Educational Administration: Theory and Practice
2024,30(5), 1402-1412
ISSN:2148-2403
https://kuey.net/

Research Article

Data Isolating Pattern Framework For Prediction Crime Sector Using Wide-Open Power System Convention Phase Algorithm

S. Imthiyas Ahamed[1*], Dr. J. Chockalingam[2], Dr. A. Shaik Abdul Khadir[3]

[1*]Research Scholar, PG & Research Department of computer science, Khadir Mohideen College, Adirampattinam, (Affiliated to Bharathidasan University, Trichy-24), Thanjavur, Tamilnadu, India.
[2]Research Supervisor, Associate professor of computer science (Retd), Khadir Mohideen College, Adirampattinam, (Affiliated to Bharathidasan University, Trichy-24), Thanjavur, Tamilnadu, India
[3]Research Co-Supervisor, Head & Associate professor of computer science, Khadir Mohideen College, Adirampattinam, (Affiliated to Bharathidasan University, Trichy-24), Thanjavur, Tamilnadu, India

Citation: S. Imthiyas Ahamed et.al (2024), Data Isolating Pattern Framework For Prediction Crime Sector Using Wide-Open Power System Convention Phase Algorithm, *Educational Administration: Theory And Practice*, *30*(5), 1402-1412
Doi: 10.53555/kuey.v30i5.3100

ARTICLE INFO

ABSTRACT

Data Mining (DM) is the process of exploring datasets in order to illustrate conclusions about the information they contain. It is a process of inspecting, cleansing, transforming, and modeling data with the goal of discovering useful information. A database contains both structured and semi-structured data. The semi organized data are from various sources. The framework is furnished with an open power framework show stage for breaking down the data from entire data holder. As per this idea, the examination is a lot of outright as opposed to other, data mining method. The fundamental target of the proposed study is to give data having better and huge viewpoint.

1. INTRODUCTION

Data Mining has become one of the most discussed topics among researchers. The data are stored in cloud storage, and that data is used to refer the enormous amount of datasets. Compared to the traditional datasets, cloud storage datasets comprised a set of unstructured data which requires more significant realtime investigation [1].The concept of cloud storage data helps us to understand the abstraction and in depth understanding behind various hidden values. Data Mining is a technology enabled strategy for gaining richer, deeper, and more accurate insights into the details of the customers, partners and the businesses and hence ultimately gaining competitive advantage. By processing a steady stream of real-time data, organizations can make time sensitive decisions faster than ever before, monitor emerging trends, course correct rapidly and jump on new business opportunities [2].

1.1. STRUCTURED DATA

Structured data is the data included in relational database system. Being structured and highly organized, it can be managed by SQL and its multiple variations developed by IBM, ADO.net, ODBC in RDBMS environments, Due to explicit semantics and structure, efficient search is possible for focused content by simple and straight forward search engines[3].

1.2. SEMI STRUCTURED DATA

Semi structured information is one kind of organized information which is not having the information display structure, Further it will accommodate a formal or unbending structure. This semi structured information does not require a composition definition as it is fairly discretionary and contains labels or different markers to isolate semantic components and authorize pecking orders of records fields inside the information. To change over the semi structured information to structured information, customary information mining systems or normal preparing dialect is relevant [4].

1.3. DATA ANALYSING

It is seen that these days, different associations are consuming huge measure of fundamental data which can be valuable in different fields, for example, checking of an articles action, sensor sending, following of data and so on. An extreme flood related with data is named as distributed storage Data, which respects what is

happening on the current foundation of Data Stockpiling the executives and the Measurable assessment of data, Comparable to circumstances emerges when an association needs to investigate its data from its own sites for dissecting the client's criticisms, redid administrations towards an item [5].Accordingly, the choices producers would convey their decisions grounded on the examination of extricated data or those data which convey some worth or weight age [6]. Data Mining is likewise planned between inconsequential traits of datasets which can be gotten from AI, database frameworks and insights. In various logical fields, network had been created with capacity, handling, and accessibility of data. A matrix is an assortment of disseminated registering assets accessible over a neighborhood or wide region network that appears to an end client or application as one enormous virtual figuring framework. It very well may be seen that a few explicit innovations and execution are expected for the cloud for giving infra-construction, stage and programming sources [7].

The assortment of advanced data as far as organized and non-organized data known as distributed storage data is quickly creating. Distributed storage data is most certainly a peculiarity with direct effect on personal satisfaction. Utilizations of data can be tracked down in versatile cloud PC frameworks, like divisions of procurement exchanges informal organizations, instructor critique, e-science and medical care frameworks. Dissecting data to sum up them and search for designs is a significant piece of each and every assessment. Systems for the investigation of the data and how the data will be orchestrated ought to be chosen at the assessment configuration stage.

By handling a constant flow of continuous data, associations can pursue time-delicate choices quicker than any time in recent memory, screen arising patterns, course right quickly and without hesitation take advantage of new business chances [8]. Data mining analyzing requirements are some of the methods that can be used to find meaning and to discover unseen relationships in Data mining. There are three major significant requirements:

- Minimizing data movement is basically all about conversing computing resources. In traditional analysis scenarios, data is brought into the computer system, processed and then sent to next destination. For example, Flipkart sales data might be extracted from e-system, transformed into a relational data type, and loaded into an operational data store structured for generating report. Whenever there is an increase in volume of data, this type of ETL (Extract, Transform and Load) architecture becomes increasingly less organized. It makes more intelligence to store and process the data in the same place.
- Using existing skills means new data or new data sources demands the need to obtain new skills. Most of the time, existing skill sets will determine where analysis can and should be done. Mostly organizations have more and more people who can analyze data using either SQL or Map Reduce, but it is very important to be able to support both type of processing.
- Attending data security is essential part for many business applications. Basically data warehouse users are habituated not only with judgment defined metrics, dimensions and attributes, but also to a reliable set of management policies and security controls. These rigorous processes are often poor with unstructured data source and open source analysis tools.

