

Machine Learning: Programming Assignment 3: Adaboost

Problem Statement:

Write a program to implement the Adaboost algorithm with decision tree as the base classifier. The decision tree implemented in Assignment 1 may be called as a function. Run Adaboost for 3 rounds. The combined classifier should be tested on test instances and the accuracy of prediction for the test instances should be printed as output. A single program should train the classifier on the training set as well as test it on the test set.

Data Set Description:

The data set is about the passengers who survived the Titanic disaster. There are three input features: Passenger class (pclass), age, and gender. Each of the features are symbolic. Only the distinct values present in the training data comprise the domain of values of the feature. There is one output to be predicted: survived (yes/no).

Training Data Filename: *data3_19.csv*, Test Data Filename: *test3_19.csv*

Submission Guidelines:

You may use one of the following languages: c/C++/Java/Python. You should name your file as `<rollnumber_3.extension>` (e.g., `15CS10001_3.c`). Your program should be standalone and should not use any *special purpose* library. Numpy/pandas may be used. You should submit the program file only and not the output/input file. The submitted single program file *should* have the following header comments:

```
# Roll          # Name          # Assignment number          # Specific compilation/execution flags (if required)
```

*Please submit the program in moodle by **October 30, 2019 midnight** (hard deadline). Copying from friends/web will lead to strict penalties.*