DEPARTMENT OF INFORMATION SCIENCE AND ENGINEERING

A MINI PROJECT REPORT ON

"IDENTIFICATION OF CRIME PRONE AREAS"

Submitted in the partial fulfilment of the requirements in the 3rd semester of

BACHELOR OF ENGINEERING

IN

INFORMATION SCIENCE AND ENGINEERING

FOR

COURSE NAME: MINI PROJECT

COURSE CODE: 20ISE391A

By

ROHAN NAGARKATTE – 1NH20IS134

KAMBOJI AKHILESH – 1NH20IS070

LOKESH PRASAD – 1NH20IS078

MOHAMED KHALID – 1NH20IS080

Under the guidance of

MRS. BILVIKA K M

Assistant Professor

NEW HORIZON
COLLEGE OF ENGINEERING
Autonomous College, Affiliated to VTU | Approved by AICTE New Delhi & UGC
Accredited by NAAC with 'A' Grade & Accredited by NBA

CERTIFICATE

Certified that the project work entitled "Identification of Crime Prone Areas" carried out by Mr. Rohan Nagarkatte bearing USN 1NH20IS134, Mr. Kamboji Akhilesh bearing USN 1NH20IS070, Mr. Lokesh Prasad bearing USN 1NH20IS078 and Mr. Mohamed Khalid bearing USN 1NH20IS080 are bona-fide students of III semester in partial fulfilment for the award of Bachelor of Engineering in Information Science & Engineering of New Horizon College of Engineering, an autonomous institute affiliated to the Visvesvaraya Technological University, Belagavi during the year 2021-22. It is certified that all corrections / suggestions indicated for Internal Assessment have been incorporated. The project report has been approved as it satisfies the academic requirements in respect of Mini Project work prescribed for the said Degree.

Name & Signature of Guide

Mrs. Bilvika K M

Name & Signature of HOD

Dr. Anandhi R J

Examiners:

	Name	**Signature**
1.	……	……..
2.	……	…..……

ABSTRACT

Crime analysis and prevention is a systematic approach for identifying and analysing patterns and trends in crime. Our system can predict the type of crime activity which have high probability for given location in terms of latitude and longitude and date and also, we can visualize crime prone areas. With the increasing introduction of automated systems, crime statistics analysts can help the Law enforcement officers to speed up the manner of fixing crimes. Using the idea of data mining we will extract previously unknown, useful information from an unstructured record. Here we have a method to develop an information mining method that can assist remedy crimes faster. Instead of focusing on reasons of crime incidence like criminal history of offender, political enmity etc we are focusing particularly on crime elements of every day.

ACKNOWLEDGEMENT

Any project is a task of great enormity and it cannot be accomplished by an individual without support and guidance. I am grateful to a number of individuals whose professional guidance and encouragement has made this project completion a reality.

We have a great pleasure in expressing our deep sense of gratitude to the beloved Chairman **Dr. Mohan Manghnani** for having provided us with a great infrastructure and well-furnished labs.

We take this opportunity to express our profound gratitude to the Principal **Dr.Manjunatha** for his constant support and management.

We are grateful to **Dr. R J Anandhi** , Professor and Head of Department of ISE, New Horizon College of Engineering, Bengaluru for her strong enforcement on perfection and quality during the course of our mini project work.

We would like to express our thanks to the guide **Mrs. Bilvika K M**, Assistant Professor, Department of ISE, New Horizon College of Engineering, Bengaluru who has always guided us in detailed technical aspects throughout our mini project.

We would like to mention special thanks to all the Teaching and Non-Teaching staff members of Information Science and Engineering Department, New Horizon College of Engineering, Bengaluru for their invaluable support and guidance.

ROHAN NAGARKATTE-1NH20IS134

KAMBOJI AKHILESH – 1NH20IS070

LOKESH PRASAD – 1NH20IS078

MOHAMED KHALID – 1NH20IS080

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1
INTRODUCTION

1.1 Introduction

As per past events and records the crime rate is increasing day by day and crimes does not occur in a systematic manner. It also does not occur randomly. The criminals are a nuisance to the society. The law and order of a place is questioned when these criminals commit crimes. The peace and tranquillity of a place is disturbed when these crimes happen. The modern advanced technologies also assist the criminals in achieving their misdeeds. According to Crime Records Bureau crimes like burglary have been decreased while crimes like murder have been increased. Even though the crime that is going to happen at a given place cannot be predicted but the probability of crime that might occur at a given place can be predicted. Though it cannot be assured of 100% accuracy, it might help the law enforcement officials across the country to constantly have a surveillance over these areas. So, a powerful analytical 24x7 monitoring system must be built that would predict the probability of a crime that might take place at a given place. For this, the crime records must be collected and evaluated.

1.2 Motivation for the project

Crimes occur at a place within a timeframe based on the vulnerability of the place. The places that are far away from the law departments are more susceptible to crimes. Such areas are to be monitored consciously and more frequently in order to maintain law and order in such places.

In order to predict the probability of occurrences of crimes in certain areas, a strong analytical tool that can process crime records is the need of the hour. To build such a tool is very challenging but also interesting. In order to achieve this objective, clusters of crimes around a given latitude and longitude needs to be found. The introduction of clustering analysis techniques provides with a new perspective on ways to conduct surveillance. For this to be done, clustering algorithm needs to be applied on a preprocessed data and the crimes need to be classified based on certain criteria. Later the output must be represented in a user-friendly pictorial representation. Hence the development of such a tool becomes the motivation for the project.

1.3 Problem Definition

To analyze historic data from emergency services and to build a solution for predictive policing that will be able to identify the crime prone areas.

1.4 Objectives and Scope of the project:

The main objective of this project is to solve the problems of the existing surveillance system that does not predict the areas which are more prone to crimes. This project is to help the law enforcement officials have an automated system to constantly identify areas that are more prone to crimes and policing-related incidents, than others. This would allow them to proactively review and plan suitable resource deployments and patrolling in those particular areas, and thereby prevent, as far as possible, untoward incidents that could result in an emergency.

- **Reliability:** The system is reliable as the user can constantly get the enhanced results based on the updated and modified information provided to the system at any point of time.
- **Automation:** The system almost does every process automatically without the need of much human intervention. Thus, the project helps in reducing time complexity by automation.
- **User-Friendly:** The system has a user-friendly interface. It represents the results in a form that is convenient to the user. It provides results in two forms that are appealing to the user. One, in the form of a tabular column and the other, in the form of a graph.
- **Extendibility:** The system can process for any amount of data inputted and still provide the accurate results without any interruption. Thus, the system's algorithm can be applied to any volume of data.
- **Accuracy:** The system provides accurate results as it works on K-means algorithm which shows output in the form of clusters and thus, it helps users to statistical analysis accurately based on the clusters.
- **Maintenance Cost**: It reduces the cost of maintenance as the algorithm need not to be constantly upgraded.

