**ASIAN COLLEGE OF ENGINEERINNG AND TECHNOLOGY**

**KONDAYAMPALAYAM,COIMBATORE-641110.**



**CRIME VISION:ADVANCED CRIME CLASSIFICATION**

Submitted by

**715320106002 G.BHARATH**

**715320106006 P.KUPPUSAMY**

**715320106007 A.MANOJ**

**715320106301 J.HARIPRIYA**

**Under The Guidance of**

**Mr.N.DURAIRAJ M.E,**

**In partial fulfillment for the award of the engineering**

**In**

**ELECTRONICS &COMMUNICATION ENGINEERING**

**Of the State Board of Technical Education**

**Government of Tamil Nadu**

**ASIAN COLLEGE OF ENGINEERINNG AND TECHNOLOGY**

**KONDAYAMPALAYAM,COIMBATORE-641110.**



**DEPARTMENT OF ELECTRONICS &COMMUNICATION ENGINEERING**

**NAAN MUDHALAVAN**

This is to certify that the project report titled **“CRIME VISION ADVANCED CRIME CLASSIFICATION’’** is the record of work done by

**Chi. /Sow. G.BHARATH             Reg. No.715320106002**

In partial fulfillment of the requirements for the award of engineering & **ECE** during the year 2022-2023

**COURSE MENTOR SPOC PRINCIPAL**

**DECLARATION**

We, **,** that this project entitled **“CRIME VISION:ADVANCED CRIME CLASSIFICATION”** is submitted to Asian College of engineering & technology, Coimbatore in project fulfillment for the award of Electronics &Communication Engineering is a record of original work done by us under the supervision and guidance of

**Mr. DURAIRAJ ME.,** Electronics &Communication Engineering, Asian College of engineering & technology, Coimbatore This project work has not been submitted by us for the award of any other Degree/diploma/Associate ship/Fellowship or any other similar degree to any other Universities.

PLACE: Coimbatore

DATE:

**G.BHARATH 715320106002**

**P.KUPPUSAMY 715320106006**

**A.MANOJ 715320106007**

**J.HARIPRIYA 715320106301**

**ACKNOWLEDGEMENT**

I am very grateful and gifted in taking up this opportunity to thank the **LORD ALMIGHTY** for showering his unlimited blessings upon me

.

At this pleasing moment of having successfully completed our project , we wish to convey our sincere thanks to our naan mudhalvan coordinator Mr. **N.DURAIRAJ .M.E**,

We   also   express   our   indebt   thanks   to our coordinator **Mr.N.DURAIRAJ .M.E**, for their sincere support in completion of this project.

I thank my family members and friends for their honorable support.

**SYNOPSIS**

Weather recognition is a common problem for many branches of industry.

CHAPTER TITLE PAGENUMBER

**1 INTRODUCTION 7**

1.1 Project Overview

1.2Purpose

**2 LITERATURE SURVEY 7**

2.1 Existing problem

* 1. References

2.3 Problem Statement Definition

**3 IDEATION & PROPOSED SOLUTION 9**

3.1 Empathy Map Canvas

3.2 Ideation & Brainstorming

3.3Proposed Solution

3.4Problem Solution fit

**4 REQUIREMENT ANALYSIS 12**

4.1 Functional requirement

* 1. Non-Functional requirements

**5 PROJECT DESIGN 12**

5.1Data Flow Diagrams

5.2 Solution & Technical Architecture

* 1. User Stories

**6 PROJECT PLANNING & SCHEDULING 14**

6.1Sprint Planning & Estimation

6.2Sprint Delivery Schedule

6.3Reports from JIRA

**7 CODING & SOLUTIONING 19**

7.1Feature 1

7.2Feature 2

7.3Database Schema (if Applicable)

**8 TESTING 20**

8.1Test Cases

8.2User Acceptance Testing

**9 RESULTS 21**

* 1. Performance Metrics

**10 ADVANTAGES & DISADVANTAGES 22**

**11 CONCLUSION 23**

**12 FUTURE SCOPE 23**

**13 APPENDIX 23**

Source Code

, GitHub & Project Demo Link

**1.INTRODUCTION**

**1.1 PROJECT OVERVIEW**

**“CRIME VISION: ADVANCED CRIME CLASSIFICATION”** The safety of a community is its top priority, and as a result, governments take necessary actions to reduce crime rates. This, in turn, ensures economic growth and quality of life. Crime analysis is a critical part of criminology that focuses on studying behavioral patterns and tries to identify the indicators of such events ([Mahmud et al., 2017](https://www.sciencedirect.com/science/article/pii/S2590291122000961" \l "bib70)). However, several complications arise during crime **“CRIME VISION:ADVANCED CRIME CLASSCIFICATION”** es

to identify either likely places of future crime scenes or

past crime perpetrators, by applying statistical predictions [29].

As a crime typically involves a perpetrator and a target and

occurs at a certain place and time, techniques of predictive

policing need to answer: a) who will commit a crime, b) who

will be offended, c) what type of crime, d) in which location

and e) at what time a new crime will take place. This work

does not focus on the victim and the offender, but on the

prediction of occurrence of a certain crime type per location

and time using past data.

1.INTRODUCTION:

1.1 Motivation

1.INTRODUCTION:

1.1 Motivation

1.INTRODUCTION:

1.1 Motivation

1.INTRODUCTION:

1.1 Motivation

1.INTRODUCTION:

1.1 Motivation

1.INTRODUCTION:

1.1 Motiva

PREDICTIVE policing is the use of

to identify either likely places of future crime scenes or

past crime perpetrators, by applying statistical predictions [29].

