

Raisin Classification Using Machine Learning Techniques

Nisha Dhital¹

¹Agricultural Engineer

Tribhuvan University, Institute of Engineering, Dharan

Abstract - Raisins are valuable and desired food product. The advancement in the field of artificial intelligence has modernized classification process significantly. Raisin classification is important as precise sorting is necessary to maintain quality and for trade purposes. Traditional way of classification using manpower is time consuming and tiring. In this study, machine learning techniques like LR, KNN, DT, RF, SVM and MLP were employed on raisin data consisting seven morphological features of 900 raisin sample, to distinguish two varieties of raisin; Besni and Kecimen. After preprocessing, a cross validation of 10 fold with 80%/20% training and testing split was used to ensure generalization. The classification achieved accuracy of 87.22% with LR, 83.89% with KNN, 82.22% with DT, 86.67 %with RF, SVM, and MLP with the highest being LR with accuracy of 87.22%. Performance of these classifiers underscores the success of study.

Key Words: Raisin, Classification Techniques, Machine Learning, Artificial Intelligence, Evaluation Metrics

1. INTRODUCTION

Raisins are dried grapes mostly obtained from different cultivars of *Vitis vinifera* L. and are extensively consumed worldwide (1). A portion of 100 g of raisins has 299 kcal energy, 3.3 g protein, 0.25 g total lipid (fat), 79.3 g carbohydrate, 4.5 g total dietary fiber, 65.2 g total sugars, 62 mg calcium, 1.79 mg iron, 36 mg magnesium, 98 mg phosphorus and 744 mg potassium (2). Because of its low cost and high satiety value, raisin plays a crucial role in human diets around the world (3).

In 2022/2023 the total raisin production worldwide was around 1.31 million metric tons (4). The expected raisin production from Turkey and USA for year 2023/2024 is 206,300 MT and 153,000 MT respectively (5). Raisins market size was valued at USD 2.2 Billion in 2022. The raisins market industry is projected to grow from USD 2.3 Billion in 2023 to USD 3.4 Billion by 2032, exhibiting a compound annual growth rate (CAGR) of 4.81% during the forecast period (2023-2032) (6).

1.1 Need For Classification

The development of automatic raisin sorting system using machine vision is essential to address the drawbacks of manual evaluation, such as high costs, drudgery and reliability issues. This technology can enhance product quality, eliminate inconsistency and reduce dependence on labour.

Classification of agricultural product is crucial for trade and marketability. It helps to maintain quality, plan logistics, plan resource allocation, set fair market value, meet food industry quality standards, and meet consumer preferences. To obtain high-quality end products, agricultural produces must be separated from the substandard ones at the initial stages. Sorting and grading are done to enhance the uniformity and commercial value of the products (7).

1.2 Artificial Intelligence (AI) in Classification

Machine learning in agriculture has progressed dramatically over the past two decades, from laboratory curiosity to a practical technology in widespread commercial use. It can be far easier to train a system by showing it examples of desired input-output behavior than to program it manually by anticipating the desired response for all possible inputs (8).

The agricultural system must become more productive in output, efficient in operation, and sustainable for future generations. Artificial intelligence and machine vision are playing a key role in the world of food safety and quality assurance. AI makes it possible for computers to learn from experience, and perform most human tasks with an enhanced degree of precision and efficiency. It offers sweeping transformation with advanced approaches that will redefine the traditional pattern and limits of agriculture (9).

Traditional methods for raisin grain classification are labor-intensive and prone to errors. Therefore artificial intelligence tools are desired in agricultural industry to develop efficient and automated techniques that can maintain product quality and align well with industry requirements. By using cameras, sensor and image processing, different features like size, morphology and color can be determined. Advanced algorithms applied to

these features can facilitate productivity by sorting accurately.

2. Literature Review

The research area of classification using machine vision has gained popularity along with advancements in machine learning. Several research works are being carried out with the help of machine learning algorithms.