CHAPTER 2

LITERATURE SURVEY

DATA CRIME APPROACHES TO CRIMINAL CAREER ANALYSIS

Paper Summary:

Crime Analysis, is a study that is done on criminals, their behaviour and their past activities. The past technique of documenting crime activity and criminal history does not do any good in controlling crime and curbing criminal activities. Finding the patterns and trends in crime from huge volumes of written records is tedious and challenging. Also, current policing strategies work towards finding the criminals, basically after the crime has occurred and thus, it does not help in preventing the crimes.

Criminal records and narrative reports can be stored digitally across individual law enforcement departments and thus maintaining a collection of crime data to process a country wide database of criminals and their crime activities. This documentation of crime data over the years provides new opportunities and possibilities to compile and analyse criminal activities through time. Skimming out only the required data from huge volumes of crime data by data mining methods like clustering would enable the police forces to statistically analyse and predict the areas that are more susceptible to crimes. The huge volume of crime datasets and also the convolution of relationships between these types of data have made criminology a relevant field for applying data mining techniques. Data mining applied in the context of law enforcement and intelligence analysis holds the promise of alleviating such problem. The reason for opting data mining and clustering technique is the availability of raw data such date, time and place for which classification technique that rely on the existing and known solved crimes will not give good predictive quality for future crimes.

Advanced technologies can help to achieve this goal of isolating necessary data to work on it. Important crime attributes like date of crime event, type of crime, latitude and longitude, and many other factors play a significant role in the analysis of criminal activities. This method yields a visual clustering of these crimes and enables the identification of crime-prone areas. The proposed method allows for several user-denied parameters. Thus, this would enable the law enforcement department to proactively review and remain vigilant on the areas that are vulnerable to crimes.

CHAPTER 3

SYSTEM REQUIREMENTS SPECIFICATIONS

3.1 Hardware Requirements

The following are needed to efficiently use the system.

Processor	-	Intel Core i3 1st Gen
Speed	-	2.5 GHz
RAM	-	4 GB (min)
Hard Disk	-	50 GB

3.2 Software Requirements

Software requirements define software resource fundamentals that need to be installed on a workstation to provide optimum working of a software. The following are required for optimal development and usage of the system.

Operating System	-	Windows 7 and above
Programming Language	-	HTML5, CSS, Python 3.10
Compiler	-	Visual Studio Code or IDLE

CHAPTER 4

SYSTEM DESIGN

4.1 System Architecture for Identification of Crime Prone Areas

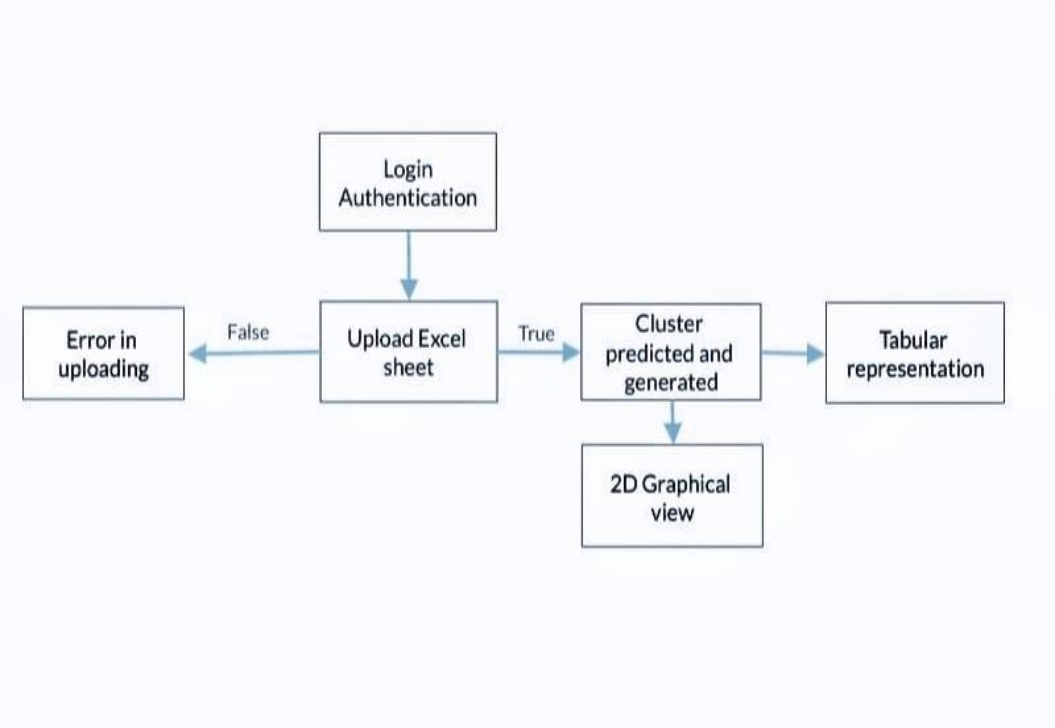


Figure 4.1: System Architecture

The user gets authenticated by entering the valid username and password. The excel sheet is then uploaded. Data from the excel sheet is processed in accordance with user's choice of input. The processed data points are plotted on a 2D graph categorizing them into clusters. The second viewing option includes tabular representation of processed data with necessary record details.