As a crime typically involves a perpetrator and a target and

occurs at a certain place and time, techniques of predictive

policing need to answer: a) who will commit a crime, b) who

will be offended, c) what type of crime, d) in which location

and e) at what time a new crime will take place. This work

does not focus on the victim and the offender, but on the

prediction of occurrence of a certain crime type per location

and time using past data.

The ultimate goal, in a policing context, is the selection of

the top areas in the city for the prioritization of law enforce-

ment resources per department. One of the most challenging

issues of police departments is to have accurate crime forecasts

to dynamically deploy patrols and other resources so as to

improve deterring of crime occurrence and police response

times

to identify either likely places of future crime scenes or

past crime perpetrators, by applying statistical predictions [29].

As a crime typically involves a perpetrator and a target and

occurs at a certain place and time, techniques of predic

to identify either likely places of future crime scenes or

past crime perpetrators, by applying statistical predictions [29].

As a crime typically involves a perpetrator and a target and

occurs at a certain place and time, techniques of predictive

policing need to answer: a) who will commit a crime, b) who

will be offended, c) what type of crime, d) in which location

and e) at what time a new crime will take place. This work

does not focus on the victim and the offender, but on the

prediction of occurrence of a certain crime type per location

and time using past data.

The ultimate goal, in a policing context, is the selection of

the top areas in the city for the prioritization of law enforce-

ment resources per department. One of the most challenging

issues of police departments is to have accurate crime forecasts

to dynamically deploy patrols and other resources so as to

improve deterring of crime occurrence and police response

times

Predictive policing is the use of analytical techniques to identify likely places of future crime scenes or past crime perpetrators, by applying statistical predictions [[**1**](https://www.mdpi.com/2571-9394/3/4/46#B1-forecasting-03-00046)]. As a crime typically involves a perpetrator and a target and occurs at a certain place and time, techniques of predictive policing need to answer: (a) who will commit a crime, (b) who will be offended, (c) what type of crime, (d) in which location and (e) at what time a new crime will take place. This work does not focus on the victim and the offender, but on the prediction of when and where a certain crime type will occur, using past data.The ultimate goal, in a policing context, is the selection of the top areas in the city for the prioritisation of law enforcement resources per department. One of the most challenging issues of police departments is to have accurate crime forecasts to dynamically deploy patrols and other resources to improve deterring of crime occurrence and police response times. Routine activity theory [[**2**](https://www.mdpi.com/2571-9394/3/4/46#B2-forecasting-03-00046)] suggests that most crimes take place when three conditions are met: a motivated offender, a suitable victim and lack of victim protection. The rational choice theory [[**3**](https://www.mdpi.com/2571-9394/3/4/46#B3-forecasting-03-00046)], suggests that prospective criminal weights the gain of successfully committing the crime against the probability of being caught and makes a rational choice whether to actually commit the crime or not. Both theories agree that a crime takes place when a person willing to commit it has an opportunity to do so. As empirical studies in near repeat victimisation [[**4**](https://www.mdpi.com/2571-9394/3/4/46#B4-forecasting-03-00046),[**5**](https://www.mdpi.com/2571-9394/3/4/46#B5-forecasting-03-00046),[**6**](https://www.mdpi.com/2571-9394/3/4/46#B6-forecasting-03-00046),[**7**](https://www.mdpi.com/2571-9394/3/4/46#B7-forecasting-03-00046)] have shown, these opportunities are not randomly distributed but follow patterns in both space and time. Traditionally, police officers use maps of an area and place a pin on the map for every reported incident. Studying these maps, they can detect these patterns and, thus, efficiently predict hotspots; the areas with the higher possibility for a crime to occur, compared to the neighbouring areas.

**1.2   PURPOSE**

Criminology literature investigates the relationship between crime and various features, developing approaches for crime forecasting. The majority of the works focus on the prediction of hotspots, which are areas of varying geographical size with high crime probability. The methods include Spatial and Temporal Analysis of Crime (STAC)[22], Thematic Mapping [42] and Kernel Density Estimation (KDE) [33]. In STAC, the densest concentrations of points on the map are detected and then fit to a standard deviational ellipse for each one. Through the study of the size and the alignment of the ellipses, the analyst can draw conclusions about the nature of the underlying crime clusters [6].

**2. LITERATUTE SURVEY**

Songnian Li et.al.in their paper “Geospatial big data handling theory and methods: A review and research challenges”, mentioned the proposed system, they made review on various geospatial theory and methods used to handle geospatial big data. Given some special attributes. Authors considered that customary data taking controlling methodologies and techniques are lacking and the following domains were recognized as in requirement for further advancement and examination in the control. This incorporates the advancements in calculations to manage real-time analytics and to support ongoing flooding data, as well as improving new spatial indexing techniques. The improvement of theoretical and methodological ways to deal with transfer of big data from illustrative and parallel research and applications to ones that investigates easygoing and illustrative connections.

**2.1 EXISTING PROBLEM**

The primary research aims to find various efficient algorithms for predicting neighborhood crimes. In our previous work [6], we used statistical analysis to predict the crimes in Newyork city. Our paper got good attention from the researchers, so we wanted to look for the efficient machine learning and deep learning approaches used in this area. We have followed a systematic approach to select the papers for this review. As part of this research, we have considered the papers from multiple databases related to predicting crime.

**2.2 REFEREMCES**

Nowadays automatic visual surveillance is the main need for security and this paper presents the first step in the direction of automatic visual weapon detection. The objective of our paper is to develop a framework for visual weapon detection for surveillance. By using the K-mean clustering algorithm the proposed framework exploits the color-based segmentation to eliminate unrelated objects from an image. Speeded up robust features (SURF) interest point detector is used to locate the weapon. Our framework is robust enough in terms of scale, rotation, and occlusion. We have implemented and tested the system over sample images of weapons. Our system performs under different appearances of images.

