IJRSET OCTOBER 2020 Volume 7, Issue 10 Pages: 9-18

ISSN 2394-739X

International Journal for Research in Science Engineering and Technology

REVIEW ON WEB SEMANTIC ANALYSIS FOR DATA MINING

[1] P. Nithya, [2] Dr. V. SATHYA
[1, 2] Assistant Professor,
[1, 2] Department of Computer Science,
[1] Jairam Arts & Science College, [2] MGR College, Hosur
[1] Salem – 636 008. [2] Tamil Nadu, India.

ABSTRACT- The integration of the two-quick creating scientific exploration regions Semantic Web and Web Mining is known as Semantic Web Mining. The tremendous expansion in the measure of Semantic Web data turned into an ideal objective for some analysts to apply Data Mining methods on it. . The thought is to improve, from one viewpoint, the consequences of Web Mining by misusing the new semantic designs in the Web; An audit of the current semantic web mining approaches is introduced in this paper. . This paper likewise gives a study of the new works nearby semantic web mining and a correlation of conventional web applications and semantic web applications along these lines giving rules to future examination nearby semantic web mining.

Keywords: [Web Mining;Semantic Web;Data Mining; SemanticWeb Mining; World Wide Web.]

1. INTRODUCTION

The World Wide Web is generally available to people while machines just have an extremely simple comprehension of its content. The original vision behind the Semantic Web (hence SW) is that computers ought to by one way or another have the option to comprehend and abuse information offered on the Web. The SW depends on two segments: (1) Formal ontology's give area explicit foundation knowledge as a jargon that is shared by a few gatherings and that depicts abstract item classes, predicate classes and their interdependencies, formalized in legitimate proclamations; (2) Annotations of web assets with explanations which can be perused and interpreted by machines through the basic ontological knowledge present started up certifiable perceptions.

Web-users exhibit various kinds of conduct contingent upon their information needs and their proposed errands. These undertakings are caught certainly by an assortment of actions taken by users during their visits to a site. For instance, in a dynamic application-based internet business Web site, user assignments might be reflected by successions of interactions with Web applications to look through a list or to make a buy. Then again, in an information-escalated site, for example, an entry or an online news source, user undertakings might be reflected in a progression of user taps on an assortment of Web pages with related content.

Semantic Web

The Semantic Web is certifiably not a different Web yet an extension of the current one, in which data is given all around

characterized importance, better-enabling computers and individuals to work in cooperation. For novices to the Semantic Web, which is regularly taken as the beginning stage for the research zone, is as great a beginning stage as any? The objective of the Semantic Web is in some sense a contrast to the Web of 2001. That Web was designed as a worldwide report archive with simple routes to access, distribute, and connect records and Web records were made to be accessed and perused by people. The Semantic Web is a machine-readable Web. As suggested over, a machine-readable Web encourages human-PC cooperation. As suitable and required, certain classes of errands can be appointed to machines and consequently handled naturally. Obviously, the plan opportunities for a machine-readable Web are exceptionally large, and a more number of plan.

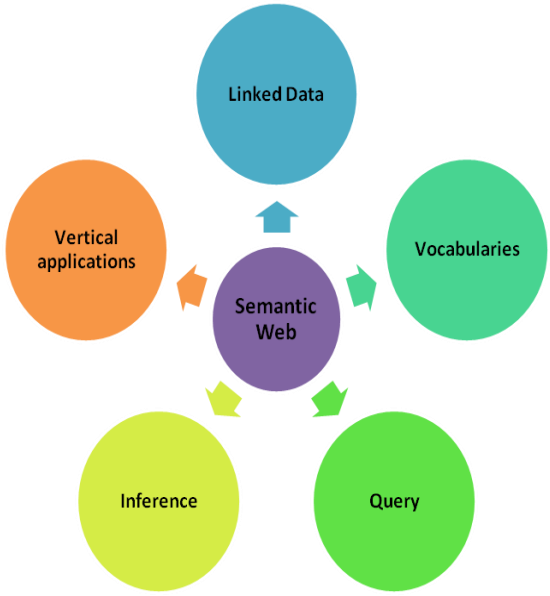

Figure 1.Semantic Web Analysis

Data mining

Information mining predominantly manages organized information coordinated in a database while text mining primarily handles unstructured information/text. Web mining lies in the middle and adapts to semi-organized information or potentially unstructured information. Web digging calls for inventive utilization of information mining or potentially text mining methods and their particular methodologies. Mining the web information is quite possibly the most difficult errands for information mining and information the board researchers in light of the fact that there are enormous heterogeneous, less organized information accessible on the web and we can without much of a stretch get overwhelmed with information.

