	Classification (Decision Tree)
In [1]:	This code imports the necessary libraries for data manipulation, model training, evaluation, and visualization in a decision tree classification task. It includes pandas for data handling, scikit-learn for model-related functions, and matplotlib for plotting the decision tree. import pandas as pd
	<pre>from sklearn.model_selection import train_test_split from sklearn.preprocessing import LabelEncoder from sklearn.tree import DecisionTreeClassifier from sklearn.metrics import accuracy_score, confusion_matrix, classification_report import matplotlib.pyplot as plt from sklearn import tree</pre>
In [2]:	Data Preparation # Load the dataset
In [3]:	<pre>df = pd.read_csv('Starbucks satisfactory survey_modified.csv') df = df.drop(columns=['Timestamp', 'Gender', 'Age']) # Separate the dataset into features and target variable x = df.drop(columns=['Continue to buy?'])</pre>
In [4]:	y = df['Continue to buy?']
	x = x.apply(le_x.fit_transform) This code initializes a LabelEncoder and applies it to the target variable y to convert its categorical labels into numerical values for model training.
In [5]:	<pre>le_y = LabelEncoder() y = le_y.fit_transform(y)</pre>
In [6]:	Modeling # Split the dataset into training and testing sets
In [7]:	<pre>X_train, X_test, y_train, y_test = train_test_split(x, y, test_size=0.2, random_state=42) # Define the classifiers clf_gini = DecisionTreeClassifier(max_depth=3, random_state=0) clf_entropy = DecisionTreeClassifier(criterion='entropy', max_depth=3, random_state=0)</pre>
In [8]:	# Fit the models clf_gini.fit(X_train, y_train) clf_entropy.fit(X_train, y_train)
Out[8]:	DecisionTreeClassifier(criterion='entropy', max_depth=3, random_state=0)
In [9]:	Model Evaluation # Predict using the models ***read minit at 216 minit and int (V to at)
In [13]:	<pre>y_pred_gini = clf_gini.predict(X_test) y_pred_entropy = clf_entropy.predict(X_test) # Evaluate the models print('Accuracy score for Gini criterion:', accuracy_score(y_test, y_pred_gini))</pre>
	<pre>print('Accuracy score for Entropy criterion:', accuracy_score(y_test, y_pred_entropy)) print('\nGini criterion') print('Training set score: {:.4f}'.format(clf_gini.score(X_train, y_train)))</pre>
	<pre>print('Test set score: {:.4f}'.format(clf_gini.score(X_test, y_test))) print('\nEntropy criterion') print('Training set score: {:.4f}'.format(clf_entropy.score(X_train, y_train))) print('Test set score: {:.4f}'.format(clf_entropy.score(X_test, y_test)))</pre>
	Accuracy score for Gini criterion: 0.84 Accuracy score for Entropy criterion: 0.84
	Gini criterion Training set score: 0.8557 Test set score: 0.8400 Entropy criterion
In [14]:	Training set score: 0.8557 Test set score: 0.8400 # Confusion matrix and classification report
In [15]:	<pre>cm_gini = confusion_matrix(y_test, y_pred_gini) cm_entropy = confusion_matrix(y_test, y_pred_entropy) print('Confusion matrix with criterion gini index: \n', cm_gini) print('Confusion matrix with criterion entropy: \n', cm_entropy)</pre>
	<pre>print('Confusion matrix with criterion entropy: \n', cm_entropy) print('Classification report with criterion gini index: \n', classification_report(y_test, y_pred_gini, zero_division=1)) print('Classification report with criterion entropy: \n', classification_report(y_test, y_pred_entropy, zero_division=1))</pre> Confusion matrix with criterion gini index:
