

Data Analytics

Project Presentation Analysis of crime against women

Name : Vijit Kumar

SRN : PES1UG20CS500

Name : Basava Prabhu

SRN : PES1UG20CS631

Name : Rhea Sudheer

SRN : PES1UG20CS680

Abstract and Scope


The purpose of this project is to analyse and predict different types of crimes against women in India. The prediction is done by following a systematic approach for identifying and analysing patterns and trends in the crime data set used.


A comparative analysis of crimes district-wise and state-wise and forecast the crime rate for each district for the future years using Random Forest.

Implementation Details

- Modules used: sklearn, pandas, seaborn, matplotlib, plotly, numpy, label encoding
- We label encoded state and district and used the user input to predict the crime rates in their state using random forest(regression), with brute force approach to estimate hyperparameters.

Project Demonstration

 Rhea Sudheer PES1UG20CS580

 Basavaprabhu PES1UG20CS631

Choose state:

KERALA

Choose district:

idukki

Predict

⚠ High rate of Rape Cases in this state

Kidnapping and abduction are less frequent here

Dowry Deaths are less frequent here

⚠ High rate of Assault on women Cases in this state

⚠ High rate of Insult to modesty of women Cases in this state

Test Plan and Strategy

We trained the various models on data from 2001 - 2012 and made predictions for the 13th year(2013), tested it against know values for 2013 compared results using normalised RMSE and R-square coefficient.

$$R^2 = 1 - \frac{RSS}{TSS}$$

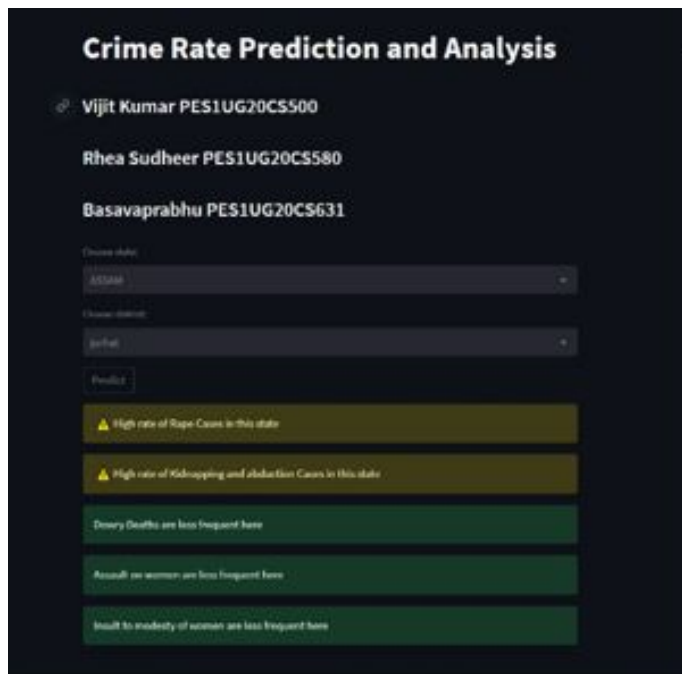
R^2 = coefficient of determination
 RSS = sum of squares of residuals
 TSS = total sum of squares

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (Predicted_i - Actual_i)^2}{N}}$$

$$NRMSE = \frac{RMSE}{y_{max} - y_{min}}$$

Results and Discussion

After trying various models for forecasting, random forests gave highest accuracy which we used to predict crime rates for future years based on the state and district name given by the user.

The screenshot shows a web application titled "Crime Rate Prediction and Analysis". It features a dark blue header with the title in white. Below the header, there are three user profile entries: "Vijit Kumar PES1UG20CS500", "Rhea Sudheer PES1UG20CS580", and "Basavaprabhu PES1UG20CS631". Each entry has a small circular icon to its left. Below the profiles, there are two dropdown menus: "Choose state:" with "ASSAM" selected, and "Choose district:" with "Jorhat" selected. A "Predict" button is located below the district dropdown. The results section consists of five horizontal bars: two yellow bars with warning icons indicating "High rate of Rape Cases in this state" and "High rate of Kidnapping and abduction Cases in this state", and three green bars indicating "Domestic Deaths are less frequent here", "Assault on women are less frequent here", and "Insult to modesty of women are less frequent here".

Crime Rate Prediction and Analysis

Vijit Kumar PES1UG20CS500

Rhea Sudheer PES1UG20CS580

Basavaprabhu PES1UG20CS631

Choose state:

ASSAM

Choose district:

Jorhat

Predict

⚠ High rate of Rape Cases in this state

⚠ High rate of Kidnapping and abduction Cases in this state

✅ Domestic Deaths are less frequent here

✅ Assault on women are less frequent here

✅ Insult to modesty of women are less frequent here

Conclusion and Future work

As you can see through the visualisation, crime in India is increasing at an exponential rate and it's pertinent that apt measures are taken to curb this growth rate. However, in the meanwhile, ensuring the safety of women is the primary concern.

Through our solution, we hope to make the citizens more aware of the places they travel to and hope to make a meaningful contribution at prevention of crime

Thank
You