4.2 Flowchart of proposed system

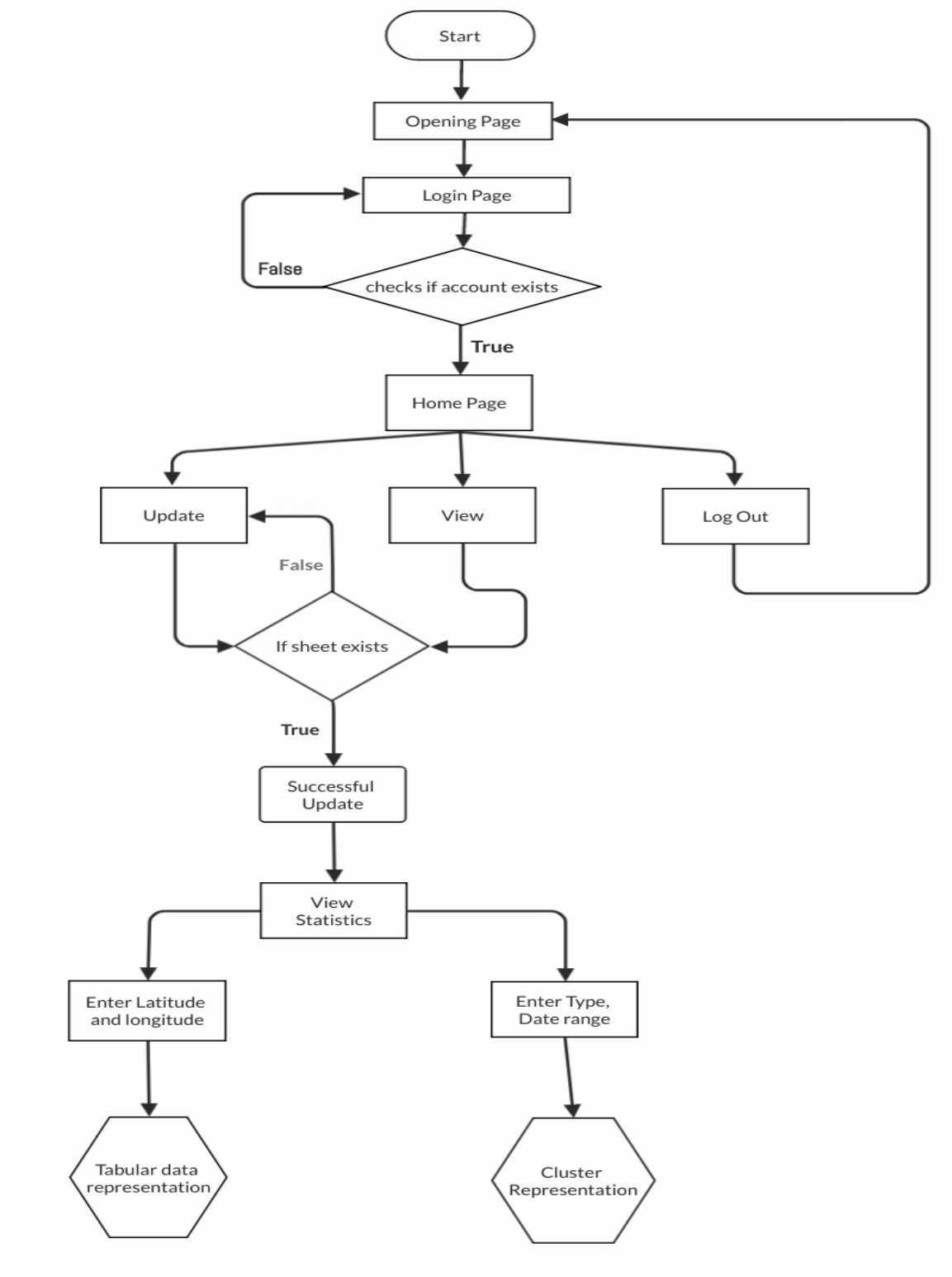


Fig 4.2: Flowchart

The above flowchart represents the systemic flow of the websites UI from login to the cluster/tabular representation**.**

CHAPTER 5

DESIGN MODULES

The project is divided into 3 modules namely “Login Module”, “Update Module” and “View Module”.

5.1 Login Module

The figure 5.1 shows how the Login module functions.

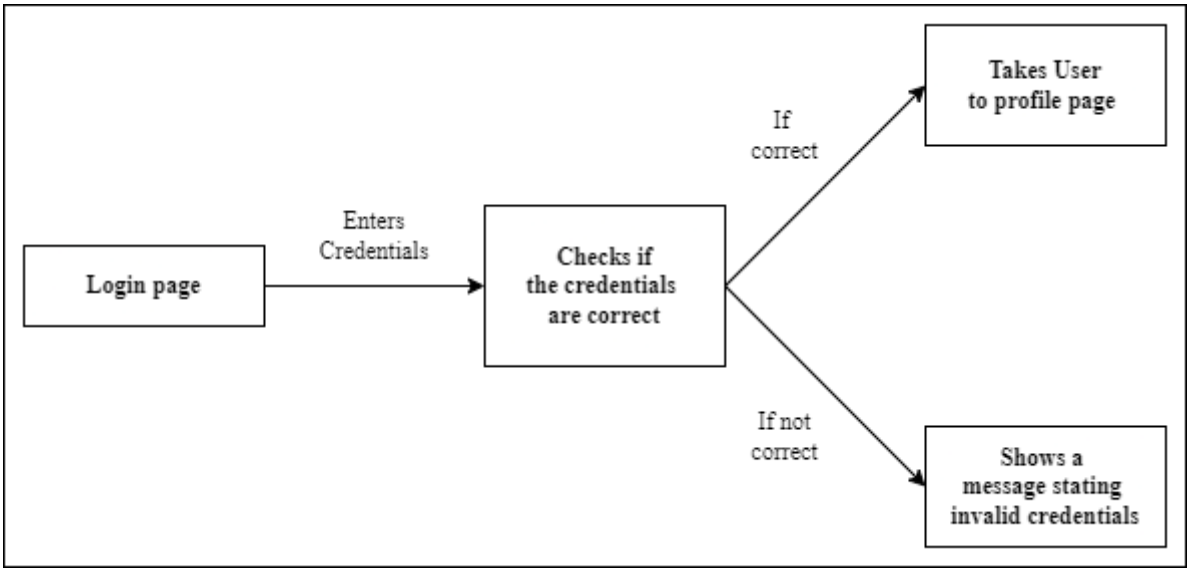


Figure 5.1: Login Module

1. In the login page the user is requested to enter 2 credentials which will be given to the user. The 2 required credentials are namely: “Username” and “Password”.
2. Once the required credentials are entered and the “LOGIN” button is hit, the program smartly checks the entered credentials.
3. If the credentials are matched, then the user is taken to a profile page with a welcome message and then the user can go to “Update” and “View” pages.
4. If the credentials do not match, an error message stating, “Please check your details and try again.” is shown to the user.

5.2 Update Module

The figure 5.2 shows how the Update module functions.