2. LITERATURE SURVEY

1. Adamov, A. (2014) et. al proposed "Data mining and analysis in depth. Case study of Qafqaz University HTTP server log analysis" The touchy development of data accessible online brings new challenges of mining and examination of the web to discover helpful information. Since Web mining is predominantly founded on text mining techniques, it works with unstructured or semi-organized data. Web mining is the way toward investigating and changing over immense volumes of pointless data into significant information by determining intriguing patterns and relations among gatherings of data. Considering the fast development and solid rivalry in the market of internet business and online services, effectiveness and comfort are key variables for progress. Following user action on the website through investigation of the log documents is a fundamental advance in determining the best utilization patterns and adjusting the user interface to the individual user's necessities.

Merits

Web mining is the process of investigating and changing over gigantic volumes of futile data into significant information by determining fascinating patterns and relations among gatherings of data.

The assortment of data types and the unstructured nature of the substance utilizes unfeasible and hard. It is the reason Web mining strategies and application have gotten one of the key research zones of data mining.

Demerits

The dangerous development of data accessible online brings new difficulties of

mining and analysis of the web to discover valuable information is somewhat difficult.

2. Singh, S. P., & Meenu. (2017) et. Al proposed "Analysis of web site using web log expert tool based on web data mining" Web Usage Mining (WUM) is important for Web Mining. There is a data mining instrument to get and find information from web data. Web usage mining utilizes the data mining measure for the research of the usage pattern from data brought from the weblog files. The web is the assortment of scholarly educational institute marry server data was dissected to help the institute further improving the terms and policies of the service they give. It helps in the assessment of the viability of the web website, and supportive in acquiring achievement in a marketing effort. Web mining likewise considers extricating patterns in data with the assistance of construction mining, content mining, and usage mining. The web analyzers investigate them cut off log files for the assurance of the framework blunder. The scholastic educational institute marry server data was dissected to help the institute further improving the terms and policies of the service they give.

Merits

Web Usage Mining is utilized for E-Learning, E-business, E-Commerce, E-Government, E-Newspapers, and Digital Libraries Making.

Personalization for a client can be performed by monitoring beforehand access pages.

Demerits

The Web mining process would be not effective if the samples are not a better representation of the bigger body of data.

3. Chen, J., Li, K., Liu, Z., Zhang, T., Wen, W., Song, Z., … Huang, T. (2019) et. Al proposed "Data Analysis and Knowledge Discovery in Web Recruitment—Based on Big Data Related Jobs"Presently, as per the operation model, the online recruitment industry can be separated into classified data recruitment websites, exhaustive recruitment websites, vertical recruitment websites, search recruitment websites, neighbourhood recruitment websites and social recruitment websites. Network data mining depends on insights, machine learning algorithm and data mining innovation to measure and dissect the huge network data, find the network data recruitment interest through modelling, and afterward foresee the future profession request pattern. Enormous data-related positions centre around back-end development, operations, and items, trailed by DBA, client, and promoting. It is clarified that the principle occupation of large data post is data-driven, giving vital choice help to key zones of big business like item system, operation strategies, client research, market pattern and client picture.

Merits

TF-IDF strategy is picked to choose the accompanying stop words: I, have, is, and so forth the chose stop words are added to the stop words rundown, and afterward the stop words list is utilized to filter the stop words.

LDA model is broadly utilized in text clustering, similitude calculation and different fields.

Demerits

Network data mining depends on statistics, machine learning algorithm and data mining technology to process and examine the enormous network data, find the network data recruitment interest through displaying.

