

CSE 572 Data Mining Project 3

Due Date: April 12th, 2018

User dependent analysis

Consider the new set of features that you obtained by multiplying the PCA output with your feature set.

Divide that new feature set into two parts for each user: a) part 1: training and b) part 2: test. Ideally keep 60% of the data for each user as training and the rest of 40% as test data. Use three types of machines: a) decision trees (fitctree in Matlab), b) support vector machines (fitcsvm in Matlab), and c) neural networks (use the neural network toolbox in Matlab).

Train each machine with the training data and then use the test data to report accuracy. Use the accuracy metrics of Precision, Recall, F1 score. Report each metric for every user.