, 7682 (2023).
- Özgen, Y., Burucu, D. & History Cultivation, and adaptation of papaver species globally. *Turkish J. Agriculture-Food Sci. Technol.* **13**, 1083–1091 (2025).
- Ajaz, R. H. & Hussain, L. Seed classification using machine learning techniques. *Seed* **2**, 1098–1102 (2015).
- Valizadeh, N., Arslan, N. & Khawar, K. M. Heterosis and heterobeltiosis studies on yield and yield components of some Turkish poppy hybrids (*Papaver somniferum* L.). *J. Appl. Res. Med. Aromatic Plants* **6**, 41–51 (2017).
- Ipek, G. *Some botanical and cultural characteristics of hybrid poppy (Papaver somniferum L.) lines with high noscapine and morphine rates* Ph.D thesis, Ankara University, (2016).
- Javani, M. *Studies on some agronomic and morphological characteristics of altered color seed with the rich noscapine in poppy (Papaver somniferum L.)* Ph.D thesis, Ankara, (2016).

24. Abramoff, M. D., Magalhaes, P. J. & Ram, S. J. "Image Processing with ImageJ". *Biophotonics*. **11**, 36–42 (2004). v. 1.54j. <https://imagej.net/ij/download.html>
25. SSPN. *Soil Analysis Report from Ankara University, Department of Soil Science and Plant Nutrition* (Ankara University, Department of Soil science and Plant Nutrition, 2016).
26. Gupta, S., Saluja, K., Goyal, A., Vajpayee, A. & Tiwari, V. Comparing the performance of machine learning algorithms using estimated accuracy. *Measurement: Sens.* **24**, 100432 (2022).
27. SAS Institute Inc. JMP® Statistical Discovery Software. v. 17.0. (2017).
28. Zhao, Y. & Cen, Y. *Data Mining Applications with R* (Academic, 2013).
29. Martinsen, O. G. & Heiskanen, A. *Bioimpedance and Bioelectricity Basics* (Elsevier, 2023).
30. Majka, M. & Majka, M. M. Package 'naivebayes'. (2020).
31. Vapnik, V. *The Nature of Statistical Learning Theory* (Springer science & business media, 1999).
32. Dimitriadou, E. et al. The e1071 package. *Misc Functions of Department of Statistics (e1071)*, TU Wien 297 (2006).
33. Schliep, K., Hechenbichler, K., Lizée, A. & Schliep, M. K. Package 'knn'. R Package version 1.4.7, CRAN R Project (2016).
34. Melin, P., Amezcua, J., Valdez, F. & Castillo, O. A new neural network model based on the LVQ algorithm for multi-class classification of arrhythmias. *Inf. Sci.* **279**, 483–497 (2014).
35. Frick, H. et al. Rsample: General Resampling Infrastructure. R Package version 1.1. 1. (2022).
36. Cebeci, Z., Tekeli, E., Tahtali, Y. & Tarım Gıda ve Yaşam bilimlerinde R ile makine Öğrenmesi ve veri Madenciliği. Volume:1, Ankara Turkey: Nobel Acad. Publishing, 668 pg (2022).
37. Breiman, L. Random forests. *Mach. Learn.* **45**, 5–32. <https://doi.org/10.1023/A:1010933404324> (2001).
38. Bischl, B. et al. Mlr: machine learning in R. *J. Mach. Learn. Res.* **17**, 1–5 (2016).
39. Breiman, L., Cutler, A., Liaw, A. & Wiener, M. Package 'randomforest'. *Univ. Calif. Berkeley: Berkeley CA USA*. **81**, 1–29 (2018).
40. Kuhn, M. & Johnson, K. *Applied Predictive Modeling* Vol. **26** (Springer, 2013).
41. Correndo, A. A. et al. MetRica: an R package to evaluate prediction performance of regression and classification point-forecast models. *J. Open. Source Softw.* **7**, 4655 (2022).
42. Rácz, A., Bajusz, D. & Héberger, K. Multi-level comparison of machine learning classifiers and their performance metrics. *Molecules* **24**, 2811 (2019).
43. Brownlee, J. Imbalanced classification with python. *Machine Learn. Mastery* **463**, 445 (2020).
44. Rumora, L., Miler, M. & Medak, D. Impact of various atmospheric corrections on sentinel-2 land cover classification accuracy using machine learning classifiers. *ISPRS Int. J. Geo-Information*. **9**, 277 (2020).
45. Zhou, T., Pan, J., Zhang, P., Wei, S. & Han, T. Mapping winter wheat with multi-temporal SAR and optical images in an urban agricultural region. *Sensors* **17**, 1210 (2017).
46. Le, T. T. & Moore, J. H. Treeheatr: an R package for interpretable decision tree visualizations. *Bioinformatics* **37**, 282–284 (2021).
47. Del Moral, P., Nowaczyk, S. & Pashami, S. Why is multiclass classification hard? *IEEE Access*. **10**, 80448–80462 (2022).
48. Németh, É. World tendencies, aims and results of poppy (*Papaver somniferum* L.) breeding. *Ser. Recent. Progress Med. Plants*. **4**, 129–141 (2002).
49. György, Z., Alam, S. & Priyanka, P. Zámoriné Németh, É. Genetic diversity and relationships of opium poppy accessions based on SSR markers. *Agriculture* **12**, 1343 (2022).
50. Schijfsma, L., Hoesbergen, M. & Nijdam, F. A study of the colour and other characters of the seed of some varieties of oil seed poppy. *Euphytica* **9**, 127–140 (1960).

Author contributions

Conceptualization, U.O. and Y.O.; methodology, U.O. and Y.O.; software, U.O.; validation, Y.O., N.B. and U.O.; formal analysis, U.O.; investigation, Y.O.; resources, Y.O.; data curation, U.O.; writing—original draft preparation, U.O. and Y.O.; writing—review and editing, U.O. and Y.O.; visualization, U.O.; supervision, N.B.

Funding

This research received no external funding.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to U.O.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2026