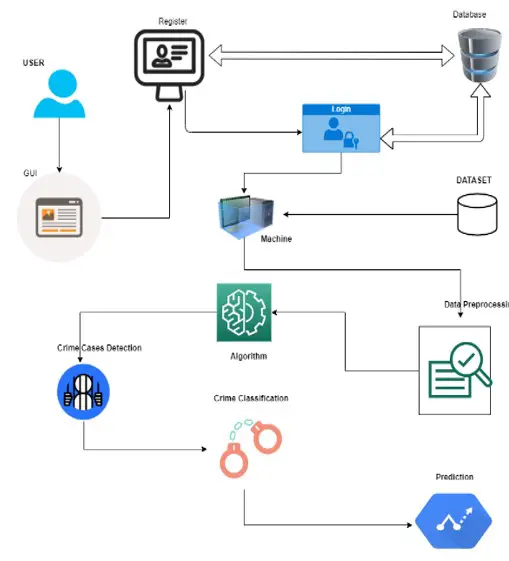
**2.3 PROBLEM STATEMENT DEFINITION**

Crimes now a days are increasing day by day and with different level ofintensity and versatility. The result is a great loss to society in terms of monitory loss,social loss and further it enhances the level of threat against the smooth livelihood inthe society. To overcome this problem, the computing era can help to reduce the crimeor even may be helpful in predicting the crime so that sufficient measures can betaken to minimize the loss to property and life. The crime rate prediction strategiescan be applied on historical data available in the police records by examining the dataat various angles like reason of crime, frequency of similar kind of crimes at specificlocation with other parameters to prepare the model crime prediction.

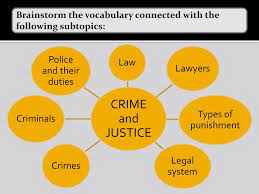
**3. IDEATION AND PROPOSED SOLUTION**

The proposed system is made on the basis of the research work that is done bygoing through various such documentations. Nearly all of the crimes are predictingbased on the location and the types of crimes that are occurring in those areas. Onsurveying previous works, Linear Regression, Decision Tree and Random Forest tendto give good accuracy so these models are used in this paper to predict crimes. Thedataset used in this paper is fromdata.world.com. The data set contains different typesof crimes that being committed in India according to the state and year respectively.

