Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib

Data Article

The systematic analysis of adults' environmental sensory tendencies dataset

Nigmet Koklu[a], Suleyman A. Sulak[b,*]

[a] *Konya Technical University, Türkiye*
[b] *Necmettin Erbakan University, Türkiye*

A R T I C L E I N F O

Article history:
Received 26 April 2024
Revised 31 May 2024
Accepted 10 June 2024
Available online 13 June 2024

Dataset link: Environmental Sensory Tendencies Dataset (Original data)

Keywords:
Artificial intelligence
Bagging
Environment interaction
Human systematic analysis
Support vector machine

A B S T R A C T

This study was conducted to investigate the profound impact of human activities on the environment, based on scientific data, recognizing the potential of environmental problems to turn into devastating crises if appropriate measures are not taken. It emphasizes the important role of education in developing environmental awareness, knowledge and sensitivity to counter adverse environmental consequences. For this purpose, a dataset was created for the emotional tendencies of university students, who represent a demographic that has the potential to influence the sustainable future of the world. A survey data including 34 different variables was collected from 388 university students in Turkey. Environmental Sensory Tendencies Dataset is intended to provide valuable guidance for the development of effective environmental education programs and policies aimed at increasing university students' awareness and participation in environmental issues. Our research underlines the vital importance of developing responsible attitudes and behaviors to effectively address environmental challenges and thereby contribute to a healthier and more sustainable global ecosystem. This study will make a significant contribution to the

* Corresponding author.
E-mail addresses: nkoklu@ktun.edu.tr (N. Koklu), sulak@erbakan.edu.tr (S.A. Sulak).

https://doi.org/10.1016/j.dib.2024.110640

literature and highlight the interconnection between human actions and environmental well-being.

Specifications Table

Subject	Social Sciences; Education
Specific subject area	Environmental Sensory Tendencies
Type of data	.xls files, zip files
Data collection	The dataset was obtained through an online survey administered by the authors to a total of 388 individuals living in Turkey. The created dataset contains 34 variables that are capable of determining sensory inclination toward the environment. The data distributions of the features in the sensory tendencies dataset are given in Table 1. As shown in Figure 2, the dataset covers a population consisting of 180 males and 208 females, with ages ranging from 18 to 45
Data source location	Institution : Necmettin Erbakan University City : Konya Country : Türkiye
Data accessibility	Direct URL to data: https://data.mendeley.com/datasets/cwbf2z7tch/1 [1]

1. Value of the Data

- This dataset provides a comprehensive overview of environmental sensitivity and behavior, provides researchers with new insights.
- These data obtained from Turkey provide an important resource for understanding the relationship of cultural differences with the environment.
- The fact that the survey is online provides an ideal data source for understanding the effects of digital transformation on environmental sensitivity.
- The dataset shows that various demographic factors (gender, age, etc.) provides a rich resource for researchers who want to study the impact on environmental sensitivity
- These data provide basic information for the effective implementation of environmental policies.

2. Background

People are an indispensable, important and at the same time risky factor for the environment. As human dominance increases in the relationship between humans and the environment, various environmental problems begin to emerge. This means that environmental problems caused by human activities have increased, spread and reached an irreversible state in a way that threatens human life [2].

Some human-caused environmental problems can turn into major disasters if the necessary measures are not taken. By the education supports, it is possible to alleviate the environmental problems of individuals only by changing attitudes and behavior [3]. It is important to integrate environmental education into all levels of education, including formal, non-formal and adult educational institutions. However, compulsory education constitutes the most important starting point for environmental education [4].

Through environmental education, individuals can become environmentally literate, make informed decisions, and develop individual behavioral principles related to environmental issues [5,6]. Environmental literacy is defined as "an individual's knowledge about the environment and environmental issues, attitudes towards the environment and environmental issues, skills, motivation to participate in activities that address environmental issues, and efforts to provide a balanced and sustainable life and environment" [7].

In order for individuals to change their environmentally harmful behaviors, their attitudes, knowledge and values related to the environment must be changed. Positive attitudes, values and responsible behavior towards the environment can be used manually with the environmental education tool [8,9].