The way toward changing over a lot of unstructured crude information, recovered from various sources to an information item valuable for associations shapes the centre of Data mining analytics. Once the data is collected, we normally have diverse data sources with different characteristics. The most immediate step would be to make these data sources homogeneous and continue to develop our data product. However, it depends on the type of data. A champion among the most basic endeavours in tremendous data examination is quantifiable showing, which implies oversaw and unsupervised portrayal or backslide issues. Once the data is cleaned and pre-prepared, available for showing, mind should be taken in surveying differing models with sensible disaster estimations and after that once the model is executed; help evaluation and results should be represented. A normal snare in judicious showing is to just complete the model and never measure its execution.

***Fig.1: Life Cycle of Data Mining* Motivation**

Privacy protection of high profile data is the challenging issue as the information can be easily hacked by unauthorized users. Data security is the major consideration of this proposed embedded security model along with the integrated data analytics model. Following measures are taken into account while developing the framework for the benefit of user.

- Providing privacy protection for high profile data
- Ensuring the data security through data analytic integrated framework ➢ Improving authentication level to protect from unauthorized users.
- Enabling security and privacy in data analytic framework.

The Research Objectives are:

- **Confidentiality (Data privacy):** Main objective of this research work is maintaining confidentiality. The data set and result analysis of data framework and the messages are concealed by encoding then. Here the encryption is by unique key formed using USB Key. When the information is needed the same can be decrypted by providing the same unique key. Hence data confidentiality is ensured.
- **Data Integrity:** Maintaining originality of the information is another important factor that is taken into account when developing the framework. The proposed work protects the information by encrypting the database with the secured key. The same database can be retrieved only when the unique key is provided. Thus the originality of the information stored is maintained.
- **Authentication:** Protecting the information from the hackers is a major concern and hence the framework is protected with the authentication using unique key. This unique key can't be altered manually as this will be generated from the firmware key number of the USB device used, password programmed and keys generated. This key is used both in encryption and decryption process.

1.5. DEPENDENCY-ORIENTED DATA

In these cases, implicit or explicit relationships may exist between data items. For example, dataset contains a set of vertices (data items) that are connected together by a set of edges (relationships). On the other hand, time series contains implicit dependencies. For example, two successive values collected from a sensor are likely to be related to one another. Therefore, the time attribute implicitly specifies a dependency between successive readings. In general, dependency-oriented data are more challenging because of the complexities created by pre-existing relationships between data items. Such dependencies between data items need to be incorporated directly into the analytical process to obtain contextually meaningful results.

In practice, the different data values may be implicitly related to each other temporally, spatially, or through explicit network relationship links between the data items. The knowledge about pre existing dependencies greatly changes the data mining process because data mining is all about finding relationships between data items. The presence of pre existing dependencies therefore changes the expected relationships in the data, and

what may be considered interesting from the perspective of these expected relationships. Several types of dependencies may exist that may be either implicit or explicit:

- **Implicit dependencies**: In this case, the dependencies between data items are not explicitly specified but are known to "typically" exist in that domain. For example, consecutive temperature values collected by a sensor are likely to be extremely similar to one another. Therefore, if the temperature value recorded by a sensor at a particular time is significantly different from that recorded at the next time instant then this is extremely unusual and may be interesting for the data mining process. This is different from multidimensional data sets where each data record is treated as an independent entity.
- **Explicit dependencies**: This typically refers to graph or network data in which edges are used to specify explicit relationships. Graphs are a very powerful abstraction that is often used as an intermediate representation to solve data mining problems in the context of other data types.

2. RELATED WORK

Chris Clifton, (2020) proposed To continue this proposed work, some survey is made on the organization of semi structured data which is recognized as one of the major uncertain problems in the information industry and data mining paradigm [16]. New data processing systems make the computing grid work by managing and pushing the data out to individual nodes, sending instructions to the networked servers to work in parallel, collecting individual results and then reassembling them to produce meaningful results. Processing the data where it resides is faster and more efficient then before analyzing transporting it to a centralized system. R.Buse, (2020) noted It will be in the form of computerized information that does not have a data model and hence are not used by data mining [15]. [14]

LaValle S, (2020) proposed the task of managing semi-structured data signifies possibly the major data management opportunity which subsequently results in managing relational data. Semi-structured data constitutes about 70% of the data collected or stored in larger organizations which are difficult to access, use or retrieve. Most existing tools generally deal with a single text corpus, or individually handle different corpus. [13] Discovered these tools may not give a full picture of ongoing events on social media. Topic Panorama was recently proposed to allow researchers to simultaneously analyze and correlate the topics of different corpora simultaneously.

Barbierato E, (2020) noted Topic Panorama is highly interactive and assists users in interacting with matched topic graphs at different granularity levels. However, Topic Panorama only handles small-scale graphs and simultaneously visualizes several corporate.

Millard (2019) proposed cloud storage Data allow researchers to decode human DNA in minutes, predict where terrorists plan to attack, determine which gene is mostly likely to be responsible for certain diseases and, of course, which ads you are most likely to respond to on Facebook. The business cases for leveraging cloud storage Data are compelling.

N. Diakopoulos,(2019)discovered The instance, Netflix mined its subscriber data to put the essential ingredients together for its recent hit House of Cards, and subscriber data also prompted the company to bring Arrested Development back from the dead.

