**A Project on**

**“Classification of IRIS Flower Species Using Machine Learning Techniques with Python “**

**What is Machine learning?**

Machine Learning is simply the ability of a machine to learn from some set of data and make prediction on the basis of that data.

**Applications of Machine Learning:**

* Youtube , which recommends videos based on our previous watched videos with the help of Machine Learning algorithm
* Autonomous helicopter, handwriting recognition, Face recognition.
* Self-customizing programs E.g., Amazon, Netflix product recommendations
* It is used in banking and financial sector, healthcare, retail, publishing and social media, etc.
* It is used by Google and Facebook to push relevant advertisements based on users search history.

**Problem Definition**

Our task is to classify iris flowers among three species (setosa, versicolor or virginica) from measurements of length and width of sepals and petals.

The iris data set contains 3 classes of 50 instances each, where each class refers to a type of iris plant. The central goal here is to design a model which makes good classifications for new data.

****

**IRIS flower Dataset:**

* The data set consists of 50 samples from each of three species of Iris (Iris setosa, Iris versicolor, , Iris virginica).
* Four features were measured from each sample (in centimetres):
  + Length of the sepals
  + Width of the sepals
  + Length of the petals
  + Width of the petals

**Sample of each data set**

|  |  |  |
| --- | --- | --- |
| 5.1,3.5,1.4,0.2,Iris-setosa  4.9,3.0,1.4,0.2,Iris-setosa  4.7,3.2,1.3,0.2,Iris-setosa  4.6,3.1,1.5,0.2,Iris-setosa | 5.0,2.0,3.5,1.0,Iris-versicolor  5.9,3.0,4.2,1.5,Iris-versicolor  6.0,2.2,4.0,1.0,Iris-versicolor  6.1,2.9,4.7,1.4,Iris-versicolor | 6.3,3.3,6.0,2.5,Iris-virginica  5.8,2.7,5.1,1.9,Iris-virginica  7.1,3.0,5.9,2.1,Iris-virginica  6.3,2.9,5.6,1.8,Iris-virginica |

**Packages to be installed:**

1.numpy

2.scipy

3.sklearn

4.matplotlib

5.pandas

To install these packages

Goto command prompt and type

pip install packagename .

For example D:/ pip install numpy

**Implementation:**

**1. Data collection**

The data set required for the classification of the iris flower species was collected from the **UCI** repository (<https://archive.ics.uci.edu/ml/machine-learning-databases/iris/iris.data>). It is also available in in sklearn package.

**2. Load iris flower data set in our program**

**#load iris dataset to our program**

from sklearn.datasets import load\_iris

iris=load\_iris()

X=iris.data #contains data

Y=iris.target #contains results

**3. Describing the data set**

#We analyze and describe the features and targets of iris dataset

Print(X) #will display all features of iris dataset

Print(Y) #will display all the target of iris data set(result)

print(X.shape) #

print(Y.shape)

**4. Splitting the dataset for training and testing**

#The dataset was splitted. (70% for training and 30% for testing). This was done by importing #train\_test\_split function .

from sklearn.model\_selection import train\_test\_split

X\_train,X\_test,Y\_train,Y\_test=train\_test\_split(X,Y,test\_size=0.3,random\_state=3)

print(X\_train.shape)

print(Y\_train.shape)

**5. Train the dataset using algorithms**

#In this step we use k-nearest neighbors algorithm to train the 70% dataset .

#Here value of k is 5

from sklearn.neighbors import KNeighborsClassifier

knn=KNeighborsClassifier(n\_neighbors=5)

knn.fit(X\_train,Y\_train)

**6. Test the model:**

Y\_pred\_knn=knn.predict(X\_test)

#Here we tested the 30% of data set to predict the targets

**7. Find accuracy**

from sklearn.metrics import accuracy\_score

knn\_accuracy=accuracy\_score(Y\_test,Y\_pred\_knn)

print("Accuracy in knn",knn\_accuracy)