Sensory tendency refers to individuals' natural responses to environmental sensory stimuli. These tendencies reflect how a person perceives, processes, and responds to specific sounds, lights, smells, tastes, or tactile inputs in their surroundings. Sensory tendencies are influenced by genetic and neurological factors, as well as by environmental interactions and experiences.

Individuals' sensory tendencies directly impact their interactions with the environment and the experiences they derive from these interactions. For example, someone with high sensory sensitivity may feel discomfort and stress in a noisy setting, while someone with a sensory-seeking tendency might find the same environment stimulating and actively seek out such stimuli. These differences determine how individuals interact with and are affected by their surroundings. Understanding sensory tendencies allows for environmental adjustments and the development of appropriate support strategies to enhance individuals' quality of life.

In order to reduce environmental problems, individuals should be equipped with environmental knowledge as well as the ability to show positive attitudes and responsible behaviors towards the environment [9,10]. In the light of all these findings, "What are the sensory tendencies of university students towards the environment?" the dataset related to the research question has been obtained.

3. Dataset Description

In this study, we have worked on numerical data related to sensory inclination towards the environment. Classification has been performed using various classification algorithms on the data prepared by us. The flow model illustrating the operation of the study is given in Fig. 1. When the structure of the figure is examined, it is necessary to obtain the data in the first step. The created data is subjected to various data preprocessing steps, and 34 different features are extracted. These extracted features are transferred to the dataset. The data, which is trained in a repetitive manner using the cross-validation method, is sent to the classification models. The trained algorithms provide output according to the specified classes.

3.1. Sensory Tendencies Dataset

With the accumulation of knowledge in the field and the conducted literature review, the parametric factors necessary to construct the dataset were determined. Subsequently, a questionnaire template containing variables that would determine sensory inclination toward the environment was created. The dataset was obtained through an online survey administered by the authors to a total of 388 individuals living in Turkey. The created dataset contains 34 variables that are capable of determining sensory inclination toward the environment.

During the data collection process, demographic questions were asked at the beginning, followed by 23 questions about Sensory Tendencies, culminating in an additional feature indicating

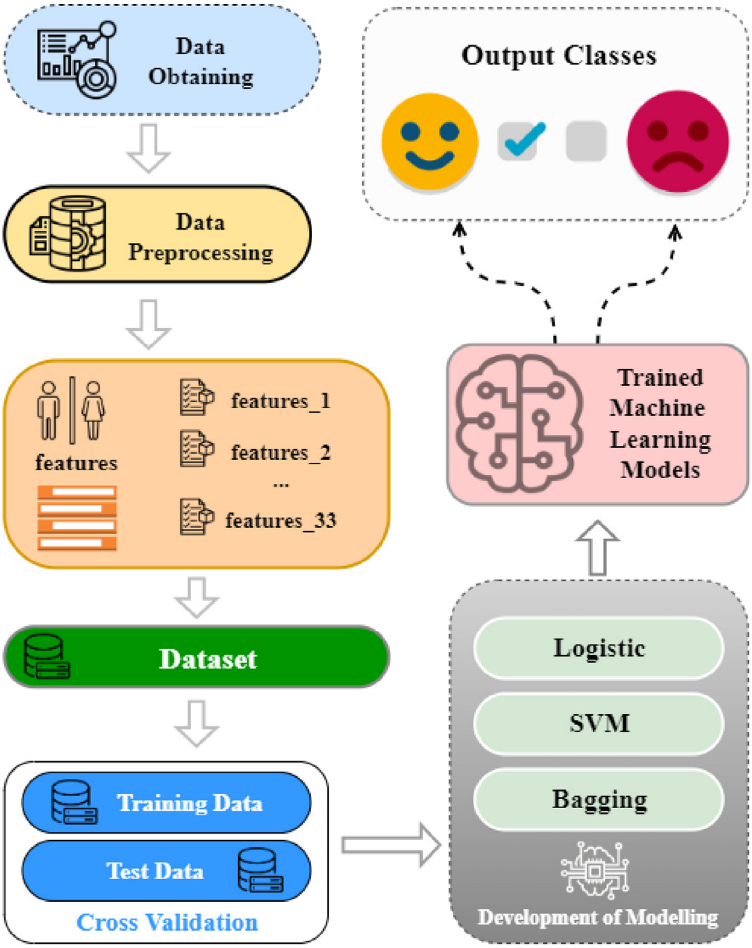

Fig. 1. Schematic representation of the working operation.

whether the Sensory Tendency is present in the individual. This resulted in a dataset comprising 34 variables. Data were collected from 388 university students from various universities across Turkey, spanning from November 6, 2023, to December 7, 2023.