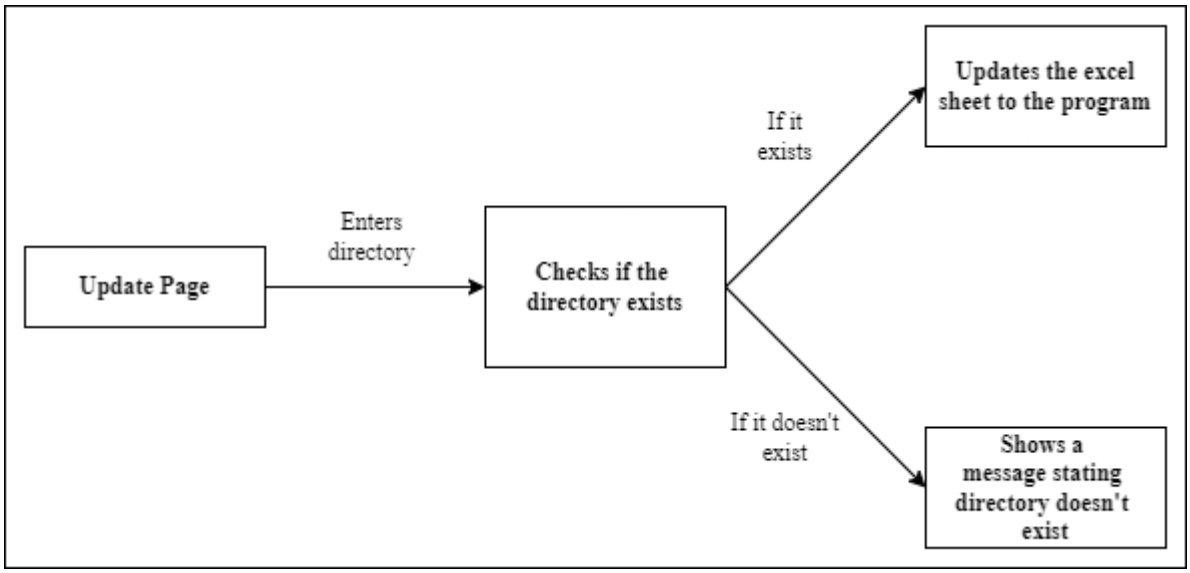


Figure 5.2: Update Module

1. In the update page the user is asked to enter the directory of the excel sheet the user wants to upload.
2. The computer again smartly checks if the entered directory exists and if it is an excel sheet.
3. If the sheet exists in the specified directory, then the excel sheet is uploaded to a local variable and now the data can be viewed in the way the user wants.
4. If the directory specified is wrong or if it doesn't exist, then an error message is shown to the user stating, "invalid file try again!!" within the Update page. So, the user can enter the directory again and try.

5.3 View Module

The figure 5.3 shows how the View module functions.

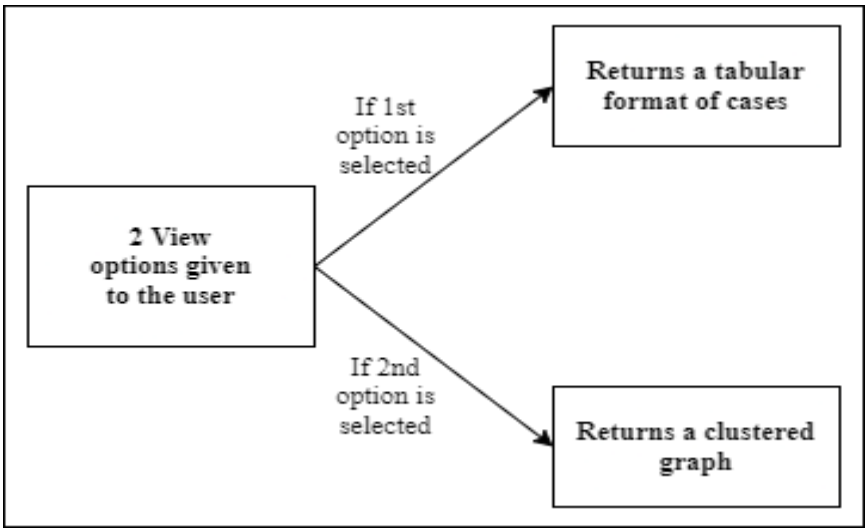


Figure 5.3: View Module

1. In the View page the user is given two options: “based on latitude and longitude” and “based on time limit and type of case”.
2. The first option shows all cases occurred at a particular point in a tabular format whose latitude and longitude are given by the user as input.
3. The second option shows all cases that occurred within a given time limit and belong to a particular type of crime (e.g., Murder, Dispute etc) which are given as inputs by the user, the output in this option is shown to the user in clustered format. The clustering is done based on the vector quantization of the cases and gives a colour for each to differentiate.
4. Once the required output is shown to the user the webpage comes back to the “profile page” from where the user is allowed to either update more crimes or view crimes in other formats the user wants.

CHAPTER 6

IMPLEMENTATION

6.1 Code Implementation

```
from flask import Flask, render_template, redirect, url_for, request, flash

import numpy as np
import pandas as pd
import statsmodels.api as sm
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.cluster import KMeans
from sklearn.preprocessing import MinMaxScaler
sns.set()

#import geopandas as gpd

main = Flask(__name__)

main.config['SECRET_KEY'] = 'miniproject'

@main.route('/')
def index():
    return render_template('index.html')

@main.route('/login')
def login():
    return render_template('login.html')

@main.route('/login', methods=['POST'])
def login_post():
    username = request.form.get('username')
    password = request.form.get('password')
    if username!='user' or password!='password':
        flash('Please check your login details and try again.')
        return redirect(url_for('login'))

    return redirect(url_for('profile'))

@main.route('/profile')
def profile():
    return render_template('profile.html')

@main.route('/update')
def update():
```

```
    return render_template('update.html')

@main.route('/update', methods=['POST'])
def update_post():
    global user, tdata
    user = request.form.get('username')
    try:
        tdata=pd.read_excel(user)
    except FileNotFoundError:
        flash("invalid file try again!!")
        return redirect(url_for('update'))
    else:
        return redirect(url_for('profile'))

@main.route('/successfull')
def successfull():
    return render_template('successfull.html')

@main.route('/successfull',methods=['POST'])
def successfull_post():
    return redirect(url_for('options'))

@main.route('/logout')
def logout():
    return redirect(url_for('index'))

@main.route('/options')
def options():
    return render_template('options.html')

@main.route('/lalo',methods=['POST'])
def lalo_post():
    global data,latitude,longitude
    la=request.form.get('latitude')
    lo=request.form.get('longitude')
    latitude=float(la)
    longitude=float(lo)
    try:
        data=tdata.loc[(tdata['Latitude'] == latitude) & (tdata['Longitude']
== longitude)]
    except NameError:
        flash("no excel sheet updated!!")
        return redirect(url_for("update"))
    else:
        return redirect(url_for("details"))

@main.route('/details')
def details():
```

```
    return
render_template('details.html',latitude=latitude,longitude=longitude,tables=[d
ata.to_html(classes='data')], titles=data.columns.values)

@main.route('/daty',methods=['POST'])
def daty_post():
    start_date = request.form.get('start_date')
    end_date = request.form.get('end_date')
    ctype=request.form.get('ctype')
    try:
        data=tdata.loc[(tdata['Event Type'] == ctype)&((tdata['Create
Date/Time'] > start_date) & (tdata['Create Date/Time'] <= end_date))]
    except NameError:
        flash("no excel sheet updated!!")
        return redirect(url_for("update"))
    else:
        latcol=data["Latitude"]
        longcol=data["Longitude"]
        maxLat=latcol.max()
        minLat=latcol.min()
        maxLong=longcol.max()
        minLong=longcol.min()
        plt.scatter(data['Longitude'],data['Latitude'])
        plt.ylim(minLat-1,maxLat+1)
        plt.xlim(minLong-1,maxLong+1)
        km=KMeans(n_clusters=3)
        y_predicted=km.fit_predict(data[['Longitude','Latitude']])
        data['cluster']=y_predicted
        data1=data[data.cluster==0]
        data2=data[data.cluster==1]
        data3=data[data.cluster==2]
        data4=data[data.cluster==3]
        plt.scatter(data1['Longitude'],data1['Latitude'],color='green')
        plt.scatter(data2['Longitude'],data2['Latitude'],color='blue')
        plt.scatter(data3['Longitude'],data3['Latitude'],color='red')
        plt.scatter(data4['Longitude'],data4['Latitude'],color='yellow')
        plt.xlabel('Longitude')
        plt.ylabel('Latitude')
        plt.legend(["red", "green","blue","yellow"], loc ="lower right")
        plt.show()
        return redirect(url_for('profile'))

if __name__ == '__main__':

      main.run()
```

