Selcuk Journal of Agriculture and Food Sciences
https://dergipark.org.tr/tr/pub/selcukjafsci
Research Article

(2025) 39(1), 95-107 DOI:10.15316/SJAFS.2025.009 e-ISSN: 2458-8377 Selcuk J Agr Food Sci

Classifying *Apis mellifera* Breeds Using Data Mining Techniques Based on Morphological Traits

Mustafa Kibar[1,2,*], **İnci Şahin Negiş**[3], **İbrahim Aytekin**[4], **İsmail Keskin**[4]

[1]Siirt University, Faculty of Agriculture, Department of Animal Science, Siirt, Türkiye

[2]Siirt University, Beekeeping Application and Research Center, Siirt, Türkiye

[3]Selçuk University, Faculty of Agriculture, Department of Plant Protection, Konya, Türkiye

[4]Selçuk University, Faculty of Agriculture, Department of Animal Science, Konya, Türkiye

HIGHLIGHTS

- The classification of bee breeds is crucial for genetic diversity and efficiency.
- The morphological traits of honeybee breeds are different.
- The random forest model achieved 99.8% success, proving highly effective for classification.
- Naïve Bayes consistently performed the worst across all evaluation metrics.

Abstract

The classification of bee breeds is significant for breeding, maintaining genetic diversity, increasing productivity and protecting the health of the bee colonies. Therefore, this study aims to classify different honeybee breeds based on their morphological traits using data mining techniques, which are cost-effective and straightforward. It were used a total of 35 colonies from a private bee farm for morphometric analysis in the study, which included seven different bee breeds and 404 bee samples. A range of data mining techniques (Support Vector Machines (SVM), Random Forest (RF), Artificial Neural Networks (ANN), Naive Bayes (NB) and k-Nearest Neighbors (k-NN)), and model fit criteria were used for the classification of bee breeds. Overall, the study shows significant differences in the morphological traits of different bee breeds, highlighting the diversity and different traits of each bee breed. In addition, the study shows that the RF model is superior in all criteria and therefore the most effective for classifying honeybee breeds. In contrast, the NB model consistently performs the worst, as evidenced by the consistently minimum values of all metrics. In conclusion, RF model exhibiting a 99.8% success rate, stands out as highly effective in the classification of bee breeds based on the morphological traits, supporting its applicability in future classification research.

Keywords: bee breeds, classification, data mining, morphological traits, performance metrics

Citation: Kibar M, Şahin Neğiş İ, Aytekin İ, Keskin İ (2025). Classifying *Apis mellifera* breeds using data mining techniques based on morphological traits. *Selcuk Journal of Agriculture and Food Sciences*, 39(1), 95-107. https://doi.org/10.15316/SJAFS.2025.009
***Correspondence:** mustafakibar@siirt.edu.tr

Received date: 14/12/2024
Accepted date: 10/02/2025

1. Introduction

Bees are essential pollinators that contribute significantly to global food production and biodiversity (Potts et al. 2010). For sustainability in these; the classification of bee breeds is significant, which has a crucial role in understanding their behavior, ecological impacts, and potential for various applications in agriculture and environmental conservation. However, traditional methods of classification have relied on manual observation and subjective judgment, often leading to inconsistencies, misclassification (Berlocher 1984) and high costs (Tapkan et al. 2016). In addition, accurately classification of bee breeds only based on their morphological traits (MTs) could be difficult due to their few variations. Accordingly, the application of data mining methods has emerged as a promising approach to overcome these challenges in recent years. Data mining methods can be defined as the process of extracting previously unknown information from large-scale databases and utilizing it effectively in the decision-making process using computational techniques, statistical analysis, and machine learning algorithms (Tapkan et al. 2016; Han et al. 2022). Because of these features, it offers a more objective and systematic approach to classification of bee breeds. The primary data mining methods Support Vector Machines (SVM), Random Forest (RF), Artificial Neural Networks (ANN), Naive Bayes (NB) and k-Nearest Neighbors (k-NN) were investigated. SVM is a powerful supervised learning technique that constructs a hyperplane (Vapnik 1999) to classify different bee breeds in a high-dimensional feature space. RF is an ensemble learning method that combines multiple decision trees to improve classification accuracy (Breiman 2001). ANN, inspired by the biological neural network, utilizes interconnected layers of artificial neurons to learn (Yegnanarayana 1999) and classify bee breeds based on their MTs. NB is based on Bayes' theorem, which assumes that the presence of a particular feature is independent of the presence of any other trait in an insect class (Antony and Pratheepa 2018). k-NN classifies based on their proximity to the nearest neighbors in the feature space (Cover and Hart 1967) and the k-NN algorithm achieves a better identification rate for species (Li and Xiong 2018). By harnessing the power of these data mining methods, researchers and beekeepers could improve the efficiency and accuracy of breed classification processes. In the classification studies, the MTs of bees include length of the wing, width of the wing, length of the tongue and other distinguishing traits (Buco et al. 1987; Rinderer et al. 1993; Crewe et al. 1994; Ftayeh et al. 1994; Diniz-Filho and Malaspina 1995; Szymula et al. 2010). However, the MTs of bees can be generally categorized into three main groups: length measurements, color measurements, and wing venation traits (Abou-Shaara et al. 2013) and the third were investigated in detail by Abou-Shaara (2013). The classification studies in terms of the MTs efficiently compile and organize a vast amount of data regarding variations (Alpatov 1929; Guler and Bek 2002). Additionally, they assess the significance of these variations for honeybee populations (Estoup et al. 1995), facilitate visual comparisons in pure breed breeding through morphometric data (Gençer 2004), and distinguish between geographically adjacent types (Ruttner et al. 1978). For example, Frunze et al. (2022) studied to regularize by six best traits (width of the abdomen, length of tergites 3 and 4, width of the head, length of the antenna, and length of the forewing). Accurate breed classification could enable more targeted management practices, selective breeding programs, informed decision-making in areas such as pollination services, disease management, and ecological conservation.

