

Assignment 1

22AIE213 Machine Learning

Last Date: March 23rd 2025 @ 11.55 PM

1.	Go through the following Tutorials on different plots available for data visualisation, https://www.kaggle.com/learn/data-visualization/lesson/1 https://www.geeksforgeeks.org/python-data-visualization-tutorial/ and generate similar applicable ports (including Venn diagrams, say to understand the ratio of samples belonging to different classes) for understanding the datasets (both features and samples) selected for your mini project.
2.	a. For the project that you have undertaken, identify a dimensionality reduction problem and report the error in reconstruction (with one and two principal components). b. Also, try LDA for the dimensionality reduction.
3.	a. Generate a toy Dataset: $(x, y) : (-1,-1), (-2,-2), (-2,-1), (-1,2), (1,2), (2,1), (1,1), (2,2)$. Fit the best regression line and compute the MSE on the training set. Also compute the R^2 matrix for the fitted line and the adjusted R^2 metric by considering 1 st and then both features together.
4.	What do you mean by training a Bayes classifier? Assume that first four samples belongs to category A and last 4 samples belongs to category B. If the features follow normal distribution, find and plot the decision boundary along with the data points? a. Compute precision, recall, F measure and accuracy (with the data used for training) b. Why is the Bayes classifier considered a generative classifier?
5.	If your project is a regression problem, perform the following regression and report results. a. Linear/polynomial regression without and with regularisation (LASSO & RIDGE) b. and report results. Use suitable charts for model and result visualisation. Else If your project is a classification problem, a. perform logistic regression, and Bayes classifier and report results b. confusion matrix, precision, recall, F Measure and AUC/ROC