Leemans et al.(2002) were able to grade two varieties of apple into four classes on the basis of color parameters, geometrical specifications and presence of defects in three steps; image acquisition, its segmentation, and classification of fruit (10). Correct classification rate was 78% for Golden Delicious variety and 72% for Jonagold variety. Results showed that healthy fruits were better graded.

A machine vision system by M.A. Shahin and S.I. Symons (2003) was used to identify the type of Canadian lentils (11). Fifteen input variables from seed size measurement combined with color attributes were used to classify five lentil varieties with accuracy of 99%.

Omid et al. (2010) combined length parameters, color values using image processing algorithm and calculation of center of gravity to grade raisins with overall accuracy of 96% (12).

A raisin sorter was designed and fabricated by M. Abbasgolipour et al. (2010) that graded raisins into two classes. System composed of conveyor belt, lighting box, controlling and processing unit along with sorting unit (13). Raisins were sorted by highly efficient algorithm developed and implemented in Visual Basic 6.0 Environment using suitable Hue, Saturation and Intensity color features and length features. The overall accuracy of apparatus in sorting raisins was 93.3%.

Research conducted by Kuo-Yi Huang in 2011 used neural networks and image processing techniques to classify areca nuts. Six geometric features, three color features and defect area was used in classification procedure (14). Image processing, detection line algorithms and back propagation neural network classifier sorted the quality of areca nuts with an accuracy of 90.9%.

Mollazade et al. (2012) graded raisins into four different classes using image processing and data mining based classifier (15). 44 features including 36 color and 8 shape factors were extracted. MLP Network with 7-6-4 topology was best classifier with accuracy of 96.33%.

Sabanci et al. (2016) classified three different varieties of apple using K Nearest Neighbor and Multi Layer Perceptron algorithms (16). Four size properties and three color properties were extracted. MLP with five hidden layers was best classifier with accuracy of 98.89%.

M. Oliveira et al. (2020) classified 1800 cocoa bean samples into four grades using image features (17). Beans were cut

lengthwise to expose cotyledon. Color and texture of exposed surface were then analyzed. Image analysis combined with random forest algorithm provides accuracy of 93%.

900 samples of Kecimen and Besni variety of raisins were classified by I. Cinar et al. (2020) using image processing and artificial intelligence methods (18). Raisins were subjected to pre-processing steps and image processing techniques to extract seven morphological features. Three models were created with Logistic Regression, Support Vector Machine and Multi Layer Perceptron and highest accuracy obtained was 86.44% with SVM.

I. Cinar and M. Koklu (2021) performed classification operations in 75,000 rice grains of five different varieties using morphological, shape and color features. Images were pre-processed using MATLAB software and 106 features were extracted (19). Different models were created using K Nearest Neighbor, Decision Tree, Logistic Regression, Multi Layer Perceptron and Support Vector Machine. Highest accuracy of 99.91% was obtained with MLP.

Hasan et al. (2021) performed classification using different machine learning techniques and deep neural networks. DBANN2 trained data provided highest accuracy of 93.44% (20).

3. Research Methodology

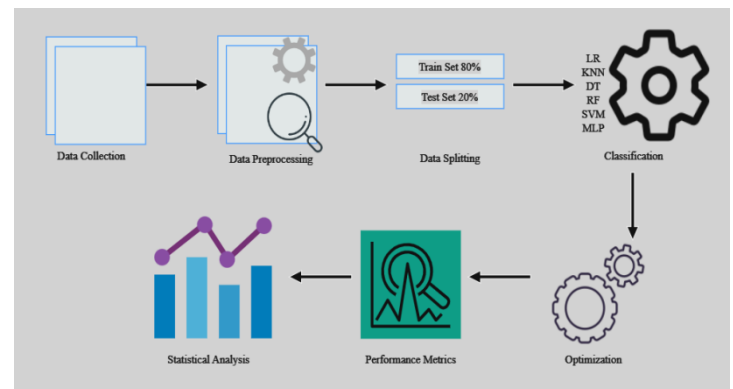


Fig -1: Work Flow Diagram

Research approach involved various steps as shown in Fig -1. First, data was acquired and pre-processed. Data were scaled and separated as test and train sets. Then data were classified using six different classifiers, optimized using grid search and by trial basis. In the final phase, performance of each classifier was evaluated using evaluation metrics.

