**PROJECT DOCUMENTATION SUBMITTED TO TECHIONARY**



PREDICTING THE HAPPENING OF A CRIME IN A REGION

PROJECT DATE: July 1st 2020-July 20th 2020

BATCH NO:13

| TEAM MEMBERS | | | | |
| --- | --- | --- | --- | --- |
| S.No | NAME | ROLE | COLLEGE | CONTACT NO |
| 1 | M.Poojya Sree | Student | VIGNAN INSTITUTE OF INFORMATION TECHNOLOGY | 7569685772 |
| 2 | H.Bharat Chandra | Student | VIGNAN INSTITUTE OF INFORMATION TECHNOLOGY | 9676085525 |
| 3 | K.Sai Divya | Student | VIGNAN INSTITUTE OF INFORMATION TECHNOLOGY | 8341481888 |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **PREPARED BY** | 1.Poojya Sree  2.Bharat Chandra  3.Sai Divya | **TITLE** | PREDICTING THE HAPPENING OF A CRIME IN A REGION | **DATE** | 01-07-2020 to 20-07-2020 |
| **GUIDED BY** | G.Pratharan Sai Rupak Reddy | **TITLE** | PREDICTING THE HAPPENING OF A CRIME IN A REGION | **DATE** | 01-07-2020 to 20-07-2020 |

Acknowledgement

I thank Techionary , for providing me an opportunity to do the project work on PREDICTING THE HAPPENING OF A CRIME IN A REGION and giving us all support and guidance which made me complete the project duly.

I am extremely thankful to Adnaan Arbaaz Ahmed,CTO and Director, Techionary for providing such a nice support, guidance and monitoring project progress accordingly by taking out the tike from his busy schedule of managing the corporate affairs.

I owe my deep gratitude to our project guide G.Pratharan Sai Rupak Reddy, who took keen interest on our project work and guided us all along, till the completion of our project work by providing all the necessary information for developing a good system.

I would not forget to remember Rajkumar Vanam, Rajesh Durgananth and Rakesh Manvada of Techionary for their encouragement and more over for their timely support and guidance till the completion of our project work.

I am thankful to and fortunate enough to get constant encouragement, support and guidance from all Teaching staffs of Information Technology and Computer Science Engineering which helped us in successfully completing our project work. Also, I would like to extend our sincere esteems to all staff in laboratory for their timely support.

Table of Contents

[PROBLEM STATEMENT 4](#_Toc45275812)

[PROJECT SCOPE 4](#_Toc45275813)

[PRODUCT OR SYSTEM FEATURES AND REQUIREMENTS 5](#_Toc45275814)

[PROJECT SUCCESS CRITERIA 5](#_Toc45275815)

[1. STATEMENT OF WORK 6](#_Toc45275816)

[1.1 SCOPE OF WORK 6](#_Toc45275817)

[1.2 LOCATION OF WORK 6](#_Toc45275818)

[1.3 PERIOD OF PERFORMANCE 6](#_Toc45275819)

[1.4 SCHEDULE 7](#_Toc45275820)

[1.5 STANDARDS FOLLOWED 8](#_Toc45275821)

[1.6 ACCEPTANCE CRITERIA FOR USERS 8](#_Toc45275822)

[1.7 ADDITIONAL REQUIREMENTS 8](#_Toc45275823)

# PROBLEM STATEMENT

Problems faced by the users while using the current system

|  |
| --- |
| Crimes now a days are increasing day by day and with different level of intensity and versatility. The result is great loss to society in terms of monitory loss, social loss and further it enhances the level of threat against the smooth livelihood in the society. To overcome this problem the computing era can help to reduce the crime or even may be helpful in predicting the crime so that sufficient measures can be taken to minimize the loss to property and life. The crime rate prediction strategies can be applied on historical data available in the police records by examining the data at various angles like reason of crime, frequency of similar kind of crimes at specific location with other parameters to prepare the machine learning model for crime prediction. It is the major challenge to understand the versatile data available with us then model it to predict the crime incidence with acceptable accuracy and further to reduce the crime rate. |

# PROJECT SCOPE

Project Purpose and Justification

|  |
| --- |
| The main scope of the project is to predict the crime rate in a particular region based on the historic data and visualize it graphically using statistical tools so that , it is easy to look and understand the data to support public safety,financial success and better outcomes.  Public safety and protection related to crime, and a better understanding of crime is beneficial in multiple ways: it can lead to targeted and sensitive practices by law enforcement authorities to mitigate crime, and more concerted efforts by citizens and authorities to create healthy neighborhood environments. With the advent of the Big Data era and the availability of fast, efficient algorithms for data analysis, understanding patterns in crime from data is an active and growing field of research.  This project presents the visualization techniques and classification algorithms that can be used for predicting the crimes and helps the law agencies. In future, there is a plan for applying other classification algorithms on the crime data and improving the accuracy in prediction. |

# PRODUCT OR SYSTEM FEATURES AND REQUIREMENTS

|  |
| --- |
| A desktop/laptop, programming knowledge with python and familiarity with machine learning concepts and data visualizations. In this project we used python-version 3.6 and the platform used is google collaboration. Google Colab allows to write and run python code interactively the same way as Jupyter Notebook does. |

# PROJECT SUCCESS CRITERIA

|  |
| --- |
| Our project aim is to predict the crime rate of a particular region.So, we thought of building the machine learning model and with minimum accuracy of 85% and data visualizations so that, it could give accurate predictions.  At first we have achieved accuracy of 85% and later on we have improved the accuracy of the model to 91% using few optimization techniques. |

