

CYCLE - II

- 1.) Build a linear regression model on the given power plant dataset. Create the training and testing set. Make predictions of the datapoints in the test set. Evaluate the model using appropriate performance matrix.
- 2.) Given the following data, which specify classifications for nine combinations of VAR1 and VAR2 predict a classification for a case where VAR1=0.906 and VAR2=0.606, using the result of k- means clustering with 3 means (i.e., 3 centroids)

VAR1	VAR2	CLASS
1.713	1.586	0
0.180	1.786	1
0.353	1.240	1
0.940	1.566	0
1.486	0.759	1
1.266	1.106	0
1.540	0.419	1
0.459	1.799	1
0.773	0.186	1

- 3.) Find the classification accuracy of K-means algorithm with MNIST dataset.
- 4.) Build a model to perform classification based on logistic regression on the given Breast cancer classification dataset. Display confusion matrix and accuracy for the test data.

- 5.) Construct a Bayesian classifier using Titanic survival prediction dataset.
Calculate the accuracy, precision, and recall for the data set.

CYCLE - III

1. Download any binary class numerical dataset from UCI repository and do the classification with
 - a. SVM
 - b. Naive Bayes classification
 - c. Find out the precision, recall and F1 score of the above classification methods with 20% of the dataset
2. Download any multi class numerical dataset from UCI repository and do the classification with
 - a. SVM with majority voting
3. Download any numerical dataset with more than 1000 features. Reduce the feature dimension with the help of PCA.
Do the following experiments
 - a. Reduce the feature dimension to 300, 400, 500
 - b. Perform the machine learning with svm for the different dimensions mentioned in a.
 - c. Find precision recall and F1 score of all the experiments said in b.
 - d. Prepare a comparison table of c.
4. Download any numerical dataset from UCI repository and do the classification process with SVM as said below
 - a. 10 fold validation
 - b. with dataset 80:20, 70:30 and 60:40
 - c. prepare a comparison table of F1 score for a and b