

CRIME ANALYSIS USING K-MEANS CLUSTERING

*project report submitted in partial fulfillment of the requirements
for the award of the Degree of*

BACHELOR OF TECHNOLOGY

In

COMPUTER SCIENCE AND ENGINEERING

Submitted by;

G.HARI KEERTHANA

D.MOUNIKA PADMA PRIYA

Regd.NO.15811A0527

Regd.No.15811A0518

A.SRI LAKSHMI

B.SWAMI TRINADH

Regd.No.15811A0501

Redg.No.15811A0512

Under the guidance of

Mr. S.V.R.VARA PRASAD M TECH *Assistant professor*

Department of Computer Science and Engineering



AVANTHI INSTITUTE OF ENGINEERING & TECHNOLOGY

(Approved by AICTE, New Delhi & permanently affiliated to JNTU Kakinada)

(Accredited by NAAC, UGC & NBA, AICTE)

MAKAVARAPALEM, NARSIPATNAM,

VISAKHAPATNAM DIST

(2015-2019)

AVANTHI INSTITUTE OF ENGINEERING & TECHNOLOGY

(Approved by AICTE, New Delhi & permanently affiliated to JNTU Kakinada)

(Accredited by NAAC, UGC & NBA, AICTE)

MAKAVARAPALEM, NARSIPATNAM,

VISAKHAPATNAM-531113



CERTIFICATE

This is to certify that the project entitled "CRIME ANALYSIS USING K-MEANS CLUSTERING" impartial fulfillment for the of degree of Bachelor of Technology in COMPUTER SCIENCE AND ENGINEERING, at AVANTHI INSTITUTE OF ENGINEERIN AND TECHNOLOGY, MAKAVARAPALEM, VISAKHAPATNAM is an work carried out by G.HARI KEERTHANA (15811A0527), D.MOUNIKA PADMA PRIYA (15811A0518), A.SRI LAKSHMI (15811A501), B.SWAMI TRINADH(15811A0512) under the guidance and supervision during 2018-2019.

svrlapad
Project Guide

G. Satyanarayana
Head of the Department

[Signature]
External Examiner

ABSTRACT

Crimes have a negative effect on any society both socially and economically. Law enforcement bodies face numerous challenges while trying to prevent crimes. We propose a Crime Data Analytic Platform (CDAP) to assist law enforcement bodies to perform descriptive, predictive, and prescriptive analysis on crime data. CDAP has a modular architecture where each component is built separate from each other. CDAP also support enabling future feature expansions. The platform can ingest any crime dataset which has the required attributes to map dataset to attributes required by the platform. It can then analyze them, train models, and then visualize data. CDAP also combines census data with crime data to achieve more comprehensive crime analysis and their impact on society. Moreover, with the combination of census data and crime data, CDAP provides process reengineering steps to optimize resource allocations of police forces. We demonstrate the utility of the platform by visualizing capabilities of the platform are demonstrated by predicting crime categories spatial and temporal relationships in a set of real-world crime datasets for which a machine learning approach is used.

To construct a model Naive Bayesian, Random Forest Classifier, and Multi-layer Perception Network classification algorithms are provided. Identification of optimized police district boundaries and allocating patrol beats are used to demonstrate the prescriptive analytics capabilities of the tool. Heuristic-based clustering approach was taken to define police district boundaries in a way that the identified districts have equitable population distribution with compact shape. The resulting districts are then evaluated on inequality of population and the compactness using Gain Coefficient and Isoperimetric Quotient. Another heuristic-based approach was taken to define new police patrol beats to optimize on equitable workload distribution, compactness, and minimizing response time for new police patrol beats.