Predicting The Crimes Based On Weather Using PYSPARK

Jamila. S [1], Ramya. K [1], Sowmiya. M [1], Mrs. M. Sindhuja [2]

[1],(Ug Student, Final Year Cse Department, Dhanalakshmisrinivasan College Of Engineering And Technology)
[2](Assistan Professor, Department Of Computer Science And Engineering
Dhanalakshmi Srinivasan College Of Engineering And Technology

Abstract: *Crime analysis is one of the most important activities of the majority of the intelligent and law enforcement organizations all over the world. Generally they collect domestic and foreign crime related data (intelligence) to prevent future attacks and utilize a limited number of law enforcement resources in an optimum manner. A major challenge faced by most of the law enforcement and intelligence organizations is efficiently and accurately analyzing the growing volumes of crime related data. The vast geographical diversity and the complexity of crime patterns have made the analyzing and recording of crime data more difficult. Data mining is a powerful tool that can be used effectively for analyzing large databases and deriving important analytical results. This paper presents an intelligent crime analysis system which is designed to overcome the above mentioned problems. The proposed system is here is we proposed location analysis along with the crime happened and we proposed Pyspark here to store large amount of data's for crime analysis. The proposed system consists of a rich and simplified environment that can be used effectively for processes of crime analysis.*

I.   Introduction

Crime analysis is one of the most important activities of the majority of the intelligent and law enforcement organizations all over the world. Generally they collect domestic and foreign crime related data (intelligence) to prevent future attacks and utilize a limited number of law enforcement resources in an optimum manner. A major challenge faced by most of the law enforcement and intelligence organizations is efficiently and accurately analyzing the growing volumes of crime related data. The vast geographical diversity and the complexity of crime patterns have made the analyzing and recording of crime data more difficult. Data mining is a powerful tool that can be used effectively for analyzing large databases and deriving important analytical results. This paper presents an intelligent crime analysis system which is designed to overcome the above mentioned problems. The proposed system is here is we proposed location analysis along with the crime happened and we proposed Pyspark here to store large amount of data's for crime analysis. The proposed system consists of a rich and simplified environment that can be used effectively for processes of crime analysis.

There are several significant reasons for crime analysis such as to identify general and specific crime trends, patterns, and series in an ongoing, timely manner, to maximize the usage of limited law enforcement resources, to access crime problems locally, regionally, nationally within and between law enforcement agencies, to be proactive in detecting and preventing crimes and to meet the law enforcement needs of the changing society. There are various crime data mining techniques available such as clustering techniques, association rule mining, sequential pattern mining, and classification and string comparison.

Several web based crime mapping systems are available on the Internet such as narcotics network in Tucson police department, but majority of them have been custom made for legislative authorities in different countries and those systems are not accessible to parties outside that particular law enforcement or legislative authorities

PROPOSED SYSTEM:

The proposed system consists of a rich and simplified environment that can be used effectively for processes of crime analysis. We proposed pyspark for classification and datasets are stored into sparksql storing and retrieval of information will be faster We proposed weather data analysis along with crime data to analysis and predict which weather takes more crime action and represented crime analysis with graphical representation.

We predicted the arrest based on year, month, week and day and top 5 crimes in 2012, 2013, 2014, 2015, 2016 in month, week and day wise and we predicted weather analysis which weather situation had more number of crimes

ADVANTAGE:

1.   Crime has been analysed with weather aspects to make crime prediction more efficient
2.   PVS park with jupyter console is used to analysis crime prediction

SYSTEM STUDY
FEASIBILITY STUDY:
The feasibility of the project is analyzed in this phase and business proposal is put forth with a very general plan for the project and some cost estimates. During system analysis the feasibility study of the proposed system is to be carried out. This is to ensure that the proposed system is not a burden to the company. For feasibility analysis, some understanding of the major requirements for the system is essential.
Three key considerations involved in the feasibility analysis are

♦ ECONOMICAL FEASIBILITY
♦ TECHNICAL FEASIBILITY
♦ SOCIAL FEASIBILITY

ECONOMICAL FEASIBILITY:
This study is carried out to check the economic impact that the system will have on the organization. The amount of fund that the company can pour into the research and development of the system is limited. The expenditures must be justified. Thus the developed system as well within the budget and this was achieved because most of the technologies used are freely available. Only the customized products had to be purchased.