CHAPTER 7

EXPERIMENTAL RESULTS

7.1 Outcome of Proposed System

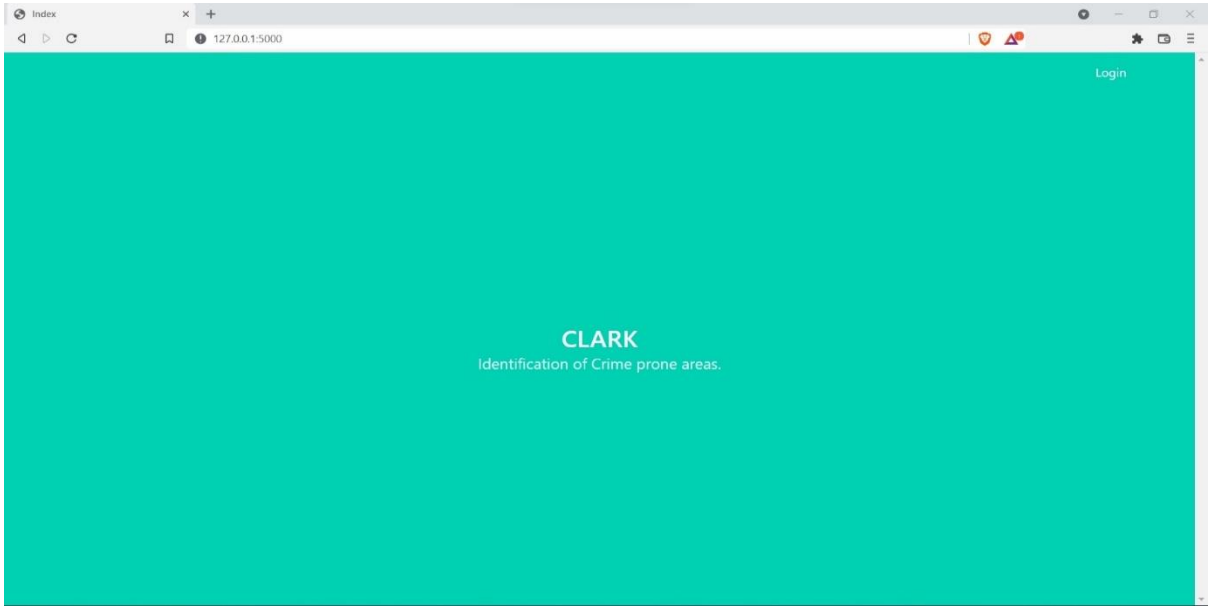


Figure 7.1: Home Page

The Figure 7.1 represents the “Home Page”, which is the first opening page of the project.

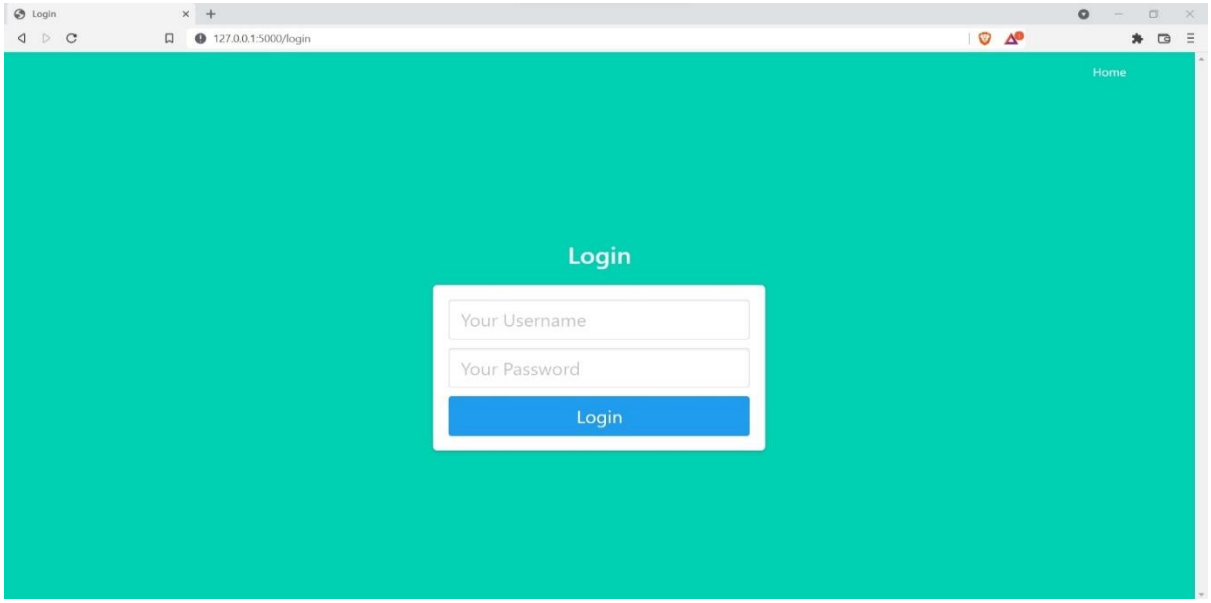


Figure 7.2: Login Page

The Figure 7.2 represents the “Login Page”, which is an authentication page where the user is requested to enter login credentials.

