

**PROJECT REPORT OF INDUSTRY ORIENTED HANDS ON EXPERIENCE (IOHE)
ON**

Crime Rate Analysis & Prediction using Machine Learning
submitted in partial fulfilment of the requirements for the award of degree of

BACHELOR OF ENGINEERING

In

COMPUTER SCIENCE AND ENGINEERING

Submitted by:

Shriyanvi Tiwari

2110991333

Supervised By:

Dr. Gurpreet Singh



DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

**CHITKARA UNIVERSITY INSTITUTE OF ENGINEERING AND TECHNOLOGY
CHITKARA UNIVERSITY, PUNJAB, INDIA**

CONTENTS

Acknowledgement	4
Abstract	5
Introduction	5
Methodology	5
Tools and Technologies	5
Implementation	6
Result and Discussion	6
Conclusion	6
Future Scope	6
References	7

DECLARATION

I hereby certify that the work which is being presented in the project report entitled “**Crime Rate Analysis and Prediction using Machine Learning**” in partial fulfilment of requirement for the award of the degree of Bachelor of Engineering (Computer Science and Engineering) submitted in the department of Computer Science and Engineering at Chitkara University Institute of Engineering and Technology, Chitkara University, Punjab, India, is an authentic record of my own work carried out under the supervision of **Dr. Gurpreet Singh**. The matter presented in this project report has not been submitted in any other university/institute for the award of any degree.

Place: Rajpura

Date: 15-05-2025

Shriyanvi Tiwari

2110991333

This is to certify that the above statement made by the candidate is correct to the best of my knowledge and belief.

Dr. Gurpreet Singh

Associate Professor and Placement Incharge

Department of Computer Science and Engineering

Chitkara University Institute of Engineering and Technology,

Chitkara University, Punjab, India

ACKNOWLEDGEMENT

I express my sincere gratitude to my mentor Dr Gurpreet Singh for their continuous guidance and support. I also thank my institution, Chitkara University, for providing the resources necessary for the successful completion of this project.

Abstract:

This project analyses crime data from the National Crime Records Bureau (NCRB) for India, spanning the years 1981 to 2022. Using machine learning techniques, specifically linear regression, the project forecasts future crime incidence. Visualizations illustrate crime trends, and a simple web dashboard was developed to interactively display the analysis and predictions. This work aims to aid in understanding crime patterns and assist policymakers in resource planning.

Introduction:

Crime analysis is vital for understanding patterns and trends in criminal activity, helping law enforcement agencies and policymakers to take proactive measures. This project uses historical crime data from NCRB reports and applies predictive analytics to forecast future crime rates. The data was extracted, cleaned, analyzed, and modelled to provide insights into India's crime trends.

Objective:

- To analyze crime data from 1981 to 2022 in India.
- To visualize trends in crime incidence and rates.
- To build a predictive model for future crime forecasting.
- To develop an interactive dashboard for easy exploration of data and predictions.

Dataset Description:

The data was sourced from the National Crime Records Bureau (NCRB) India Crime Reports 2022. The raw data was extracted from PDF volumes and converted into CSV format for analysis. The dataset contains yearly crime incidence and crime rates at the national level.

Tools and Technologies:

- **Programming Language:** Python
- **Libraries:** pandas, matplotlib, seaborn, scikit-learn, streamlit
- **Data Source:** NCRB Crime Reports (PDF converted to CSV)
- **Environment:** Jupyter Notebook, Streamlit Web App

Methodology:

1. **Data Extraction:** Converted NCRB PDF reports into CSV files.
2. **Data Cleaning:** Removed irrelevant rows/columns, fixed column headers, and converted data types.
3. **Exploratory Data Analysis:** Visualized crime incidence and crime rates over the years.
4. **Predictive Modeling:** Trained a linear regression model on historical data to forecast future crimes.
5. **Dashboard Development:** Created a Streamlit- based web dashboard to visualize and predict crime trends interactively

Implementation:

- Data was loaded using pandas and cleaned for analysis.
- Visualizations such as line charts were generated to show crime trends.
- A linear regression model predicted crime incidence for 2023 and 2024.
- The Streamlit app allows users to input a year and get predicted crime numbers.

Results and Discussion:

- The total crime incidence in India showed a rising trend from 1981 to 2022.
- The crime rate per 100,000 population fluctuated but generally increased.
- The model achieved a good fit with an R^2 score above 0.9.
- Predictions suggest a continued rise in crime incidence for 2023 and 2024.

Conclusion:

This project demonstrated how machine learning and data visualization can provide meaningful insights into crime data. The predictive model and dashboard can assist authorities in better understanding and preparing for future crime trends.

Future Scope:

- Incorporate stat-wise and category-wise crime data for granular analysis.
- Used advanced models like LSTM for better time-series forecasting.
- Integrate real-time data sources for dynamic prediction and monitoring.

References:

- National Crime Records Bureau (NCRB) India Crime Reports 2022
- Python Official documentation
- Scikit-learn documentation
- Streamlit documentation