3.1 Data Collection

The 'Raisin' dataset was collected from University of California, Irvine Machine Learning Repository that was

extracted by I. Cinar et al. (2020) from raisin sample images for raisins classification using machine vision and artificial intelligence methods. The dataset is available at following link:

<https://archive.ics.uci.edu/dataset/850/raisin>

3.2 Data Description

The dataset contains two different raisin varieties grown in Turkey. 900 samples of raisin grains were used, including 450 samples from each Besni and Kecimen variety. Images of these samples were subjected to image processing techniques and various pre-processing steps to extract 7 morphological features. There was no missing value in dataset. The morphological features and their description are given below.

Area: Gives the number of pixels within the boundaries of the raisin.

Perimeter: It measures the environment by calculating the distance between the boundaries of the raisin and the pixels around it.

MajorAxisLength: Gives the length of the main axis, which is the longest line that can be drawn on the raisin.

MinorAxisLength: Gives the length of the small axis, which is the shortest line that can be drawn on the raisin.

Eccentricity: It gives a measure of the eccentricity of the ellipse, which has the same moments as raisins.

ConvexArea: Gives the number of pixels of the smallest convex shell of the region formed by the raisin.

Extent: Gives the ratio of the region formed by the raisin to the total pixels in the bounding box.

3.3 Data Preprocessing

Different preprocessing techniques were applied to raisin dataset. Firstly, raisin data was checked for any missing value. There was no missing value. Counts of both classes were same indicating dataset was balanced. Standardization was employed to uniformly distribute features. Furthermore, two classes with categorical value were label encoded to get numerical value.

3.4 Classification

The classification process involved utilization of six distinct classifiers, applied to the raisin data. To ensure generalization of the models, a 10-fold cross-validation was used. Optimization and tuning of parameters were carefully carried with the help of grid search and on a trial basis to enhance the model's performance.

3.5 Evaluation Metrics

Raisin data was divided into 80% training set and 20% testing set. A 10 fold cross validation strategy was

employed on data for robust evaluation and generalization of classifier performance. A higher cross validation score indicates better generalization.

Classification algorithm performance can be evaluated using confusion matrix and roc-auc score. Confusion matrix is used to calculate different values such as accuracy, sensitivity, specificity, precision and f1-score. These values are essential to evaluate how well a classification algorithm performs. Confusion matrix helps to distinguish between correct and incorrect classifications using four parameters True Positives (tp), True Negatives (tn), False Positives (fp) and False Negatives (fn).

Table -1: Confusion Matrix

Confusion Matrix		
	Predicted Positive	Predicted Negative
Actual Positive	True Positive	False Negative
Actual Negative	False Positive	True Negative

The metrics that were used to evaluate classification models are listed below:

Accuracy: Accuracy measures the ratio of correct predictions over total predictions.

$$\text{Accuracy} = \frac{tp+tn}{tp+fp+tn+fn}$$

Precision: Precision measures the positive patterns that are correctly predicted from the total predicted patterns in a positive class.

$$\text{Precision} = \frac{tp}{tp+fp}$$

Recall: Recall measures the true positive rate.

$$\text{Recall} = \frac{tp}{tp+fn}$$

F1-Score: F1-Score is the harmonic mean of precision and recall.

$$\text{F1-Score} = \frac{2tp}{2tp+fp+fn}$$

$$(21)$$

ROC AUC Score: The ROC AUC score is the area under the Receiver Operating Characteristic curve. It is used to evaluate overall performance of classification model. A higher score indicates better classifier performance.