# STATEMENT OF WORK

## SCOPE OF WORK

|  |
| --- |
| The main scope of the project is to predict the crime rate in a particular region based on the historic data and visualize it graphically using statistical tools so that , it is easy to look and understand the data to support public safety,financial success and better outcomes.  As per problem statement there are the methods to be followed:  Data collection,Data cleaning, Data handling, Predictive modeling ,Model selection, Prediction, Visualization, Conclusion. |

## STANDARDS FOLLOWED

|  |
| --- |
| To predict the crime of a particular region some of the standards should be followed.In this model the standards we followed are mentioned as below:   * Collecting the dataset from kaggle and exploring the dataset as per requirement. In this dataset 1994 areas are present. Size of the dataset is 1994 rows and 128 columns. The datasetfeatures are: the occurrence month, the occurrence day of the week, the occurrence time and the crime location. * Importing the required modules into the jupyter notebook and implementing data cleaning, analysing correlation(relationship between variables using heatmap)on the data set are done. * After that, we have used machine learning algorithms like regression and decision tree classifier to train the dataset and build machine learning model to get good accuracy. * The accuracy rate is 85% with Decision tree classifier and 0.65 R2 score with Linear Regression.To increase the previous accuracy rate tensor-flow GPU is used and Decision tree classifier algorithm is optimized.   After following these standards the accuracy increased to 91%. |

## LOCATION OF WORK

|  |
| --- |
| We completed this project in online mode. |

## PERIOD OF PERFORMANCE

|  |
| --- |
| The project is accomplished in the span of 14 days .It includes four phases:  Planning, Execution, Monitoring, Closing.  First phase of the project i.e., planning is done in three days.  Second phase of the project i.e.,execution is finished in five days.  Third phase i.e., monitoring is done in the period of five days and fourth phase i.e.,closing done with one day. |

## SCHEDULE

|  |  |  |  |
| --- | --- | --- | --- |
| **CATEGORY/TASK** | **WORK DONE** | **START DATE** | **END DATE** |
| Phase 1 - Planning |  |  |  |
| **Task A**-Planning for kind of data set required and planning for documentation. | Planned for the content required for documentation and kind of dataset required. | 01-07-2020 | 02-07-2020 |
| **Task B**-Planning for way of approach like required python modules and required machine learning algorithms. | Planned for way of approach. | 02-07-2020 | 03-07-2020 |
| Phase 2 - Execution |  |  |  |
| **Task A**-Collecting the dataset from internet and exploring the dataset as per requirement. | Imported libraries, loaded dataset,Data cleaning, standardization. | 04-07-2020 | 06-07-2020 |
| **Task B**-Importing the required modules into the jupyter notebook and implementing data cleaning, Analysing correlation(relationship between variables using heatmap),Applying ML algorithms like Linear Regression and Decision tree classification. | Finally after implementing all the ML algorithms the accuracy rate is 85% with Decision tree classification and 0.65 R2 score with Linear Regression. | 06-07-2020 | 09-07-2020 |
| Phase 3 - Monitoring |  |  |  |
| **Task A**-Making changes in the documentation as per the modifications in code. | Documentation is prepared. | 09-07-2020 | 11-07-2020 |
| **Task B-**Making progress to increase the previous accuracy rate like using tensor-flow GPU and optimizing Decision tree classifier algorithm. | After the following procedure the accuracy increases to 91%. | 11-07-2020 | 13-07-2020 |
| Phase 4 - Closing |  |  |  |
| **Task A- ------** | ------ |  |  |
| **Task B- ------** | We are predicting crime at this accuracy rate is 91% using decision tree classifier. | 13-07-2020 | 14-07-2020 |

## ACCEPTANCE CRITERIA FOR USERS

|  |
| --- |
| Implementing data cleaning, analysing correlation the accuracy rate is reached to 85% with Decision tree classifier and 0.65 R2 score with Linear Regression.To increase the previous accuracy rate tensor-flow GPU is used and Decision tree classifier algorithm is optimized. This made performance high in our project i.e.,91% accuracy. |

## ADDITIONAL REQUIREMENTS

|  |
| --- |
|  |

**CONCLUSION**

With the help of machine learning technology, it has become easy to find out relation and patterns among various data’s. The work in this project mainly revolves around predicting the crime of region where it has occurred. Using the concept of machine learning we have built a model using data set that have undergone data cleaning. The model predicts the crime of a particular area with accuracy of 91%. Data visualization helps in analysis of data set.

**REFERENCES**

1. Yadav, S., Timbadia, M., Yadav, A., Vishwakarma, R., & Yadav, N. (2017, April).Crime pattern detection, analysis & prediction. In Electronics,Communication and Aerospace Technology (ICECA), 2017 International conference of (Vol. 1, pp. 225- 230). IEEE.

2. Shamsuddin, N. H. M., Ali, N. A., & Alwee, R. (2017, May). An overview on crime prediction methods. In Student Project Conference (ICT-ISPC), 2017 6th ICT International (pp. 1-5) . IEEE.

3. Sivaranjani, S., Sivakumari, S., & Aasha, M. (2016, October). Crime prediction and forecasting in Tamilnadu using clustering approaches. In Emerging Technological Trends (ICETT) International Conference on (pp. 1-6). IEEE.

4. Sathyadevan, S., & Gangadharan, S. (2014, August). Crime analysis and prediction using data mining. In Networks & Soft Computing (ICNSC), 2014 First International Conference on (pp. 406-412). IEEE.

5. Nath, S. V. (2006, December). Crime pattern detection using data mining. In Web intelligence and intelligent agent technology workshops, 2006. wi-iat 2006 workshops. 2006 ieee/wic/acm international conference on (pp. 41-44). IEEE.

6. Zhao, X., & Tang, J. (2017, November). Exploring Transfer Learning for Crime Prediction. In Data Mining Workshops (ICDMW), 2017 IEEE International Conference on (pp. 1158-1159). IEEE.