**Expt-1**

**Aim: To access and observe the samples from**

1. **iris dataset**
2. **digits dataset**
3. **diabetes dataset**

**1.a Program to access and observe the samples from iris dataset.**

**from sklearn import datasets**

**import pandas as pd**

**print('THIS IS  THE PROGRAM TO ACCESS  IRIS DATASET')**

**iris=datasets.load\_iris()**

**print(" To print the description of Iris Dataset")**

**print(iris.DESCR)**

**print('\n\n\n\n')**

**# df will fold dataset as a table**

**df=pd.DataFrame(**

**iris.data,**

**columns=iris.feature\_names**

**)**

**#labels are assigned to df[target] table or array**

**df['target']=pd.Series(**

**iris.target**

**)**

**df['target\_names']=df['target'].apply(lambda y:iris.target\_names[y])**

**print('To display First 5 samples')**

**# df.head(5) will return the first five samples in the dataset**

**print(df.head(5))**

**print('To display randomply 5 samples')**

**#df.sample(5) will return randomly five samples from the dataset**

**print(df.sample(5))**

**# Train Test Split Ratio**

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

**df\_train,df\_test=train\_test\_split(df,test\_size=0.3) # For 70: 30 Split**

**print('The total number of samples in the dataset = ',df.shape[0])**

**print('The number of samples  in training set = ',df\_train.shape[0])**

**print('The number of samples  in testing set = ',df\_test.shape[0])**

**print('The first five samples of training set')**

**print(df\_train.head(5))**

**print('\n\nThe first five samples of testing set')**

**print(df\_test.head(5))**

**OUTPUT**

**THIS IS THE PROGRAM TO ACCESS IRIS DATASET**

**To print the description of Iris Dataset**

**.. \_iris\_dataset:**

**Iris plants dataset**

**--------------------**

**\*\*Data Set Characteristics:\*\***

**:Number of Instances: 150 (50 in each of three classes)**

**:Number of Attributes: 4 numeric, predictive attributes and the class**

**:Attribute Information:**

**- sepal length in cm**

**- sepal width in cm**

**- petal length in cm**

**- petal width in cm**

**- class:**

**- Iris-Setosa**

**- Iris-Versicolour**

**- Iris-Virginica**

**:Summary Statistics:**

**============== ==== ==== ======= ===== ====================**

**Min Max Mean SD Class Correlation**

**============== ==== ==== ======= ===== ====================**

**sepal length: 4.3 7.9 5.84 0.83 0.7826**

**sepal width: 2.0 4.4 3.05 0.43 -0.4194**

**petal length: 1.0 6.9 3.76 1.76 0.9490 (high!)**

**petal width: 0.1 2.5 1.20 0.76 0.9565 (high!)**

**============== ==== ==== ======= ===== ====================**

**:Missing Attribute Values: None**

**:Class Distribution: 33.3% for each of 3 classes.**

**:Creator: R.A. Fisher**

**:Donor: Michael Marshall (**[**MARSHALL%PLU@io.arc.nasa.gov**](mailto:MARSHALL%25PLU@io.arc.nasa.gov)**)**

**:Date: July, 1988**

**The famous Iris database, first used by Sir R.A. Fisher. The dataset is taken**

**from Fisher's paper. Note that it's the same as in R, but not as in the UCI**

**Machine Learning Repository, which has two wrong data points.**

**This is perhaps the best known database to be found in the**

**pattern recognition literature. Fisher's paper is a classic in the field and**

**is referenced frequently to this day. (See Duda & Hart, for example.) The**

**data set contains 3 classes of 50 instances each, where each class refers to a**

**type of iris plant. One class is linearly separable from the other 2; the**

**latter are NOT linearly separable from each other.**

**.. topic:: References**

**- Fisher, R.A. "The use of multiple measurements in taxonomic problems"**

**Annual Eugenics, 7, Part II, 179-188 (1936); also in "Contributions to**

**Mathematical Statistics" (John Wiley, NY, 1950).**

**- Duda, R.O., & Hart, P.E. (1973) Pattern Classification and Scene Analysis.**

**(Q327.D83) John Wiley & Sons. ISBN 0-471-22361-1. See page 218.**

**- Dasarathy, B.V. (1980) "Nosing Around the Neighborhood: A New System**

**Structure and Classification Rule for Recognition in Partially Exposed**

**Environments". IEEE Transactions on Pattern Analysis and Machine**

**Intelligence, Vol. PAMI-2, No. 1, 67-71.**

**- Gates, G.W. (1972) "The Reduced Nearest Neighbor Rule". IEEE Transactions**

**on Information Theory, May 1972, 431-433.**

**- See also: 1988 MLC Proceedings, 54-64. Cheeseman et al"s AUTOCLASS II**

**conceptual clustering system finds 3 classes in the data.**

**- Many, many more ...**

**To display First 5 samples**

**sepal length (cm) sepal width (cm) petal length (cm) petal width (cm) \**

**0 5.1 3.5 1.4 0.2**

**1 4.9 3.0 1.4 0.2**

**2 4.7 3.2 1.3 0.2**

**3 4.6 3.1 1.5 0.2**

**4 5.0 3.6 1.4 0.2**

**target target\_names**

**0 0 setosa**

**1 0 setosa**

**2 0 setosa**

**3 0 setosa**

**4 0 setosa**

**To display randomply 5 samples**

**sepal length (cm) sepal width (cm) petal length (cm) petal width (cm) \**

**3 4.6 3.1 1.5 0.2**

**94 5.6 2.7 4.2 1.3**

**55 5.7 2.8 4.5 1.3**

**107 7.3 2.9 6.3 1.8**

**139 6.9 3.1 5.4 2.1**

**target target\_names**

**3 0 setosa**

**94 1 versicolor**

**55 1 versicolor**

**107 2 virginica**

**139 2 virginica**

**The total number of samples in the dataset = 150**

**The number of samples in training set = 105**

**The number of samples in testing set = 45**

**The first five samples of training set**

**sepal length (cm) sepal width (cm) petal length (cm) petal width (cm) \**

**104 6.5 3.0 5.8 2.2**

**23 5.1 3.3 1.7 0.5**

**18 5.7 3.8 1.7 0.3**

**52 6.9 3.1 4.9 1.5**

**75 6.6 3.0 4.4 1.4**

**target target\_names**

**104 2 virginica**

**23 0 setosa**

**18 0 setosa**

**52 1 versicolor**

**75 1 versicolor**

**The first five samples of testing set**

**sepal length (cm) sepal width (cm) petal length (cm) petal width (cm) \**

**55 5.7 2.8 4.5 1.3**

**124 6.7 3.3 5.7 2.1**

**111 6.4 2.7 5.3 1.9**

**74 6.4 2.9 4.3 1.3**

**76 6.8 2.8 4.8 1.4**

**target target\_names**

**55 1 versicolor**

**124 2 virginica**

**111 2 virginica**

**74 1 versicolor**

**76 1 versicolor**

**Assignment**

1. **Show that in two different executions same samples will not be present training and testing set**
2. **Give reason why target names appear as setosa, versicular and virginica?**

**Ans: The following command gives target names.**

**iris.target\_names**

**OUTPUT: array(['setosa', 'versicolor', 'virginica'], dtype='<U10')**

1. **Execute the following commands and give reason for the outputs obtained.**

**print(df.shape)**

**print(df\_train.shape)**

**print(df\_test.shape)**

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**1.b to access and observe the samples from iris dataset.**