TECHNICAL FEASIBILITY:
This study is carried out to check the technical feasibility, that is, the technical requirements of the system. Any system developed must not have a high demand on the available technical resources. This will lead to high demands on the available technical resources. This will lead to high demands being placed on the client. The developed system must have a modest requirement, as only minimal or null changes are required for implementing this system.

SOCIAL FEASIBILITY:
The aspect of study is to check the level of acceptance of the system by the user. This includes the process of training the user to use the system efficiently. The user must not feel threatened by the system, instead must accept it as a necessity. The level of acceptance by the users solely depends on the methods that are employed to educate the user about the system and to make him familiar with it. His level of confidence must be raised so that he is also able to make some constructive criticism, which is welcomed, as he is the final user of the system.

INPUT DESIGN AND OUTPUT DESIGN
INPUT DESIGN:
The input design is the link between the information system and the user. It comprises the developing specification and procedures for data preparation and those steps are necessary to put transaction data in to a usable form for processing can be achieved by inspecting the computer to read data from a written or printed document or it can occur by having people keying the data directly into the system. The design of input focuses on controlling the amount of input required, controlling the errors, avoiding delay, avoiding extra steps and keeping the process simple. The input is designed in such a way so that it provides security and ease of use with retaining the privacy. Input Design considered the following things:

➢ What data should be given as input?
➢ How the data should be arranged or coded?
➢ The dialog to guide the operating personnel in providing input.
➢ Methods for preparing input validations and steps to follow when error occur.

OBJECTIVES:
1. Input Design is the process of converting a user-oriented description of the input into a computer-based system. This design is important to avoid errors in the data input process and show the correct direction to the management for getting correct information from the computerized system.
2. It is achieved by creating user-friendly screens for the data entry to handle large volume of data. The goal of designing input is to make data entry easier and to be free from errors. The data entry screen is designed in such a way that all the data manipulates can be performed. It also provides record viewing facilities.
3. When the data is entered it will check for its validity. Data can be entered with the help of screens. Appropriate messages are provided as when needed so that the user will not be in maize of instant. Thus the objective of input design is to create an input layout that is easy to follow.

OUTPUT DESIGN:

A quality output is one, which meets the requirements of the end user and presents the information clearly. In any system results of processing are communicated to the users and to other system through outputs. In output design it is determined how the information is to be displaced for immediate need and also the hard copy output. It is the most important and direct source information to the user. Efficient and intelligent output design improves the system's relationship to help user decision-making.

1. Designing computer output should proceed in an organized, well thought out manner; the right output must be developed while ensuring that each output element is designed so that people will find the system can use easily and effectively. When analysis design computer output, they should Identify the specific output that is needed to meet the requirements.
2. Select methods for presenting information.
4. Create document, report, or other formats that contain information produced by the system.

The output form of an information system should accomplish one or more of the following objectives.
❖ Convey information about past activities, current status or projections of the
❖ Future.
❖ Signal important events, opportunities, problems, or warnings.
❖ Trigger an action.
❖ Confirm an action.

ER DIAGRAM:
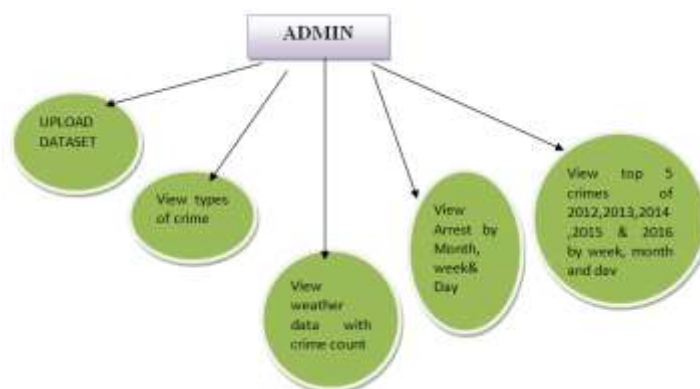
II. Conclusion

The project's objective to analyze the crime data and provide fruitful suggestions to the department of security to protect the precincts and beats where the crime rates are high. Here, we used the different regression, clustering, classification and frequent growth patterns in order to organize and arrange the data in a ordered fashion. Thus I could ascertain that the research if applied on the appropriate data would result in generating patterns that would help to identify the crime rate early.