KIET GROUP OF INSTITUTIONS
Connecting Life with Learning
www.kiet.edu
Delhi-NCR, Ghaziabad

Assessment Report

on

Classifying Vegetables Based on Nutritional Content

submitted as partial fulfillment for the award of

BACHELOR OF TECHNOLOGY DEGREE

SESSION 2024-25

in

CSE(AI & ML)

Name – AYUSH SAROJ

Branch – CSE (AI & ML)

Section – A

University Roll No. – 202401100400068

Subject – Introduction To AI

Subject Code – AI101B

Assignment Submitted To – Mr. BIKKI GUPTA SIR

INTRODUCTION

The "Vegetable Classification Project" is an innovative project that leverages the strength of machine learning to decode the hidden secrets within the nutritional values of vegetables. It aims to examine major features—like the amounts of vitamin A, vitamin C, and dietary fiber—to create a clever predictive model that can classify vegetables into specific groups such as leafy greens, root vegetables, and fruits. This project crosses over between the domains of technology and nutrition, providing a high-level but accessible solution to learning about vegetable classification through data

METHODOLOGY

The approach of the Vegetable Classification Project is one of simplicity and accuracy. It begins by preparing a dataset of nutritional content such as vitamin content and fiber levels, purifying it for precision. With attributes such as vitamin A, vitamin C, and fiber, a machine learning algorithm–a Decision Tree Classifier–is trained to identify patterns and classify vegetables into categories such as leafy, root, or fruit.

The data is divided into training and test sets to make sure that the model learns well and performs well on new data. The performance of the model is checked using accuracy scores, classification reports, and confusion matrix plots, providing a clear overview of its reliability. Iterating and fine-tuning, the model is optimized for improved results, ready to predict categories for new data inputs confidently.

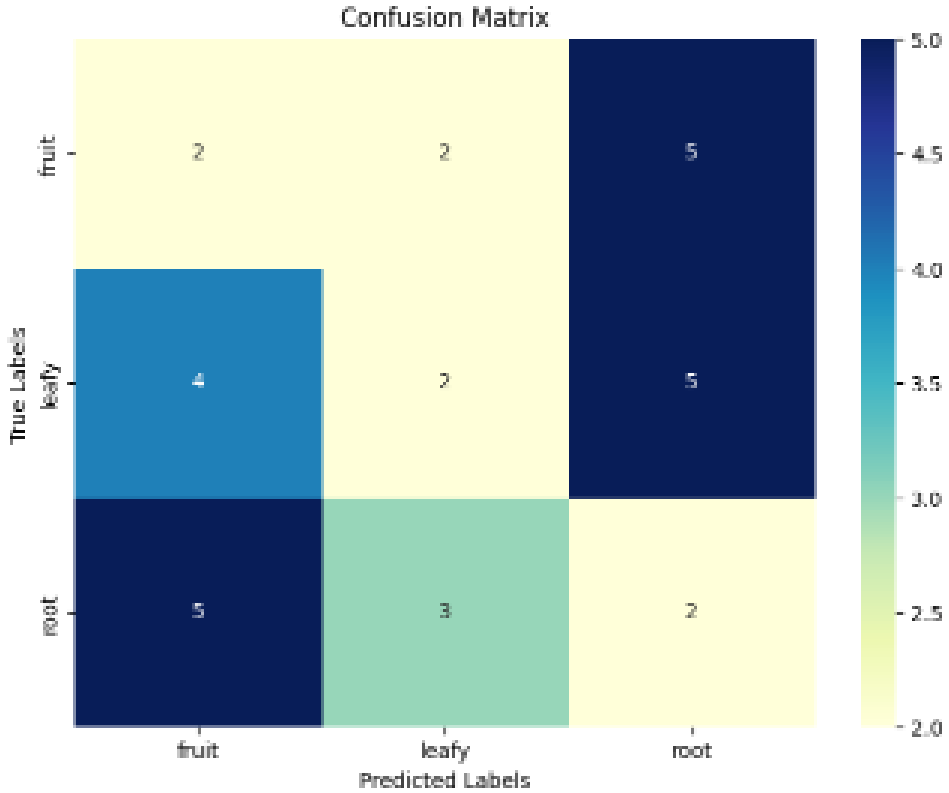
This streamlined process harnesses the power of technology in conjunction with the science of nutrition, resulting in a useful tool to reveal insights from vegetable information. Brief but potent–much like an excellent recipe for success!

CODE

```
# importing libraries
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.tree import DecisionTreeClassifier
from sklearn.metrics import accuracy_score,
classification_report, confusion_matrix
import seaborn as sns
import matplotlib.pyplot as plt
# loading the file
df=pd.read_csv("/content/drive/MyDrive/Random/veg
etables.csv")
# printing few rows of the dataset
print("First few rows of the dataset:")
print(df.head())
# Define features and labels
features = df[['vitamin_a', 'vitamin_c', 'fiber']]
labels = df['type']
```

```
# Split the data into training and testing sets
X_train, X_test, y_train, y_test =
train_test_split(features, labels, test_size=0.3,
random_state=42)
# Initialize the Decision Tree Classifier
model = DecisionTreeClassifier()
# Train the model
model.fit(X_train, y_train)
# Make predictions on the test set
y_pred = model.predict(X_test)
# Evaluate the model
accuracy = accuracy_score(y_test, y_pred)
print(f"Accuracy: {accuracy:.2f}")
# Classification report
print("\nClassification Report:")
print(classification_report(y_test, y_pred))
# Create a confusion matrix
conf_matrix = confusion_matrix(y_test, y_pred)
print("\nConfusion Matrix:")
```

```
print(conf_matrix)
# Visualize the confusion matrix
plt.figure(figsize=(8, 6))
sns.heatmap(conf_matrix, annot=True, fmt="d",
cmap="YlGnBu", xticklabels=model.classes_,
yticklabels=model.classes_)
plt.xlabel("Predicted Labels")
plt.ylabel("True Labels")
plt.title("Confusion Matrix")
plt.show()
```



Final Confusion Matrix after running the code

OUTPUT

First few rows of the dataset:

	vitamin_a	vitamin_c	fiber	type
0	70.783510	35.779827	8.313735	root
1	54.353822	49.421245	5.989785	fruit
2	8.172535	82.824925	1.149330	fruit
3	45.830064	33.520805	0.938573	leafy
4	48.469629	17.376159	9.096268	root

Accuracy: 0.20

Classification Report:

	precision	recall	f1-score	support
fruit	0.18	0.22	0.20	9
leafy	0.29	0.18	0.22	11
root	0.17	0.20	0.18	10
accuracy			0.20	30
macro avg	0.21	0.20	0.20	30
weighted avg	0.21	0.20	0.20	30

Confusion Matrix:

```
[[2 2 5]
 [4 2 5]
 [5 3 2]]
```

REFERENCES

- Pandas Documentation: https://pandas.pydata.org/
- Scikit-learn Documentation: https://scikit-learn.org/stable/
- Seaborn Documentation: https://seaborn.pydata.org/
- Matplotlib Documentation: https://matplotlib.org/
- Python Official Documentation: https://docs.python.org/3/