4. S. Zhou, X. Zhang, X. Li, G. Zheng and G. Zhang (2017) et. Al proposed "Design and Implementation of Data Mining and Analysis Platform Based on Web Service Techniques," Taking into account the straightforwardness of conventional data mining and analysis stage application situations, a high data mining and analysis stage dependent on B/S Architecture was planned, which was called HSP-DMA. The stage incorporated the calculation service with the application service and utilized the REST Web Service to distribute as a segment service to encourage the outsider system call. HSP-DMA embraces the intuitive modelling strategy, which is not difficult to work and decreases the trouble of modelling. It joined the business situation to

build the application model and fulfilled the real business needs of the power grid. It incredibly improved the application degree and diminished advancement costs. The HSP-DMA stage is planned dependent on Web Service, and every module has been delivered by segment service, which is advantageous for the outsider system call. The stage incorporates an assortment of develop data mining items and supports an assortment of data files, standard database, with internet modelling, model management, and model publishing and other data mining coordinated management capacities.

Merits

Web service is a platform-independent, low-coupled, self-contained, web-based application. It is regularly used to develop circulated, interoperable applications.

Data decrease strategies can be utilized to acquire the data set protocol said it is a lot smaller yet at the same time close to keep up the uprightness of the original data.

Demerits

A definitive users of the platform and they accomplish their every day work through the browser interface is hard for different users.

5. Hidayat, W., & Yaqin, A. (2019) et. Al proposed"Business Trends Based on News Portal Websites for Analysis of Big Data Using K-Means Clustering"Business analysis is performed to decide the business that is well known, in Indonesia with text mining can take information from a few news portals in Indonesia. Text preprocessing is utilized to change the text and tags on the news to be changed over into loads. The heaviness of the information will be handled utilizing the K-Means algorithm to be assembled into clusters and each cluster will be pictured utilizing Word Cloud so that words that regularly show up as famous word recognizable proof are known. Testing utilizes the Silhouette Coefficient to calculate the quality of every part against the cluster. Moreover, every part will be deciphered by the test outcomes. The analysis is completed each month in 2018 with a sum of 995 information with a month to month normal of 6 clusters, in January were the most well known business as indicated by the quantity of individuals from 64 information shaped 6 clusters, the most part clusters were cluster 1 the Silhouette Coefficient test results are solid 0.00%, medium 65.22%, feeble 30.43%, not generous 4.35%, Word Cloud framed was a cowhide pack business. In the wake of realizing the number k moreover decide the centroid of the cluster haphazardly and calculate the distance utilizing the Euclidean equation. Besides, the assurance of the centroid of each gathering is taken from the normal (mean) of all information esteems in each cluster and recalculated the distance between information to centroid until the cluster part has no change.

Merits

Testing utilizes the Silhouette Coefficient to calculate the quality of every part against the cluster.

Text mining from title and tag in news entries at that point processed utilizing text pre-processing and standardized to defeat a lot of distance to the load between documents.

Demerits

The more weight the relevant data builds the nature of clusters and the more similar words will explain, Word Cloud not to be changed after the analysis process.

6. Annamoradnejad, R., Annamoradnejad, I., Safarrad, T., & Habibi, J. (2019) et. Al proposed "Using Web Mining in the Analysis of Housing Prices: A Case study of Tehran" There have been numerous past attempts to decide the determinants of lodging prices. These works depended on a moderately little arrangement of information, generally gathered with the assistance of land organizations. In this work, we utilized web mining strategies to produce a big, coordinated dataset from a famous national financier site. The dataset contains underlying attributes of in excess of 139,000 condos, close by their location and cost. We gave our full dataset to the article with the goal that different researchers can

repeat our results or conduct further analyses. Utilizing this dataset, we examined the lodging prices of Tehran to recognize its significant determinants. To this point, we analyze the elements of lodging prices at the district levels of Tehran utilizing the Hedonic Price model. In view of our accumulated dataset, this paper expects to inspect the elements of the real estate market at the district levels of Tehran, as a special multi-dimensional city. By giving this dataset as advantageous information, different researchers can replicate the results of this investigation or conduct further analyses.

Merits

The Hedonic model is widely used to analyze the price of a good, by assuming that the final price is a function of the good's characteristics.

Hedonic Price model in order to assess the structural, location and environmental determinants of housing prices in the city of Tehran.

Demerits

Collecting more number of variables and housing models are not possible. It's a limited data collecting model.