O. T. Menzies, (2019) discovered Hashing technique was applied to evaluate the Mining process. This method satisfies volume and velocity, where by better performance was observed for 7285documents. For evaluation, these documents were splinting to 72% of training data 28% test data, respectively. The actual sizes of these two datasets were 7285 and 18,846. This method is feasible for different sorts of data like optical characters, speech audio, and document scripts. Sparse hashing technique is not suit-able for velocity, since it considers fixed data size. Moreover, it does not provide anything related to data accuracy [9].

3. PROBLEM IDENTIFICATION

A large portion of the current framework gives just investigating the data which is unstructured and there is no question intricacy. The unstructured/semi organized data are overall persistently comes from different sources like satellite pictures, sensor readings, email messages, virtual entertainment, web logs, review results, sound, recordings etc. Hence our proposed open matrix rule creating calculation gives the data into outright sifted data. Prior to sifting, the data are planned and decreased to a level, which is more proficient in giving rough required data [10].

3.1. PROPOSED METHODOLOGY

The data analysis concept of cloud storage Data gives analytical methods which can be applied to analyze traditional datasets which includes analytical architecture and software requirement for exploration of cloud storage data [11]. Open power system convention phase is one of the most essential stages of the cloud storage data value chain where the main objective is to extract the meaningful information and providing suggestions and decisions. Different types of possible and gravitational values can be produced through the several stages of analysis in different fields [12].

3.1.1. USER INTERACTION ISSUES

There can be user interaction issues as follows:

- **Mining different kinds of knowledge in databases**: Different users may be interested in different kinds of knowledge. Therefore it is necessary for data mining to cover a broad range of knowledge discovery task.
- **Interactive mining of knowledge at multiple levels of abstraction**: The data mining process needs to be interactive because it allows users to focus the search for patterns, providing and refining data mining requests based on the returned results.
- **Incorporation of background knowledge**: To guide discovery process and to express the discovered patterns, the background knowledge can be used. Background knowledge may be used to express the discovered patterns not only in concise terms but at multiple levels of abstraction.
- **Data mining query languages and ad hoc data mining**: Data mining query language that allows the user to describe ad hoc mining tasks should be integrated with a data warehouse query language and optimized for efficient and flexible data mining.
- **Presentation and visualization of data mining results:** Once the patterns are discovered it needs to be expressed in high level languages, and visual representations. These representations should be easily understandable.
- **Handling noisy or incomplete data**: The data cleaning methods are required to handle the noise and incomplete objects while mining the data regularities. If the data cleaning methods are not there then the accuracy of the discovered patterns will be poor.
- **Pattern evaluation**: The patterns discovered should be interesting because either they represent common knowledge or lack novelty.

3.1.2. PERFORMANCE ISSUES

There can be performance-related issues as follows:

- **Efficiency and scalability of data mining algorithms**: In order to effectively extract the information from huge amount of data in databases, data mining algorithm must be efficient and scalable.
- **Parallel, distributed, and incremental mining algorithms:** The factors such as huge size of databases, wide distribution of data, and complexity of data mining methods motivate the development of parallel and distributed data mining algorithms. These algorithms divide the data into partitions which is further processed in a parallel fashion. Then the result from the partitions is merged. The incremental algorithms, update databases without mining the data again from scratch.

3.1.4. OPEN POWER SYSTEM CONVENTION PHASE

Data Mining is typically performed using specialized soft-ware tools and applications for predictive Mining, data mining, text mining, and forecasting and data optimization. Collectively these processes are handled with open grid rule generating which shows highly integrated functions of high performance Mining. This enables an organization to process extremely large volumes of data that a business has collected to determine which data is relevant and can be analyzed to drive better business decisions.

- Creating the data set with a new template for importing the data from the database which contains both the structured and semi structured data. When the data are imported it shows in the raw form, what is to be used, within the new data set template.datx (data analytic tool extension).
- The data are set into its framework after importing for the purpose of mapping, there is a raw data loaded in the framework. Hence, it is taken for the purpose of reducing the attributes.
- The raw data are reduced for the process of mapping that is assigning the attributes list into the absolute form. Since the mapped data are processed in the database.
- Data filtered based on threshold value i.e support and confidence level to reduce the redundancy and duplication of data. The history of data and log files can be retrieved.

A flexible and reconfigured grid along with the cloud storage data preprocessing enhancement, mapping and reducing, data filtering, and data parallelization schemes can be more effective approaches for extracting more meaningful knowledge from the given datasets.

Wide-Open Power Convention is one of the most essential stages of the big data value chain where the main objective is to extract the meaningful information and providing suggestions and decisions. Different types of possible and gravitational values can be produced through the several stages of analysis in different fields. The system is provided with rule generation algorithm for analyzing the data from whole data container. According to this concept, the analysis is much absolute than the other data mining techniques. The main objective of the proposed study is to provide a better and significant perspective data and an overview of data filtering.

The following figure 2 shows the process of data Mining with the open power system convention phase. Since the process is more effective, itshows the data to a higher level with high efficiency [17]. All the raw data are cleared and can be retrieved whenever the user needs.