The aim of this study is to determine whether bee breeds can be effectively classified using various data mining techniques such as SVM, RF, ANN, NB and k-NN, based on the analysis of certain MTs (e.g., scutellum color (SC), second abdominal tergit color (TC-2), fourth abdominal tergit color (TC-4), length of the tibia (LTI), length of the wing (LW), width of the wing (WW), cubital index (CI), width of the vein-a (WV-a), width of the vein-b (WV-b), wing vein angles (J10, B4, A4, E9), length of the tongue (LTO), length of the glossa (LG), length of the prementum (LPR) and length of the postmentum (LPM)) obtained from a diverse set of bee breeds including Pure Anatolia, Anatolia Hybrid, Carniolan Hybrid, Carpathian Hybrid, Caucasian hybrid, Italian Hybrid, and Kerkuk bees. The findings of this study may have practical implications for beekeepers, researchers, interested in improving bee breed identification as well as understanding the ecological roles and behaviors of different bee breeds.

2. Materials and Methods

This study was conducted in *Apis mellifera L.* (*Hymenoptera*: *Apidae*) honeybees raised Bozkır district (Latitude: 37.171860 and Longitude: 32.216846), Konya province, Türkiye. In the study, some MTs of seven different honeybee breeds in thirty-five colonies were used for breed classification. For the Carniolan hybrid: 1 colony and 27 bee samples; for the Kerkuk: 1 colony and 10 bee samples; for the Anatolian hybrid: 2 colony and 21 bee samples; for the Italian hybrid: 2 colony and 20 bee samples; for the Pure Anatolian: 1 colony and 23 bee samples; for the Carpathian hybrid: 3 colonies and 32 bee samples; and finally, for the Caucasian hybrid: 25 colonies and 271 bee samples were used. Totally, 35 colonies, 7 different breeds and 404 bee samples were used for the study. No morphometric analyzes and/or any practices related to the MTs were performed prior to the study to determine the breeds of used in the study; The colonies and bee samples were randomly selected. The collection of bee samples and testing techniques followed the animal welfare guidelines outlined in Article 9 of the “Veterinary Services, Plant Health, Food, and Feed Law” of Türkiye. The colonies used in the study were raised in ten-frame Langstroth hives in Konya during the summer months (May-September) and Antalya province during the winter months (October-April). The colonies were fed with sugar syrup and bee cake, except in January and February when temperatures are not suitable for feeding. In the current apiary, pure queens naturally mate with different drones and form hybrid colonies. For example, by mating pure Caucasian queens with different drones, Caucasian hybrids were formed and new queens were bred from these hybrid colonies. Colonies with pure queens were formed through purchases. Also, Kerkuk bees were used in the study to strengthen morphological distinction.

Bee samples from the colonies were taken in the summer of 2020 and were placed in alcohol and stored at +4 °C until morphological analysis. The specimens were initially preserved in 70% ethanol until the dissection. Subsequently, the forewings, proboscis and legs of honeybees were precisely severed and sequentially immersed in 60%, 70%, 80%, 90% and 100% ethanol solutions for gradual hydration. For all parts measurements, after immersing in alcohol series, were maintained in clove oil at least for 24 hours to facilitate tissue softening and enhance resolution. Following the dissection, the forewings, proboscis, and legs were carefully positioned in the microscope slides with hoyer medium (Distilled water, Gum Arabic, Chloral hydrate, and Glycerin) developed by Anderson (1954). The MTs scutellum color (SC); second abdominal tergit color (TC-2); fourth abdominal tergit color (TC-4); length of the tibia (LTI); length of the wing (LW); width of the wing (WW); cubital index (CI); width of the vein-a (WV-a); width of the vein-b (WV-b); wing vein angles (J10, B4, A4 and E9); length of the tongue (LTO); length of the glossa (LG); length of the prementum (LPR); and length of the postmentum (LPM) of honeybees were used. CI was calculated following equation reported by Özbakır (2011).

$$CI = \left(\frac{WV - a}{WV - b}\right) \quad (1)$$

In the equation, WV-a is the width of the vein-a and WV-b is the width of the vein-b. While SC was determined using a 5-point color scale from 0 to 4, TC-2 and TC-4 were determined using a 10-point color scale from 0 to 9 (Dodoloğlu 2000). Each bee sample was scored by two assessors and the average value was used for the evaluation. Optical microscopic observations and photography were conducted using an integrated 16-megapixel digital camera (AmScope MU1603, Irvine, CA, USA) under 20X magnification. All comparative morphological observations and the preparation of slides for colony morphology were carried out with the assistance of a Nikon SMZ 745T stereo microscope, specifically the stereomicroscope equipped with a G-AL-2X objective. This microscope model offered essential magnification capabilities and stereo vision, facilitating in-depth examination, including color detection, and precise documentation of MTs in the specimens under investigation. To conduct measurements, the 3-point circle measurement, arc measurement, and line measurement tools available in AmScope software version x64, 4.8 were utilized. These tools provided a comprehensive means of quantifying and analyzing various aspects of the captured images (Figure 1).