As shown in Fig. 2, the dataset covers a population consisting of 180 males and 208 females, with ages ranging from 18 to 45.

The data distributions of the features in the Sensory Tendencies dataset, including the values for the demographic questions and the answers to the Sensory Tendencies questions, are presented in Table 1.

Table 1
Sensory tendencies dataset features.

Attributes	Values
Sex	1- Male; 2-Female
Age	Values in integers
Grade	1; 2; 3; 4
Have you taken an environmental education course?	1-Yes; 2-No
Have you taken a global warming and climate course?	1-Yes; 2-No
Have you taken an Environmental Literacy course?	1-Yes; 2-No
Income status	1:0-7500; 2.7501-10000; 3.10001-15000; 4.15001-more:
Did you go to pre-school education?	1-Yes; 2-No
Where you live	1-Town; 2-District; 3-City
Your number of siblings	0; 1; 2; 3; 4; 5. more
1) I believe that I can contribute to the protection of the environment with my individual behavior.	1-Never; 2-Rarely; 3-Sometimes; 4-Usually; 5-Always
2) I believe it is no use trying to influence my family or friends on environmental issues.	1-Never; 2-Rarely; 3-Sometimes; 4-Usually; 5-Always
3) If I had more knowledge about environmental issues, I would integrate my sensitivity towards the environment into my daily habits.	1-Never; 2-Rarely; 3-Sometimes; 4-Usually; 5-Always
4) Damage to the environment can be reduced by laws.	1-Never; 2-Rarely; 3-Sometimes; 4-Usually; 5-Always
5) Every person has a responsibility to protect the environment.	1-Never; 2-Rarely; 3-Sometimes; 4-Usually; 5-Always
6) Since other people cause great harm to the environment, my individual efforts such as saving water, saving energy, using environmentally friendly products will not have any effect on protecting the environment.	1-Never; 2-Rarely; 3-Sometimes; 4-Usually; 5-Always
7) It is the responsibility of every teacher to include environmental issues and values in education.	1-Never; 2-Rarely; 3-Sometimes; 4-Usually; 5-Always
8) Every teacher candidate should take at least one environmental course during their education.	1-Never; 2-Rarely; 3-Sometimes; 4-Usually; 5-Always
9) Environmental issues should be prioritized over existing issues on the national agenda.	1-Never; 2-Rarely; 3-Sometimes; 4-Usually; 5-Always
10) Penalties cannot prevent the damage people cause to the environment.	1-Never; 2-Rarely; 3-Sometimes; 4-Usually; 5-Always
11) Nowadays, environmental awareness is not at a sufficient level.	1-Never; 2-Rarely; 3-Sometimes; 4-Usually; 5-Always
12) It is a human right to benefit from nature's resources only to meet people's basic needs.	1-Never; 2-Rarely; 3-Sometimes; 4-Usually; 5-Always
13) It is important that the education system includes environmental issues.	1-Never; 2-Rarely; 3-Sometimes; 4-Usually; 5-Always
14) Factories that harm the environment should be punished.	1-Never; 2-Rarely; 3-Sometimes; 4-Usually; 5-Always
15) I think that individual environmental actions will not be taken into account by the authorities.	1-Never; 2-Rarely; 3-Sometimes; 4-Usually; 5-Always
16) People who harm the environment should be punished.	1-Never; 2-Rarely; 3-Sometimes; 4-Usually; 5-Always
17) It is important to organize environmentally related activities (such as trips and observations) in schools.	1-Never; 2-Rarely; 3-Sometimes; 4-Usually; 5-Always
18) Too much open space in the natural environment has been unnecessarily used for highways	1-Never; 2-Rarely; 3-Sometimes; 4-Usually; 5-Always
19) Industrial enterprises should be forced to reduce their emissions of pollutants.	1-Never; 2-Rarely; 3-Sometimes; 4-Usually; 5-Always
20) In order to protect nature in Turkey, construction (building, construction, etc.) is unnecessarily prohibited in many places.	1-Never; 2-Rarely; 3-Sometimes; 4-Usually; 5-Always
21) The development of renewable energy sources (such as solar, wind energy, etc.) should be supported in Turkey.	1-Never; 2-Rarely; 3-Sometimes; 4-Usually; 5-Always
22) The construction of buildings such as hotels and summer houses on the seashore should be reduced.	1-Never; 2-Rarely; 3-Sometimes; 4-Usually; 5-Always
23) The value of living creatures in nature is determined only by how useful they are to humanity.	1-Never; 2-Rarely; 3-Sometimes; 4-Usually; 5-Always
Sensory Tendencies class	0-No (This person has no sensory tendencies towards the environment); 1-Yes (This person has sensory tendencies towards the environment)