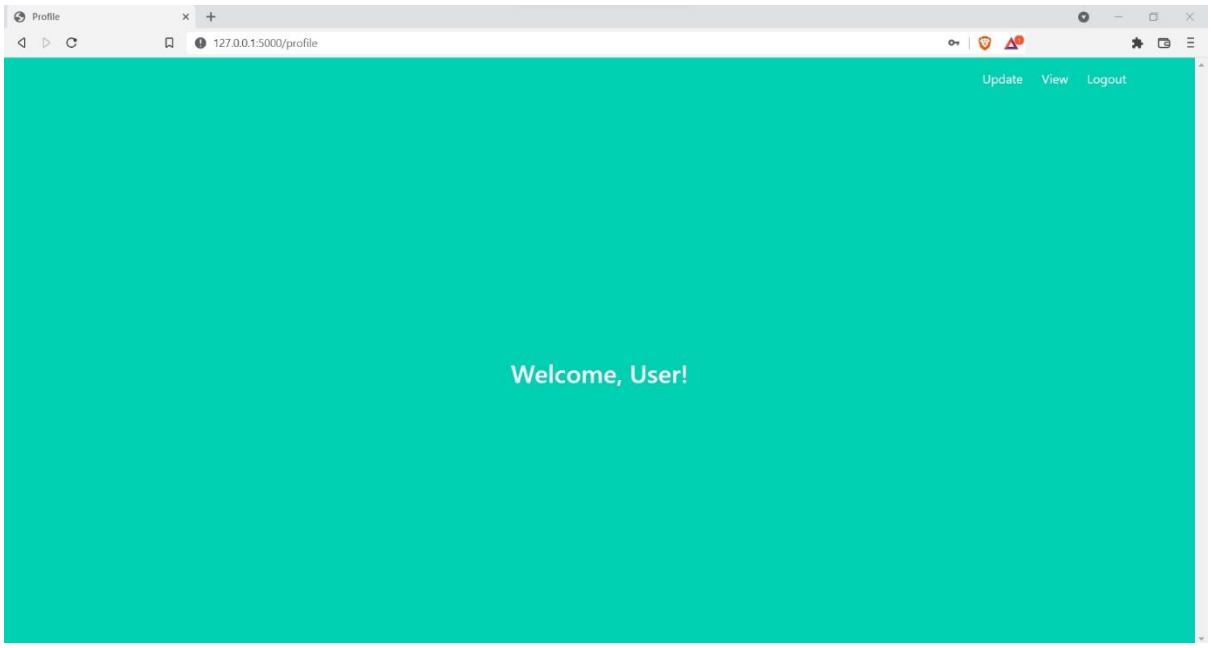


Figure 7.3: Welcome Page

The Figure 7.3 represents the “Welcome Page” which is shown to the user after successfully logging in.

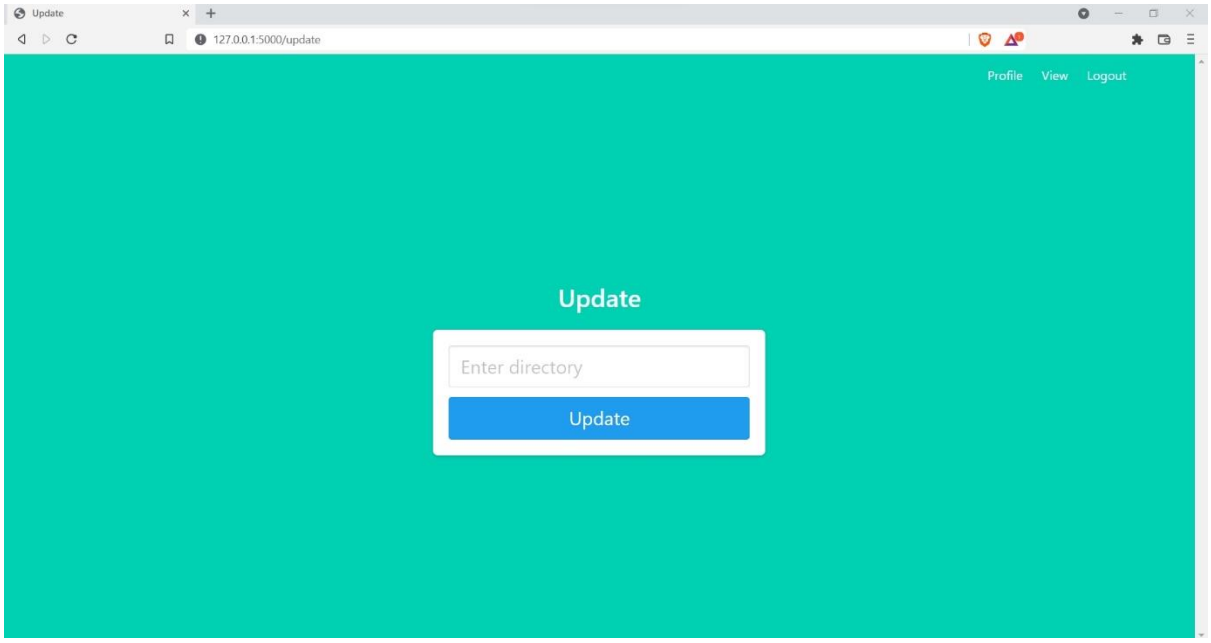


Figure 7.4: Update Page

The Figure 7.4 shows the “Update Page” where the user is asked to enter a valid excel file directory.