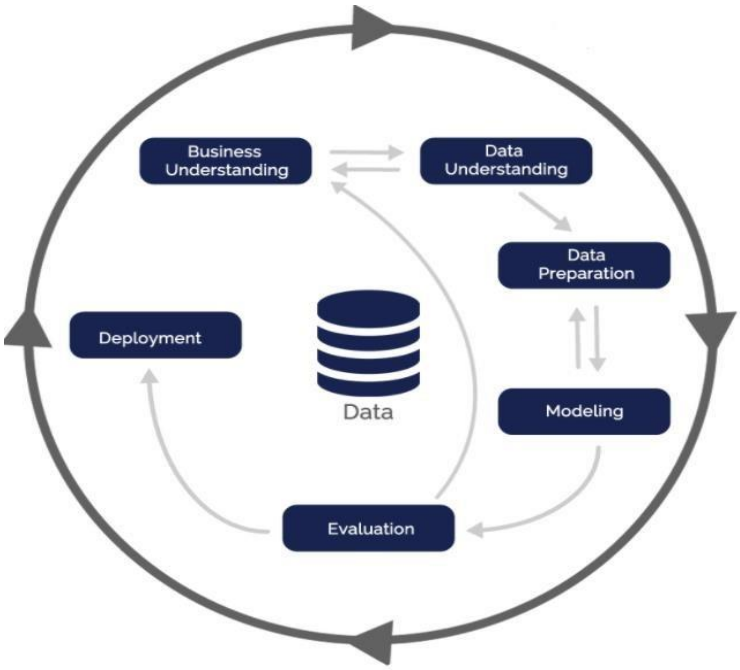


Fig.2: Mining Processing Framework.

ALGORITHM

ArraylistLk←NewArraylistRegexR ← NewRegex
ForI=0 → L.Count
String [] Subl1 ← R.Split(L[I].Tostring.ForJ=I+1→L.Count String[]Subl2←R.Split (L[J].Tostring.
// Comparing Two ItemsStringTemp→L[J].Tostring
//Store The Two Key Sets.ForM=0→SubL1.Length
Boolean Subl1mlnsubl2 = FalseForN=0 →Sub L2.Length
If (Subl1 [M] ← Subl2 [N])Subl2mln Subl2=True
If Subl1mlnsubl2 == FalseTemp=Temp+",”+Subl1 [M]String[] Subtemp→R.Split(Temp)
If (Subtemp.Length ← Subl1.Length + 1BoolsIs Exists = False
For M=0 → Lk.CountBoolIsContained =True
ForN=0→Subtemp.Length
If(!Lk[M].Tostring().Contains(Subtemp[N]))IsContained=False.
If(Iscontained==True)IsExists=TrueIf(Exists ==False) LkAdd(Temp)Return Lk.

Efficiency calculation

Preliminariesmi–initialtime.
md – destination timeat–averagetime pt–approximatetimemt– meantime
OPSC Algorithm process executionMeantimemt= mi+md/at at←avgtime at←mt [mi+md]
// avg time for the attributept← at[mt/2] [mt→mi+md/at]at[mi+md/at]/2 mt←mi+md/at
// here, the mean time value is identified.Pt ← at[mt/2]
Ptnotequalto mt
//buttheapproximate valueissimilarwithattributemeantime. pt→at[mi+md]
//approximate valueforthosecrimedataPt isunequal tomt

Every attribute has the specific mean time as the user specifies. The initial time for each attribute may vary due to its possibility, such in case the mean time for attribute changes, but it has only some difference with the average time provided for those attributes. In such a case, the approximate time for each attribute is valued by their mean time and its initial time [18].The rough estimation of the investigation procedure all through the essential sort and its weight. By the Process, the data put in the database and the crude data are cleared from the data compartment and it demonstrates the over seed information delivered for diminishing the rough data from the unstructured shape and makes it into the sorted out data with the capable course of action with the way toward sifting [19,10,12,22]. Data Mining is a technology-enabled strategy for gaining richer, deeper, and more accurate insights into customers, partners and the business and ultimately gaining competitive advantage. In the present procedure, the data are inspected through a such essential arrangement and the separated data are needy upon 55%.Here the examination over through 95% induced with high compelling and sifted information are more estimated.

- The first step is creating the data set with a new template for importing the data from the database which contains both the structured and semi-structured data. When the data are imported it shows in the raw form, what is to be used, within the new data set template .datx (data analytic tool extension).
- The data are set into its framework after importing the raw data for the purpose of mapping, there is a raw data loaded in the framework. It is taken for the purpose of reducing the attributes.
- The raw data are reduced for the process of mapping that is, assigning the attributes list into the absolute form, since the mapped data are processed in the database.
- Data filtered based on threshold value that is support and confidence level to reduce the redundancy and duplication of data. The history of data and log files can be retrieved.

4. RESULTS AND DISCUSSIONS

By the open power system convention phase process, the information in the database are broken down up to the superior level. The crude information are cleared from the information compartment and the manage is produced for decreasing the crude information from the unstructured shape and makes it into the organized information with the proficient arrangement with the rule of mapping. In the current strategy, the information is examined through such basic configuration and the broke down information are dependent upon 60%. Here the examination over through 95%surmised with high efficiency.

