Python Programming

Machine Learning Assignment : 14

There is one data set of wether conditions.
That dataset contains information as wether and we have to decides whether to play or not.
Data set contains the target variable as Play which indicates whether to play or not.

Consider below Marvellous Infosystems Play Predictor Dataset as

Marvellous Infosystems Play Predictor

	Wether	Temperature	Play
1	Sunny	Hot	No
2	Sunny	Hot	No
3	Overcast	Hot	Yes
4	Rainy	Mild	Yes
5	Rainy	Cool	Yes
6	Rainy	Cool	No
7	Overcast	Cool	Yes
8	Sunny	Mild	No
9	Sunny	Cool	Yes
10	Rainy	Mild	Yes

According to above dataset there are two features as

1. Wether
2. Temperature

We have two labels as

1. Yes
2. No

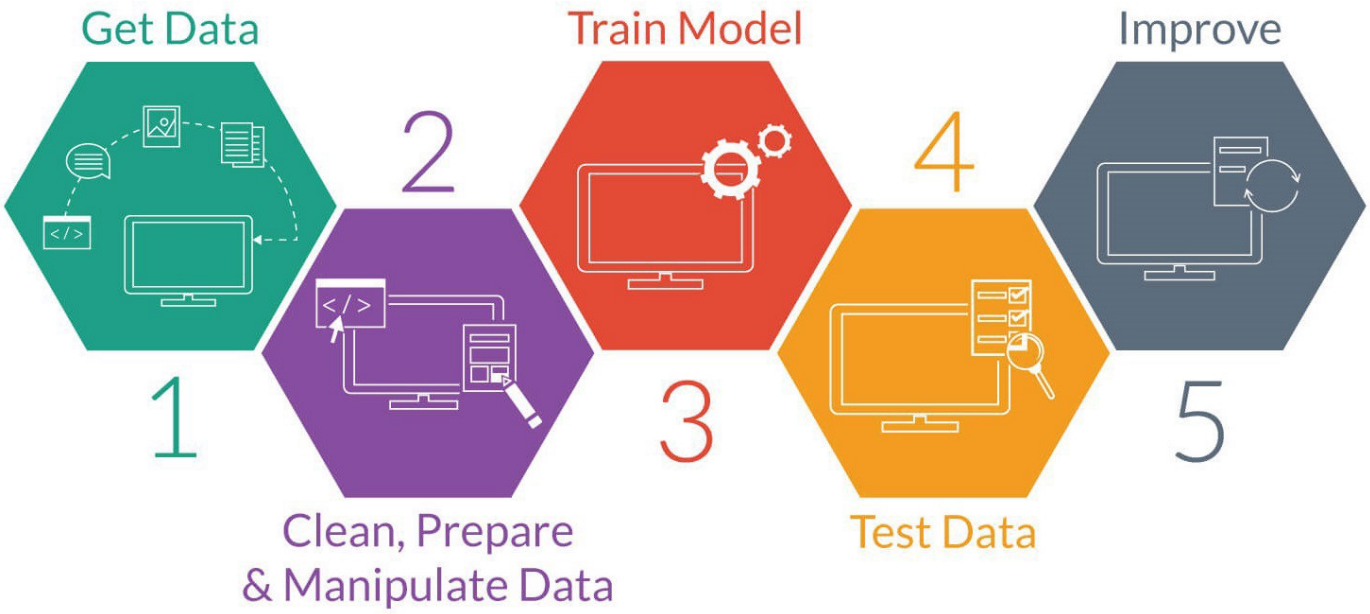
There are three types of different entries under Wether as

1. Sunny
2. Overcast
3. Rainy

There are three types of different entries under Temperature as

1. Hot
2. Cold
3. Mild

We have to design Machine Learning application which uses Classification technique.



Design machine learning application which follows below steps as

Step 1:
Get Data
Load data from MarvellousInfosystems_PlayPredictor.csv file into python application.

Step 2:
Clean, Prepare and Manipulate data
As we want to use the above data into machine learning application we have prepare that in the format which is accepted by the algorithms.
As our dataset contains two features as Wether and Temperature. We have to replace each string field into numeric constants by using LabelEncoder from processing module of sklearn.

Step 3:
Train Data
Now we want to train our data for that we have to select the Machine learning algorithm. For that we select K Nearest Neighbour algorithm.
use fit method for training purpose. For training use whole dataset.

Step 4:
Test Data
After successful training now we can test our trained data by passing some value of wether and temperature.
As we are using KNN algorithm use value of K as 3.
After providing the values check the result and display on screen.
Result may be Yes or No.

Step 5:
Calculate Accuracy
Write one function as CheckAccuracy() which calculate the accuracy of our algorithm.
For calculating the accuracy divide the dataset into two equal parts as Training data and Testing data.
Calculate Accuracy by changing value of K.