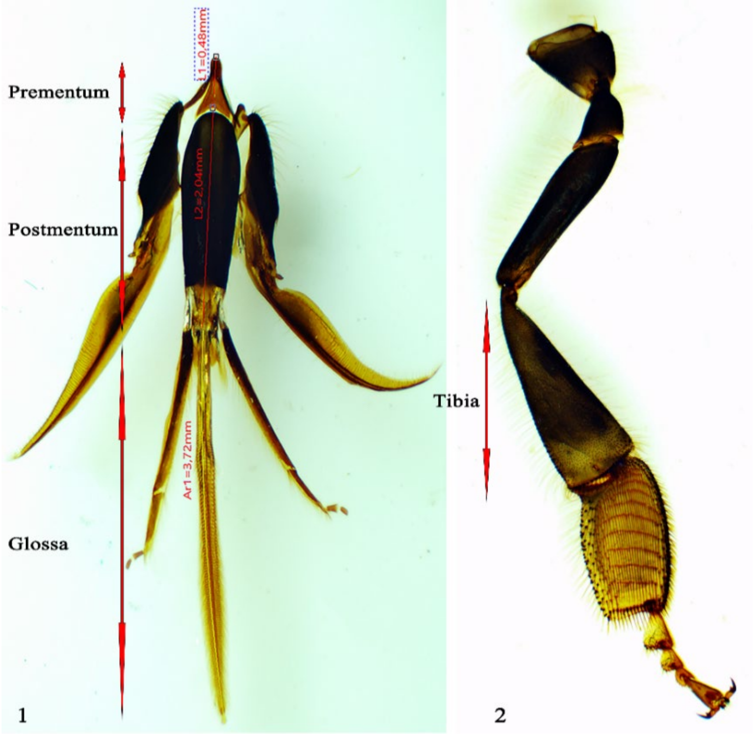

Figure 1. The analyzed some morphological traits. The honeybee tongue parts: (1) the prementum, the postmentum, the glossa. The honeybee leg part: (2) the tibia

Statistical analyzes were made using the ORANGE program version 3.3.1. (Demsar et al. 2013). Five different data mining algorithms SVM, RF, ANN, NB and k-NN models were used for classification based on the MTs of seven different bee breeds. A 70% training – 30% test ratio was used in ANN. Classification methods allow test data to be accurately assigned to specific classes by learning a model from the training data set (Bhavsar and Ganatra 2012). The data mining classification methods (SVM, RF, ANN, NB and k-NN) and model fit criteria (AUC, CA, F1, Precision, Recall and Specificity) used are explained below. SVM belongs to the family of generalized linear classifiers and is a supervised learning method for classification and regression. In other words, SVM is a classification and regression prediction method that leverages machine learning theory to maximize predictive power (Khan et al. 2023). Breiman (2001) developed the RF method in which different classification assignment algorithms are used together. RF combines multiple decision trees, and the majority vote based on each tree's predictions together determines the final prediction (Breiman 2001). Additionally, it is quite user-friendly in that it has only two parameters, the number of variables in the random subset at each node and the number of trees in the forest, and is generally not very sensitive to their values (Liaw and Wiener 2002). ANN, an information processing system, simulates the thinking system of the human brain. The fact that ANN is non-linear and designed to perform its tasks in a similar manner to the human brain makes ANN unique, and the sigmoid (logistic) activation function was used in this study. Therefore, it is a suitable method for processing the signals sent by various sensors and communication devices (Pan et al. 2023). NB is a well-known probabilistic classification algorithm that is simple but efficient and has a wide range of real-world applications including product recommendations, medical diagnostics, and autonomous vehicle control (Wickramasinghe and Kalutarage 2021). The nearest neighbor rule assigns an unclassified sample point the classification of the closest point from a set of previously classified points (Cover and Hart 1967). Given an unlabeled example, k-NN classifier searches for the k objects in the sample space that are closest to it and assigns the class based on the most common class label among them (Bhavsar and Ganatra 2012). If the value of k is too small, the k-NN classifier may be prone to overfitting due to noise in the training data set. On the other hand, if k is too large, the nearest neighbor classifier might misclassify the test example because the nearest neighbor list may contain data points that are far from the neighborhood. A value of k=1 is called nearest neighbor classification (Jadhav and Channe 2016). Model comparison was performed using the six following criteria. First, Area Under the Curve (AUC) is a criterion for measuring the performance of classification models. It represents the area under the Receiver Operating Characteristic (ROC) curve. AUC takes a value between 0 and 1 and indicates how well the classifier is performing. A value close to 1 indicates high accuracy, while a value close to 0.5 corresponds to random classification and indicates low model performance. Second, Classification Accuracy (CA) is a commonly used criterion for measuring the accuracy of classification models and represents the model's ability to classify correctly. It is calculated by dividing the total number of correctly classified examples by the total number of examples, ranges from 0% to 100%. A CA of 0% means that the model has not made any correct classifications, while a CA of 100% indicates that the model has classified all instances correctly. In general, a CA value above 80% is considered sufficient. Third, F1 score is the harmonic mean of a classification model's precision and recall measurements. Precision measures the ratio of true positive predictions to the overall prediction, and recall measures the rate of true positive predictions detected. The F1 score effectively measures the accuracy of the model by calculating these two criteria together. The F1 score takes a value between 0 and 1. The higher the value, the better the performance of the model. Fourth, Precision measures the ratio of truly positive predictions to the overall prediction. This evaluates how accurately the classifier identifies positive examples. Precision takes a value between 0 and 1, and a higher value indicates how accurately the model predicts positive classes. Fifth, Recall is a criterion for measuring the performance of a classification model and indicates how accurately the model detects true positives. Recall takes a value between 0 and 1, and a higher value indicates how accurately the model detects true positives. Finally, specificity typically is defined as the ability of a screening test to detect a true negative, based on the true negative rate, and to correctly exclusion of irrelevant MTs (Trevethan 2017).