4.1. DATA TRANSFER

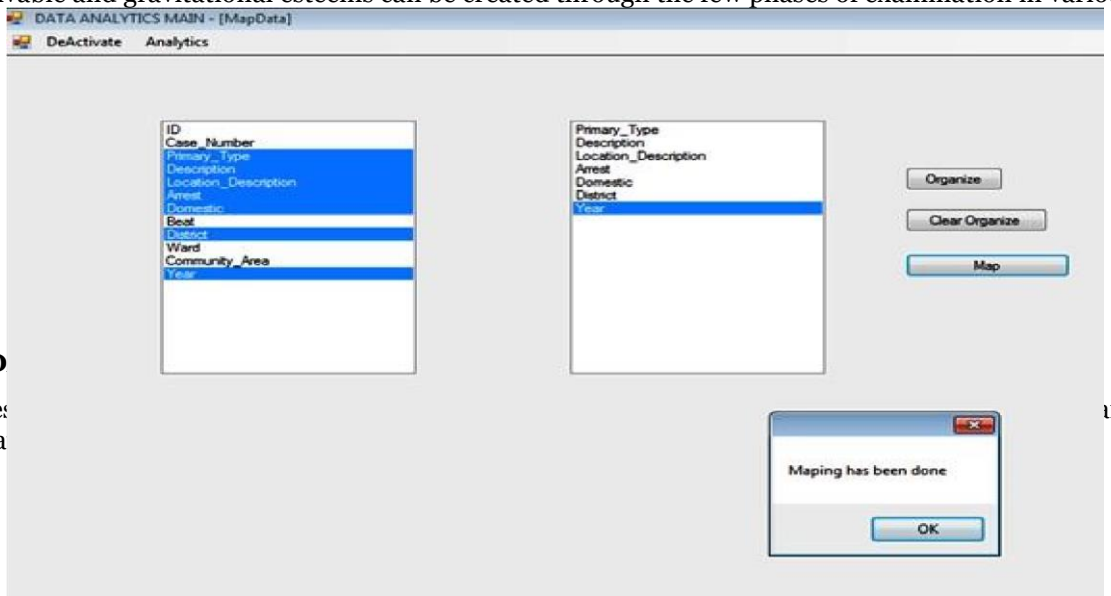
Transmission of the sensor data through unsecured frameworks should be guaranteed. Characterization and respectability should be ensured for any data trade. Mystery is securing tricky data against a noxious customer and genuineness is shielding the trustworthiness of the data. Cryptography or VPN strategies are a part of the normally used philosophies for securely trading data.

4.2. DATA STORAGE AND PROCESSING

Data set away with before long identifiable information (or identifiers) in an external group is a honest to goodness peril to data security. Individual and semi identifiers delineate before long identifiable information. These characteristics can direct or in-particularly reveal singular information. Securing the data used for examination/mining as determined above can achieve this.

4.3. DATA ACCESS

Access to the system should be ensured through suitable confirmation and endorsement. The system should be configurable to dole out rights to execute examination/mining work to legitimate customers and access the created comes to fruition. Among various procedures, the role based access control (RBAC) has been extensively recognized in light of its straightforwardness, versatility in getting dynamic essentials and support for the run of least advantage and profitable advantage organization. The information examination idea of Data mining gives logical techniques which can be connected to break down customary datasets which incorporates systematic engineering and programming necessity for investigation of enormous information. Open network lead age is a standout amongst the most basic phases of the enormous information esteem chain where the primary goal is to remove the significant data and giving recommendations and choices. Diverse sorts of conceivable and gravitational esteems can be created through the few phases of examination in various fields.



4.5. D

The re ıring the raw da

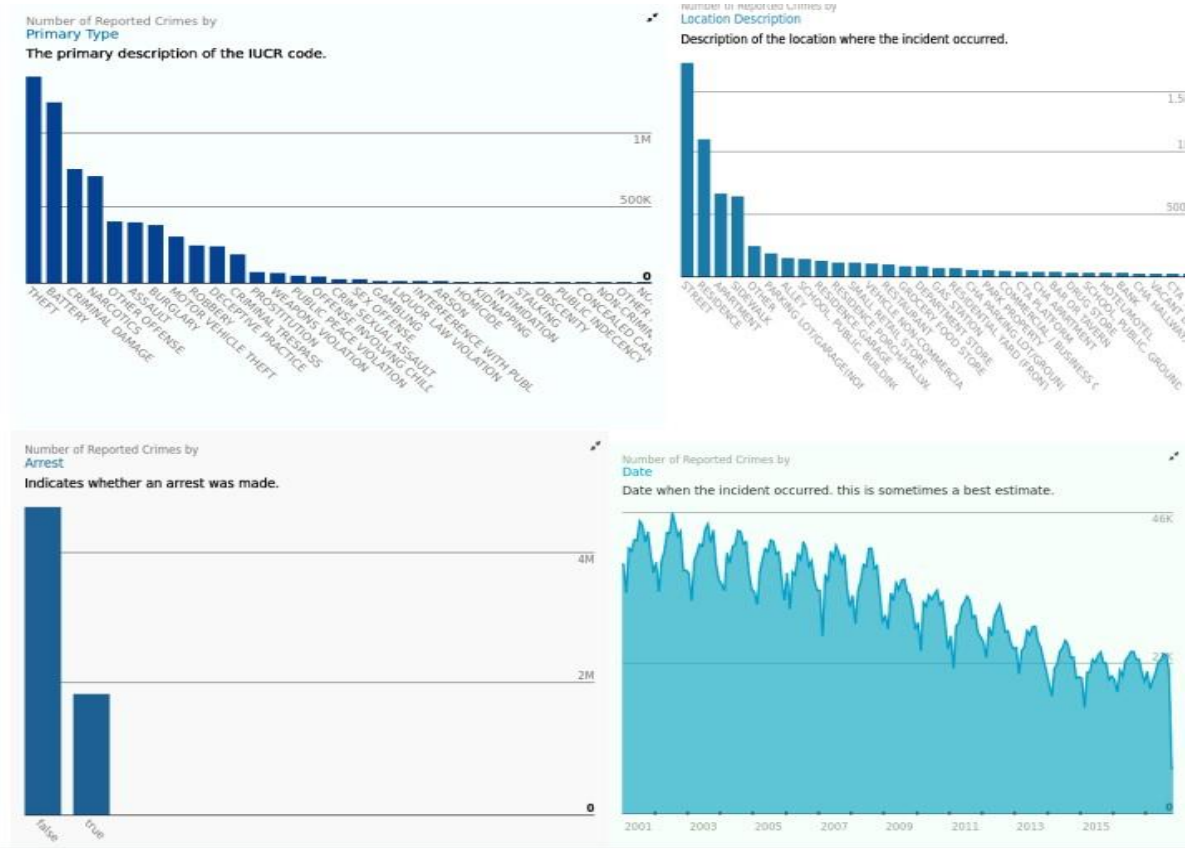


Fig.3.Visualization of Crime Dataset

4.4. MAPPING DATA

The mapping is the process of assigning the data which is reduced in a level which can be retrieved from it whenever the user needs. The following figure shows the process of mapping the dataset