4. Classification Models

4.1 Logistic Regression (LR)

Logistic Regression (LR) is one of the most statistical and data mining techniques employed by statisticians and researchers for the analysis and classification of binary and proportional response data sets. Some of the main advantages of LR are that it can naturally provide probabilities and extend to multiclass classification problems. Another advantage is that most of the methods

used in LR model analysis follow the same principles used in linear regression (22). There is no need to create normal distribution of variables in LR. Because the values envisaged in the LR are probabilities, LR is limited to 0 and 1. This is because LR predicts its probability, not itself, in the results (19). It is well suited for describing and testing hypothesis about relationships between a categorical outcome and one or more categorical or continuous predictor variables (23).

4.2 K-Nearest Neighbor (KNN)

It is a non-parametric method used for classification and regression. Given N training samples, KNN algorithm identifies the k-nearest neighbors of an unknown data whose class is to be identified (24). However to apply KNN we need to choose appropriate value for k, and the success of classification is very much dependent on this value. There are many ways of choosing the k value, but a simple one is to run the algorithm many times with different k values and choose the one with the best performance (25).

When decision about class of new data is needed, distance between query data and training samples is calculated. Based on the defined value of k, k samples with least distances are selected and the case with more samples inbound is the result (26). In this study, the number of nearest neighbors was selected as 5.

4.3 Decision Tree (DT)

Decision Tree is a graph to represent choices and their results in form of a tree. The nodes in the graph represent an event or choice and the edges of the graph represent the decision rules or conditions. Each tree consists of nodes and branches. Each node represents attributes in a group that is to be classified and each branch represents a value that the node can take (27).

Classification of an instance starts at the root node called the decision node. Based on the value of node, the tree traverses down along the edge. The edge corresponds to the value of the output of feature test. This process continues in the sub-tree headed by the new node at the end of previous edge. Finally, the leaf node signifies the classification categories or the final decision (24).

DT use an architecture of branching choices, beginning with the main question for a specific problem which needs to be answered to solve that problem, later a secondary question must be answered to continue dis-aggregating the data and classify the outcomes (28). The classification rules are derived from the decision tree in the form of -if then else. These rules are used to classify the records with unknown value for class label (29).

The decision tree comprises two parts. First, in growing phase, based on recursive process and local optimal criteria, the training set is split until all or most of the outcomes of each partition have the same class label. The

drawback of this method for building a tree is that over fitting may happen (30).

4.4 Random Forest (RF)

A random forest (RF) is an ensemble classifier which consist many decision trees to classify new instance by the majority vote. Generally, the user sets the number of trees in trial and error basis (31).

Significant improvements in classification accuracy have resulted from growing an ensemble of trees and letting them vote for the most popular class. In order to grow these ensembles, often random vectors are generated that govern the growth of each tree in ensemble (32).

RF algorithms have three main hyper parameters, which need to be set before training. These include node size, the number of trees, and the number of features sampled. From there, RF classifier can be used to solve for regression or classification problems (33).

4.5 Support Vector Machine (SVM)

SVM has been employed widely in different classification and regression problems because of its effectiveness in working with linearly non-separable and high dimensional data sets (34).

SVM training algorithm builds a model of data points in space so that the data points of the separate categories are divided by a clear gap that is as wide as possible. New data are then mapped into that space and predicted to belong to category based on which side of the gap they fall on (35).

Maximizing the margin and thereby creating the largest possible distance between the separating hyper plane and the instances on either side of it has been proven to reduce an upper bound on the expected generalization error (36).

By using different kernel functions, varying degrees of non-linearity and flexibility can be included in the model (37). In this study, radial basis function (rbf) kernel was used with 0.1 regularization parameter ($c=0.1$).

4.6 Multi Layer Perceptron (MLP)

A Multi Layer Perceptron (MLP) is a feedforward network of simple neurons that maps sets of input data onto a set of outputs. The fundamental component of MLP is the neuron. In MLP, a pair of neurons is connected in two adjacent layers, using weighted edges. MLP comprises at least three layers of neurons, including one input layer, one or more hidden layers, and one output layer (38).

On most occasions, the signals are transmitted within the network in one direction; from input to output. There is no

loop; the output of each neuron does not affect the neuron itself (39).