3. Results

Morphological breed discrimination in honeybees is important for improving productivity, disease resistance, climate adaptation, and behavioral traits in beekeeping. Also, it is vital in determining genetic diversity and maintaining biological diversity worldwide. The means and standard deviations for seventeen different MTs of seven bee breeds used in the study were presented (Table 1). Table 1 shows an extensive comparison of different traits in different bee breeds. The fact that the scutellum color (3.60±0.52 pt) and TC-2 (9.0±0.0 pt) are higher in the Kerkuk breed than in other breeds indicates a distinctive trait of this breed. Conversely, TC-4 was lower in the Caucasian hybrid (1.19±0.89 pt) and Pure Anatolian (1.2±0.78 pt) breeds. It was determined that Italian hybrid bees had higher values in terms of LW (8.65±0.17 mm) and WW (2.97±0.08 mm). These various values may indicate specific traits or deficiencies of these breeds compared to others. Overall, the data shows significant differences in the traits of different bee breeds, highlighting the diversity and different traits of each bee breed.

In the present study, it can be seen that there are differences between the breeds in terms of the mean values of the MTs. It is noteworthy that SC and TC-2 are higher in the Kerkuk breed compared to other breeds. The observed differences in LTI, wing dimensions and CI suggest differences in flight abilities and foraging behavior between the breeds. Additionally, WV and wing vein angles are important traits that indicate the structural integrity and aerodynamics of the wings and are critical for efficient flight and maneuverability. The measurements of LTO and LG as well as LPR and LPM provide insights into the nutritional mechanisms and possible flowering preferences of the breeds. Morphological distinctions between bee breeds utilized in the current population could be regarded as an important gene pool for pollination efficiency, nectar collection, maintenance of ecological balance and sustainability of biodiversity. Gençer and Günbey (2020) reported that there were statistically significant differences between bees in the Caucasian, Yığılca and Korgan groups in terms of LTO, WW, LW and CI (P<0.001). Same researchers stated that the Caucasian breed is higher in LTO, but lower in other traits than other groups. Kambur and Kekeçoğlu (2018) found that the average of J10, B4, A4, and E9 vein angles were 53.86±0.09, 101.75±0.11, 33.69±0.05 and 19.22±0.03, respectively.

Table 1. The means and standard deviations of morphological traits of breeds

Traits	Breeds							Mean
	Carniolan hybrid	Pure Anatolian	Caucasian hybrid	Italian hybrid	Anatolian hybrid	Carpathian hybrid	Kerkuk	
SC (pt)	1.58±1.31	1.33±1.18	1.53±1.28	2.35±1.54	1.056±1.06	1.27±1.34	3.60±0.52	1.58±1.33
TC-2 (pt)	6.92±2.39	2.67±2.19	4.8±3.35	6.82±3.21	4.17±3.26	4.83±3.70	9.0±0.0	4.99±3.39
TC-4 (pt)	1.75±1.06	1.2±0.78	1.19±0.89	1.94±1.44	1.5±0.99	2.03±1.07	1.6±0.70	1.37±0.98
LTI (mm)	3.08±0.06	3.07±0.06	3.06±0.09	3.1±0.10	3.05±0.08	3.09±0.08	2.9±0.08	3.06±0.09
LW (mm)	8.64±0.10	8.56±0.12	8.52±0.20	8.65±0.17	8.47±0.24	8.43±0.14	8.49±0.21	8.52±0.19
WW (mm)	2.98±0.09	2.97±0.05	2.93±0.10	2.97±0.08	2.92±0.07	2.9±0.06	3.01±0.08	2.94±0.09
CI (mm)	2.51±0.36	2.26±0.39	2.25±0.41	2.36±0.53	2.2±0.30	2.36±0.37	2.74±0.61	2.29±0.42
WV-a (mm)	0.53±0.05	0.52±0.05	0.52±0.05	0.52±0.06	0.51±0.03	0.53±0.04	0.55±0.03	0.52±0.05
WV-b (mm)	0.21±0.02	0.23±0.03	0.23±0.03	0.23±0.03	0.24±0.03	0.23±0.02	0.21±0.04	0.23±0.03
J10 (°)	54.69±5.16	51.47±3.38	52.6±3.99	52.09±4.29	50.55±2.97	50.63±3.64	52.25±1.78	52.28±3.96
B4 (°)	105.94±5.31	97.75±5.58	102.44±6.65	102.99±5.11	99.56±5.38	105.19±6.97	106.65±4.07	102.62±6.58
A4 (°)	31.23±1.77	33.81±2.44	31.37±3.10	31.14±2.54	32.5±2.09	31.51±5.77	30.95±1.43	31.53±3.31
E9 (°)	20.96±1.25	17.9±1.29	19.42±1.53	19.84±1.33	19.72±1.23	20.47±1.46	22.08±1.38	19.63±1.62
LTO (mm)	6.01±0.56	5.67±0.65	5.9±0.60	6.01±0.53	6.03±0.58	6.1±0.45	5.7±0.55	5.92±0.58
LG (mm)	3.47±0.49	3.25±0.52	3.41±0.50	3.48±0.41	3.56±0.48	3.55±0.38	3.32±0.44	3.43±0.48
LPR (mm)	0.48±0.06	0.53±0.04	0.48±0.13	0.5±0.05	0.51±0.04	0.51±0.05	0.41±0.09	0.49±0.11
LPM (mm)	2.06±0.10	1.89±0.16	2.01±0.19	2.04±0.14	1.97±0.15	2.04±0.12	1.98±0.11	2.01±0.17

SC: Scutellum color; TC-2: Second abdominal tergit color; TC-4: Fourth abdominal tergit color; LTI: Length of the tibia; LW: Length of the wing; WW: Width of the wing; CI: Cubital index; WV-a: Width of the vein-a; WV-b: Width of the vein-b; J10, B4, A4 and E9: Wing vein angles; LTO: Length of the tongue; LG: Length of the glossa; LPR: Length of the prementum; LPM: Length of the postmentum; pt: Point; mm: Millimeter; °: Angle degree

