

Chaz Hurd

#### Rapid Miner Process of Lab4

##### Lab 4 Process:

For the first line of operators I used was on the training data. I used the 'Select Attributes' operator to remove function from the data set. After that I 'Set Role' on Localization to label which is made for learning operators. I then applied k-Nearest Neighbor operator which is used for classification which is what I need (I weighted the vote as well). I then applied the model on the test set. With the model ready, I removed duplicated values as the GeneID was duplicated many times. Using the key I joined the model using inner join. Now the key and model is joined, the next operator was 'Set Role' with the attribute name Localization for the next operator to use. Lastly I chose 'Performance(Categorization)' to have a statistical evaluation of the training data model.