	Confusion matrix with criterion gini index: [[2 2] [2 19]] Confusion matrix with criterion entropy: [[2 2]
	[2 19]] Classification report with criterion gini index: precision recall f1-score support
	0 0.50 0.50 4 1 0.90 0.90 0.90 21 accuracy 0.84 25 macro avg 0.70 0.70 25
	weighted avg 0.84 0.84 0.84 25 Classification report with criterion entropy: precision recall f1-score support
	0 0.50 0.50 4 1 0.90 0.90 0.90 21 accuracy 0.84 25
	macro avg 0.70 0.70 0.70 25 weighted avg 0.84 0.84 25
In [17]:	<pre># Visualize the decision trees plt.figure(figsize=(50,10)) tree.plot_tree(clf_gini, feature_names=x.columns, filled=True) plt.show()</pre>
	Rate the price range <= 1.5 gini = 0.372 samples = 97 value = [24, 73] Rate the product quality <= 1.5
	gini = 0.498 samples = 34 value = $[18, 16]$ Rate the ambiance $<= 0.5$ Rate the service of staff $<= 2.5$ gini = 0.172 samples = 63 value = $[6, 57]$ Has membership card $<= 0.5$
	$\begin{array}{c ccccccccccccccccccccccccccccccccccc$
In [18]:	<pre>value = [1, 1]</pre>
	<pre>plt.show()</pre> Rate the price range <= 1.5
	$\begin{array}{c} \text{entropy} = 0.807 \\ \text{samples} = 97 \end{array}$
	Rate the product quality <= 1.5 entropy = 0.998 samples = 34 Rate the ambiance <= 1.5 entropy = 0.454 samples = 63
	samples = 97 value = [24, 73] Rate the product quality <= 1.5 entropy = 0.998 Rate the ambiance <= 1.5 entropy = 0.454
	Rate the ambiance $<= 0.5$ entropy $= 0.454$ samples $= 63$ value $= [18, 16]$ Rate the ambiance $<= 0.5$ entropy $= 0.469$ samples $= 10$ Rate the service of staff $<= 2.5$ entropy $= 0.469$ samples $= 24$ Rate the ambiance $<= 0.5$ entropy $= 0.469$ samples $= 24$ Rate the service of staff $<= 2.5$ entropy $= 0.05$ samples $= 2.5$ entropy $= 0.349$ samples $= 61$
In [19]:	Rate the product quality <= 1.5 entropy = 0.988 samples = 34 value = [18, 16] Rate the ambiance <= 0.5 entropy = 0.454 samples = 20.5 entropy = 0.459 samples = 10 value = [9, 1] entropy = 1.0 samples = 20 value = [9, 1] entropy = 1.0 samples = 8 value = [1, 7] Part of the ambiance <= 1.5 entropy = 0.454 samples = 63 value = [6, 57] entropy = 0.0 samples = 20 value = [2, 0] entropy = 1.0 samples = 10 value = [4, 57] entropy = 0.0 samples = 20 value = [
In [19]:	Rate the product quality <= 1.5 entropy = 0.998 samples = 34 value = [18, 16] Rate the ambiance <= 0.5 entropy = 0.454 samples = 20 value = [9, 11] entropy = 1.0 samples = 8 value = [1, 1] entropy = 0.544 samples = 8 value = [1, 1] Pata Preparation # Load the dataset again for balancing df = pd.read_csv('Starbucks satisfactory survey_modified.csv') df = df.drop(columns=['Timestamp', 'Gender', 'Age']) #Separate the dataset into two DataFrames based on the class label
	Rate the ambiance <= 0.5 entropy = 0.498 samples = 34 value = [18, 16] Rate the ambiance <= 0.5 entropy = 0.494 samples = 63 value = [6, 57] Rate the ambiance <= 0.5 entropy = 0.494 samples = 63 value = [6, 57] Rate the service of staff <= 2.5 entropy = 0.494 samples = 63 value = [6, 57] Rate the membiance <= 0.5 entropy = 0.494 samples = 2 value = [2, 0] entropy = 0.05 samples = 2 value = [2, 0] entropy = 0.05 samples = 2 value = [2, 0] entropy = 0.05 samples = 2 value = [1, 1] Parameters in card <= 0.5 entropy = 0.05 samples = 2 value = [2, 0] entropy = 0.065 samples = 2 value = [2, 0] entropy = 0.605 samples = 2 value = [1, 0] entropy = 0.605 samples = 2 value = [2, 0] entropy = 0.605 samples = 2 value