The number of input neurons depends on the dimensions of the input features. The number of output neurons is determined by the number of classes. The number of hidden layers and the number of neurons in each hidden layer depend on the type of problem being solved. Fewer neurons result in inefficient learning and a larger number of neurons result in inefficient generalization. MLP uses a supervised learning technique called backpropagation for training the network (38).

First, the network is trained on a set of paired data to determine input-output mapping. The weights of connections between neurons are then fixed and the network is used to determine the classifications of a new set of data (36).

The alteration of the standard linear perceptron, MLP is capable of distinguishing data which are not linearly separable (16). MLP parameters used in study are given in Table 2.

Table -2: Parameters used in MLP

Parameters	
Activation Function	relu
Hidden Layers	1 with 7 neurons
Solver	adam
Max iterations	1000

5. Result Evaluation

After applying different machine learning algorithms, different evaluation metrics were used. Confusion matrix of all classifiers was observed which is provided in Table 3. Precision, recall and F1 score were collected using classification report. Accuracy, precision, recall, F1 score and ROC AUC score are presented in Table 4.

Table -3: Confusion Matrix of All Classifier

Confusion Matrix											
LR		KNN		DT		RF		SVM		MLP	
85	14	79	20	83	16	81	18	81	18	83	16
9	72	9	72	16	65	6	75	6	75	8	73

Table -4: Evaluation Metrics of All Classifier

Evaluation Metrics					
	Accuracy	Precision	Recall	F1-Score	ROC AUC Score
LR	87.22	87.50	87.50	87.50	87.37
KNN	83.89	84.90	83.93	83.97	84.34
DT	82.22	82.50	82.50	82.50	82.04
RF	86.67	87.17	86.81	86.86	87.21
SVM	86.67	87.15	86.81	86.86	87.21
MLP	86.67	87.45	86.74	86.81	86.98

The bar graph showing accuracy of all six classifiers is given in Figure 1.

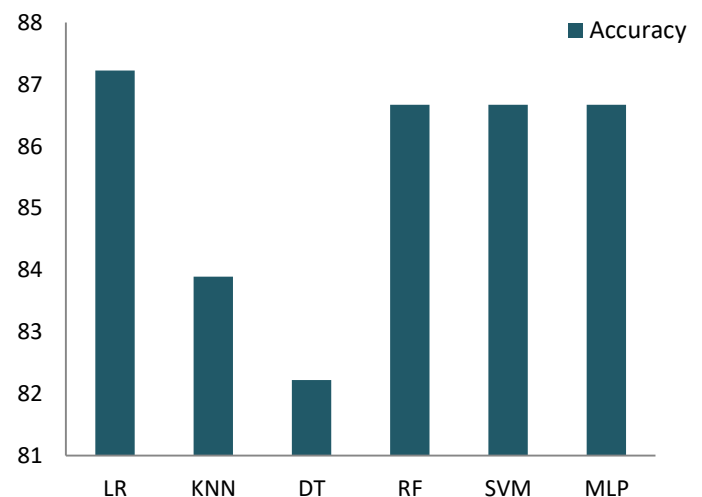


Fig -2: Bar graph Plot for Classifier vs. Accuracy

After observing all evaluation metrics, the highest accuracy of 87.22% was achieved was by LR classifier, highest F1 score of 87% and also the highest ROC AUC score of 87.37%. The cross validation score for LR was observed to be 85.67%. The accuracy of KNN classifier was 83.89% and that of DT being lowest at 82.22%. The accuracy of RF, SVM, and MLP was equal at 86.67% with ROC AUC for RF and SVM being 87.21% while MLP being 86.98%. Looking at these scores in tabular and graphical representation, the high performing model was LR.

6. Discussion

Considering the sample size, the classifier accuracy achieved was satisfactory. Higher performance can be achieved by increasing database and by addition of color

and shape parameters. With same database, performance can be improved by tuning of hyper-parameters.

This study is limited to use of six classification model, future studies could explore more by using hybrid models and deep neural network. There is scope of study using suitable features only. Classification with selective features decreases time and memory requirement leading to efficient design.