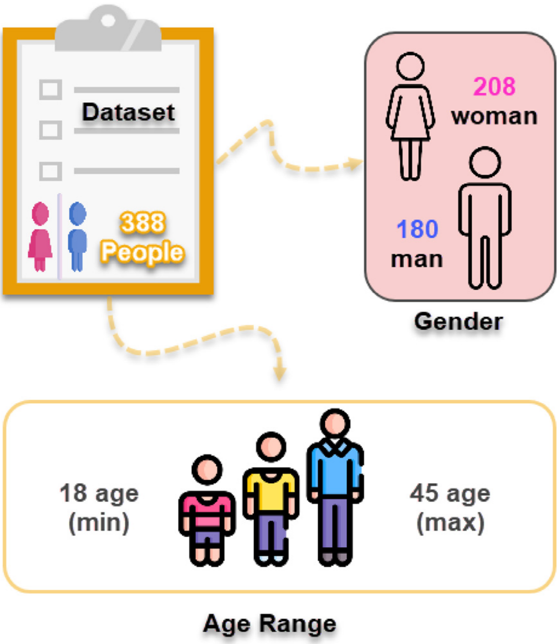

Fig. 2. Information about the study group that created the dataset.

4. Experimental Design, Materials and Methods

4.1. Performance Measures

Confusion matrix is a table commonly used to evaluate the performance of a classification model by comparing the predicted results with the actual results. It provides a useful tool to analyze the strengths and weaknesses of an artificial intelligence model and can help identify areas for improvement [11,12]. Performance measurements of the classification model are calculated using the data shown in Table 2, as illustrated.

To evaluate the trained model, Precision (P), Recall (R), F1-score (F), and Accuracy (AC) metrics are commonly used. These metrics measure the model's ability to correctly classify positive and negative classes, as well as its ability to minimize false positives and false negatives. The calculations of these metrics are performed using the four different measures given in Fig. 3 [12,13].

Table 2
Evaluation criteria affecting the two-class Confusion matrix.

Evaluation Criteria	Definition
True Positive (TP)	TP refers to the cases where the model correctly predicts the class and the actual class is also positive.
False Positive (FP)	FP refers to the cases where the model predicts the class as positive, but the actual class is negative.
True Negative (TN)	TN refers to the cases where the model correctly predicts the class as negative and the actual class is also negative.
False Negative (FN)	FN refers to the cases where the model predicts the class as negative, but the actual class is positive.

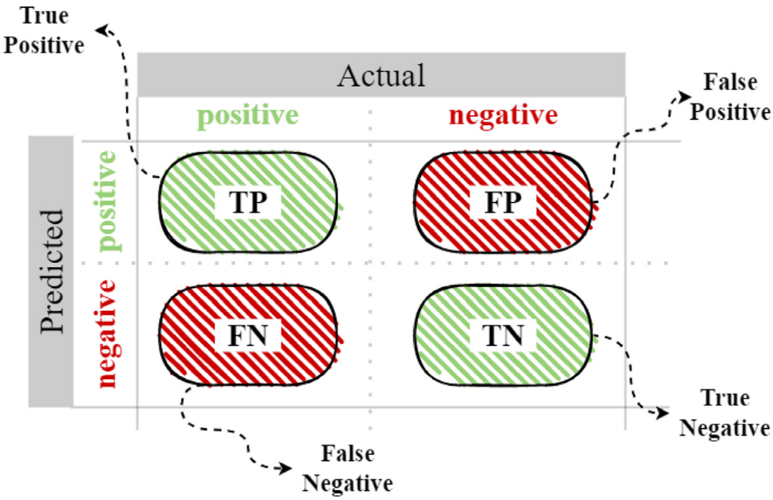

Fig. 3. Complexity matrix used in the classification.