**from sklearn import datasets**

**import pandas as pd**

**print('THIS IS  THE PROGRAM TO ACCESS  DIGITS DATASET')**

**digit=datasets.load\_digits()**

**print(" To print the description of Digits Dataset")**

**print(digit.DESCR)**

**print('\n\n\n\n')**

**# df will fold dataset as a table**

**df=pd.DataFrame(**

**digit.data,**

**columns=digit.feature\_names**

**)**

**#labels are assigned to df[target] table or array**

**df['target']=pd.Series(**

**digit.target**

**)**

**df['target\_names']=df['target'].apply(lambda y:digit.target\_names[y])**

**print('To display First 5 samples')**

**# df.head(5) will return the first five samples in the dataset**

**print(df.head(5))**

**print('To display randomply 5 samples')**

**#df.sample(5) will return randomly five samples from the dataset**

**print(df.sample(5))**

**# Train Test Split Ratio**

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

**df\_train,df\_test=train\_test\_split(df,test\_size=0.3) # For 70: 30 Split**

**print('The total number of samples in the dataset = ',df.shape[0])**

**print('The number of samples  in training set = ',df\_train.shape[0])**

**print('The number of samples  in testing set = ',df\_test.shape[0])**

**print('The first five samples of training set')**

**print(df\_train.head(5))**

**print('\n\nThe first five samples of testing set')**

**print(df\_test.head(5))**

**OUTPUT**

**THIS IS THE PROGRAM TO ACCESS DIGITS DATASET**

**To print the description of Digits Dataset**

**.. \_digits\_dataset:**

**Optical recognition of handwritten digits dataset**

**--------------------------------------------------**

**\*\*Data Set Characteristics:\*\***

**:Number of Instances: 1797**

**:Number of Attributes: 64**

**:Attribute Information: 8x8 image of integer pixels in the range 0..16.**

**:Missing Attribute Values: None**

**:Creator: E. Alpaydin (alpaydin '@' boun.edu.tr)**

**:Date: July; 1998**

**This is a copy of the test set of the UCI ML hand-written digits datasets**

[**https://archive.ics.uci.edu/ml/datasets/Optical+Recognition+of+Handwritten+Digits**](https://archive.ics.uci.edu/ml/datasets/Optical+Recognition+of+Handwritten+Digits)

