Plant Leaf Classification Using Weka Tool

Midhi V Mohan
Amal Jyothi College of Engineering Kanjirappally,
Kottayam
midhivmohan@mca.ajce.in

Sona Maria Sebastian
Asst.Professor
Amal Jyothi College of Engineering Kanjirappallly,
Kottayam
sonasebastian@amaljyothi.ac.in

***Abstract*—Leafclassificationisanapplicationensuinginthe field of agriculture and is especially significant to the biology diversity research. the leaves contain useful features, conditions that are mainly used for recognizing various types of plants, so the feature can be applied by automatic algorithms leaf classification to classify various plant leaftypes.**

Keywords— leaf classification, cross-validation, cluster methods

I. INTRODUCTION

The biodiversity is constantly decreasing as a result of various d human powerful activities that deals with various recognition techniques and leaf classification have recently beenusedtodealwiththisproblemPlantsplayimportantrole in preserving earth's ecology and surrounding environment by maintaining a healthy atmosphere and provides shelter to innumerable insect and animalspecies.

The leaves have several features that are used for identification such as leaf type (simple or compound), condition, no of species etc., They classify plant by its category Aim of this research paper is to concentrate on the plant leaf classification based on the types.

Plants are important for man. Herbs, various plant types are identified by practitioners based on previous experiences through personal sensory. Extraction is performed by contour-based.

In this research paper we consider the real-world problem of Leaf Classification using weka tool. Machine Learning uses logical steps to perform the decision or to produce best result.

II. MACHINELEARNING

This models creates models and train themselves to improve,perceivethecomplexpatterns,andfindsolutionsto the new problems based on the previousdata.

Machine learning (ML) is the method of delving algorithms that perform a specific task without being explicitly programmed. web search engine like Google,Bing isusedtoaccessvariousdetailsthroughinternet.Theboonof using machine is that once we understand the algorithm clearly, we know what to do with the data and works automatically.

III. METHODOLOGY

A. Data Collection

This study was achieved through given dataset including features like LeafType ,Condition, NoofSpecies. Results are generated based on the given dataset. Weka tool, are used for identify different leaves based on their type and other features.

Dataset	Leaf Type	Condition	NoofSpecies
Flavia	Colour	Isolated	32
Foliage	Colour	Isolated	30
Folio	Colour	Isolated	32
LeafsnapFie	Colour	Field	183
LeafsnapLat	Colour	Lab	183
ShapeCN	Contour	NA	30
SwedishLea	Colour	Lab	15
100-Leaves	Segmentation	NA	100
Geranium	circular	light	430
Egg Plant	ovate	Best	87
Ficus	fiddle	indirect	900
Croton	oval	moist	150

Fig no:1- Given dataset

B. Cross-Validation

Cross-validation is re-sampling method which used to assess the ability of predictive models and prevent over-fitting.

C. k-fold random subsampling

In this method the single hold-out method repeated k times, that is k pairs of D train, j and D test, j = 1.k, are generated. The performance is calculated as average of 65 k test sets in this method there is no common so, Dtrain, j ∩ Dtest, j = ∅ . 2 training sets may overlap.

D. k-fold cross-validation

It is the task of reiterating sub- sampling method, the sampling is done in a special way that two test sets don't overlap each other.

Thesampleispartitionedinto*K*subsamples.singlesampleis assigned as validation and remaining $K-1$ subsamples are taken as trainingdata.

DOI: 10.5281/zenodo.6181179
ISBN: 978-93-5607-317-3 @2022 MCA, Amal Jyothi College of Engineering Kanjirappally, Kottayam

This process is repeated upto K times. Tenfold cross-validation is used here.

E. Clustered Instances

Clustering is a adaptive procedure in which similiar objects aregroupedtogether.Weclassifythegivendatasetie,(plant Leaf classification) by using several number ofclusters.

F. CFS SubsetEvaluator

It means Correlation based Feature Selection. It is an automatic algorithm—so it does not require user tospecify thresholds because both are simple to incorporate if desired. The features are selected byCFS.

IV. IMPLEMENTATION

Step 1: Install Weka
Step 2: Open Weka Tool
Step 3: Open the data Dataset
Step 4: Select and Run an Algorithm
Step 5: Review Results

WEKA Tool is a collection of various methods and it includes: pre-processing data, classification and clustering.