Although the raisin samples were of Besni and Kecimen variety from Turkey, this model could be adapted for other varieties as well with addition of sample data from different varieties and different features. This could subsequently help in design of instant and efficient classification application.

7. Conclusion

In this research study, classification of raisin was performed using six distinct machine learning techniques; LR, KNN, DT, RF, SVM, and MLP. Looking at all the data obtained from confusion matrix, classification report and ROC AUC and performance analysis, LR model was best performing classifier with accuracy of 87.22%. And ROC AUC score of 87.37% further emphasizes the algorithms effectiveness in distinguishing two varieties of raisins.

REFERENCES

- [1] Is Eating Raisins Healthy? Alexandra Olmo-Cunillera, Danilo Escoba-Avello, Andy J. Perez, Maria Marhuenda-Munoz, Rosa M Lamuela-Raventos, Anna Vallverdu-Queralt. 2020, MDPI, p. 1. DOI: 10.3390/nu12010054.
- [2] USDA. [Online] <https://fdc.nal.usda.gov/fdc-app.html#/food-details/168165/nutrients>.
- [3] Prediction modeling using deep learning for the classification of grape type dried fruits. Md Nurul Raihen, Sultana Akter. 2023, p.1. DOI: 10.2478/ijmce-2024-0001
- [4] [Online] <https://www.statista.com/statistics/205021/global-raisin-production/>.
- [5] [Online] <https://inc.nutfruit.org/raisins-sultanas-and-currants-global-statistical-review-2/#:~:text=At%20the%20time%20of%20writing,130%2C000%20MT%20and%20140%2C000%20MT..>
- [6] [Online] <https://www.marketresearchfuture.com/reports/raisins-market-6793>.
- [7] Sorting operations for the classification of agricultural crops. Sourav Garg, Venkat Saicharan Kolli, Shivanand S. Shirkole. 2022, p. 1. (Abstarct). DOI: 10.1016/B978-0-12-818572-8.00011-5.
- [8] Machine learning: Trends. M. I. Jordan, T. M. Mitchell. 2015.
- [9] Artificial Intelligence (AI) in Agriculture. Liu, Simon Y. 2020, pp. 14-15. DOI: 10.1109/MITP.2020.2986121.
- [10] On-line Fruit Grading according to their External Quality using Machine Vision. V. Leemans, H. Magein, M.-F. Destain. 2002. DOI: 10.1006/bioe.2002.0131.
- [11] Lentil type identification using machine vision. M. A. Shahin, S.J. Symons. 2003.
- [12] Implementation of an Efficient Image Processing Algorithm for Grading Raisins. M. Omid, M. Abbasgolipour, A. Keyhani and S.S. Mohtasebi. 2010.
- [13] Sorting Raisins by Machine Vision System. Mahdi Abbasgolipour, Mahmoud Omid, Alireza Keyhani, Seyed Saeid Mohtasebi. 2010. DOI: 10.5539/mas.v4n2p49.
- [14] Detection and classification of areca nuts with machine vision. Huang, Kuo-Yi. 2012. DOI: 10.1016/j.camwa.2011.11.041.
- [15] Comparing data mining classifiers for grading raisins based on visual features. Kaveh Mollazade, Mahmoud Omid, Arman Aref. 2011.
- [16] Different Apple Varieties Classification Using kNN and MLP Algorithms. Kadir Sabanci, Muhammed Fahri Ünlerşen. 2016, p. 167. DOI: 10.18201/ijisae.2016Special%20Issue-146967.
- [17] Classification of fermented cocoa beans (cut test) using computer vision. Marciano M. Oliveira, Breno V. Cerqueira, Sylvio Barbon Jr., Douglas F. Barbin. 2021. DOI: 10.1016/j.jfca.2020.103771.
- [18] Classification of Raisin Grains Using Machine Vision and Artificial Intelligence Methods. İlkey Çinar, Murat Koklu, Sakir Tasdemir. 2020. DOI: 10.24432/C5660T.
- [19] Identification of Rice Varieties Using Machine Learning Algorithms. İlkey Çinar, Murat Koklu. p. 10. DOI: 10.15832/ankutbd.862482.
- [20] A Deep Neural Network for Multi-class Dry Beans Classification. Md. Mahadi Hasan, Muhammad Usama Islam, Muhammad Jafar Sadeq. 2021. DOI: 10.1109/ICCIT54785.2021.9689905.
- [21] Dalianas, H. Clinical Text Mining. 2018. pp. 46-48. DOI: 10.1007/978-3-319-78503-5_6.
- [22] Logistic regression in data analysis: An overview. Maalouf, Maher. 2011, p. 1. DOI: 10.1504/IJDATS.2011.041335.
- [23] An Introduction to Logistic Regression. Chao-Ying Joanne Pengg, Kuk Lida Lee, Gary M. Ingersoll. 2002, p. 4. DOI: 10.1080/00220670209598786.
- [24] Machine Learning from Theory to Algorithms: An Overview. Jafar Alzubi, Anand Nayyar, Akshi Kumar. 2018, p. 11. DOI: 10.1088/1742-6596/1142/1/012012.
- [25] KNN Model-Based Approach in Classification. Gongde Guo, Hui Wang, David Bell, Yaxin Bi, Kieran Greer. 2004.
- [26] Comparison of two classifiers; K-nearest neighbor and artificial neural network, for fault diagnosis on a main engine journal-bearing. A. Moosavian, H. Ahmadi, A. Tabatabaefar and M. Khazae. 2012, pp. 266-267. DOI 10.3233/SAV-2012-00742.
- [27] Machine Learning Algorithms-A Review. Mahesh, Batta. 2019, p. 382. DOI:10.21275/ART20203995.
- [28] A review of Machine Learning (ML) algorithms used for modeling travel mode choice. Pineda-Jaramillo, Juan D. 2019, p. 5. DOI: 10.15446/dyna.v86n211.79743.
- [29] Analysis of Feature Selection with classification: Breast Cancer Datasets. D.Lavanya, Dr.K.Usha Rani.