Table1:LegendsKDDProcess:KnowledgeDiscoveryDatabases,HDFS

S.N O	Primary type	Year	KDD process	HDFS Process	OPSC
1	Drug store	2011	60%	55%	97%
2	Narcotics'	2012	67%	62%	99%
3	Theft	2016	55%	66%	92%
4	Parking slot	2016	64%	45%	98%
5	Railway	2017	56%	56%	100%

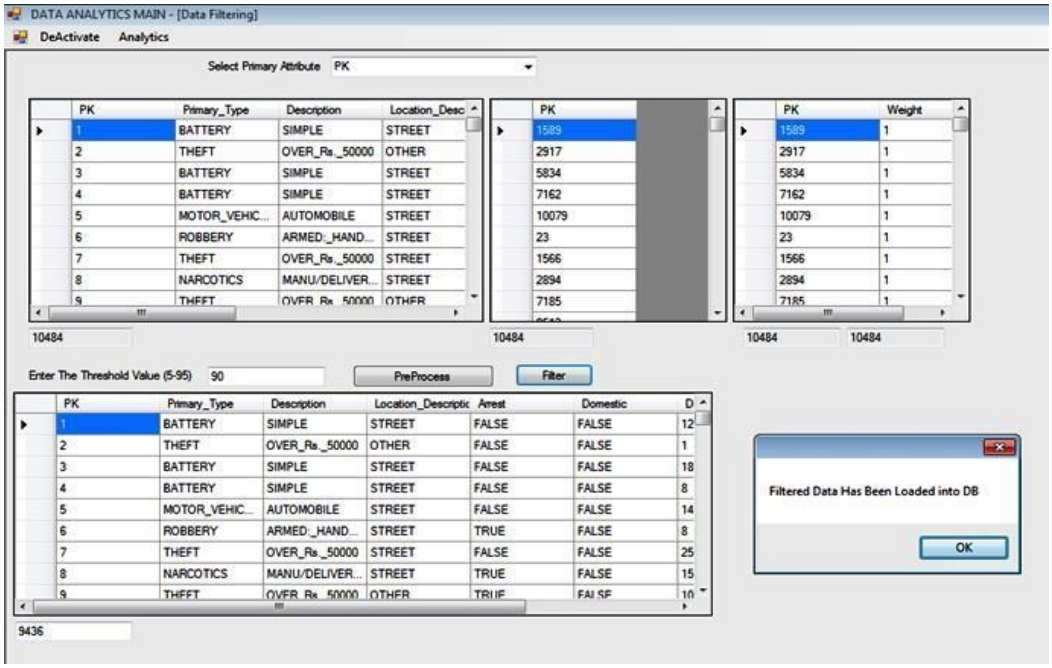
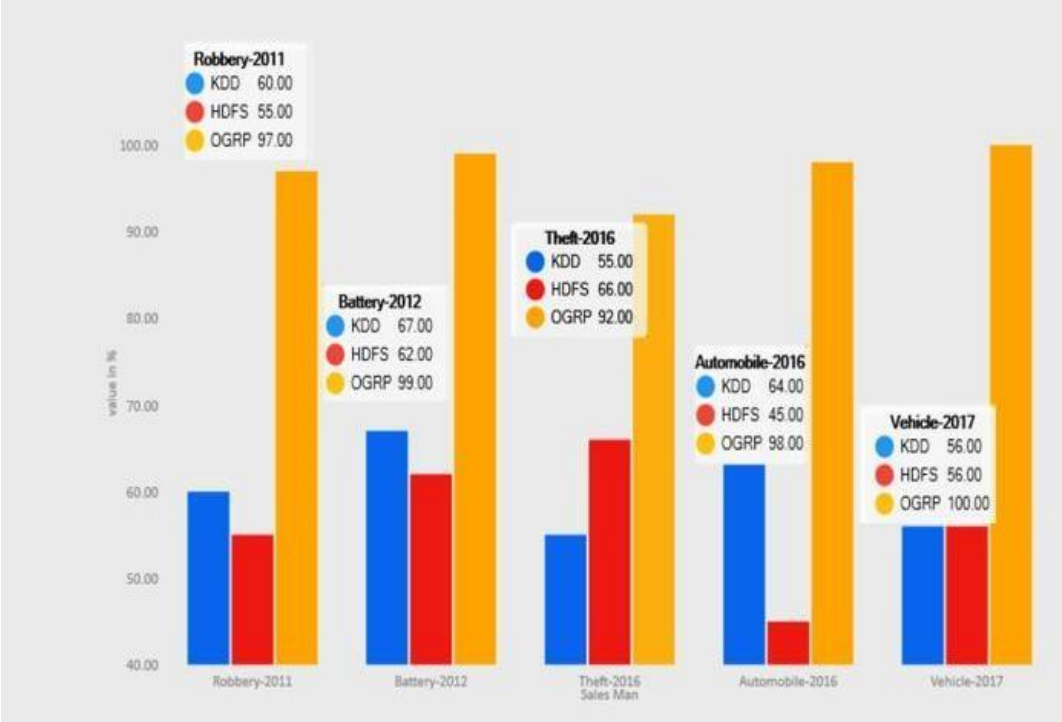


