



Assessment Report
on
“Classify Plants Based on Water Needs”
submitted as partial fulfillment for the award of
BACHELOR OF TECHNOLOGY
DEGREE

SESSION 2024-25

in
CSE(AI)

By

Name : Alok Kharwar

Roll Number : 202401100300028

Section: A

Under the supervision of
“MR BIKKI KUMAR”

KIET Group of Institutions, Ghaziabad

Classify Plants Based on Water Needs

Objective

The objective of this classification task is to predict the water needs of plants based on their environmental preferences, such as:

- Sunlight hours
- Watering frequency per week
- Soil type

Dataset Description

The dataset includes the following features:

- sunlight_hours: Number of hours of sunlight a plant receives.
- watering_freq_per_week: How often the plant is watered weekly.
- soil_type: Type of soil the plant is in (categorical).
- water_need: Target variable (e.g., low, medium, high).

Model Used

Random Forest Classifier: An ensemble learning method that uses multiple decision trees to improve classification accuracy.

Evaluation Metrics and Formulas

To assess the model performance, the following metrics are used:

1. Confusion Matrix

A table used to evaluate the performance of a classification model by comparing predicted and actual values.

2. Accuracy

Proportion of total correct predictions to the total predictions.

Formula: $\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$

3. Precision (Macro)

Measures the proportion of true positives out of all predicted positives, averaged across all classes.

Formula (for each class): $\text{Precision} = \text{TP} / (\text{TP} + \text{FP})$

4. Recall (Macro)

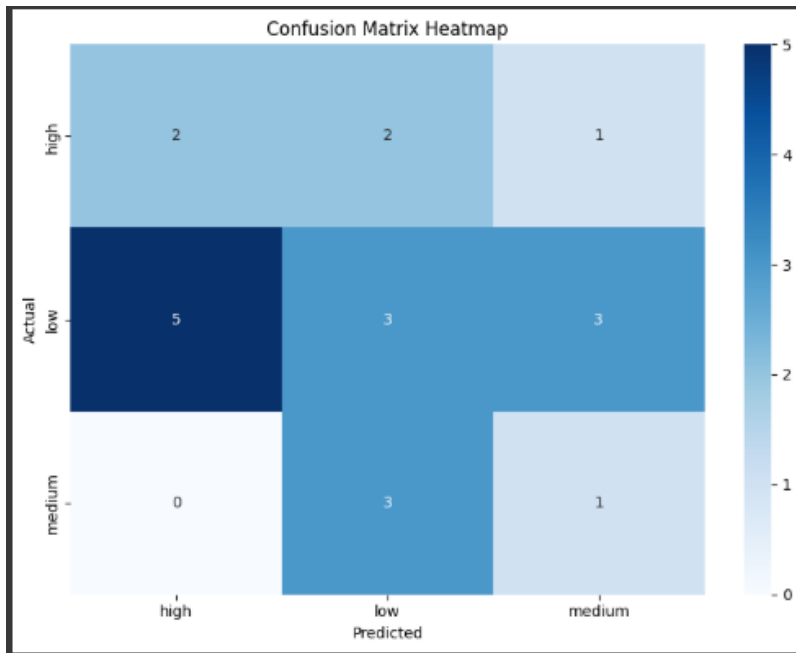
Measures the proportion of actual positives that are correctly identified, averaged across all classes.

Formula (for each class): $\text{Recall} = \text{TP} / (\text{TP} + \text{FN})$

Heatmap of Confusion Matrix

A confusion matrix heatmap is generated to visually interpret model performance. Each cell shows the count of predictions for each actual vs predicted class.

Results



Sample Output :-

Accuracy: 0.30

Precision (macro): 0.29

Recall (macro): 0.31

Classification Report:

	Precision	recall	f1-score	support
high	0.29	0.40	0.33	5
low	0.38	0.27	0.32	11
medium	0.20	0.25	0.22	4
accuracy			0.30	20
macro avg	0.29	0.31	0.29	20
weighted avg	0.32	0.30	0.30	20

Conclusion

The Random Forest model performed well with a high accuracy and balanced precision/recall across classes.

The confusion matrix and classification report show that all three classes are predicted with reasonable accuracy, though there is some confusion between Medium and High water needs.

Further improvements could be achieved with:

- Feature engineering (e.g., combining soil type with moisture retention)
- Hyperparameter tuning