• Command Line Interface: it is very useful forscripting large jobs.
Weka tool is an implementation of several algorithms. Main aim is to analyze the basic concepts and application of algorithms in real world.
FEATURES

- Preprocessing Filters: it has the number of preprocessing tools.
- Learning Schemes: - instance-based learning, provides support distancefunctions.
- User Interfaces: the graphical start pointredesign provides access to various userinterfaces.

RESULTS

The performance has been examined by experiments conducted on the real datasets collected for supporting leaf typeclassification.theresultsareobtainedbasedonthegiven datasets for clusters it is compared to wholedatasets..

GUI of weka

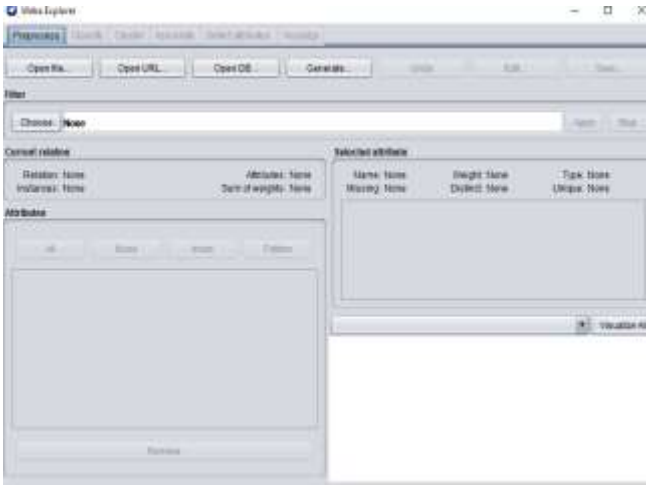
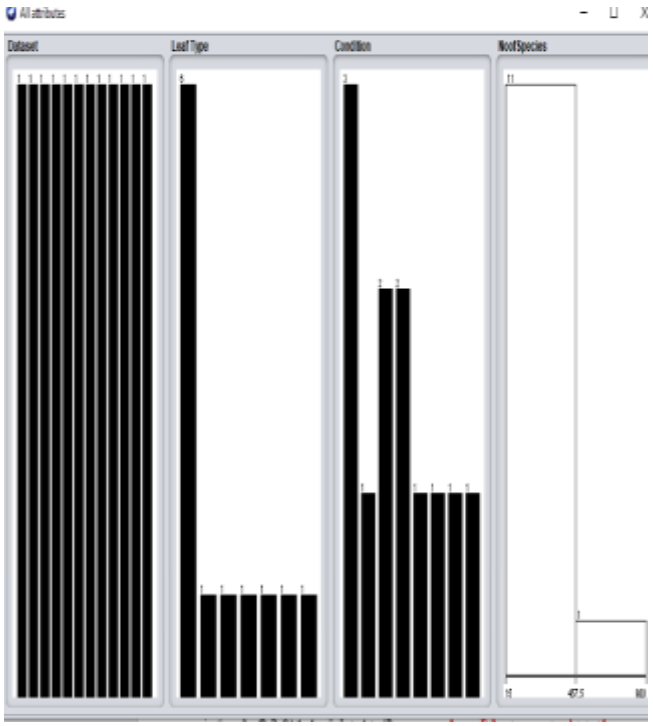


Fig no:2 Select explorer and the Interface

Fig no:3 Select the dataset

DOI: 10.5281/zenodo.6181179
ISBN: 978-93-5607-317-3 @2022 MCA, Amal Jyothi College of Engineering Kanjirappally, Kottayam