**The data set contains images of hand-written digits: 10 classes where**

**each class refers to a digit.**

**Preprocessing programs made available by NIST were used to extract**

**normalized bitmaps of handwritten digits from a preprinted form. From a**

**total of 43 people, 30 contributed to the training set and different 13**

**to the test set. 32x32 bitmaps are divided into nonoverlapping blocks of**

**4x4 and the number of on pixels are counted in each block. This generates**

**an input matrix of 8x8 where each element is an integer in the range**

**0..16. This reduces dimensionality and gives invariance to small**

**distortions.**

**For info on NIST preprocessing routines, see M. D. Garris, J. L. Blue, G.**

**T. Candela, D. L. Dimmick, J. Geist, P. J. Grother, S. A. Janet, and C.**

**L. Wilson, NIST Form-Based Handprint Recognition System, NISTIR 5469,**

**1994.**

**.. topic:: References**

**- C. Kaynak (1995) Methods of Combining Multiple Classifiers and Their**

**Applications to Handwritten Digit Recognition, MSc Thesis, Institute of**

**Graduate Studies in Science and Engineering, Bogazici University.**

**- E. Alpaydin, C. Kaynak (1998) Cascading Classifiers, Kybernetika.**

**- Ken Tang and Ponnuthurai N. Suganthan and Xi Yao and A. Kai Qin.**

**Linear dimensionalityreduction using relevance weighted LDA. School of**

**Electrical and Electronic Engineering Nanyang Technological University.**

**2005.**

**- Claudio Gentile. A New Approximate Maximal Margin Classification**

**Algorithm. NIPS. 2000.**

**To display First 5 samples**

**pixel\_0\_0 pixel\_0\_1 pixel\_0\_2 pixel\_0\_3 pixel\_0\_4 pixel\_0\_5 \**

**0 0.0 0.0 5.0 13.0 9.0 1.0**

**1 0.0 0.0 0.0 12.0 13.0 5.0**

**2 0.0 0.0 0.0 4.0 15.0 12.0**

**3 0.0 0.0 7.0 15.0 13.0 1.0**

**4 0.0 0.0 0.0 1.0 11.0 0.0**

**pixel\_0\_6 pixel\_0\_7 pixel\_1\_0 pixel\_1\_1 ... pixel\_7\_0 pixel\_7\_1 \**

**0 0.0 0.0 0.0 0.0 ... 0.0 0.0**

**1 0.0 0.0 0.0 0.0 ... 0.0 0.0**

**2 0.0 0.0 0.0 0.0 ... 0.0 0.0**

**3 0.0 0.0 0.0 8.0 ... 0.0 0.0**

**4 0.0 0.0 0.0 0.0 ... 0.0 0.0**

**pixel\_7\_2 pixel\_7\_3 pixel\_7\_4 pixel\_7\_5 pixel\_7\_6 pixel\_7\_7 target \**

**0 6.0 13.0 10.0 0.0 0.0 0.0 0**

**1 0.0 11.0 16.0 10.0 0.0 0.0 1**

**2 0.0 3.0 11.0 16.0 9.0 0.0 2**

**3 7.0 13.0 13.0 9.0 0.0 0.0 3**

**4 0.0 2.0 16.0 4.0 0.0 0.0 4**

**target\_names**

**0 0**

**1 1**

**2 2**

**3 3**

**4 4**

**[5 rows x 66 columns]**

**To display randomply 5 samples**

**pixel\_0\_0 pixel\_0\_1 pixel\_0\_2 pixel\_0\_3 pixel\_0\_4 pixel\_0\_5 \**

**527 0.0 0.0 6.0 13.0 0.0 0.0**

**95 0.0 0.0 0.0 11.0 16.0 8.0**

**162 0.0 5.0 16.0 16.0 16.0 11.0**

**1120 0.0 0.0 1.0 11.0 14.0 5.0**

**1295 0.0 0.0 4.0 15.0 13.0 3.0**

**pixel\_0\_6 pixel\_0\_7 pixel\_1\_0 pixel\_1\_1 ... pixel\_7\_0 pixel\_7\_1 \**

**527 0.0 0.0 0.0 0.0 ... 