Table 2. Classes assigned to bee breeds by different classification models

Models		Breed	Predicted							Total
			1	2	3	4	5	6	7	
SVM	Actual	1	2	0	25	0	0	0	0	27
		2	0	0	23	0	0	0	0	23
		3	0	0	271	0	0	0	0	271
		4	0	0	20	0	0	0	0	20
		5	0	0	21	0	0	0	0	21
		6	0	0	29	0	0	3	0	32
		7	0	0	4	0	0	0	6	10
	Total		2	0	393	0	0	3	6	404
RF	Actual	1	22	0	5	0	0	0	0	27
		2	0	22	1	0	0	0	0	23
		3	0	0	271	0	0	0	0	271
		4	0	0	5	15	0	0	0	20
		5	0	0	3	0	18	0	0	21
		6	0	0	6	0	0	26	0	32
		7	0	0	1	0	0	0	9	10
	Total		22	22	292	15	18	26	9	404
ANN	Actual	1	12	0	15	0	0	0	0	27
		2	1	7	15	0	0	0	0	23
		3	2	2	264	0	2	0	1	271
		4	0	0	18	2	0	0	0	20
		5	0	0	16	0	4	1	0	21
		6	0	0	23	0	2	7	0	32
		7	0	0	2	0	0	0	8	10
	Total		15	9	353	2	8	8	9	404
NB	Actual	1	20	1	2	0	0	0	4	27
		2	0	18	3	1	0	1	0	23
		3	15	29	135	26	26	13	27	271
		4	0	3	1	10	1	2	3	20
		5	0	2	2	1	15	1	0	21
		6	1	2	3	4	5	13	4	32
		7	0	0	0	0	0	0	10	10
	Total		36	55	146	42	47	30	48	404
k-NN	Actual	1	17	0	9	0	0	0	1	27
		2	1	9	13	0	0	0	0	23
		3	7	7	256	0	0	1	0	271
		4	2	1	13	3	0	0	1	20
		5	3	2	14	0	2	0	0	21
		6	1	1	25	0	0	5	0	32
		7	1	0	3	1	0	0	5	10
	Total		32	20	333	4	2	6	7	404

Kekeçoğlu et al. (2023) found that the average J10, B4, A4, and E9 vein angles were 53.63±0.80, 99.63±1.16, 33.78±0.49 and 19.87±0.33, respectively. Cariveau et al. (2018) found that the average LG and LPR were 2.05 mm and 1.57 mm. The results of the current study were found to be generally similar to Gençer and Günbey (2020) and different from Kambur and Kekeçoğlu (2018). However, the current study was found to be higher than Kekeçoğlu et al. (2023) and Cariveau et al. (2018) in terms of B4 vein angle and LG, respectively, and lower than Cariveau et al. (2018) in terms of LPR. Furthermore, Souza et al. (2002) reported that there were statistically significant correlations between tibia width (0.46), LTI (0.55), and LG (0.41) with honey production. Zemskova et al. (2020) stated that the necessity for thorough monitoring of honeybee MTs arises from the declining bee populations and the reduced marketability of apiaries. When the results of the current study are evaluated together with the literature, it becomes evident that identifying and classifying the MTs of bees is important for the sustainability of the bee industry and genetic resources, as well as for increasing productivity per colony.

In addition, the classification of bee breeds was carried out and presented using some data mining algorithms (Table 2). A matrix showed how each model classified the different bee breeds (numbered 1 to 7). In these matrices, the actual and model-predicted breed class is in the rows and columns, respectively. The values in each cell of the matrix represent the frequency with which the model predicted each actual bee breed class as a certain class. For instance, in the case of the SVM model, out of the bees that truly belonged to breed class "1", the model predicted "1" for two bees and "3" for twenty-five bees. Comparing the number of samples correctly assigned to each class shows the classification success of the models. This table can be considered as an accuracy matrix that shows how accurately different data mining algorithms classify bee breeds. This analysis can be used to determine which model is more effective in automatically classifying bee breeds.

After the classification analysis, model fit criteria were used to evaluate the success of the models. The evaluation metrics for comparison of the models are presented in Table 3. Table 3 compares the performance metrics of various machine learning models SVM, RF, ANN, NB, and k-NN by their AUC, CA, F1, Precision, Recall and Specificity values.

In addition, Figure 2 is provided to facilitate a deeper comprehension of the classification performance of the models. Figure 2 indicate the models' performances and predictions. RF model is found to be most efficient, with the highest AUC values of 0.998, indicating superior ability in distinguishing between breeds. Additionally, it has the CA of 0.948 and an impressive F1 score of 0.946, highlighting its balanced precision and recall. The excellence of this model indicates its robustness in tackling complex classification tasks in beekeeping research. Conversely, NB model is identified as the least effective model, as evidenced by the lowest CA of 0.547 and a modest AUC of 0.853. Despite its simplicity and ease of implementation, the performance of the NB model highlights limitations in addressing the intricacies of bee classification, particularly in scenarios that require fine-grained differentiation between similar species. The SVM and k-NN models exhibit commendable performance with AUC values of 0.950 and 0.919, respectively, positioning them as viable alternatives for certain applications. However, they do not exceed the overall effectiveness of the RF model. Although the ANN model shows potential with an AUC of 0.905, it falls short of expectations compared to RF in terms of CA and F1 score, suggesting that further optimization is required to improve classification accuracy.

