**HASHANALYTICS DATA ANALYSIS INTERNSHIP-Charity Wanjiru**

**ASSIGNMENT 4**

**Decision Tree- Iris Dataset**

Installations needed

* pandas
* seaborn
* sklearn-kit

The first step after installing the needed packages, Create new project and inside the project create new file where all the code goes. Then import the packages you will need. For example import pandas as pd

To get started you need to import all the packages like so

import pandas as pd

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

from sklearn.tree import DecisionTreeClassifier

from sklearn import metrics

Read read your data from csv file

iris\_dataset=pd.read\_csv('iris.csv')

Dividing the dataset into dependent and independent variables

X=iris\_dataset.iloc[:,:-1].values #feature variable

Y=iris\_dataset.iloc[:,4].values #Target

Splitting dataset into training and testing(70/30)respectively

X\_train,X\_test,Y\_train,Y\_test=train\_test\_split(X,Y,test\_size=0.3,random\_state=1)

Create decision tree object and Train Decision Tree Classifier

classifier=DecisionTreeClassifier()

classifier=classifier.fit(X\_train,Y\_train)

Y\_predict=classifier.predict(X\_test)

Calculating the accuracy of the model using metrics

#model Accuraccy

d=metrics.accuracy\_score(Y\_test, Y\_predict)