- s.l.: Indian Journal of Computer Science and Engineering (IJCSE), 2011, p. 758.
- [30] Comparative Analysis of Classification Approaches for Breast Cancer. Asfaw, Temesgen Abera. 2019, p. 13.
- [31] How Many Trees in a Random Forest? Thais Mayumi Oshiro, Pedro Santoro Perez, and Jos'e Augusto Baranauskas. s.l.: Springer, Berlin, Heidelberg, 2012, p. 155. DOI: 10.1007/978-3-642-31537-4_13.
- [32] Random Forests. Breiman, Leo. 2001, p. 2.
- [33] IBM. [Online] <https://www.ibm.com/topics/random-forest#:~:text=Random%20forest%20is%20a%20commonly,both%20classification%20and%20regression%20problems..>
- [34] Assessment of the effects of training data selection on the landslide susceptibility mapping: a comparison between support vector machine (SVM), logistic regression (LR) and artificial neural networks (ANN). Bahareh Kalantar, Biswajeet Pradhan, Seyed Amir Naghibi, Alireza Motevalli. 2017, p. 57. DOI: 10.1080/19475705.2017.1407368.
- [35] Data classification using Support vector Machine (SVM), a simplified approach. S Amarappa, Dr. S V Sathyanarayana. International Journal of Electronics and Computer Science Engineering, p. 436.
- [36] Supervised Machine Learning: A Review of Classification. Kotsiantis, S. B. 2007, p. 260.
- [37] Logistic regression and artificial neural network classification models: a methodology review. Stephan Dreiseitl, Lucila Ohno-Machado. 2002, p. 353. DOI: 10.1016/S1532-0464(03)00034-0.
- [38] Mariette Awad, Rahul Khanna. Efficient Learning Machines. pp. 25-26.
- [39] Multilayer Perceptron and Neural Networks. Marius-Constantin Popescui, Valentina E. Balas, Liliana Perescu-Popescu, Nikos Mastorakis. 2009, p. 579.