

MACHINE LEARNING PRACTICE ASSIGNMENT

Scenario: Breast cancer is one of the mostly occurred cancers all over the world. It is important to detect and cure it. If this cancer is identified in time, then it could be prevented by treatment. In order to do so, machine learning can help to make prediction to know whether patient has Breast Cancer or not.

Problem statement: Build a predictive model to classify whether patient has cancer or not.

Sample question:

Consider the dataset given.

1. Check any missing values are there or not. If there are any, take necessary action.
2. Find correlation between input data using pearson correlation coefficient.
3. Build classifier using KNN to predict breast cancer. Send normalized data to the model
4. a. Build linear and logistic regression and find out in this case which regressor is better.
b. Find r^2 score.
c. Find MSE
5. Build a tree-based model to build a classifier. Draw confusion matrix
6. Perform ensemble-based model ie Random forest classifier and check whether there is any improvement in the accuracy of the model. Show precision, recall.
7. Draw necessary plots to visualize how data is fitted in the model.
8. Build any other classifier to get a model with better performance.