**ITCS 6156: Project Proposal**

Team members:

1. Ashwin Venkatesh Prabhu
2. Febin Zachariah
3. Varun Krishnan
4. Yateen Kedare

Dataset to be used: Dataset from Murder Accountability Project

Link to the dataset: <https://www.kaggle.com/murderaccountability/homicide-reports>

<http://www.murderdata.org/p/data-docs.html>

The Murder Accountability Project is a nonprofit group organized in 2015 and dedicated to educate Americans on the importance of accurately accounting for unsolved homicides within the United States. Information is obtained from federal, state and local governments about unsolved homicides and to publish this information. The Murder Accountability Project has the most complete database of homicides in the United States currently available. This dataset has a data of over 638,454 recorded crimes over a period of 35 years (1980 to 2014) in the United States. This dataset includes attributes like age, race, sex, ethnicity of victims and perpetrators, in addition to the relationship between the victim and perpetrator and weapon used.

Our primary objective here is to predict the perpetrators sex and age based on the features available. We will also try to discover crime patterns over the period of time from this dataset and understand the correlation among different attributes of victims and perpetrators.

We plan on trying out the following algorithms for the purpose of predictions:

1. Artificial Neural Networks.
2. Support Vector Machines.
3. Gradient Boosting.

Proposed plan:

Ashwin, Varun: Data cleaning - It is important to remove outliers, and impute NA values before creating a model for regression

Varun, Ashwin: Correlation Analysis - Identify highly correlated features for feature selection

Febin, Yateen: Pattern extraction - To identify few patterns among the data (like success rate on solving crimes in counties. Identify increase/decrease in different types of crime based on region)

Everyone: Predictive analysis - Use ANN, SVM and GBM to predict age and gender

Related Papers:

1. <http://www.psychology.nottingham.ac.uk/staff/ddc/c8cxpa/further/Dissertation_examples/Buttrick_09.pdf>
2. <http://repository.cmu.edu/cgi/viewcontent.cgi?article=1137&context=hsshonors>
3. <https://www.ncjrs.gov/pdffiles1/nij/249895.pdf>