

ANALYZING INDIAN CRIME DATA

Abstract

Problem Statement

With the chosen dataset objective is to build a predictive model which can predict the total number of specific crimes and total number of total IPC crimes for a given year, state, district/area or any combination of these three provided as input. Other than this, time series analysis and possible set of crimes committed frequently in India on the basis of given data. And to group different districts(treated as instances within a year).

Our project as a solution

Analyzing the Indian crime data set to produce some useful inferences: 01_District_wise_crimes_committed_IPC_2001_2012.csv (available at data.gov.in). In this project the above mentioned data set will be used and other transformed and processed forms of the dataset to produce useful data visualization and also have applied different machine learning algorithms namely, Linear regression, Multiple linear regression, Polynomial regression, support vector regression, decision tree regression and random forest regression to build some useful regression models. Also implementation of clustering algorithms k means and hierarchical clustering with the given data set will be done. Time series analysis and association rule mining algorithms, namely, eclat and apriori algorithms on the transformed form of the data set will also be performed.