7. Thomas, D. M., & Mathur, S. (2019) "Data Analysis by Web Scraping using Python" The standard data examination is built on the root and impact relationship, molded as an illustration microscopic assessment, subjective and quantitative assessment, the objectivity approach of making extrapolation assessment. The Web Scraper's scheming morals and procedures are compared, it clarifies the working of how the scraper is planned. The procedure of it is dispensed into three fragments: the web scraper draws the ideal connections from the web, and afterward the data is extracted to get the data from the source joins lastly stowing that data into a CSV document. The Python language is executed for the doing. Thusly, connecting all these with the ethical information on libraries and working expertise, we can have a sufficient Scraper in our grasp to deliver the ideal outcome. Because of a huge local area and library assets for Python and the perfection of coding stylish of python language, it is the most proper one for Scraping wanted data from the ideal website.

Merits

The scraper tool for the web is used for got information from the web have and as a bit of employments utilized for web orders.

For clients, the upside of RDB and SQL is to can isolate their information on interest. It made the methodology to get information basic and spread database use.

Demerits

1. Scraper's was delete some data suddenly during

8. Spiridonov, R. E., Cvetkov, V. D., & Yurchik, O. M. (2017) et. Al "Data mining for social networks open data analysis" After social networks promotion, marketing has changed, following the prerequisites of the business. Presently it isn't sufficient to gather information and assemble deals estimate graphs – you need to discover fundament for it, and to lessen the marketing spending you should cooperate with the business crowd in the best manner. The assignments of processing and examining a lot of information taken from client profiles in social networks permit us to become acquainted with the crowd better compared to companions think about them. This article depicts one of the potential solutions for examining information removed from the social networks profiles on Instagram with a view to their further translation to improve the effectiveness of Internet marketing in the social organization. The solution of the depicted issue toward the start of the article, the undertaking of looking and extending the crowd in the social organization, isn't restricted to grouping profile's information by boundaries, yet additionally requires tackling numerous different errands identified with separating the necessary informational indexes from the social organization, distinguishing their qualities and discovering relationships.

Merits

The assignments of classifier and investigating a lot of data taken from user profiles in social networks permit us to become more acquainted with the crowd better compared to friends think about them.
The classifier will be utilized in a system that separates the expansion of the crowd given to it to analyze the social network account, with the areas and interests of this crowd as a main priority.

Demerits

The task of searching and expanding the audience in the social network is not limited to clustering profile's data by parameters.

9. Nemeth, M., & Michalconok, G. (2017) et. Al "The initial analysis of failures emerging in production process for further data mining analysis" Data from the manufacturing process can be acquired effectively, as a rule, in light of the fact that the data are naturally created by the sensors which are introduced on gadgets and monitor the activity of the creation process. This data is put away in the pre-organized database from where it would then be able to be traded utilizing a proper interface in different organizations, as indicated by client needs. Understanding the data from which the already unknown knowledge is to be found a vital advance to get a right portrayal of this new knowledge. It is in this way important to comprehend the design of the gathered data and to realize the data kinds of the data contained in the data set. This underlying data examination is crucial to have the option to survey the knowledge capability of the given data. In light of this, we can call attention to intriguing issues that can be additionally settled with KDD (knowledge discovery from databases) techniques.

Merits

To meet the expressed goal, it is feasible to utilize a few unique techniques. This relies upon the design of the data and on the necessities of a specific technique for data mining on the information data.
Friedman has portrayed the data mining process as a bunch of techniques used to find connections in data in huge databases. The process of data mining covers somewhat with man-made brainpower, AI, and example acknowledgment and data representation.

Demerits

Look at and optimize algorithms to discover tedious patterns in the information about consolidating disappointments in the framework. These patterns can thusly distinguish gathering and request of disappointments that are arising consistently isn't possible in this method.

10. Anoopkumar M, & Rahman, A. M. J. M. Z. (2016)et. Al proposed "A Review on Data Mining techniques and factors used in Educational Data Mining to predict student amelioration"Educational Data Mining (EDM) is an interdisciplinary ingenuous research region that handles the improvement of strategies to investigate data arising in scholastic fields. Computational methodologies utilized by EDM are to inspect scholastic data to examine educational inquiries. Therefore, it gives characteristic information on the instructing and learning measure for powerful schooling arranging. In the genuine world, foretelling the presentation of the faculties is a difficult errand. Be that as it may, the significant test of advanced education is the abatement in the achievement pace of faculties. An early forecast of faculties' exhibition may assist the administration with giving ideal activities just as preparing to build the achievement rate. The educational plans changes and the various styles in the academic interaction is additionally a major worry in the achievement paces of faculties.