Table 3. The evaluation metrics used for model comparison

Model	AUC	CA	F1	Precision	Recall	Specificity
SVM	0.950	0.698	0.589	0.633	0.698	0.385
RF	0.998	0.948	0.946	0.952	0.948	0.894
ANN	0.905	0.752	0.703	0.766	0.752	0.549
NB	0.853	0.547	0.584	0.744	0.547	0.922
k-NN	0.919	0.735	0.688	0.750	0.735	0.607

AUC: Criterion for measuring the performance of classification models; CA: Criterion for measuring the accuracy of classification models; F1: Harmonic mean of a classification model's precision and recall measurements; Precision: Measures the ratio of truly positive predictions to the overall prediction; Recall: Criterion indicates how accurately the model detects true positives; Specificity: Criterion indicates how accurately the model detects true negatives

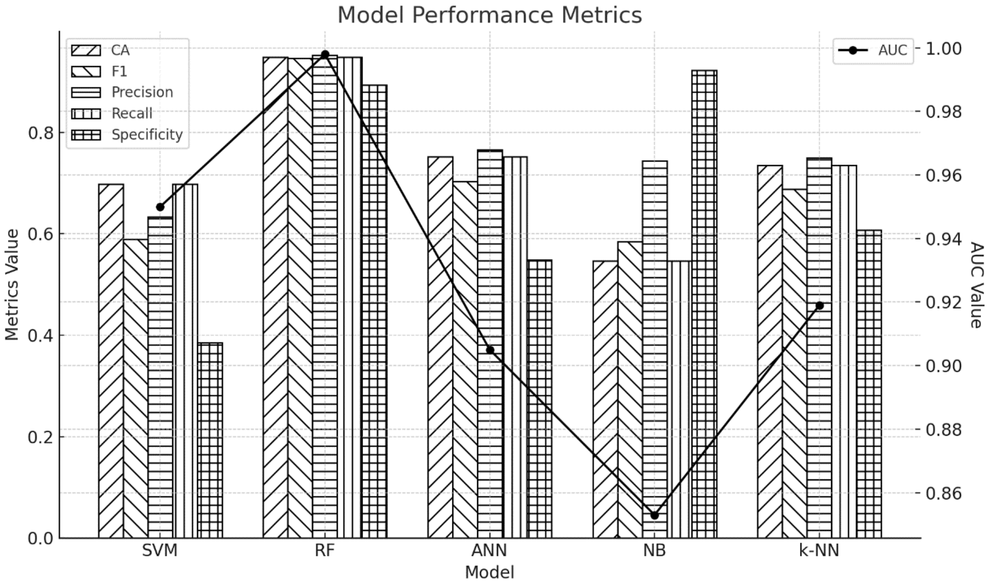

Figure 2. Evaluation of machine learning models in bee classification

AUC represents the area under the ROC curve, indicating the balance between sensitivity and specificity of a classification model. A higher AUC value, closer to 1, signifies a better balance and, consequently, a more accurate classification. Among these models, RF exhibited the highest AUC value (0.998), indicating its superior performance in classifying bee breeds based on their MTs. Conversely, NB demonstrated the lowest AUC value (0.853), suggesting a comparatively lower accuracy in classifying bee breeds. These results highlight that the RF was the most effective model for classifying bee breeds. CA calculates the ratio of correct predictions to the total number of samples and could be misleading for imbalanced class data. Since CA considers total correct predictions rather than distinguishing between classes, it can lead to erroneous interpretations. These values indicate that, in classifying bee breeds based on their MTs, RF demonstrated superior accuracy, outperforming other models, whereas NB exhibited comparatively lower accuracy. In this study, despite the CA determined by the SVM model being lower than the ANN model, the AUC value was found to be higher for the ANN model. This situation underscores the importance of not relying solely on CA as a singular criterion for evaluation. Although the SVM model made fewer correct predictions, it better explained the differences between bee breeds. This observation highlights that the performance of the model is not solely contingent on the number of correct predictions. The F1 score is a balanced measure that takes precision and recall into account and is calculated using the harmonic mean of these two metrics. A high F1 score plays a crucial role in distinguishing true positives (correct identification of the MTs under consideration) from true negatives (correct exclusion of irrelevant MTs). When false positives (misclassifying non-relevant traits as relevant) and false negatives (misclassifying relevant traits as not relevant) are of equal importance, the F1 score becomes a significant metric. In the context of beekeeping, the F1 score is of great importance, especially in applications such as selection processes that influence the efficiency of bee breed. According to the F1 score, RF exhibited the highest classification accuracy, while NB demonstrated the lowest performance among the models. Therefore, when it is essential to precisely differentiate between the MTs of bee breeds, choosing the model with the highest F1 score is crucial. The precision metric signifies the proportion of samples predicted as positive (possessing a specific feature) by the model that are actually positive. This measure holds particular importance in scenarios where false positive predictions (predicting a feature that is not actually present as present) carry significant implications. In the context of apiculture, when selecting colonies based on the MTs of disease-resistant bee breeds, misclassifying non-disease-resistant bees as disease-resistant (resulting in false positive predictions) can markedly diminish the efficacy of the selection process. Consequently, RF exhibited the highest accuracy in identifying colonies with the desired MTs, whereas SVM demonstrated the lowest performance in this aspect. The recall metric indicates the ratio of true positives (correctly identified traits) to all actual positives. If identifying the presence of a specific trait (false negative prediction) holds significance, the recall value becomes crucial. For instance, in apiculture, when selecting hygienic colonies based on MTs, including non-hygienic colonies in the selection can increase the colony's susceptibility to diseases, making false negative predictions important in this context. Therefore, it is essential to consider a model with a high recall value. Consequently, the RF model, with the highest recall value, will be more effective in selecting hygienic colonies. The NB model, which yielded the lowest recall value, should not be considered due to the high likelihood of misclassifying non-hygienic colonies as hygienic ones. Significant differences have been found in the data mining algorithms used to classify breeds of bees based on their MTs. When all criteria used in the classification were evaluated together, it was found that the RF model achieved the highest values for all criteria and performed the best classification. The RF model is followed by ANN. On the other hand, the lowest ranking was generally achieved in the NB model, which consistently had the lowest values. The RF model has demonstrated superior performance in key metrics such as AUC, CA, F1 score, precision, and recall, establishing itself as an effective method for classifying bee breeds. However, other algorithms such as SVM, ANN, NB, and k-NN have shown lower and varying degrees of success in classification compared to the RF model. Rodrigues et al. (2022) reported that the SVM algorithm classified the forewings with an average accuracy of 86.6±6.9% across the 26 subspecies represented in the test data set, although the number of images used for classification was small and many of them were low quality. The same researchers found classification accuracy of 98.9%, 97.7%, 91.1%, 96.4%, and 95% for the Carniolan, Caucasian, Iberiensis (Spanish bee), Italian bee, and Mellifera subspecies, respectively. In the current study, the RF model showed the highest accuracy in classifying bees and was different from Rodrigues et al. (2022).