0.0 0.0**

**95 0.0 0.0 0.0 0.0 ... 0.0 0.0**

**162 1.0 0.0 0.0 4.0 ... 0.0 2.0**

**1120 0.0 0.0 0.0 0.0 ... 0.0 0.0**

**1295 0.0 0.0 0.0 4.0 ... 0.0 0.0**

**pixel\_7\_2 pixel\_7\_3 pixel\_7\_4 pixel\_7\_5 pixel\_7\_6 pixel\_7\_7 \**

**527 6.0 16.0 16.0 16.0 16.0 12.0**

**95 0.0 12.0 16.0 15.0 0.0 0.0**

**162 15.0 16.0 9.0 0.0 0.0 0.0**

**1120 2.0 13.0 16.0 9.0 0.0 0.0**

**1295 5.0 15.0 16.0 5.0 0.0 0.0**

**target target\_names**

**527 1 1**

**95 6 6**

**162 5 5**

**1120 1 1**

**1295 8 8**

**[5 rows x 66 columns]**

**The total number of samples in the dataset = 1797**

**The number of samples in training set = 1257**

**The number of samples in testing set = 540**

**The first five samples of training set**

**pixel\_0\_0 pixel\_0\_1 pixel\_0\_2 pixel\_0\_3 pixel\_0\_4 pixel\_0\_5 \**

**4 0.0 0.0 0.0 1.0 11.0 0.0**

**53 0.0 0.0 4.0 8.0 16.0 5.0**

**1187 0.0 0.0 9.0 14.0 15.0 6.0**

**1686 0.0 0.0 8.0 14.0 12.0 3.0**

**1727 0.0 0.0 6.0 11.0 16.0 13.0**

**pixel\_0\_6 pixel\_0\_7 pixel\_1\_0 pixel\_1\_1 ... pixel\_7\_0 pixel\_7\_1 \**

**4 0.0 0.0 0.0 0.0 ... 0.0 0.0**

**53 0.0 0.0 0.0 0.0 ... 0.0 0.0**

**1187 0.0 0.0 0.0 2.0 ... 0.0 0.0**

**1686 0.0 0.0 0.0 6.0 ... 0.0 0.0**

**1727 5.0 0.0 0.0 2.0 ... 0.0 0.0**

**pixel\_7\_2 pixel\_7\_3 pixel\_7\_4 pixel\_7\_5 pixel\_7\_6 pixel\_7\_7 \**

**4 0.0 2.0 16.0 4.0 0.0 0.0**

**53 6.0 16.0 12.0 1.0 0.0 0.0**

**1187 6.0 14.0 5.0 0.0 0.0 0.0**

**1686 7.0 16.0 16.0 8.0 0.0 0.0**

**1727 5.0 14.0 11.0 6.0 0.0 0.0**

**target target\_names**

**4 4 4**

**53 8 8**

**1187 0 0**

**1686 9 9**

**1727 3 3**

**[5 rows x 66 columns]**

**The first five samples of testing set**

**pixel\_0\_0 pixel\_0\_1 pixel\_0\_2 pixel\_0\_3 pixel\_0\_4 pixel\_0\_5 \**

**415 0.0 0.0 3.0 14.0 10.0 3.0**

**740 0.0 0.0 0.0 7.0 14.0 16.0**

**67 0.0 0.0 5.0 14.0 0.0 0.0**

**768 0.0 0.0 4.0 12.0 16.0 8.0**

**463 0.0 0.0 13.0 14.0 3.0 0.0**

**pixel\_0\_6 pixel\_0\_7 pixel\_1\_0 pixel\_1\_1 ... pixel\_7\_0 pixel\_7\_1 \**

**415 0.0 0.0 0.0 0.0 ... 0.0 0.0**

**740 6.0 0.0 0.0 0.0 ... 0.0 0.0**

**67 0.0 0.0 0.0 0.0 ... 0.0 0.0**

**768 0.0 0.0 0.0 5.0 ... 0.0 0.0**

**463 0.0 0.0 0.0 4.0 ... 0.0 0.0**

**pixel\_7\_2 pixel\_7\_3 pixel\_7\_4 pixel\_7\_5 pixel\_7\_6 pixel\_7\_7 target \**

**415 7.0 12.0 14.0 14.0 6.0 0.0 9**

**740 0.0 9.0 6.0 0.0 0.0 0.0 7**

**67 4.0 14.0 16.0 12.0 7.0 0.0 6**

**768 3.0 13.0 13.0 10.0 1.0 0.0 8**

**463 11.0 12.0 14.0 14.0 6.0 0.0 2**

**target\_names**

**415 9**

**740 7**

**67 6**

**768 8**

**463 2**

**[5 rows x 66 columns]**

**Assignment**

1. **Show that in two different executions same samples will not be present training and testing set.**
2. **Give reason why target name is not appearing as One , two, three**