Merits

The found knowledge can be utilized for organization of schedule, expectation with respect to enrolment of students in a specific program, estrangement of conventional classroom teaching model, recognition of out of line implies utilized in online examination, location of strange qualities in the outcome sheets of the students.

Demerits

The field of Data Mining (DM) is worried about finding beginning examples in astronomically massive amounts of data. DM is an innovation utilized in various controls to test for consequential connections among factors in the astronomically colossal dataset; it's the most troublesome process in DM.

11. Abuomar, O., King, R., & Younan, N. (2015) et. Al proposed"Data stream mining formultitemporal remote sensing data" One of the applications of data stream mining is multi-temporal remote sensing where numerous algorithms and techniques have been advancing to oblige continuous data and analysis for information revelation and ongoing control. Mining time-series multi-temporal remote sensing is viewed as the vital achievement here. In this paper, a few parts of mining time-series remote sensing databases are inspected, with an attention on pattern analysis and likeness search. A time-series database is likewise a succession database. Nonetheless, an arrangement database is any database that comprises of successions of requested occasions, with or without solid thoughts of time. Time series pattern analysis consolidates demonstrating time-series and guaging time series. In likeness search in time series analysis, there are two sorts of comparability look: aftereffect coordinating and entire arrangement coordinating. Aftereffect coordinating with finds the arrangements that contain aftereffects that are like a given inquiry grouping, while entire succession coordinating with tracks down a bunch of successions that are like one another.

Merits

For efficient accessing, a multidimensional file can be developed utilizing the initial not many Fourier coefficients. At the point when a closeness query is submitted to the framework, the file can be utilized to recover the sequences that are all things considered a specific little separation away from the query grouping.

Demerits

Where stream mining algorithms have been executed for synchrophasor information to meet fast dynamic necessity of future situational mindfulness applications in power systems and smart grids.

12. Pandey, K. K., & shukla, D. (2018)et. Al proposed "Challenges of Big Data to Big Data Mining with their Processing Framework"Big data is an emerging pattern and need of ventures, sciences, and engineering zone since all territories are having a great deal of data and these data have given a result for a specific problem. This result is exceptionally useful for making a decision for developing to possess zone and this decision is taken by big data investigation and big data mining process. At the point when an association is playing out a big data mining process then it faces a ton of challenges and these challenges have given a ton of dangers to settle on a decision. This paper presents the big data and data mining idea with big data mining challenges which are identified with data, process, and management. Data challenges are identified with big data, Process challenges are identified with data mining and Management challenges are identified with the common factor of big data and data mining. In the last section, this paper proposed a theoretical big data mining process system which is useful for taking care of challenges and plan big data mining calculation. Big data mining challenges which are identified with big data, data model, management, mining process and so on and the last section proposed to big data mining system which is process model of the distinctive sort of challenges in big data mining. This proposed big data mining structure uphold relationship among data challenges, Process challenges, and management challenges.

Merits

Systems of support and analysis Exabyte or more prominent then Exabyte extent data in day by day for dynamic or developing a client support system.

on the off chance that two diverse datasets size is something very similar yet data is not

the same as analytics or mining technique is distinctive utilized methods data volume isn't relying upon a data or mining technique

Demerits

The process of data mining and Big Data Mining process is a similar just contrast is some new difficulties is meet in each progression and need to grow new data mining motor or data mining algorithm and technique in Big Data Mining.

13. Sharma, S., Sharma, A. K., & Soni, D. (2017) et. Al proposed "Enhancing DBSCAN algorithm for data mining"Today data mining is generally utilized by organizations with a solid purchaser center like retail, financial, communication and advertising associations. Here actually data mining is the process of extraction of required information from gigantic databases. It permits clients to dissect data from various measurements or points, classify it and sum up the connections distinguished. A definitive objective of this paper is to propose a procedure for the improvement in the DB-SCAN algorithm to improve clustering accuracy. The proposed improvement depends on a back propagation algorithm to figure Euclidean distance in a powerful way. Additionally, this paper shows the acquired consequences of carried out proposed and existing techniques and thinks about the outcomes regarding their execution time and accuracy. Clustering can be performed by its different algorithms among which some depend on thickness which is called thickness based clustering algorithms, DBSCAN is likewise a thickness based clustering algorithm that is utilized in this paper for the process of data mining.

