

PROJECT

Title:

Construct base-line model for training and evaluating models using ML Tools.

Tools, Equipment and Materials:

- 1 Windows PC with Python editor, Python compiler, MSOffice installed, Web Browser.

Instructions:

You have to construct and test a ML (Machine Learning) model.

- 1 Browse to kaggle.com website to search for 'binary categorical data' where the target is a binary value.
<https://www.kaggle.com/datasets?search=Binary+Categorical+Data>
- 2 Choose and download one of the listed datasets from the search results to predict the binary target.
- 3 Write a script to cleanse the downloaded datasets for the following from at least one of the columns:
 - a. Invalid data
 - b. Null values
 - c. Unknown values
- 4 Choose and convert one of the categorical column data to numeric values.
- 5 Create datasets from csv files using the required features and target.
- 6 Choose between supervised and unsupervised method for machine learning for the downloaded dataset.
- 7 Choose an appropriate machine learning algorithm to predict the binary target.
- 8 Build the model using the dataset and specify the appropriate parameters for the algorithm chosen in Step 7.
- 9 Select training data and test data from the dataset.

PROJECT

- 10 Train the machine learning model using the selected algorithm and dataset.
- 11 Use the test data selected in Step 9 to evaluate the model.
- 12 Demonstrate the results using the following parameters.
 - a. Accuracy of prediction
 - b. False Positives
 - c. False Negatives
 - d. True Positives
 - e. True Negatives

- END -