ZUHA AHMAD

DS 670: Capstone: Big Data & Business Analytics

Dr. Jaume

11 May 2017

# LAB 9 – Apply Aggregation & Group Operations

The reference that I am trying to outperform is from an IEEE Journal called '*Crime forecasting using data mining techniques*'. The goal of my competitor's article was to explore a methodology for reliably predicting the location, time, and/or likelihood of future residential burglary in a year's time. They had several classifiers/algorithms to choose such as SVM, Decision Trees, Neural Networks, Naïve Bayes, and 1NN. We both had wanted to tackle crime forecasting in the United States. However, our focus is different. We are focusing on one algorithm, K means-clustering and decided to build an interactive application in R Shiny from R. The deployment and implementation of visual application for crime forecasting is beneficial and an improvement from our competitors. In the aggregation process, the target would be to show distinction of quality and geographical state of improvement that can be implemented in Portland, Oregon.

| **COMPETITOR** | **MY CONTRIBUTION** |
|---|---|
| ▪ Our datasets are from the United States National Institute of Justice | |
| The research is made up of the following consecutive sections:<br>**'Abstract',**<br>**'Introduction',**<br>**'Data Generation'** (describes the data set);<br>**'Approach Architecture'** (details the feature construction and data manipulation);<br>**'Experimental Results'** (explore our analysis);<br>**'Conclusion'** (reviews our research findings);<br>**'Deployment'** (discusses the motivation for our research and its intended use); | The research is made up of the following sections:<br>**'Abstract',**<br>**'Work by Competitors',**<br>**'Contribution',**<br>**'Data',**<br>**'Method',**<br>**'Results',**<br>**'Discussions',**<br>**'Conclusion'** |
| Classifiers:<br>SVM<br>Decision Trees<br>Naïve Bayes | Classifiers<br>K-Means Clustering |