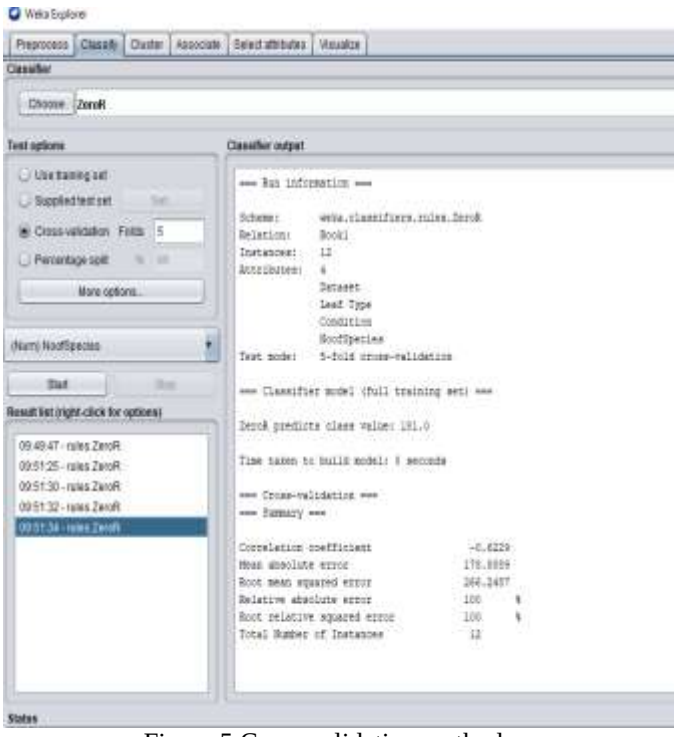
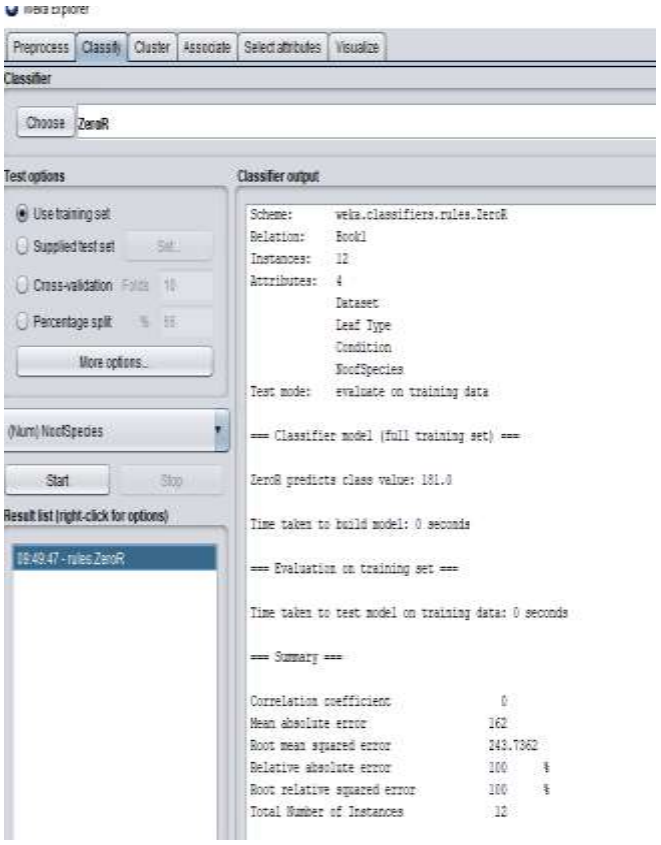


Fig no:5 Cross-validation method



Classify Data-Test options

Fig No: -4 use training Data

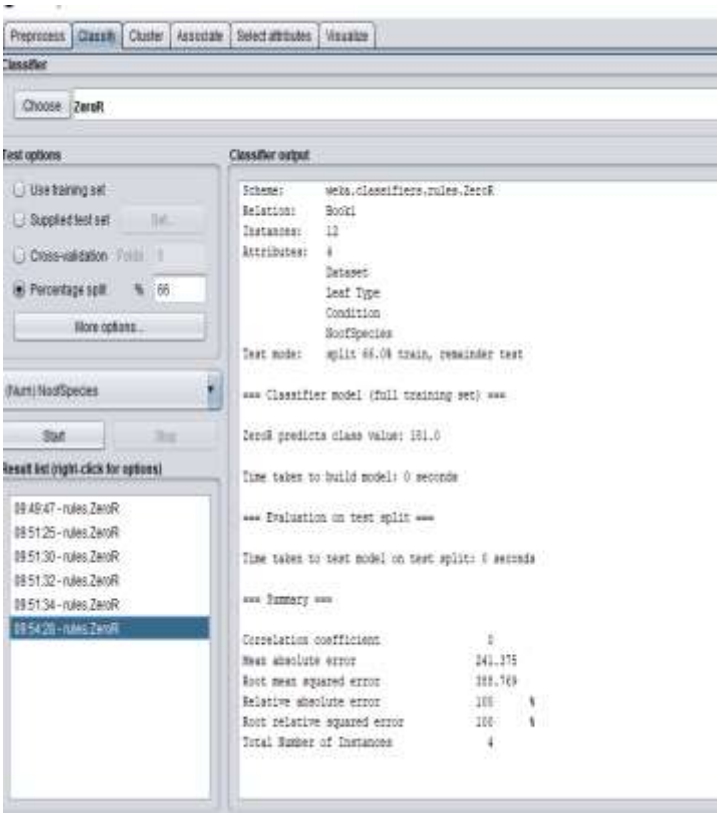


Fig no:6 Percentage split

DOI: 10.5281/zenodo.6181179
ISBN: 978-93-5607-317-3 @2022 MCA, Amal Jyothi College of Engineering Kanjirappally, Kottayam

