

Course Title: Artificial Intelligence

Course Code: CSE 475

Lab-1

Submitted By

Kazi Md Nafiz Fuad

ID:2021-2-60-089

Submitted to

Dr. Raihan Ul Islam

Associate Professor

Department of Computer Science & Engineering

Introduction:

I conducted an analysis to evaluate how the accuracy of the Random Forest classifier changes with different numbers of trees. Using a dataset of mushrooms, our goal was to classify them as either edible or poisonous. We tested the Random Forest algorithm with varying numbers of trees (1, 50, 100, 150, 200, and 250) to see how this affects accuracy.

Data set:

The data set consists of various features describing mushrooms, such as cap shape, color, odor, and many others. Each mushroom is labeled as either edible or poisonous. Since all features are categorical, they need to be converted into numerical values for the model to process.

Methodology

Data Preparation:

I used label encoding to convert categorical features into numerical values. This involves assigning a unique number to each category.

Splitting the Data: The data set was divided into training and testing sets, with 80% used for training and 20% for testing.

Model Training and Evaluation

Random Forest Classifier: We trained the Random Forest classifier with different numbers of trees (n\_estimators).

Measuring Accuracy: For each value of n\_estimators, we measured how accurately the model predicted whether mushrooms were edible or poisonous on the test set.

Results

Here’s what we found:

1 Tree: Accuracy was 99.94%

50 Trees: Accuracy reached 100%

100 Trees: Accuracy remained at 100%

150 Trees: Accuracy was still 100%

200 Trees: Accuracy continued to be 100%

250 Trees: Accuracy stayed at 100%

Insights

**Decision Tree Performance:** Even with just one tree, the model performed remarkably well, achieving an accuracy of 99.94%. This suggests that the dataset is well-suited for decision tree-based models.

**Random Forest Performance:** Increasing the number of trees to 50 or more resulted in perfect accuracy (100%). This shows that Random Forest effectively uses multiple trees to enhance model stability and accuracy.

**Conclusion:**

1. Decision Tree Accuracy - 98.2%

2. Random Forest Accuracy - 99.22%

Hence, Ramdom Forest is comparitively better

The Random Forest classifier demonstrated exceptional performance on the mushroom data set. Even with a small number of trees, the model achieved high accuracy, and with 50 or more trees, it reached perfect classification. This confirms the robustness and reliability of the Random Forest algorithm for this type of classification problem.