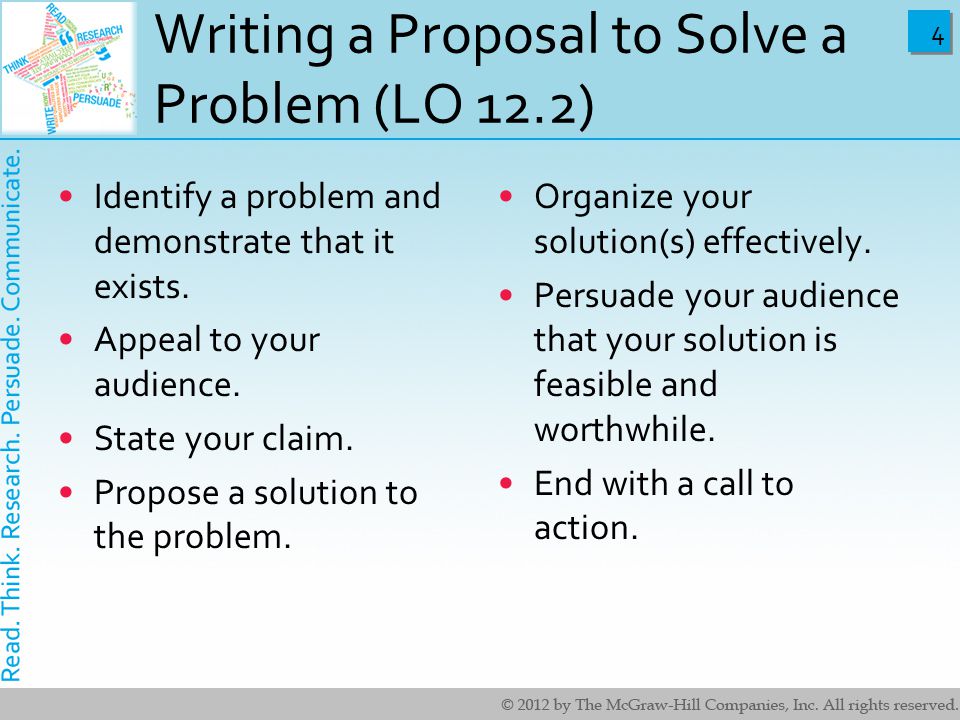
**3.1EMTATHY MAP CANVAS**



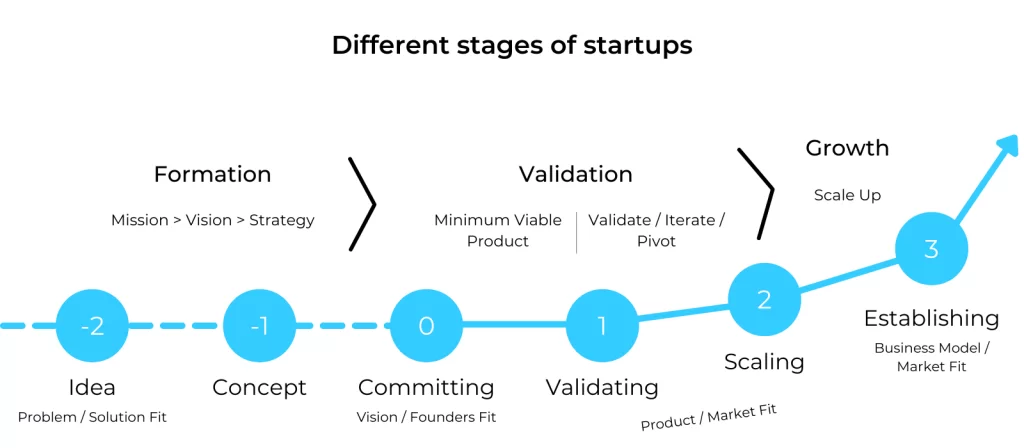
**3.2 IDEATION &BRAIN STORMING**



**3.3 PROPOSED SOLUTION**



**3.4 PROMBLEM SOLUTION FIT**

****

**4. REQUIREMENT ANALYSIS**

1.Data Collection

2.Data Pre processing

3.Training and Testing

4.Modiling

5.Predicting

**4.1 FUNCTIONAL REQUIREMENT**

A functional requirement is a statement of how a system must behave. It defines what the system should do in order to meet the user's needs or expectations. Functional requirements can be thought of as features that the user detects.

**4.2 NON-FUNCTIONAL REQUIREMENT**

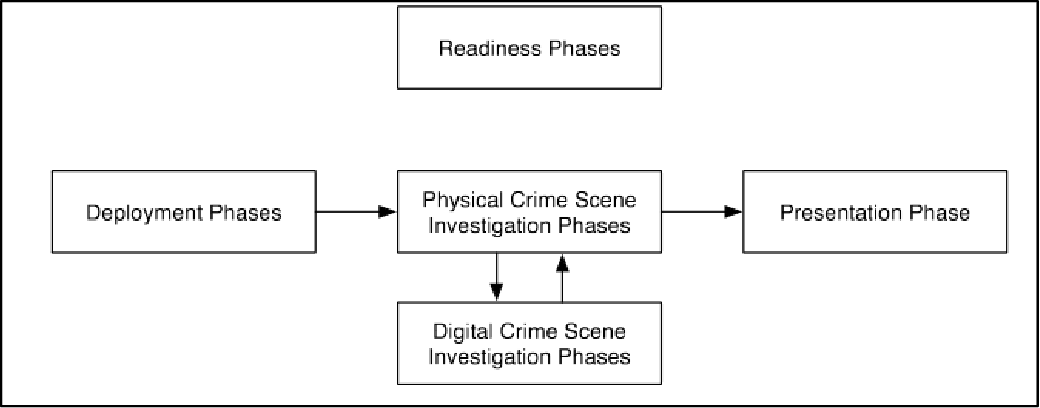
**Packing slips shall be printed on both sides of 4”x 6” white paper, the standard size for packing slips used by local printers.**

**5.PROJECT DESGIN**

****

A process of outlining all of a project's stages and creating a project plan . It includes strategizing, developing ideas, gathering resources and creating processes to achieve goals and keep within a budget and deadline.

**6. PROJECT PLANNING & SCHEDULING**

****

**Seven Steps to Successful Project Planning**

Think of your plan as a roadmap for stakeholders.

Break the project into a list of deliverables.

Talk to your team.

Identify risks.

Create a budget.

Add milestones.

Set progress reporting guidelines.

**7.CODING & SOLUTIONING(explain the features added in the project along with code )**

train\_set=image\_dataset\_from\_directory(

train\_dir,

label\_mode="categorical",

batch\_size=BATCH\_SIZE,

image\_size=IMG\_SHAPE,

shuffle=True,

seed=seed,

validation\_split=0.2,

subset="training",

)

val\_set=image\_dataset\_from\_directory(

train\_dir,

label\_mode="categorical",

batch\_size=BATCH\_SIZE,

image\_size=IMG\_SHAPE,

shuffle=True,

seed=seed,

validation\_split=0.2,

subset="validation",

)

test\_set=image\_dataset\_from\_directory(

test\_dir,

label\_mode="categorical",

class\_names=None,

batch\_size=BATCH\_SIZE,

image\_size=IMG\_SHAPE,

shuffle=False,

seed=seed,

)

Found 1266345 files belonging to 14 classes.

Using 1013076 files for training.

**8.TESTING**

8.1Test Cases

2022-01-10 17:23:02.203850: I tensorflow/stream\_executor/cuda/cuda\_gpu\_executor.cc:937] successful NUMA node read from SysFS had negative value (-1), but there must be at least one NUMA node, so returning NUMA node zero

2022-01-10 17:23:02.204962: I tensorflow/stream\_executor/cuda/cuda\_gpu\_executor.cc:937] successful NUMA node read from SysFS had negative value (-1), but there must be at least one NUMA node, so returning NUMA node zero

2022-01-10 17:23:02.205600: I tensorflow/stream\_executor/cuda/cuda\_gpu\_executor.cc:937] successful NUMA node read from SysFS had negative value (-1), but there must be at least one NUMA node, so returning NUMA node zero

2022-01-10 17:23:02.208281: I tensorflow/core/platform/cpu\_feature\_guard.cc:142] This TensorFlow binary is optimized with oneAPI Deep Neural Network Library (oneDNN) to use the following CPU instructions in performance-critical operations: AVX2 AVX512F FMA

To enable them in other operations, rebuild TensorFlow with the appropriate compiler flags.