**1.c to access and observe the samples from diabetis dataset.**

**from sklearn import datasets**

**import pandas as pd**

**print('THIS IS  THE PROGRAM TO ACCESS  DIABETES DATASET')**

**dbt=datasets.load\_diabetes()**

**print(" To print the description of Digits Dataset")**

**print(dbt.DESCR)**

**print('\n\n\n\n')**

**# df will fold dataset as a table**

**df=pd.DataFrame(**

**dbt.data,**

**columns=dbt.feature\_names**

**)**

**#labels are assigned to df[target] table or array**

**df['target']=pd.Series(**

**dbt.target**

**)**

**print('To display First 5 samples')**

**# df.head(5) will return the first five samples in the dataset**

**print(df.head(5))**

**print('To display randomply 5 samples')**

**#df.sample(5) will return randomly five samples from the dataset**

**print(df.sample(5))**

**# Train Test Split Ratio**

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

**df\_train,df\_test=train\_test\_split(df,test\_size=0.3) # For 70: 30 Split**

**print('The total number of samples in the dataset = ',df.shape[0])**

**print('The number of samples  in training set = ',df\_train.shape[0])**

**print('The number of samples  in testing set = ',df\_test.shape[0])**

**print('The first five samples of training set')**

**print(df\_train.head(5))**

**print('\n\nThe first five samples of testing set')**

**print(df\_test.head(5))**

**OUTPUT**

**THIS IS THE PROGRAM TO ACCESS DIABETES DATASET**

**To print the description of Digits Dataset**

**.. \_diabetes\_dataset:**

**Diabetes dataset**

**----------------**

**Ten baseline variables, age, sex, body mass index, average blood**

**pressure, and six blood serum measurements were obtained for each of n =**

**442 diabetes patients, as well as the response of interest, a**

**quantitative measure of disease progression one year after baseline.**

**\*\*Data Set Characteristics:\*\***

**:Number of Instances: 442**

**:Number of Attributes: First 10 columns are numeric predictive values**

**:Target: Column 11 is a quantitative measure of disease progression one year after baseline**

**:Attribute Information:**

**- age age in years**

**- sex**

**- bmi body mass index**

**- bp average blood pressure**

**- s1 tc, total serum cholesterol**

**- s2 ldl, low-density lipoproteins**

**- s3 hdl, high-density lipoproteins**

**- s4 tch, total cholesterol / HDL**

**- s5 ltg, possibly log of serum triglycerides level**

**- s6 glu, blood sugar level**

**Note: Each of these 10 feature variables have been mean centered and scaled by the standard deviation times the square root of `n\_samples` (i.e. the sum of squares of each column totals 1).**

**Source URL:**

[**https://www4.stat.ncsu.edu/~boos/var.select/diabetes.html**](https://www4.stat.ncsu.edu/~boos/var.select/diabetes.html)

**For more information see:**

**Bradley Efron, Trevor Hastie, Iain Johnstone and Robert Tibshirani (2004) "Least Angle Regression," Annals of Statistics (with discussion), 407-499.**