4. Conclusions

As a result, RF model is the most effective tool for classifying bee breeds within the evaluated dataset and offers a promising approach for researchers seeking to improve the sustainability and productivity of beekeeping through genetic and morphological analyses. While NB model is valuable for its computational efficiency, it requires careful application in this context due to its comparatively low performance. In addition, the present study demonstrates the possible applications and limitations of each algorithm in the context of classifying bee breeds based on their MTs.

Author Contributions: Investigation, Writing – original draft, Writing – review and editing, Resources, Funding acquisition, Visualization, MK.; Methodology, Supervision, Validation, İŞN.; Conceptualization, Writing – original draft, Project administration, Validation, İA.; Data curation, Formal Analysis, Software, İK. All authors have read and agreed to the published version of the manuscript

Funding: This research received no external funding.

Data Availability Statement: Data generated or analyzed during this study are available from the corresponding author upon reasonable request.

Acknowledgments: Many thanks to veterinarian Kamil Kibar for providing the bee samples used in this study.

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

References

Abou-Shaara HF (2013). Wing venation characteristics of honey bees. *Journal of Apicultural* 28: 79-86.

Abou-Shaara HF, Al-Ghamdi AA, Mohamed AA (2013). Body morphological characteristics of honey bees. *Agricultura,* 10: 45-49.

Alpatov WW (1929). Biometrical studies on variation and races of the honey bee (*Apis mellifera L.*). *The Quarterly Review of Biology,* 4: 1-58. https://www.jstor.org/stable/2808231

Anderson LE (1954). Hoyer's solution as a rapid permanent mounting medium for bryophytes. *The Bryologist,* 57: 242-244. https://doi.org/10.2307/3240091

Antony JC, Pratheepa M (2018). A Bayesian classification approach for predicting Gesonia gemma Swinhoe population on soybean crop in relation to abiotic factors based on economic threshold level. *Journal of Biological Chemistry,* 32: 68-73. https://doi.org/10.18311/jbc/2018/16309

Berlocher SH (1984). Insect molecular systematics. *Annual Review of Entomology,* 29: 403-433. https://doi.org/10.1146/annurev.en.29.010184.002155

Bhavsar H, Ganatra A (2012). A comparative study of training algorithms for supervised machine learning. *International Journal of Soft Computing and Engineering,* 2: 74-81.

Breiman L (2001). Random forests. *Journal of Machine Learning Research,* 45: 5-32. https://doi.org/10.1023/A:1010933404324

Buco SM, Rinderer TE, Sylvester HA, Collins AM, Lancaster VA, Crewe RM (1987). Morphometric differences between South American Africanized and South African (*Apis mellifera scutellata*) honey bees. *Apidologie,* 18: 217-222. https://doi.org/10.1051/apido:19870301

Cariveau DP, Nayak GK, Bartomeus I, Zientek J, Ascher JS, Gibbsand J, Winfree R (2018). The allometry of bee proboscis length and its uses in ecology. *PloS One,* 13: Article e0207900. https://doi.org/10.1371/journal.pone.0207900

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

Crewe RM, Hepburn HR, Moritz RFA (1994). Morphometric analysis of 2 southern African races of honey bee. *Apidologie,* 25: 61-70. https://doi.org/10.1051/apido:19940107

Demsar J, Curk T, Erjavec A, Gorup C, Hocevar T, Milutinovic M, Mozina M, Polajnar M, Toplak M, Staric A, Stajdohar M, Umek L, Zagar L, Zbontar J, Zitnik M, Zupan B (2013). Orange: Data Mining Toolbox in Python. *Journal of Machine Learning Research,* 14: 2349–2353.

Diniz-Filho JAF, Malaspina O (1995). Evolution and population structure of Africanized honey bees in Brazil: Evidence from Spital analysis of morphometric data. *Evolution,* 49: 1172-1179. https://doi.org/10.1111/j.1558-5646.1995.tb04444.x

Dodoloğlu A (2000). Kafkas ve Anadolu balarısı (*Apis smellifera L.*) ırkları ile karşılıklı melezlerinin morfolojik, fizyolojik ve davranış özellikleri (Publication No. 96437). PhD Thesis, Atatürk University (Unpublished), Türkiye.