2022-01-10 17:23:02.209404: I tensorflow/stream\_executor/cuda/cuda\_gpu\_executor.cc:937] successful NUMA node read from SysFS had negative value (-1), but there must be at least one NUMA node, so returning NUMA node zero

2022-01-10 17:23:02.210062: I tensorflow/stream\_executor/cuda/cuda\_gpu\_executor.cc:937] successful NUMA node read from SysFS had negative value (-1), but there must be at least one NUMA node, so returning NUMA node zero

2022-01-10 17:23:02.210694: I tensorflow/stream\_executor/cuda/cuda\_gpu\_executor.cc:937] successful NUMA node read from SysFS had negative value (-1), but there must be at least one NUMA node, so returning NUMA node zero

2022-01-10 17:23:03.954678: I tensorflow/stream\_executor/cuda/cuda\_gpu\_executor.cc:937] successful NUMA node read from SysFS had negative value (-1), but there must be at least one NUMA node, so returning NUMA node zero

2022-01-10 17:23:03.955790: I tensorflow/stream\_executor/cuda/cuda\_gpu\_executor.cc:937] successful NUMA node read from SysFS had negative value (-1), but there must be at least one NUMA node, so returning NUMA node zero

2022-01-10 17:23:03.956569: I tensorflow/stream\_executor/cuda/cuda\_gpu\_executor.cc:937] successful NUMA node read from SysFS had negative value (-1), but there must be at least one NUMA node, so returning NUMA node zero

2022-01-10 17:23:03.957257: I tensorflow/core/common\_runtime/gpu/gpu\_device.cc:1510] Created device /job:localhost/replica:0/task:0/device:GPU:0 with 15385 MB memory: -> device: 0, name: Tesla P100-PCIE-16GB, pci bus id: 0000:00:04.0, compute capability: 6.0

Found 1266345 files belonging to 14 classes.

Using 253269 files for validation.

Found 111308 files belonging to 14 classes.

8.2User Acceptance Testing

model=create\_model()

model.compile(optimizer="adam",

loss='categorical\_crossentropy',

metrics = [tf.keras.metrics.AUC()])

Downloading data from https://storage.googleapis.com/tensorflow/keras-applications/densenet/densenet121\_weights\_tf\_dim\_ordering\_tf\_kernels\_notop.h5

29089792/29084464 [==============================] - 0s 0us/step

29097984/29084464 [==============================] - 0s 0us/step

Model: "sequential"

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Layer (type) Output Shape Param #

=================================================================

densenet121 (Functional) (None, 2, 2, 1024) 7037504

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

global\_average\_pooling2d (Gl (None, 1024) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dense (Dense) (None, 256) 262400

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dropout (Dropout) (None, 256) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dense\_1 (Dense) (None, 512) 131584

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dropout\_1 (Dropout) (None, 512) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dense\_2 (Dense) (None, 1024) 525312

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

classification (Dense) (None, 14) 14350

=================================================================

Total params: 7,971,150

Trainable params: 6,386,894

Non-trainable params: 1,584,256

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

fig, c\_ax = plt.subplots(1,1, figsize = (15,8))

print('ROC AUC score:', multiclass\_roc\_auc\_score(y\_true , y\_pred , average = "micro"))

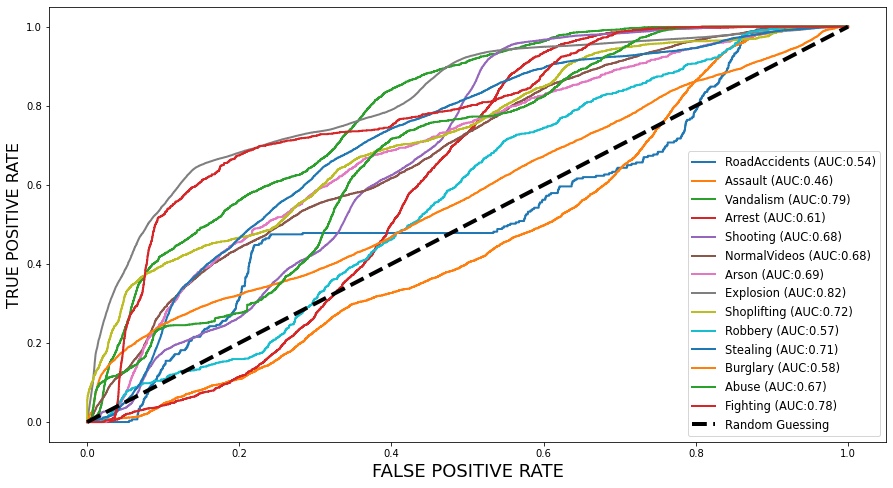
plt.xlabel('FALSE POSITIVE RATE', fontsize=18)

plt.ylabel('TRUE POSITIVE RATE', fontsize=16)

plt.legend(fontsize = 11.5)

plt.show()

ROC AUC score: 0.8298142844970892

****

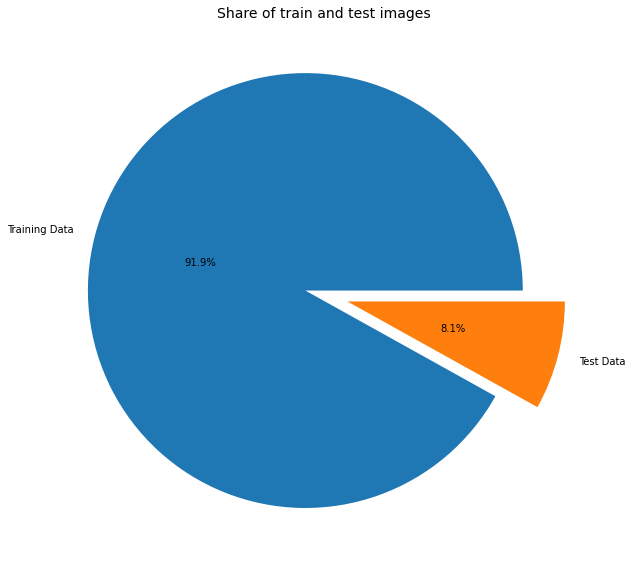
**9.RESULT**

**9.1 Performance Metrics**

plt.figure(figsize=(15, 10))

plt.pie(x=np.array([train,test]), autopct="**%.1f%%**", explode=[0.1, 0.1], labels=["Training Data", "Test Data"], pctdistance=0.5)

plt.title("Share of train and test images ", fontsize=14);



**10.ADVANTAGES & DIS-ADVANTAGES**

**Advantages**

As such crime has no advantages in perspective of society but for the in perspective of criminal the advantages are

* Criminal may get recognition in the society, though it will be bad recognition.
* Criminal may possess physical goods through burglary

**Disadvantages**

* The criminal has to face various punishments set by the law department.
* The person will get hatred from the society.
* He may feel remorseful after doing that and may get trouble sleeping.