Merits

The role of the clustering algorithms is to distinguish clusters of POIs and afterwards utilize the clusters to naturally portray geographic districts.

Density-put together clustering is performed with respect to enormous datasets; Synthetic datasets are utilized for experimental assessment which shows that the new clustering calculation is quicker and more scalable than the first DBSCAN.

Demerits

Accuracy of clustering EPS values is calculated in the dynamic manner which leads to the clustering of the points which are remained unclustered.

14. F. Jiang, C. K. Leung and A. G. M. Pazdor (2016) et. Al proposed "Big data mining of social networks for friend recommendation," In the current period of big data, high volumes of important data can be effortlessly gathered and created. Social networks are instances of producing wellsprings of this big data. Users in these social networks are frequently connected by some interdependency like friendship. As these big social networks continue to develop, there are circumstances in which an individual client needs to discover mainstream gatherings of friends with the goal that he can prescribe similar gatherings to different users. In this paper, we present a big data analytic arrangement that utilizes the Map-Reduce model in mining these big social networks for finding gatherings of habitually associated users for companion recommendation. Assessment results show the productivity and practicality of our data analytic arrangement in mining big social networks, finding mainstream users, and suggesting friends. big data analytic arrangement—called BigPFM—which directs big social organization mining and examination by means of mainstream friendship mining with the MapReduce model. BigPFM encourages social organization users to find gatherings of habitually associated client from big social networks by utilizing the MapReduce model. Assessment results show the effectiveness and practicality of BigPFM in mining big social networks, finding famous users, and suggesting friends.

Merits

Subsequent to finding frequently associated friends, BigPFM at that point structures rules by utilizing these friend records. The resulting rules would then be able to be utilized for friend recommendation.

Demerits

More data mining technique involve reducing the map also its take more time for processing.

15. Vignesh, V., Pavithra, D., Dinakaran, K., & Thirumalai, C. (2017) et. Al proposed "Data analysis using box and whisker plot for stationary shop analysis" In statistical analysis, we have a collection of information, with the utilization of these information, we need to do analysis dependent on our requirements. With the collection of information utilizing Statistical analysis, we bargain collection, analysis, introduction and sorting out the information. With the assistance of statistical analysis, we can discover underlying patterns, connections, and patterns between information tests. The R system for statistical figuring is a climate for information analysis and designs. Here we will execute the boxplot technique and control diagram strategies for deals analysis in a fixed shop dataset. With the assistance of a boxplot, we can without much of a stretch make relations among tests and we can discover the exceptions. Outwardly of the crate in a boxplot, two more vertical lines are drawn, one vertical close to the upper quartile is called upper stubble and another line close to the lower quartile is called lower hair. The motivation behind this paper is to utilize the "case and hair plot" technique for visualizing the examples of the dataset and from those outcomes, we can without much of a stretch make connections between the attributes.

Merits

Boxplot helps us to easily compare the different data sets in the Word.

"Box and whisker plot" method for visualizing the samples of the dataset and from that results we can easily make relationships between the attributes.

Demerits

Finding box plot is more time consumption process.

CONCLUSION

In this paper, we have given a survey on the usage of Semantic Web data, most prominently Linked Open Data, for data mining and knowledge discovery. We show the flexibility of this framework in characterizing various relationships among users, user tasks and Web objects. logical inference with inductive procedures that exploit regularities in the data. In this chapter we have presented a comprehensive discussion the Web personalization process viewed as an application of data mining which must therefore be supported during the various phases of a typical data mining cycle. Web Data Mining is perhaps still in its infancy and much research is being carried out in the area. The major challenge is to develop more automatic semantic data mining algorithms and systems by utilizing the full strength of formal ontology that has well defined representation language, formal semantics, and reasoning tools for logic inference and consistency checking.

REFERENCES

[1]. A. Adamov, "Data mining and analysis in depth. case study of Qafqaz University HTTP server log analysis," 2014 IEEE 8th International Conference on Application of Information and Communication Technologies (AICT), Astana, Kazakhstan, 2014, pp. 1-4, doi: 10.1109/ICAICT.2014.7035947.