In [20]:	Rate the product quality <= 1.5 entropy = 0.998 samples = 34 value = [18, 16] Rate the ambiance <= 0.5 entropy = 0.454 samples = 20 samples = 10 value = [9, 11] entropy = 1.05 samples = 8 value = [1, 1] Path and the dataset again for balancing df = pd. read_csv('Starbucks satisfactory survey_modified.csv') df = df.drop(columns=['Timestamp', 'Gender', 'Age']) #Separate the dataset into two DataFrames based on the class label df_majority = df[df['Continue to buy?'] == "No"] df_minority = df[df['Continue to buy?'] == "Yes"]
In [20]: In [21]: In [22]:	Rate the product quality <= 1.5 entropy = 0.998 samples = 30 value = [24, 73] Rate the ambiance <= 0.5 entropy = 0.998 samples = 30 value = [6, 57] Rate the service of staff <= 2.5 entropy = 0.00 value = [6, 57] Rate the service of staff <= 2.5 entropy = 0.00 value = [6, 57] Rate the service of staff <= 2.5 entropy = 0.00 value = [6, 57] Rate the service of staff <= 2.5 entropy = 0.00 value = [6, 57] Rate the service of staff <= 2.5 entropy = 0.00 value = [6, 57] Rate the service of staff <= 2.5 entropy = 0.00 value = [6, 57] Rate the service of staff <= 2.5 entropy = 0.00 value = [6, 57] Rate the service of staff <= 2.5 entropy = 0.00 value = [6, 57] Rate the service of staff <= 2.5 entropy = 0.00 value = [6, 57] Rate the service of staff <= 2.5 entropy = 0.00 value = [6, 57] Rate the service of staff <= 2.5 entropy = 0.00 value = [6, 57] Rate the service of staff <= 2.5 entropy = 0.00 value = [6, 57] Rate the service of staff <= 2.5 entropy = 0.00 value = [6, 57] Rate the service of staff <= 2.5 entropy = 0.00 value = [6, 57] Rate the service of staff <= 2.5 entropy = 0.00 value = [6, 57] Rate the service of staff <= 2.5 entropy = 0.00 value = [6, 57] Rate the service of staff <= 2.5 entropy = 0.00 value = [6, 57] Rate the service of staff <= 2.5 entropy = 0.00 value = [6, 57] Rate the service of staff <= 2.5 entropy = 0.00 value = [6, 57] Rate the service of staff <= 2.5 entropy = 0.00 value = [6, 57] Rate the service of staff <= 2.5 entropy = 0.00 value = [6, 57] Rate the service of staff <= 2.5 entropy = 0.00 value = [6, 57] Rate the service of staff <= 2.5 entropy = 0.00 value = [6, 57] Rate the service of staff <= 2.5 entropy = 0.00 value = [6, 57] Rate the service of staff <= 2.5 entropy = 0.00 value = [6, 57] Rate the service of staff <= 2.5 entropy = 0.00 value = [6, 57] Rate the service of staff <= 2.5 entropy = 0.00 value = [6, 57] Rate the service of staff <= 2.5 entropy = 0.00 value = [6, 57] Rate the service of staff <= 2.5 entropy = 0.00 value = [6, 57] Rate the se
In [20]: In [21]: In [22]: In [23]: In [24]:	Rate the product quality <= 1.5 entropy = 0.936 value = [24, 73] Rate the product quality <= 1.5 entropy = 0.434 value = [18, 16] Rate the ambience <= 0.5 entropy = 0.434 value = [18, 16] Rate the ambience <= 0.5 entropy = 0.434 value = [18, 16] Rate the service of staff <= 2.5 entropy = 0.434 value = [16, 57] Rate the mabience <= 0.5 entropy = 0.434 value = [16, 57] Has memberable carried the samples = 10 value = [16, 57] Has memberable carried the samples = 24 value = [1, 17] Has memberable carried the samples = 24 value = [1, 17] Has memberable carried the samples = 24 value = [1, 17] Has memberable carried the samples = 24 value = [1, 17] Has memberable carried the samples = 10 value = [1, 17] Has memberable carried the samples = 10 value = [1, 17] Has memberable carried the samples = 10 value = [1, 17] Has memberable carried the samples = 10 value = [1, 17] Has memberable carried the samples = 10 value = [1, 17] Has memberable carried the samples = 10 value = [1, 17] Has memberable carried the samples = 10 value = [1, 17] Has memberable carried the samples = 10 value = [1, 17] Has memberable carried the samples = 10 value = [1, 17] Has memberable carried the sample = 10 value = [1, 17] Has memberable carried the sample = 10 value = [1, 17] Has memberable carried the sample = 10 value = [1, 17] Has memberable carried the sample = 10 value = [1, 17] Has memberable carried the sample = 10 value = [1, 17] Has memberable carried the sample = 10 value = [1, 17] Has memberable carried the sample = 10 value = [1, 17] Has memberable carried the sample = 10 value = [1, 17] Has memberable carried the sample = 10 value = [1, 17] Has memberable carried the sample = 10 value = [1, 17] Has memberable carried the sample = 10 value = [1, 17] Has memberable carried the sample = 10 value = [1, 17] Has memberable carried the sample = 10 value = [1, 17] Has memberable entropy = 0.55 value = [1, 17] Has memberable entropy = 0.55 value = [1, 17] Has memberable entropy = 0.55 value = [1, 17] Ha
In [20]: In [21]: In [22]: In [23]:	Rate the product quality <= 1.5 entropy = 0.936 value = [24, 73] Rate the product quality <= 1.5 entropy = 0.434 value = [18, 16] Rate the ambience <= 0.5 entropy = 0.434 value = [18, 16] Rate the ambience <= 0.5 entropy = 0.434 value = [18, 16] Rate the service of staff <= 2.5 entropy = 0.434 value = [16, 57] Rate the mabience <= 0.5 entropy = 0.434 value = [16, 57] Has memberable carried the samples = 10 value = [16, 57] Has memberable carried the samples = 24 value = [1, 17] Has memberable carried the samples = 24 value = [1, 17] Has memberable carried the samples = 24 value = [1, 17] Has memberable carried the samples = 24 value = [1, 17] Has memberable carried the samples = 10 value = [1, 17] Has memberable carried the samples = 10 value = [1, 17] Has memberable carried the samples = 10 value = [1, 17] Has memberable carried the samples = 10 value = [1, 17] Has memberable carried the samples = 10 value = [1, 17] Has memberable carried the samples = 10 value = [1, 17] Has memberable carried the samples = 10 value = [1, 17] Has memberable carried the samples = 10 value = [1, 17] Has memberable carried the samples = 10 value = [1, 17] Has memberable carried the sample = 10 value = [1, 17] Has memberable carried the sample = 10 value = [1, 17] Has memberable carried the sample = 10 value = [1, 17] Has memberable carried the sample = 10 value = [1, 17] Has memberable carried the sample = 10 value = [1, 17] Has memberable carried the sample = 10 value = [1, 17] Has memberable carried the sample = 10 value = [1, 17] Has memberable carried the sample = 10 value = [1, 17] Has memberable carried the sample = 10 value = [1, 17] Has memberable carried the sample = 10 value = [1, 17] Has memberable carried the sample = 10 value = [1, 17] Has memberable carried the sample = 10 value = [1, 17] Has memberable carried the sample = 10 value = [1, 17] Has memberable entropy = 0.55 value = [1, 17] Has memberable entropy = 0.55 value = [1, 17] Has memberable entropy = 0.55 value = [1, 17] Ha
In [20]: In [21]: In [22]: In [23]: In [24]:	Rate the amblance <= 0.5
In [20]: In [21]: In [22]: In [23]: In [24]:	Pate the arbitine = 0.5 Samples = 21 Samples = 0.5 Sampl