**11.CONCLUSION**

:It is clear that basic details of criminal activities in a neighborhoodcontain indicators that will be employed by machine learning agents to classify acriminal activity given a location and date. The training agent suffers fromimbalanced categories of the dataset, it had been ready to overcome the problemby oversampling and under-sampling the dataset. This paper presents a crimedata prediction by taking the types of crimes as input and giving are in whichthese crimes are committed as output using Jupyter notebook having python as acore language and python provide inbuilt libraries such as Pandas and Numpythrough which the work will be completed faster and Scikit provides all theprocesses of how to use different libraries providing by the python.

**12.FUTURE SCOPE**

From the encouraging results, we believe that crime data mining has a promising future for increasing the effectiveness and efficiency of criminal and intelligence analysis. Visual and intuitive criminal and intelligence investigation techniques can be developed for crime pattern. As we have applied clustering technique of data mining for crime analysis we can also perform other techniques of data mining such as classification.

**S13. APPENDIX**

import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

%matplotlib inline

import seaborn as sns

crime = pd.read\_csv('../input/crime-classifcication/Crime1.csv',usecols=['Dates','Category','Descript','DayOfWeek','PdDistrict','Resolution','Address'])

crime.head()

Out[3]:

|  | Dates | Category | Descript | DayOfWeek | PdDistrict | Resolution | Address |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 0 | 5/13/2015 23:53 | WARRANTS | WARRANT ARREST | Wednesday | NORTHERN | ARREST, BOOKED | OAK ST / LAGUNA ST |
| 1 | 5/13/2015 23:53 | OTHER OFFENSES | TRAFFIC VIOLATION ARREST | Wednesday | NORTHERN | ARREST, BOOKED | OAK ST / LAGUNA ST |
| 2 | 5/13/2015 23:33 | OTHER OFFENSES | TRAFFIC VIOLATION ARREST | Wednesday | NORTHERN | ARREST, BOOKED | VANNESS AV / GREENWICH ST |
| 3 | 5/13/2015 23:30 | LARCENY/THEFT | GRAND THEFT FROM LOCKED AUTO | Wednesday | NORTHERN | NONE | 1500 Block of LOMBARD ST |
| 4 | 5/13/2015 23:30 | LARCENY/THEFT | GRAND THEFT FROM LOCKED AUTO | Wednesday | PARK | NONE | 100 Block of BRODERICK ST |

crime.info()

<class 'pandas.core.frame.DataFrame'>

RangeIndex: 835 entries, 0 to 834

Data columns (total 7 columns):

Dates 835 non-null object

Category 835 non-null object

Descript 835 non-null object

DayOfWeek 835 non-null object

PdDistrict 835 non-null object

Resolution 835 non-null object

Address 835 non-null object

dtypes: object(7)

memory usage: 45.8+ KB

crime.dtypes

Out[5]:

Dates object

Category object

Descript object

DayOfWeek object

PdDistrict object

Resolution object

Address object

dtype: object

crime.describe()

Out[6]:

|  | Dates | Category | Descript | DayOfWeek | PdDistrict | Resolution | Address |
| --- | --- | --- | --- | --- | --- | --- | --- |
| count | 835 | 835 | 835 | 835 | 835 | 835 | 835 |
| unique | 379 | 28 | 165 | 7 | 10 | 7 | 592 |
| top | 4/30/2015 18:00 | LARCENY/THEFT | GRAND THEFT FROM LOCKED AUTO | Saturday | SOUTHERN | NONE | 800 Block of BRYANT ST |
| freq | 16 | 223 | 76 | 313 | 144 | 600 | 15 |

crime['Dates'] = pd.to\_datetime(crime['Dates'])

In [8]:

crime.dtypes

Out[8]:

Dates datetime64[ns]

Category object

Descript object

DayOfWeek object

PdDistrict object

Resolution object

Address object

dtype: object

crime['Category'].value\_counts()

Out[9]:

LARCENY/THEFT 223

NON-CRIMINAL 102

OTHER OFFENSES 98

ASSAULT 60

VEHICLE THEFT 54

VANDALISM 40

BURGLARY 39

SUSPICIOUS OCC 35

MISSING PERSON 27

WARRANTS 26

DRUG/NARCOTIC 15

ROBBERY 14

SECONDARY CODES 14

FRAUD 14

PROSTITUTION 13

TRESPASS 11

WEAPON LAWS 11

SEX OFFENSES FORCIBLE 10

DRUNKENNESS 6

DRIVING UNDER THE INFLUENCE 5

KIDNAPPING 4

STOLEN PROPERTY 3

ARSON 3

DISORDERLY CONDUCT 2

LIQUOR LAWS 2

FORGERY/COUNTERFEITING 2

BRIBERY 1

EMBEZZLEMENT 1

Name: Category, dtype: int64

In [10]:

crime['DayOfWeek'].value\_counts()

Out[10]:

Saturday 313

Thursday 207

Monday 117

Tuesday 94

Sunday 77

Friday 17

Wednesday 10

Name: DayOfWeek, dtype: int64

In [11]:

crime['PdDistrict'].value\_counts()

Out[11]:

SOUTHERN 144

NORTHERN 126

MISSION 116

CENTRAL 87

BAYVIEW 84

TARAVAL 75

INGLESIDE 66

RICHMOND 49

TENDERLOIN 47

PARK 41

Name: PdDistrict, dtype: int64

In [12]:

crime['Resolution'].value\_counts()

Out[12]:

NONE 600

ARREST, BOOKED 214

UNFOUNDED 12

ARREST, CITED 4

CLEARED-CONTACT JUVENILE FOR MORE INFO 2

JUVENILE BOOKED 2

EXCEPTIONAL CLEARANCE 1

Name: Resolution, dtype: int64

crime\_category = crime.groupby('Category')['Category'].count().sort\_values(ascending=False)

In [14]:

crime\_category

Out[14]:

Category

LARCENY/THEFT 223

NON-CRIMINAL 102

OTHER OFFENSES 98

ASSAULT 60

VEHICLE THEFT 54

VANDALISM 40

BURGLARY 39

SUSPICIOUS OCC 35

MISSING PERSON 27

WARRANTS 26

DRUG/NARCOTIC 15

FRAUD 14

SECONDARY CODES 14

ROBBERY 14

PROSTITUTION 13

WEAPON LAWS 11

TRESPASS 11

SEX OFFENSES FORCIBLE 10

DRUNKENNESS 6

DRIVING UNDER THE INFLUENCE 5

KIDNAPPING 4

ARSON 3

STOLEN PROPERTY 3

DISORDERLY CONDUCT 2

FORGERY/COUNTERFEITING 2

LIQUOR LAWS 2

EMBEZZLEMENT 1

BRIBERY 1

Name: Category, dtype: int64

In [15]:

plt.figure(figsize=(10,8))

crime\_category.plot(kind='barh')

plt.xlabel('Count')

plt.title('Number of times each Crime Category took place')

plt.show()

In [16]:

print('We can see from the above plot that LARCENY/THEFT is the most common crime category')

We can see from the above plot that LARCENY/THEFT is the most common crime category

In [17]:

crime.head()

Out[17]:

|  | Dates | Category | Descript | DayOfWeek | PdDistrict | Resolution | Address |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 0 | 2015-05-13 23:53:00 | WARRANTS | WARRANT ARREST | Wednesday | NORTHERN | ARREST, BOOKED | OAK ST / LAGUNA ST |
| 1 | 2015-05-13 23:53:00 | OTHER OFFENSES | TRAFFIC VIOLATION ARREST | Wednesday | NORTHERN | ARREST, BOOKED | OAK ST / LAGUNA ST |
| 2 | 2015-05-13 23:33:00 | OTHER OFFENSES | TRAFFIC VIOLATION ARREST | Wednesday | NORTHERN | ARREST, BOOKED | VANNESS AV / GREENWICH ST |
| 3 | 2015-05-13 23:30:00 | LARCENY/THEFT | GRAND THEFT FROM LOCKED AUTO | Wednesday | NORTHERN | NONE | 1500 Block of LOMBARD ST |
| 4 | 2015-05-13 23:30:00 | LARCENY/THEFT | GRAND THEFT FROM LOCKED AUTO | Wednesday | PARK | NONE | 100 Block of BRODERICK ST |

plt.figure(figsize=(10,8))

sns.countplot(crime['DayOfWeek'])

plt.title('Day of the week on which most crimes take place')

plt.show()

In [19]:

linkcode

print('On SATURDAY most crimes take place')

On SATURDAY most crimes take place

plt.figure(figsize=(12,8))

sns.countplot(crime['PdDistrict'])

plt.show()

In [21]:

linkcode

print('SOUTHERN district is famous in terms of crimes')

SOUTHERN district is famous in terms of crimes

larseny\_descript = crime.loc[crime['Category']=='LARCENY/THEFT','Descript'].value\_counts()

In [23]:

larseny\_descript

Out[23]:

GRAND THEFT FROM LOCKED AUTO 76

PETTY THEFT OF PROPERTY 30

GRAND THEFT FROM UNLOCKED AUTO 20

PETTY THEFT FROM LOCKED AUTO 16

GRAND THEFT OF PROPERTY 13

PETTY THEFT FROM A BUILDING 12

GRAND THEFT FROM A BUILDING 10

GRAND THEFT FROM PERSON 10

GRAND THEFT PICKPOCKET 9

PETTY THEFT FROM UNLOCKED AUTO 7

PETTY THEFT SHOPLIFTING 6

ATTEMPTED THEFT FROM LOCKED VEHICLE 3

GRAND THEFT BICYCLE 3

GRAND THEFT SHOPLIFTING 2

THEFT, GRAND, OF FIREARM 1

ATTEMPTED THEFT FROM UNLOCKED VEHICLE 1

ATTEMPTED SHOPLIFTING 1

LOST PROPERTY, PETTY THEFT 1

PETTY THEFT BICYCLE 1

ATTEMPTED GRAND THEFT PURSESNATCH 1

Name: Descript, dtype: int64

*# grouping the data set on the basis of 'Category' and 'Resolution' columns*

category\_resolution = crime.groupby(['Category','Resolution'])['Category'].count()

In [25]:

category\_resolution

Out[25]:

