Data Mining Project 2

Alay Parikh (1001774636)

Vianny D'souza (1001770237)

Report For NN

Q-1 Describe the Nearest Neighbors method.

NN was one of the first Algorithm used to solve difficult algorithms such as Travelling sales man problem. It classifies a sample on the bases of the category of its Nearest neighbor. It is based on the classifier which uses some or all the patterns available in the training set to classify a test pattern.

Q-2 Explain what was your criteria for selecting the three attributes.

Top attributes can be chosen in multiple ways

One being correlation matrix where in the attributes that correlate most with survived are chosen.

Similarly relation between attributes can be determined by other available methods aswell.

Q-3 Visualizations of the classifier in a 2D projection, and write your observations.

Highest values as per our observations are for the attributes: sibsp and parch.

Q-4 Interpret and compare the results.

We have used 4 models: 1 for Euclidian distance and # different numbers for neighbors

2. First we used 3 neighbors and the accuracy came out to be 0.73

3. Second we used 5 neighbors and the accuracy came out to be 0.7564

4. Third we used 7 neighbors and the accuracy came out to be 0.7478