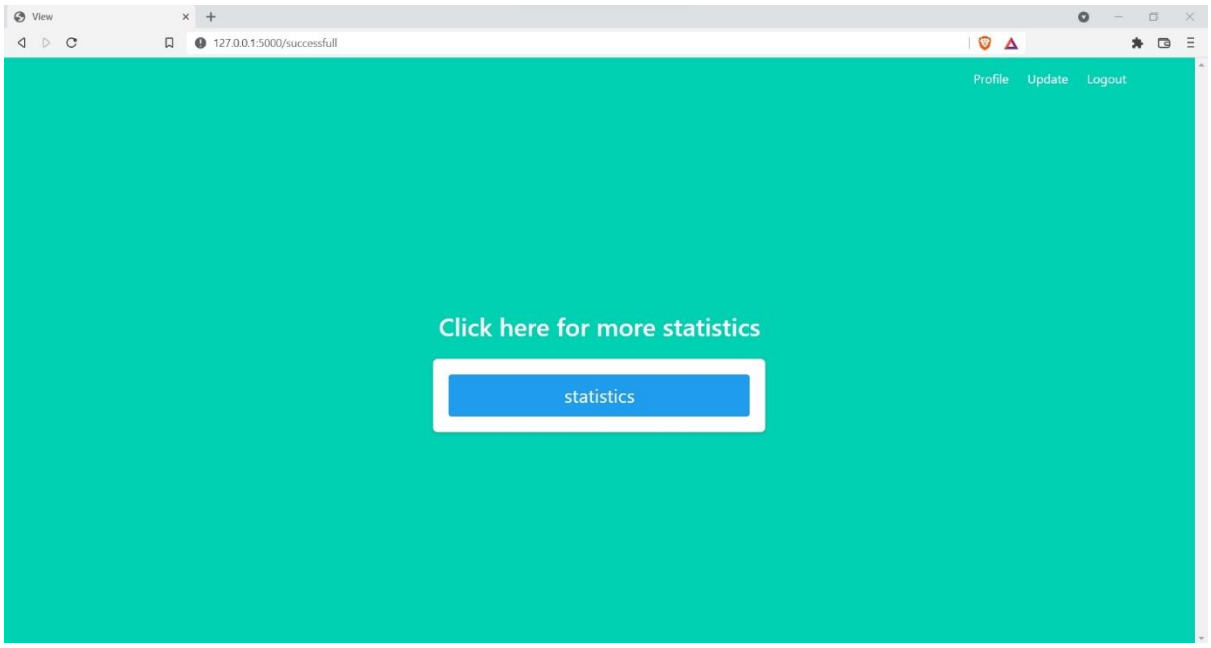


Figure 7.5: Statistics Page

The Figure 7.5 shows the “Statistics Page” which has a button to view results/statistics.

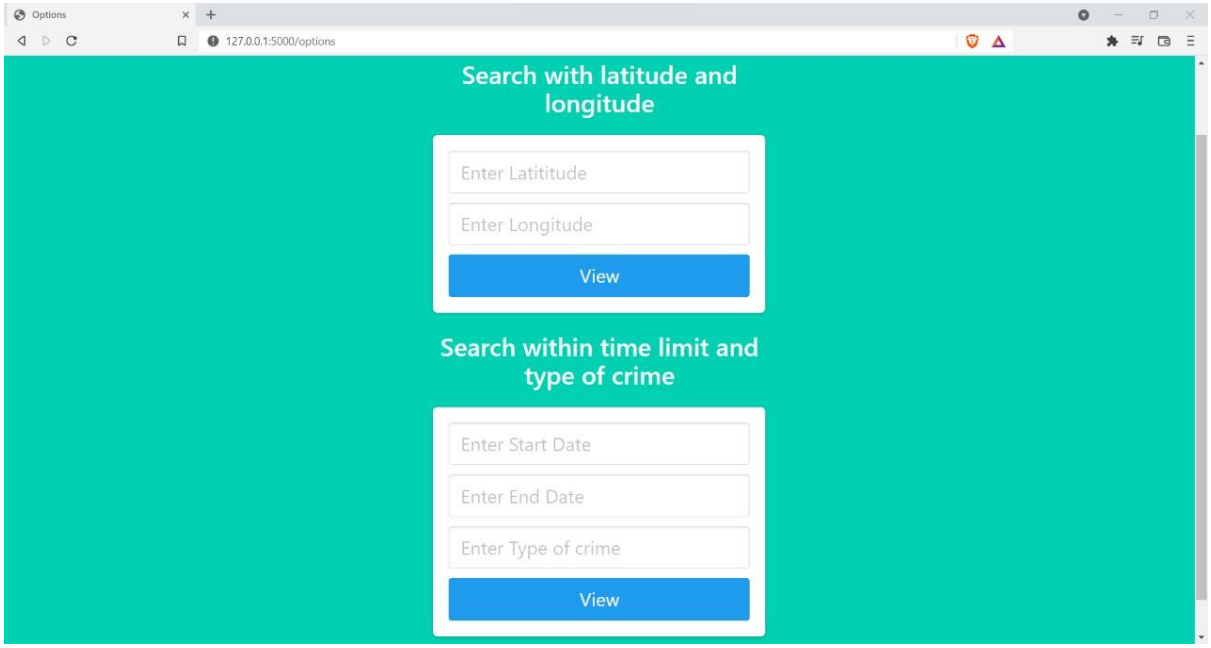
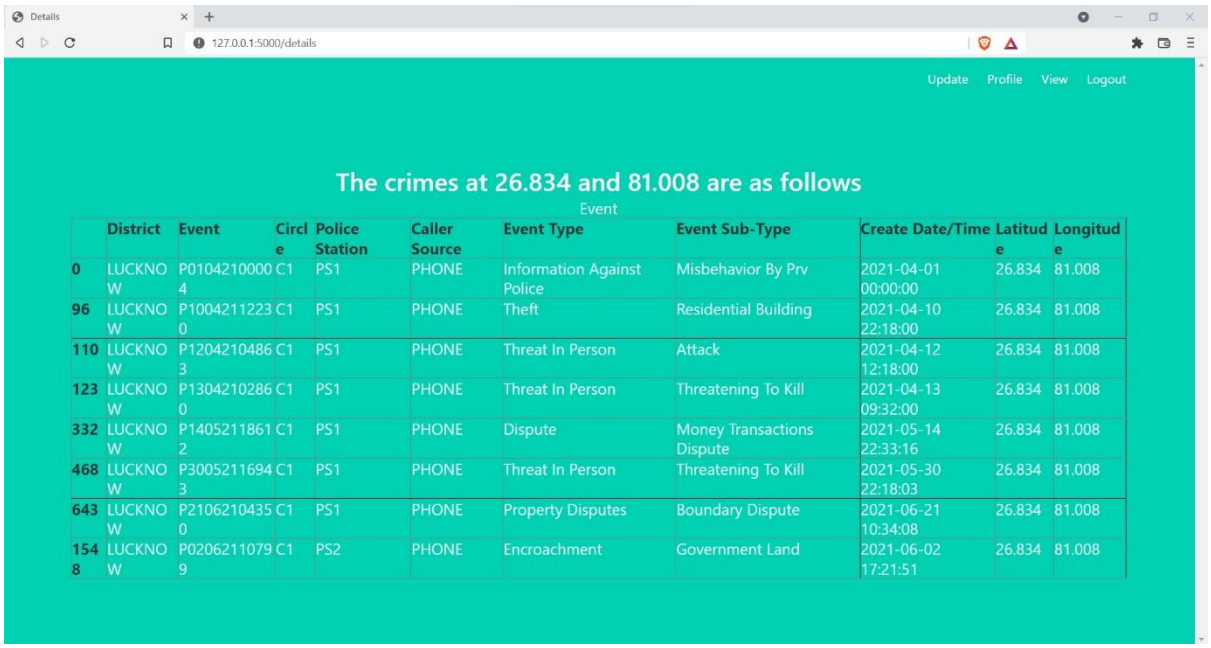


Figure 7.6: Option Entry Page

The Figure 7.6 shows the “Option Entry Page” in which there are two possibilities for the user to view the results. First being a tabular representation based on latitude and longitude and the other being a cluster representation based on time range and type of crime.