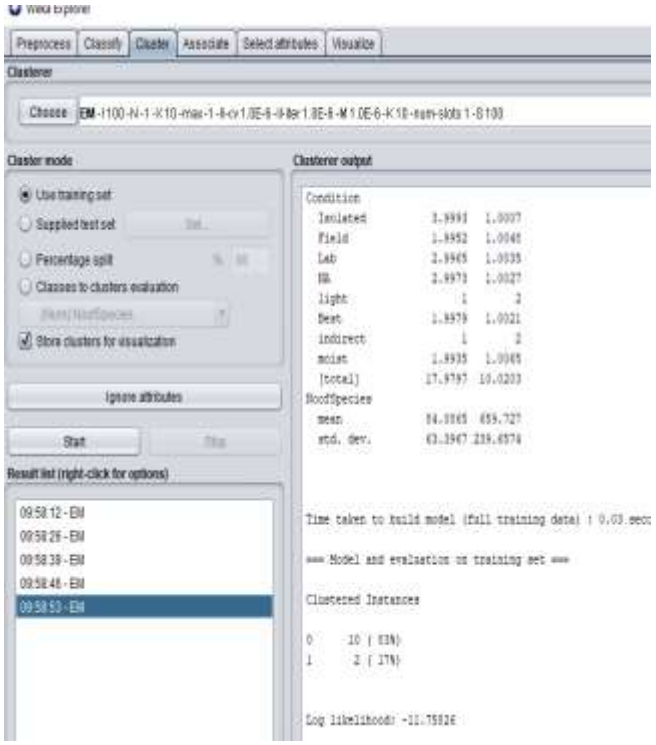


Fig No;7 Cluster

Use training set

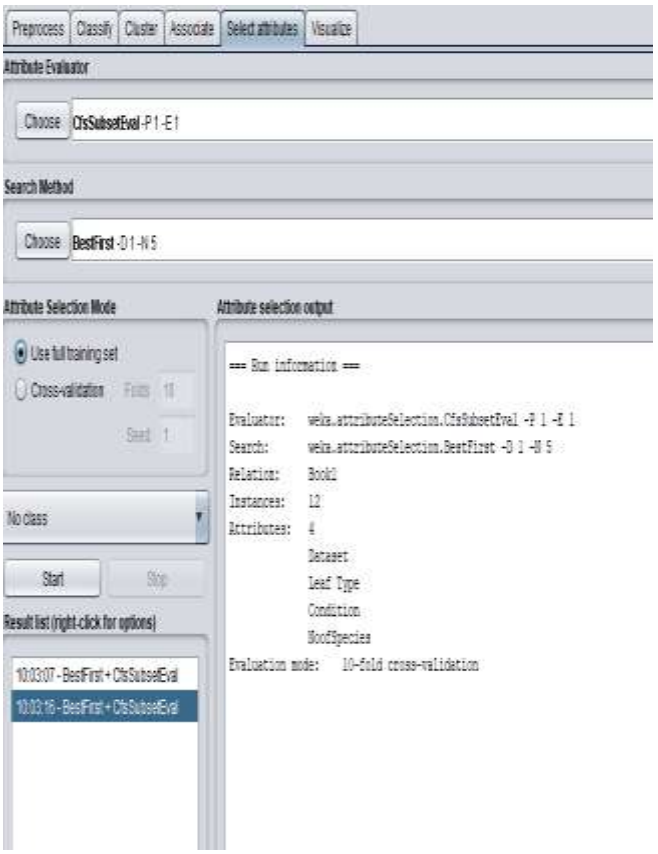


Fig No: -9 Select attributes-use full training set

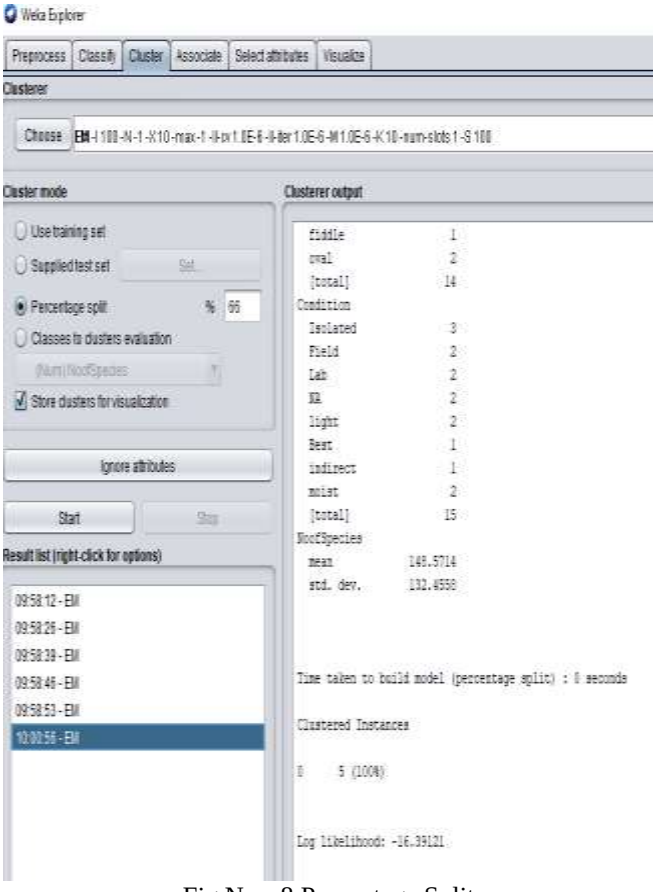


Fig No: -8 Percentage Split

Fig No: -10 Cross-Validation

DOI: 10.5281/zenodo.6181179
ISBN: 978-93-5607-317-3 @2022 MCA, Amal Jyothi College of Engineering Kanjirappally, Kottayam

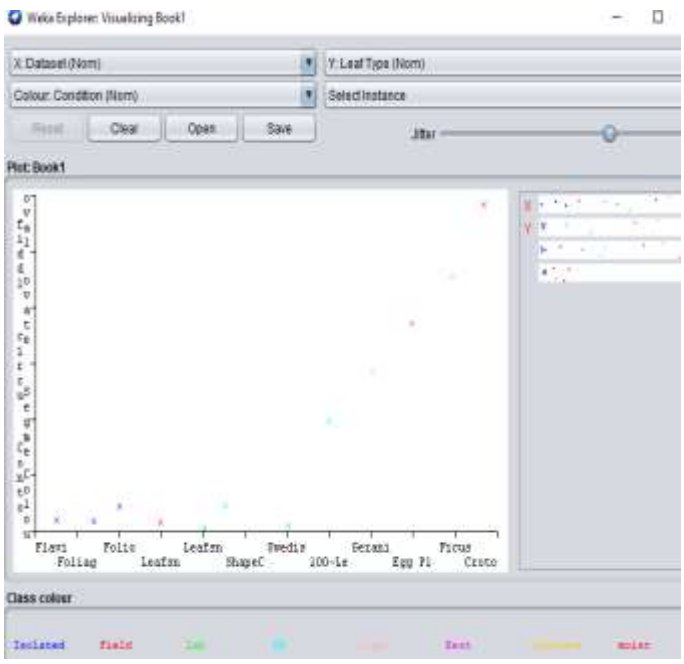


Fig No: -11 Visualize

ACKNOWLEDGMENT

The Department of Computer Applications at Amal Jyothi College of Engineering contributed to this research in taking part. This report and the research that would not have been feasible without my supervisor, Miss Sona Maria Sebastian, unwavering support. Her intelligence and encouragement which assisted me in finishing the work.

CONCLUSION

From given dataset we can conclude that it gives better result.

REFERENCES

[1] Sabu review of image features and classifiers used in leaf-based plant recognition approach.I

[2] Munisami, M.; Kishnah S. Plant Leaf Recognition Using Leaf Type with K-nearest Neighbour Classifiers. ProcediaComputer.

[3] Senthil Aranamani Plant Leaf Classification", International Journal of Computer Science,Vol 4, Issue 8, June 2011,pp.6-9.

[4] T. Beghin, J. SShape and texture-based plant leaf classification ssAdvanced Concepts for Intelligent Vision Systems (ACVIS), Vol 6475,2010,pp.45-353

DOI: 10.5281/zenodo.6181179
ISBN: 978-93-5607-317-3 @2022 MCA, Amal Jyothi College of Engineering Kanjirappally, Kottayam