ASSIGNMENT-5

AKSHAY SRIRANGAM (AXS156630)

VISHNU KUMAR REDDY CHINAREDDY (VXC151130)

ANUPAMA VIJAYAN (AXV146030)

DATASET-1

Banknote Authentication:

* K-fold validation technique is used for selecting the best model that fits the given dataset.
* The accuracy values after using the K-fold validation technique are given below:

1. Lm fit : 98.9%
2. Glm fit : 99.2%
3. Lda fit : 97.6%
4. Qda fit : 97.6%
5. K-nn fit : 98.4%

* From the above accuracy value, Glm fit is a better fit compared to others.
* Since, the accuracy values are good for linear models, we did not use any cross terms or higher degree polynomial terms.
* 5-fold validation is used for initial splitting of data into training and test.

DATASET-2

Yacht Hydrodynamics:

* K-fold validation technique is used for selecting the best model that fits the given dataset.
* The accuracy values after using the K-fold validation technique are given below:

1. Lm fit : 81.945 (MSE)

2. Glm fit : 99.02%

3. Lda fit : 76.32 (MSE)

4. Qda fit : 98.3%

5. K-nn fit : 100%

* From the above accuracy values, K-NN fit is a better fit compared to others.
* Since, the accuracy values are good for linear models, we did not use any cross terms or higher degree polynomial terms.
* 5-fold validation is used for initial splitting of data into training and test.
* MSE is found for lm and lda models as it provides much more precise information about the accuracy.

DATASET-3

Concrete Strength:

* K-fold validation technique is used for selecting the best model that fits the given dataset.
* The accuracy values after using the K-fold validation technique are given below:

1. Lm fit : 109.434 (MSE)

2. Glm fit : 109.43 (MSE)

3. Lda fit : 7165.604 (MSE)

4. Qda fit : 93.3%

5. K-nn fit : 96.7%

* From the above accuracy values, K-NN fit is a better fit compared to others.
* Since, the accuracy values are good for linear models, we did not use any cross terms or higher degree polynomial terms.
* 5-fold validation is used for initial splitting of data into training and test.
* MSE is found for lm, glm and lda models as it provides much more precise information about the accuracy.