Fig.5: Data Filtering

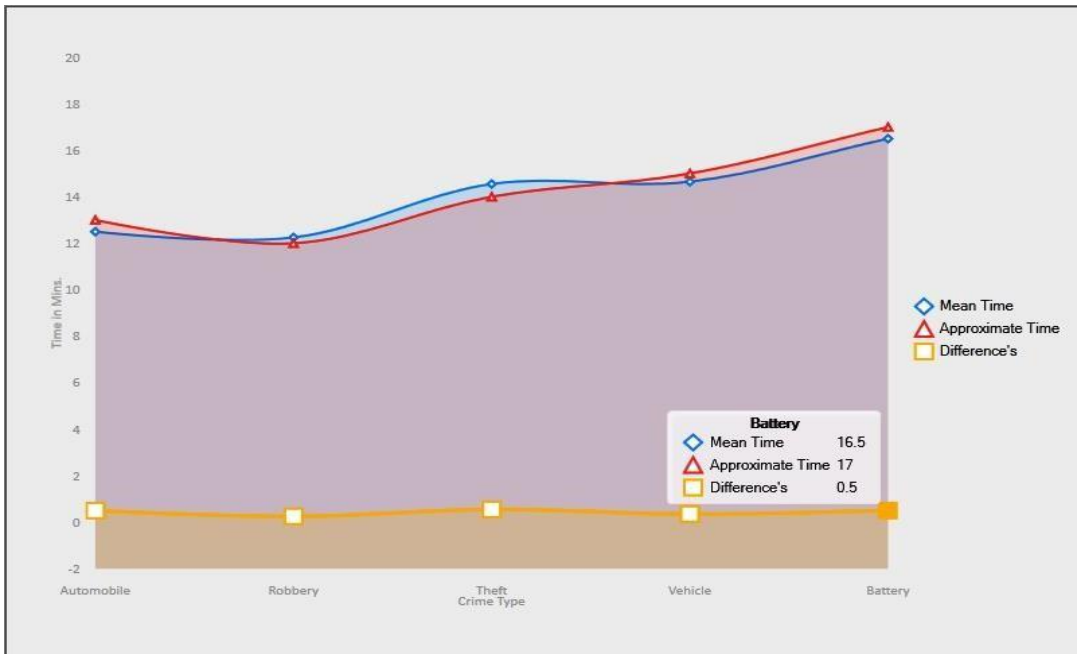


Graph.1:Efficiency Comparison.

The above graph compares the produced open power system convention phase algorithm with Data Mining KDD process and Hadoop HDFS algorithm. It is clear that the proposed system has produced high efficiency information with the analysis process; there is a huge difference between the existing and the proposed algorithm.

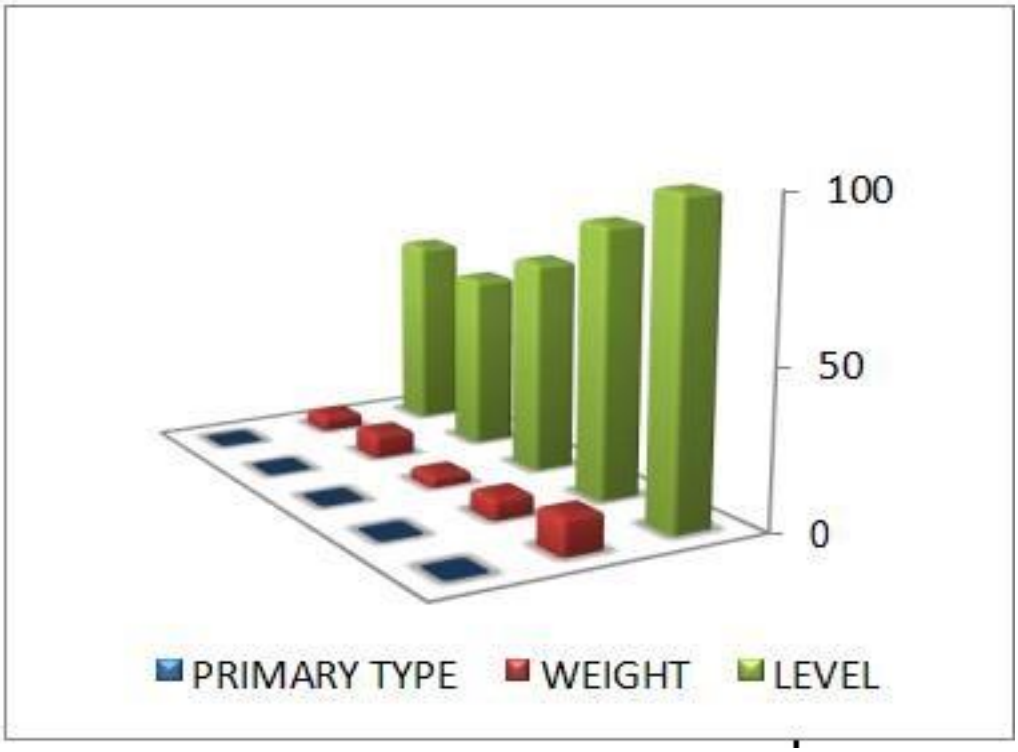
Table .2: The Time Comparison of the Attributes by OPSC

S.no	Attribute	Mean Time	Approximate Value	Difference
1.	Robbery	12.5	13	0.1
2.	Battery Theft	12.25	12	0.4
3.	Assault	14.55	14	0.6
4.	Motor Vehicle Theft	14.65	17	0.5
5.	Burglary	16.5	15	0.3



Graph2:Time Estimation Of Crime Detection

The above graph states the time estimation of crime detection by OPSC algorithm by showing the mean value for the attributes and our process shows the better approximate value.



Graph.3:Attribute Comparison Chart.

The above graph estimates the analyzing process of the attributes by their weightage. Data Mining is a technology enabled strategy for gaining richer, deeper, and more accurate insights about customers, partners and business environment. This process is handled by clearing the raw data from the data base. When a data mapping process is handled it is reviewed that data has been reduced and there is no data redundancy.