Estoup A, Garnery L, Solignac M, Cornuet JM (1995). Microsatellite variation in honey bee (*Apis mellifera L.*) populations: hierarchical genetic structure and test of the infinite allele and stepwise mutation models. *Genetics,* 140: 679-695. https://doi.org/10.1093/genetics/140.2.679

Frunze O, Kim DW, Kang EJ, Kim K, Park BS, Choi YS (2022). The accuracy of morphometric characteristic analysis depends on the type of the assessed traits of honey bees (*Apis cerana F.* and *Apis mellifera L.*). *Journal of Asia-Pacific Entomology,* 25: Article 101991. https://doi.org/10.1016/j.aspen.2022.101991

Ftayeh A, Meixner M, Fuchs S (1994). Morphometrical investigation in Syrian honey bees. *Apidologie,* 25: 396-401. https://doi.org/10.1051/apido:19940406

Gençer HV (2004). The graphic evaluation of morphological characters in honey bees (*Apis mellifera L.*) by Chernoff Faces. *Journal of Agricultural Sciences,* 10: 245-249. https://doi.org/10.1501/Tarimbil_0000000901

Gençer HV, Günbey B (2020). The morphological characteristics of distinctive honey bee (*Apis mellifera L.*) genotypes in Black Sea Region. *Journal of Animal Science and Products,* 3: 40-53.

Guler A, Bek Y (2002). Forewing angles of honey bee (*Apis mellifera*) samples from different regions of Turkey. *Journal of Apicultural Research,* 41: 43-49. https://doi.org/10.1080/00218839.2002.11101067

Han J, Pei J, Tong H (2022). Data mining: concepts and techniques. Morgan kaufmann. Amsterdam, Elsevier, p.735.

Jadhav SD, Channe HP (2016). Comparative study of K-NN, naive Bayes and decision tree classification techniques. *International Journal of Science and Research,* 5: 1842-1845. https://www.ijsr.net/getabstract.php?paperid=NOV153131

Kambur M, Kekeçoğlu M (2018). The current situation of Turkey Honey Bee (*Apis mellifera L.*) biodiversity and conservations studies. *Biological Diversity and Conservation,* 11: 105-119. https://doi.org/10.13140/RG.2.2.12203.54568

Kekeçoğlu M, Bir S, Acar MK (2023). Determination of Anatolian honeybee biodiversity by wing characters. *Sociobiology,* 70: Article e8333. https://doi.org/10.13102/sociobiology.v70i3.8333

Khan SN, Khan SU, Aznaoui H, Şahin CB, Dinler ÖB (2023). Generalization of linear and non-linear support vector machine in multiple fields: a review. *Computer Science and Information Technology,* 4: 226-239. https://doi.org/10.11591/csit.v4i3.p226-239

Li F, Xiong Y (2018). Automatic identification of butterfly species based on HoMSC and GLCMoIB. *The Visual Computer,* 34: 1525-1533. https://doi.org/10.1007/s00371-017-1426-1

Liaw A, Wiener M (2002). Classification and regression by random Forest. *R News,* 2: 18-22. http://CRAN.R-project.org/doc/Rnews/

Özbakır GÖ (2011). Morphological characterization of honey bee populations (*Apis mellifera L.*) along the southeast border of Turkey (Publication No. 299725). PhD Thesis, Ankara University(Unpublished), Türkiye.

Pan X, Wang Y, Qi Y (2023). Artificial neural network model and its application in signal processing. *Asian Journal of Advanced Research and Reports,* 17: 1-8. https://doi.org/10.9734/AJARR/2023/v17i1459

Potts SG, Biesmeijer JC, Kremen C, Neumann P, Schweiger O, Kunin WE (2010). Global pollinator declines: trends, impacts and drivers. *Trends in Ecology and Evolution,* 25: 345-353. https://doi.org/10.1016/j.tree.2010.01.007

Rinderer TE, Buco SM, Rubink WL, Daly HV, Stelszer JA, Rigio RM, Baptista C (1993). Morphometric identification of Africanized and European honey bees using large reference populations. *Apidologie,* 24: 569-585. https://doi.org/10.1051/apido:19930605

Rodrigues PJ, Gomes W, Pinto MA (2022). DeepWings©: automatic wing geometric morphometrics classification of honey bee (*Apis mellifera*) subspecies using deep learning for detecting landmarks. *Big Data and Cognitive Computing,* 6: 70. https://doi.org/10.3390/bdcc6030070

Ruttner F, Tassencourt L, Louveaux J (1978). Biometrical-statistical analysis of the geographic variability of *Apis mellifera L.* I. Material and methods. *Apidologie,* 9: 363-381. https://doi.org/10.1051/apido:19780408

Souza D, Cruz CD, Campos L, Regazzi AJ (2002). Correlation between honey production and some morphological traits in africanized honey bees (*Apis melifera*). *Ciência Rural,* 32: 869-872. https://doi.org/10.1590/S0103-84782002000500020

Szymula J, Skowronek W, Bienkowska M (2010). Use of various morphological traits measured by microscope or by computer methods in the honeybee taxonomy. *Journal of Apicultural Science,* 54: 91-97.

Tapkan P, Özbakır L, Kulluk S, Baykasoğlu A (2016). A cost-sensitive classification algorithm: BEE-Miner. *Knowledge-Based Systems,* 95: 99-113. https://doi.org/10.1016/j.knosys.2015.12.010

Trevethan R (2017). Sensitivity, specificity, and predictive values: foundations, pliabilities, and pitfalls in research and practice. *Frontiers in Public Health,* 5: 307. https://doi.org/10.3389/fpubh.2017.00307

Vapnik V (1999). The nature of statistical learning theory. *Springer Science and Business Media*, pp.314. https://doi.org/10.1007/978-1-4757-3264-1

Wickramasinghe I, Kalutarage H (2021). Naive Bayes: applications, variations and vulnerabilities: a review of literature with code snippets for implementation. *Soft Computing,* 25: 2277-2293. https://doi.org/10.1007/s00500-020-05297-6

Yegnanarayana B (1999). Artificial neural networks. PHI Learning Pvt. Ltd.: New Delhi, India, pp. 476.

Zemskova NE, Sattarov VN, Skvortsov AI, Semenov VG (2020). Morphological characteristics of honey bees of the Volga region. BIO Web Conf 17, 35. EDP Sciences. https://doi.org/10.1051/bioconf/20201700035