Table 3
Calculation formulas of model evaluation metrics.

Evaluation Metric	Formula	Formula Number
Accuracy (%)	$\frac{TP + TN}{TP + FP + FN + TN} \times 100$	(1)
Precision (%)	$\frac{TP}{TP + FP} \times 100$	(2)
Recall (%)	$\frac{TP}{TP + FN} \times 100$	(3)
F1-Score (%)	$\frac{2 \times TP}{2\times TP + FP + FN} \times 100$	(4)

Performance metrics are the preferred evaluation criteria in artificial intelligence to assess the performance of models. These metrics are used to analyze, evaluate, and identify areas that need improvement in how well an artificial intelligence model performs a task [12,14,15]. Commonly used performance metrics, along with their formulas, are provided in Table 3, and detailed explanations are given in the continuation of the table.

Accuracy (AC) *measures the proportion of correctly classified samples in a dataset. It is a commonly used performance metric for classification problems.*

Precision (P) *measures the ratio of true positives among the samples predicted as positive. It is a widely used performance metric in cases where false positives are costly.*

Recall (R) *measures the rate at which true positives are correctly identified. It is a performance metric often used in cases where false negatives are costly.*

F1-Score (F) *combines precision and recall rates into a single metric. It is frequently used in situations where both precision and recall are important.*

4.2. Cross Validation

Cross-validation is a technique used in artificial intelligence and statistical modeling to evaluate the performance of a predictive model on a limited amount of data. It is preferred to estimate the generalization ability of a model on new data. By repeatedly training and evaluating the model on different subsets of the data, it helps mitigate the problem of overfitting or underfitting patterns in the data that the model may have missed. This allows us to improve the generalization performance of the model and make more accurate predictions on new, unseen

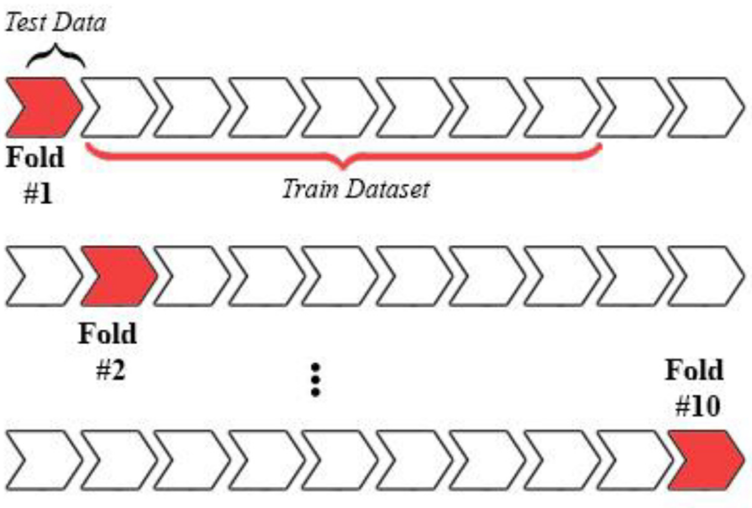

Fig. 4. The complexity matrix is used in the classification of the dataset.

data. As shown in Fig. 4, the 10-fold cross-validation strategy is commonly used for the cross-validation model. Cross-validation is a crucial tool in artificial intelligence to assess and improve the generalization performance of predictive models [16,17].

4.3. Development of Modelling

The dataset was modelled using Logistic Regression [18,19], Support Vector Machine [12,20,21], and Bagging [22–25] algorithms.

4.4. Classification Results with Logistic, SVM and Bagging Models

In Table 4, accuracy rates for Logistic, SVM and Bagging models are calculated based on the confusion matrix. The results are shown in Table 5 and Fig. 5. It can be seen that the logistic model has an accuracy of 96.38%. When the accuracy rate for the SVM model is calculated, it is observed to have a classification success of 93.29%. For the bagging model, it can be seen that it has an accuracy of 83.76%. When these accuracy rates are examined, it can be concluded that logistic and SVM models have high classification success. The classification performance of the bagging model is acceptable, but it can be concluded that it is not at a very high level.