In [20]: In [21]: In [22]: In [23]: In [24]: In [26]:	Again the product stally x = 1.5 white = (24, 23) white =
<pre>In [20]: In [21]: In [22]: In [23]: In [24]: In [25]: Out[27]:</pre>	And the product suply < 15 service - 97 state 15 ser
<pre>In [20]: In [21]: In [22]: In [23]: In [24]: In [25]: Out[27]:</pre>	Let me provert make = 10. Serving = 2.03 Ser
In [20]: In [21]: In [22]: In [23]: In [24]: In [25]: Out [27]: In [28]:	Data Preparation Load the dataset again for behances Load the dataset again Load the dataset ag
In [20]: In [21]: In [22]: In [23]: In [24]: In [25]: Out [27]: In [28]:	### Control of the American Control of the State Control of the Co
In [20]: In [21]: In [22]: In [23]: In [24]: In [25]: Out [27]: In [28]:	Data Preparation For any Public Processing Company and Company an
In [20]: In [21]: In [22]: In [23]: In [24]: In [25]: Out [27]: In [28]:	Controlled to the pasted gapley = 15 Secretary Controlled to the pasted gapley = 15 Secretary Controlled to the pasted gapley = 16 Secretary Controlled to the pasted gapley Controlled gapley Control
In [20]: In [21]: In [22]: In [23]: In [24]: In [25]: In [27]: Out [27]: In [28]: In [29]:	Data Preparation Compared to the content pains of the content pains o
In [20]: In [21]: In [22]: In [23]: In [24]: In [26]: In [27]: In [27]: In [27]: In [28]:	Data Preparation ***Commence of the Commence
In [20]: In [21]: In [22]: In [23]: In [24]: In [26]: In [27]: Out [27]: In [28]:	Data Proparation ***Application above again for information and an application and applicatio
In [20]: In [21]: In [22]: In [23]: In [24]: In [26]: In [27]: In [27]: In [27]: In [28]:	Data Preparation Figure 1 of 10 and
In [20]: In [21]: In [22]: In [23]: In [24]: In [26]: In [27]: In [27]: In [27]: In [28]:	Data Proparation June 1997 -
In [20]: In [21]: In [22]: In [23]: In [24]: In [26]: In [27]: In [27]: In [27]: In [28]:	Data Preparation Final Properties and the second s
In [20]: In [21]: In [22]: In [23]: In [24]: In [26]: In [27]: In [27]: In [27]: In [28]:	Data Preparation **Continue to Continue t
In [20]: In [21]: In [22]: In [23]: In [24]: In [26]: In [27]: In [27]: In [27]: In [28]:	Data Preparation Data Prepara
In [20]: In [21]: In [23]: In [24]: In [27]: Out [27]: In [28]: In [32]: In [33]:	Data Preparation Figure 1. The control of the cont
In [20]: In [21]: In [23]: In [24]: In [27]: Out [27]: In [28]: In [32]: In [33]:	Date Preparation The and the place again for the black of the place o
In [20]: In [21]: In [23]: In [24]: In [27]: Out [27]: In [28]: In [32]: In [33]:	Decar Preparation Figure 2 and Section Sectio
In [20]: In [21]: In [23]: In [24]: In [26]: In [27]: Out[27]: In [28]: In [32]:	Data Propuration Service Propuration Data Prop
In [20]: In [21]: In [23]: In [24]: In [26]: In [27]: Out [27]: In [28]: In [32]:	Dotal Programme of the Control of th

Rate the product quality <= 1.5 entropy = 0.702 samples = 21 value = [17, 4]

> entropy = 0.971 samples = 5 value = [2, 3]

entropy = 0.0 samples = 8 value = [8, 0] Rate the service of staff <= 2.5 entropy = 0.89 samples = 13 value = [9, 4]

entropy = 0.544 samples = 8 value = [7, 1] Has membership card <= 0.5 entropy = 0.755 samples = 23 value = [5, 18]

> entropy = 0.592 samples = 7 value = [1, 6]

entropy = 0.0 samples = 9 value = [0, 9]

Rate the ambiance <= 2.5 entropy = 0.94 samples = 14 value = [5, 9]

entropy = 0.985 samples = 7 value = [4, 3]