**(**[**https://web.stanford.edu/~hastie/Papers/LARS/LeastAngle\_2002.pdf**](https://web.stanford.edu/~hastie/Papers/LARS/LeastAngle_2002.pdf)**)**

**To display First 5 samples**

**age sex bmi bp s1 s2 s3 \**

**0 0.038076 0.050680 0.061696 0.021872 -0.044223 -0.034821 -0.043401**

**1 -0.001882 -0.044642 -0.051474 -0.026328 -0.008449 -0.019163 0.074412**

**2 0.085299 0.050680 0.044451 -0.005670 -0.045599 -0.034194 -0.032356**

**3 -0.089063 -0.044642 -0.011595 -0.036656 0.012191 0.024991 -0.036038**

**4 0.005383 -0.044642 -0.036385 0.021872 0.003935 0.015596 0.008142**

**s4 s5 s6 target**

**0 -0.002592 0.019907 -0.017646 151.0**

**1 -0.039493 -0.068332 -0.092204 75.0**

**2 -0.002592 0.002861 -0.025930 141.0**

**3 0.034309 0.022688 -0.009362 206.0**

**4 -0.002592 -0.031988 -0.046641 135.0**

**To display randomply 5 samples**

**age sex bmi bp s1 s2 s3 \**

**261 0.048974 -0.044642 -0.041774 0.104501 0.035582 -0.025739 0.177497**

**263 -0.074533 0.050680 -0.077342 -0.046985 -0.046975 -0.032629 0.004460**

**337 0.019913 0.050680 -0.012673 0.070072 -0.011201 0.007141 -0.039719**

**140 0.041708 0.050680 0.014272 0.042529 -0.030464 -0.001314 -0.043401**

**363 -0.049105 0.050680 -0.024529 0.000079 -0.046975 -0.028245 -0.065491**

**s4 s5 s6 target**

**261 -0.076395 -0.012909 0.015491 103.0**

**263 -0.039493 -0.072133 -0.017646 116.0**

**337 0.034309 0.005386 0.003064 91.0**

**140 -0.002592 -0.033246 0.015491 118.0**

**363 0.028405 0.019196 0.011349 58.0**

**The total number of samples in the dataset = 442**

**The number of samples in training set = 309**

**The number of samples in testing set = 133**

**The first five samples of training set**

**age sex bmi bp s1 s2 s3 \**

**168 0.001751 0.050680 0.059541 -0.002228 0.061725 0.063195 -0.058127**

**295 -0.052738 0.050680 0.039062 -0.040099 -0.005697 -0.012900 0.011824**

**100 0.016281 -0.044642 0.017506 -0.022885 0.060349 0.044406 0.030232**

**269 0.009016 -0.044642 -0.032073 -0.026328 0.042462 -0.010395 0.159089**

**316 0.016281 0.050680 0.014272 0.001215 0.001183 -0.021355 -0.032356**

**s4 s5 s6 target**

**168 0.108111 0.068986 0.127328 268.0**

**295 -0.039493 0.016307 0.003064 85.0**

**100 -0.002592 0.037236 -0.001078 128.0**

**269 -0.076395 -0.011897 -0.038357 87.0**

**316 0.034309 0.074966 0.040343 220.0**

**The first five samples of testing set**

**age sex bmi bp s1 s2 s3 \**

**435 -0.012780 -0.044642 -0.023451 -0.040099 -0.016704 0.004636 -0.017629**

**271 0.038076 0.050680 0.008883 0.042529 -0.042848 -0.021042 -0.039719**

**62 -0.027310 0.050680 -0.007284 -0.040099 -0.011201 -0.013840 0.059685**

**74 0.012648 0.050680 0.002417 0.056301 0.027326 0.017162 0.041277**

**413 -0.052738 -0.044642 -0.000817 -0.026328 0.010815 0.007141 0.048640**

**s4 s5 s6 target**

**435 -0.002592 -0.038460 -0.038357 64.0**

**271 -0.002592 -0.018114 0.007207 127.0**

**62 -0.039493 -0.082379 -0.025930 52.0**

**74 -0.039493 0.003709 0.073480 85.0**

**413 -0.039493 -0.035816 0.019633 113.0**

**Assignment**

1. **Show that in two different executions same samples will not be present training and testing set.**
2. **Give a reason , why the target name is not present?**