Table 4
Confusion matrix table of the classification made by the Logistic, SVM and Bagging method.

		Logistic				SVM				Bagging	
		Actual Class				Actual Class				Actual Class	
		1	0			1	0			1	0
Predicted Class	1	146	7	Predicted Class	1	136	17	Predicted Class	1	117	36
	0	7	228		0	9	226		0	27	208

Table 5
Performance rates tables of the dataset.

Algorithm	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Logistic	**96.39**	96.40	96.40	96.40
SVM	**93.29**	93.30	93.30	93.30
Bagging	**83.76**	83.70	83.80	83.70

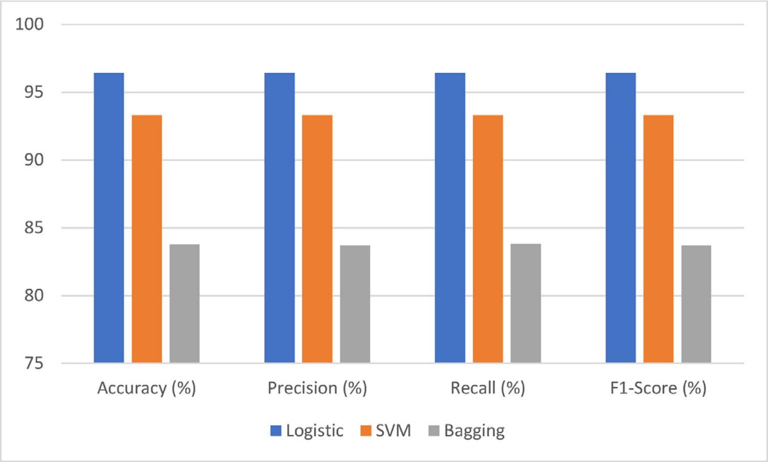

Fig. 5. Performance rates and graphs of the dataset.

4.5. Classification Results and Evaluation Charts of Models in EST Dataset

By using the confusion matrix data of all models, the accuracy, precision, recall and F1 Score values of each model were obtained. Classification result values are shown graphically in Table 5 and data of models in Fig. 5.

Upon examining the results in Fig. 5, it can be observed that the Logistic model has the highest classification performance. The Bagging model, on the other hand, has the lowest classification performance. Additionally, according to Table 4, in parallel with the classification accuracy values, the model with the highest values for metrics such as precision, recall, and f1-score is the Logistic model. Similarly, the Bagging model has the lowest values for these metrics as well.

Limitations

No limitations.

Ethics Statement

This is to confirm that the relevant informed consent was obtained from the respondents. During the distribution of questionnaire, respondents were explicitly informed of the voluntary nature of their participation. They were also made aware that the data provided would be exclusively used for academic purposes, specifically for the dissemination of findings through publication in scholarly journals. Furthermore, prior to initiating the data collection process, ethical approval from the Necmettin Erbakan University Social and Behavioral Sciences Institutional Review Board Ethics Committee was granted. The ethical approval bears the protocol number

2023-02 and is accessible in the supplementary files section, which accompanies this submission.

Data Availability

Environmental Sensory Tendencies Dataset (Original data) (Mendeley Data).

CRediT Author Statement

Nigmet Koklu: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing – original draft, Writing – review & editing; **Suleyman A. Sulak:** Conceptualization, Methodology, Software, Validation, Resources, Writing – original draft, Writing – review & editing, Visualization, Supervision.

Data availability

The data used to support the findings of this study are available on Mendeley Data.

Koklu, Nigmet & Sulak, Suleyman Alpaslan (2024), "Environmental Sensory Tendencies Dataset", Mendeley Data, V1, doi: 10.17632/cwbf2z7tch.1.