[2]. S. P. Singh and Meenu, "Analysis of web site using web log expert tool based on web data mining," 2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS), Coimbatore, India, 2017, pp. 1-5, doi: 10.1109/ICIIECS.2017.8275961.

[3]. J. Chen et al., "Data Analysis and Knowledge Discovery in Web Recruitment—Based on Big Data Related Jobs," 2019 International Conference on Machine Learning, Big Data and Business Intelligence (MLBDBI), Taiyuan, China, 2019, pp. 142-146, doi: 10.1109/MLBDBI48998.2019.00033.

[4]. S. Zhou, X. Zhang, X. Li, G. Zheng and G. Zhang, "Design and Implementation of

Data Mining and Analysis Platform Based on Web Service Techniques," 2017 International Conference on Computer Technology, Electronics and Communication (ICCTEC), Dalian, China, 2017, pp. 545-549, doi: 10.1109/ICCTEC.2017.00124.
[5]. W. Hidayat and A. Yaqin, "Business Trends Based on News Portal Websites for Analysis of Big Data Using K-Means Clustering," 2019 International Conference on Information and Communications Technology (ICOIACT), Yogyakarta, Indonesia, 2019, pp. 445-450, doi: 10.1109/ICOIACT46704.2019.8938413.
[6]. R. Annamoradnejad, I. Annamoradnejad, T. Safarrad and J. Habibi, "Using Web Mining in the Analysis of Housing Prices: A Case study of Tehran," 2019 5th International Conference on Web Research (ICWR), Tehran, Iran, 2019, pp. 55-60, doi: 10.1109/ICWR.2019.8765250.
[7]. D. M. Thomas and S. Mathur, "Data Analysis by Web Scraping using Python," 2019 3rd International conference on Electronics, Communication and Aerospace Technology (ICECA), Coimbatore, India, 2019, pp. 450-454, doi: 10.1109/ICECA.2019.8822022.
[8]. R. E. Spiridonov, V. D. Cvetkov and O. M. Yurchik, "Data mining for social networks open data analysis," 2017 IEEE II International Conference on Control in Technical Systems (CTS), St. Petersburg, Russia, 2017, pp. 395-396, doi: 10.1109/CTSYS.2017.8109578.
[9]. M. Nemeth and G. Michalconok, "The initial analysis of failures emerging in production process for further data mining analysis," 2017 21st International Conference on Process Control (PC), Strbske Pleso, Slovakia, 2017, pp. 210-215, doi: 10.1109/PC.2017.7976215.
[10]. Anoopkumar M and A. M. J. M. Z. Rahman, "A Review on Data Mining techniques and factors used in Educational Data Mining to predict student amelioration," 2016 International Conference on Data Mining and Advanced Computing (SAPIENCE), Ernakulam, India, 2016, pp. 122-133, doi: 10.1109/SAPIENCE.2016.7684113.
[11]. O. Abuomar, R. King and N. Younan, "Data stream mining for multitemporal remote sensing data," 2015 8th International Workshop on the Analysis of Multitemporal Remote Sensing Images (Multi-Temp), Annecy, France, 2015, pp. 1-4, doi: 10.1109/Multi-Temp.2015.7245745.
[12]. K. K. Pandey and D. shukla, "Challenges of Big Data to Big Data Mining with their Processing Framework," 2018 8th International Conference on Communication Systems and Network Technologies (CSNT), Bhopal, India, 2018, pp. 89-94, doi: 10.1109/CSNT.2018.8820282.
[13]. S. Sharma, A. K. Sharma and D. Soni, "Enhancing DBSCAN algorithm for data mining," 2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS), Chennai, India, 2017, pp. 1634-1638, doi: 10.1109/ICECDS.2017.8389724.
[14]. F. Jiang, C. K. Leung and A. G. M. Pazdor, "Big data mining of social networks for friend recommendation," 2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), San Francisco, CA, USA, 2016, pp. 921-922, doi: 10.1109/ASONAM.2016.7752349.
[15]. V. Vignesh, D. Pavithra, K. Dinakaran and C. Thirumalai, "Data analysis using box and whisker plot for stationary shop analysis," 2017 International Conference on Trends in Electronics and Informatics (ICEI), Tirunelveli, India, 2017, pp. 1072-1076, doi: 10.1109/ICOEI.2017.8300874.