

i) Use the California housing dataset (either fetch it using sklearn or if you are comfortable download it from another source and use it), apply linear regression the data using sklearn function.

Also apply ridge and lasso regression on the same dataset for different values of  $\lambda$ .

Plot a graph of cost vs  $\lambda$  for both ridge & lasso regression.

For all 3 models, you must divide data into 80% training and 20% test data. Your final results must be on test data. [10 points]

ii) Load the wine dataset.

It has 3 classes and a total of 178 datapoints. Divide it into 70% training data and 30% test data. Apply the following classifiers

i) Logistic Regression

ii) SVM with RBF kernel

iii) Decision tree

iv) Random forests

v) 2 different feed forward neural network architectures of your design

The models are to be trained on the training data with appropriate choice of parameters.

Report the accuracy, F1 score and the confusion matrix on the test data for each of the models [20 points]

Notes: Use Google Colab for coding.

Refer to appropriate documentation, codes we used in class to learn about loading datasets and different models to be used. You can find good examples in the sklearn documentation. Feel free to use generative AI also to 'polish' your code or get a first draft of your code.