5. CONCLUSION

Data Mining has the potential to transform the way which shows sophisticated technologies to gain insight from their clinical and other data repositories and make informed decisions. This system was handled before by using many principles, but it does not give the clearance, OPSC algorithm gives. It is proved that this proposed system has produced more efficient result than comparing KDD and HDFS process. The data filtered is with high efficiency and there is no redundancy. The raw data can be also retrieved whenever the user needs, the data is more efficient upto 95% of approximate. This is the main reason for proper implementation of the data analytics methodology. Exploratory data analytics is a very important topic in the field of data analysis. This research work is in the domain of analytics but the foundation concepts pertaining to the exploratory data analysis form the main content of the research work.

6. REFERENCES

[1] Maluf, A. David, Bell, G. David, Knight, Chris,Tran, Peter, La,Tracy, Lin, Jenessa, McDermott,Bill, Pell, Barney,"NASA-XDB-IPG: ExtensibleDatabase - Information Grid" Global Grid Forum8,2020.
[2] S. Kaisler, F. Armour, and J. A. Espinosa, "Introduction to cloud storage da-ta:Challenges, opportunities, and realities minitrack," in 2019 47thHawaii InternationalConferenceon,pp.728-728,2019.
[3] Motoi Iwashita, Ken Nishimatsu, Shinsuke Shimogawa. Semantic Analysis Method for Unstructured Data inTelecomSer-vices.IEEE13. IEEE 13th International Conference on Data Mining Workshops;p.789795,2008.
[4] Ohlhorst F: cloud storage Data Mining: Turning cloud storage Data into cloud storage Money.USA:JohnWiley&Sons;2020.
[5] LaValle S, Lesser E, Shockley R, Hopkins MS, Kruschwitz N: cloud storage data, Mining and the path from insights to value. MITSloanManagRev,52:20–32,2020.
[6] Gärtner M, Rauber A, Berger H, Bridging structured and unstructured data viahybridsemanticsearchandinteractive ontology-enhancedqueryformulation. KnowlInfSyst 1–32,2021.
[7] Barbierato E, Gribaudo M, Iacono M, Performance evaluation ofNoSQL big-data applications using multi-formalism models. FutureGenerComputSyst37:345–353,2019.

[8] N. Diakopoulos, M. Naaman, and F. Kivran-Swaine, “Diamonds inthe rough: Social media visual Mining for journalistic inquiry,” inProceedings of IEEE Conference on Visual Mining Science andTechnology, pp.115–122,2019.
[9] Rajesh, M., and J. M. Gnanasekar. "Path Observation Based Physi-cal Routing Protocol for WirelessAd Hoc Networks." WirelessPersonalCommunications97,no.1:1267-1289. (2020)
[10] Chris Clifton, Murat kartarcioglu, Jaideepvaidya,Xiaodong Lin andMichael Y.Zhu, Tools for Privacy Preserving Distributed Data min-ing, SIGKDDEplor.Vol4,Issue2, 2021.
[11] Mansuri I.R. Sarawagi S. “Integrating Unstructured Data into Rela-tional Databases” Data Engineering. ICDE '06. Proceedings of the22nd InternationalConference,IITBombay2020
[12] DafangZhuang,WenYuan,JiyuanLiu,DongshengQiu,TaoMing,„The UnstructuredDataSharingSystemforNatural Re-sources and Environment Science Data of the Chinese Academy ofScience“ in Data Science Journal, Volume 6, Supplement, 20 Octo-ber2019.
[13] T.Menzies,"Beyonddatamining;towardsideaengineering,"inProceedings of the 9thInternational Conference onPredictiveModelsinSoftwareEngineering,2021,pp.1-6.
[14] hu, h., wen, y., chua, t-s., li, "toward scalable systems for cloud storage dataMining: a technology tutorial," access, ieee , vol.2, no., pp.652-687,2019.
[15] Chun-Wei Tsai,Chin-Feng Lai,Han-Chieh Chao, Athanasios V. Va-silakos,cloud storage data Mining:asurvey,,Journalofcloud storage data,Springer, December2021
[16] V. Srilakshmi, V.LakshmiChetana , T.P.AnnThabitha ,A Study oncloud storage DataTechnologies,InternationalJournalofInnovativeRe-search in Computer and Communication Engineering, Vol. 4, Issue6, June2020
[17] D. Assunçãoa, Rodrigo N. Calheiros b ,cloud storage Data computing andclouds: Trends and future directions Marcos, Elseiver, 27 August2019
[18] Samiya Khan, KashishAraShakil, MansafAlam, “Cloud Based cloud storage Data Mining: A Survey of Current Research and Future Directions”, JournalofContemporaryPsychotherapy,2021.
[19] Chang RM, Kauffman RJ, Kwon Y (2019) Understanding the paradigm shift to computational social science in the presence of cloud storage data.Decision SupportSyst63:67–80.
[20] Agrawal, D., Das, S., El Abbadi, A.: cloud storage data and cloud computing:current state and future opportunities. In: Proceedings of the 14thInternationalConferenceonExtendingDatabaseTechnology(EDBT/ICDT’11),pp.530–533(2020)
[21] Chen, C.L.P., Zhang, C. Y.: Data-intensive applications, challenges,techniques andtechnologies: a survey on cloud storage data. Inf. Sci. 275,314–347 (2019)
[22] Zikopoulos PC, DeRoos D, Parasuraman K, Deutsch T, Corrigan D,GilesJ:HarnessthePowerofcloud storage Data.McGraw-Hill:TheIBMcloud storage DataPlatform;2021.