Acknowledgements

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Supplementary Materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.dib.2024.110640.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Nigmet Koklu, S.A. Sulak, Environmental sensory tendencies dataset", Mendeley Data V1 (2024) Access Date:22.04.2024, doi:10.17632/cwbf2z7tch.1.
[2] F. Doğan, Y. Keleş, Environmental awareness and environmental behavior in secondary and high school students, NEÜ Ereğli Eğitim Fakültesi Dergisi 2 (1) (2020) 80–90.
[3] S. Sinha, N.K. Jangira, S. Das, Environmental education: Module for pre-service training of social science teachers and supervisors for secondary schools, UNESCO, Paris, 1985.
[4] P. Neal, J. Palmer, The handbook of environmental education, Routledge, 2003.
[5] A. Cutter-Mackenzie, R. Smith, Ecological literacy: the 'missing paradigm'in environmental education (part one), Environment (2003).
[6] S. Pe'er, D. Goldman, B. Yavetz, Environmental literacy in teacher training: Attitudes, knowledge, and environmental behavior of beginning students, J. Environ. Educ. 39 (1) (2007) 45–59.
[7] C.E. Roth, Environmental literacy: its roots, evolution and directions in the 1990s.Quinlan, J. R. (1992). Learning with continuous classes, 5th Australian joint conference on artificial intelligence, 1992.

[8] H.R. Hungerford, T.L. Volk, Changing learner behavior through environmental education, J Environ. Educ. 21 (3) (1990) 8–21.
[9] A. Kollmuss, J. Agyeman, Mind the gap: why do people act environmentally and what are the barriers to pro-environmental behavior? Environ. Educ. Res. 8 (3) (2002) 239–260.
[10] J.M. Hines, H.R. Hungerford, A.N. Tomera, Analysis and synthesis of research on responsible environmental behavior: A meta-analysis, J. Environ. Educ. 18 (2) (1987) 1–8.
[11] D. Chicco, G. Jurman, The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation, BMC Genom. 21 (2020) 1–13.
[12] V. Vapnik, The nature of statistical learning theory, Springer science & business media, 1999.
[13] D.C. Cireşan, A. Giusti, L.M. Gambardella, J. Schmidhuber, Mitosis detection in breast cancer histology images with deep neural networks, in: Medical Image Computing and Computer-Assisted Intervention–MICCAI 2013: 16th International Conference, Nagoya, Japan, 2013 September 22-26, 2013, Proceedings, Part II 16.
[14] I. Cinar, M. Koklu, Classification of rice varieties using artificial intelligence methods, Int. J. Intell. Syst. Appl. Eng. 7 (3) (2019) 188–194.
[15] M. Sokolova, N. Japkowicz, S. Szpakowicz, Beyond accuracy, F-score and ROC: a family of discriminant measures for performance evaluation, in: AI 2006: Advances in Artificial Intelligence: 19th Australian Joint Conference on Artificial Intelligence, Hobart, Australia, 2006 December 4-8, 2006. Proceedings 19.
[16] M.W. Browne, Cross-validation methods, J. Math. Psychol. 44 (1) (2000) 108–132.
[17] I. Guyon, A. Elisseeff, An introduction to variable and feature selection, J. Mach. Learn. Res. 3 (Mar) (2003) 1157–1182.
[18] N. Landwehr, M. Hall, E. Frank, Logistic model trees, Mach. Learn. 59 (2005) 161–205.
[19] E. Frank, Y. Wang, S. Inglis, G. Holmes, I.H. Witten, Using model trees for classification, Mach. Learn. 32 (1998) 63–76.
[20] J. Cervantes, F. Garcia-Lamont, L. Rodríguez-Mazahua, A. Lopez, A comprehensive survey on support vector machine classification: Applications, challenges and trends, Neurocomputing 408 (2020) 189–215.
[21] N. Cristianini, J. Shawe-Taylor, An introduction to support vector machines and other kernel-based learning methods, Cambridge university press, 2000.
[22] L. Breiman, Bagging predictors, Mach. Learn. 24 (1996) 123–140.
[23] A.M. Prasad, L.R. Iverson, A. Liaw, Newer classification and regression tree techniques: bagging and random forests for ecological prediction, Ecosystems 9 (2006) 181–199.
[24] T. Hothorn, B. Lausen, A. Benner, M Radespiel-Tröger, Bagging survival trees, Stat. Med. 23 (1) (2004) 77–91.
[25] B. Ghimire, J. Rogan, V.R. Galiano, P. Panday, N. Neeti, An evaluation of bagging, boosting, and random forests for land-cover classification in Cape Cod, Massachusetts, USA, GIScience Remote Sens. 49 (5) (2012) 623–643.