Category Resolution

ARSON ARREST, BOOKED 1

NONE 2

ASSAULT ARREST, BOOKED 30

NONE 30

BRIBERY ARREST, BOOKED 1

BURGLARY ARREST, BOOKED 5

NONE 34

DISORDERLY CONDUCT NONE 2

DRIVING UNDER THE INFLUENCE ARREST, BOOKED 5

DRUG/NARCOTIC ARREST, BOOKED 12

ARREST, CITED 2

NONE 1

DRUNKENNESS ARREST, BOOKED 6

EMBEZZLEMENT NONE 1

FORGERY/COUNTERFEITING NONE 2

FRAUD ARREST, BOOKED 1

NONE 13

KIDNAPPING ARREST, BOOKED 2

NONE 2

LARCENY/THEFT ARREST, BOOKED 3

NONE 220

LIQUOR LAWS ARREST, BOOKED 2

MISSING PERSON CLEARED-CONTACT JUVENILE FOR MORE INFO 2

JUVENILE BOOKED 1

NONE 23

UNFOUNDED 1

NON-CRIMINAL ARREST, BOOKED 6

EXCEPTIONAL CLEARANCE 1

NONE 88

UNFOUNDED 7

OTHER OFFENSES ARREST, BOOKED 65

ARREST, CITED 2

JUVENILE BOOKED 1

NONE 30

PROSTITUTION ARREST, BOOKED 12

NONE 1

ROBBERY ARREST, BOOKED 2

NONE 12

SECONDARY CODES ARREST, BOOKED 4

NONE 10

SEX OFFENSES FORCIBLE ARREST, BOOKED 4

NONE 5

UNFOUNDED 1

STOLEN PROPERTY ARREST, BOOKED 3

SUSPICIOUS OCC ARREST, BOOKED 2

NONE 31

UNFOUNDED 2

TRESPASS ARREST, BOOKED 5

NONE 6

VANDALISM ARREST, BOOKED 5

NONE 35

VEHICLE THEFT ARREST, BOOKED 6

NONE 47

UNFOUNDED 1

WARRANTS ARREST, BOOKED 25

NONE 1

WEAPON LAWS ARREST, BOOKED 7

NONE 4

Name: Category, dtype: int64

In [26]:

*# filtering out 'LARCENY/THEFT' cases*

larceny\_cases = category\_resolution['LARCENY/THEFT']

In [27]:

larceny\_cases

Out[27]:

Resolution

ARREST, BOOKED 3

NONE 220

Name: Category, dtype: int64

In [28]:

larceny\_cases.plot(kind='bar')

plt.show()

In [29]:

linkcode

print('There was no resolution for majority of LARCENY/THEFT cases')

There was no resolution for majority of LARCENY/THEFT cases

*# grouping the data based on 'Category' and 'DayOfWeek' columns*

category\_day = crime.groupby(['Category','DayOfWeek'])['Category'].count()

In [31]:

category\_day

Out[31]:

Category DayOfWeek

ARSON Monday 1

Saturday 1

Tuesday 1

ASSAULT Friday 1

Monday 6

..

WARRANTS Wednesday 1

WEAPON LAWS Monday 2

Saturday 6

Sunday 1

Thursday 2

Name: Category, Length: 109, dtype: int64

In [32]:

*# filtering out 'LARCENY/THEFT' cases*

larceny\_day = category\_day['LARCENY/THEFT']

In [33]:

larceny\_day

Out[33]:

DayOfWeek

Friday 8

Monday 25

Saturday 82

Sunday 18

Thursday 61

Tuesday 24

Wednesday 5

Name: Category, dtype: int64

In [34]:

plt.figure(figsize=(10,8))

larceny\_day.plot(kind='bar')

plt.ylabel('Count')

plt.show()

larseny\_address = crime.groupby(['Category','PdDistrict','Address'])['Address'].count()['LARCENY/THEFT']['SOUTHERN'].sort\_values(ascending=False)

In [36]:

larseny\_address

Out[36]:

Address

800 Block of BRYANT ST 3

800 Block of MISSION ST 2

800 Block of FOLSOM ST 2

1100 Block of MARKET ST 2

1200 Block of MARKET ST 2

400 Block of THE EMBARCADEROSOUTH ST 2

800 Block of MARKET ST 2

400 Block of BRANNAN ST 1

400 Block of 10TH ST 1

3RD ST / BRYANT ST 1

1ST ST / MISSION ST 1

1900 Block of MARKET ST 1

THE EMBARCADEROSOUTH ST / HOWARD ST 1

1300 Block of MARKET ST 1

1300 Block of FOLSOM ST 1

1200 Block of FOLSOM ST 1

1100 Block of HOWARD ST 1

10TH ST / MISSION ST 1

100 Block of HAWTHORNE ST 1

400 Block of FOLSOM ST 1

4TH ST / TOWNSEND ST 1

4TH ST / SHIPLEY ST 1

SHERMAN ST / HARRISON ST 1

500 Block of BRANNAN ST 1

5TH ST / MARKET ST 1

700 Block of MARKET ST 1

700 Block of NATOMA ST 1

900 Block of HOWARD ST 1

BRANNAN ST / 3RD ST 1

FOLSOM ST / 11TH ST 1

FREELON ST / 4TH ST 1

HARRISON ST / 10TH ST 1

HARRISON ST / 6TH ST 1

MARKET ST / 10TH ST 1

MARKET ST / 5TH ST 1

100 Block of BRANNAN ST 1

Name: Address, dtype: int64

In [37]:

print('No specific address')

No specific address

Demo link:

https://drive.google.com/file/d/1OUJ80CLnrh7FlUkWJEsNmw-EmZc6KjHp/view?usp=share\_link

THANK YOU

**SUMMITED BY:**

**TEAM WORK**

to identify either likely places of future crime scenes or

past crime perpetrators, by applying statistical predictions [29].

As a crime typically involves a perpetrator and a target and

occurs at a certain place and time, techniques of predictive

policing need to answer: a) who will commit a crime, b) who

will be offended, c) what type of crime, d) in which location

and e) at what time a new crime will take place. This work

does not focus on the victim and the offender, but on the

prediction of occurrence of a certain crime type per location

and time using past