	District	Event	Circle	Police Station	Caller Source	Event Type	Event Sub-Type	Create Date/Time	Latitude	Longitude
0	LUCKNOW	P01042100004	C1	PS1	PHONE	Information Against Police	Misbehavior By Prv	2021-04-01 00:00:00	26.834	81.008
96	LUCKNOW	P10042112230	C1	PS1	PHONE	Theft	Residential Building	2021-04-10 22:18:00	26.834	81.008
110	LUCKNOW	P12042104863	C1	PS1	PHONE	Threat In Person	Attack	2021-04-12 12:18:00	26.834	81.008
123	LUCKNOW	P13042102860	C1	PS1	PHONE	Threat In Person	Threatening To Kill	2021-04-13 09:32:00	26.834	81.008
332	LUCKNOW	P14052118612	C1	PS1	PHONE	Dispute	Money Transactions Dispute	2021-05-14 22:33:16	26.834	81.008
468	LUCKNOW	P30052116943	C1	PS1	PHONE	Threat In Person	Threatening To Kill	2021-05-30 22:18:03	26.834	81.008
643	LUCKNOW	P21062104350	C1	PS1	PHONE	Property Disputes	Boundary Dispute	2021-06-21 10:34:08	26.834	81.008
1548	LUCKNOW	P02062110799	C1	PS2	PHONE	Encroachment	Government Land	2021-06-02 17:21:51	26.834	81.008

Figure 7.7: Table Page

The Figure 7.7 shows the “Table Page” in which all the crime records belonging to a particular latitude and longitude are shown in a tabular format.

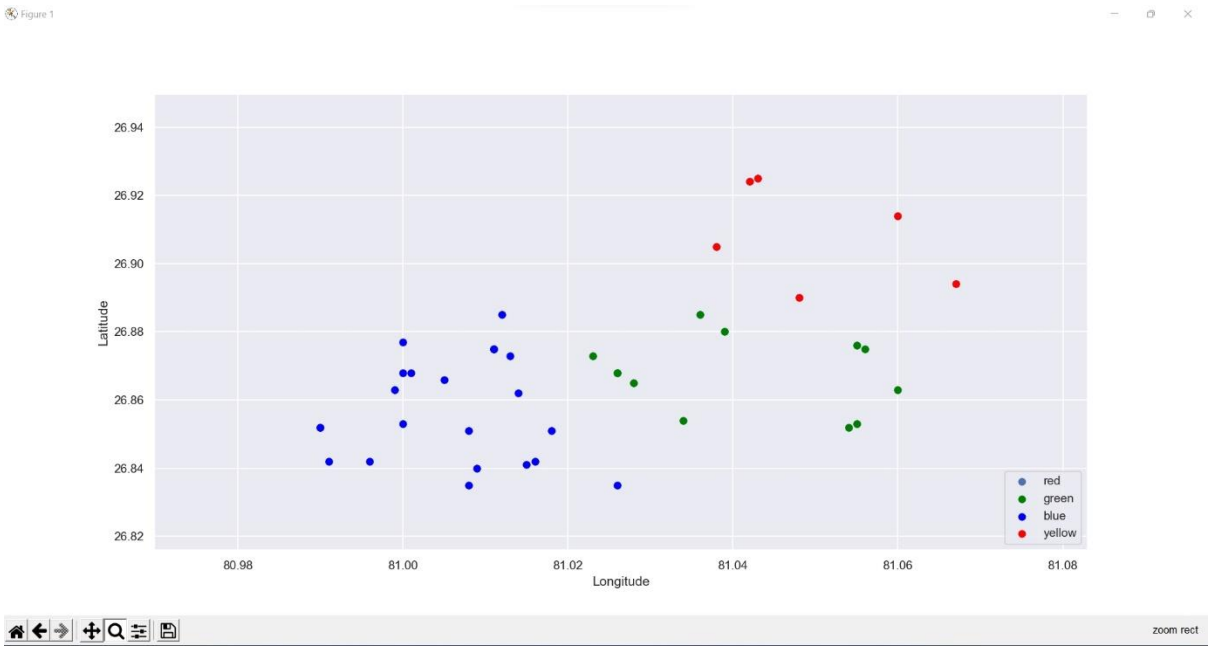


Figure 7.8: Graph Result Page

The Figure 7.8 shows the “Graph Result Page” in which the crimes occurred within a given time range and type of crime are shown in a clustered format where the clusters represent the densities of the crime.

CHAPTER 8

CONCLUSION AND FUTURE ENHANCEMENT

8.1 Conclusion:

The system takes required details of an area where the crime has taken place and preprocessing provides the frequent patterns of that place. Corresponding to each place a model is built by training on these frequent patterns. Crime patterns cannot be static since patterns change over time. So, the system is trained based on some particular inputs so that machine automatically learns the converting patterns in crime and provides the necessary outputs in the form of clusters. Since only some limited factors are considered, full accuracy cannot be achieved. For getting better results in prediction, more crime attributes of places have to be found. Now the system is trained using certain attributes. Thus, the system helps in overcoming the barriers present in the current available systems. Though it has some limitations, the implementation of this system will surely benefit the law enforcement department.

8.2 Future Enhancement:

This application has much potential for future enhancement. More algorithms can be implemented to show crime rates in order to do more statistical analysis on the results. The project can also be enhanced by integrating plot graph with geographical map to show a more user appealing representation.

REFERENCES

1. Data Analytics using Python by Bharti Motwani
2. CSS and HTML Web Design by Craig Grannell
3. Data Clustering: Algorithms and Applications by Charu C Aggarwal
4. Building Web Apps with Python and Flask