Techniques and Softwares for Social Media Data Mining

Dr Komarasamy G [1], Dr. Sanjiv Kumar Jain [2]

[1] VIT Bhopal University, India
[2] Ballari Institute of Technology and Management, Ballari, Karnatak, India
Corresponding Author E-Mail: [1] gkomarasamy@gmail.com

Abstract

Data mining refers to a framework by which a company can easily enhance their performance. In this study, social media data mining process is critically evaluated with help of several types of data sets. Techniques and software's are also mentioned in this study, by which a firm maintains this data mining process in a significant way. Importance of social media data mining is to attract more customers. Every company has a goal and objective in global market to enhance their performance. Social media data mining process helps a company to achieve those objectives and goals. Predictive analysis, improved revenue, lower costs and creating awareness are essential benefits of this particular process related to a company. Privacy and security related issues of data are faced by a firm with help of this data mining process. This process requires more amount of money initially, for this reason, small organisations cannot be able to maintain this social media data mining process. Association, clustering, classification, machine learning process, data cleaning and data visualisation are unique and relevant techniques of data mining. Usage of these techniques are briefly discussed in this study..

Keywords

Data mining, social media, clustering, association..

INTRODUCTION

Process of obtaining big data from user generated content on social media sites is known as social media data mining. A process of extracting useful information from huge volumes of data is also referred to as a data mining process. This process helps to enhance cost cutting, revenues and improve customer relationships. Two types of data mining process are available such as: predictive and descriptive data mining analysis. Predictive data mining analysis helps users to predict future trends of a business activity [1]. In order to find numerical output, prediction data mining analysis is used to gather relevant and authentic data related to work process. A model helps to predict continuous valued function to maintain a proper business activity. Several types of predictive data mining processes are available such as: classification, regression, prediction and time serious analysis.

Descriptive data mining process refers to summarising several types of data into relevant information. Proper description of work process is gained with help of this descriptive data mining process. Frequency, correlation and cross tabulation are used to maintain descriptive data mining analysis. Customer preference plays a positive impact on economic growth of a business activity. This particular data mining analysis helps to identify those customer preferences in a significant way. In case a company identifies demand of customers and markets, this firm can easily earn high profit. Several types of descriptive data mining are available such as: sequence discovery, summarisation, clustering and association rules analysis [2]. These should be helpful for a company to maintain their organisational and financial performance. In order to create meaningful objectives this particular analysis is helpful. Objectives into predefined classes are mentioned with help of this clustering data mining analysis. Several types of data related to work are available, company management team has a responsibility to summarise all data and provide a concussion related to work.

MATERIALS AND METHODS

Entire process of a project work is maintained with help of this materials and methods segment. This chapter plays an essential role to gather, interpret and summarise all gathered data in a simple way. Several types of steps are followed by a researcher for this study to finish their work within a given deadline. Research design, type, philosophy and data collection methods are followed by a researcher for this study. Overall process of a research work is maintained by this research design [3]. Several types of questions related to a project work are available, researcher can easily answer all questions with help of this research design. In this study, researcher uses "cross sectional" research design to maintain a proper process and strategy of work. Observational study design mainly refers to this cross-sectional research design. Several types of variables are available related to this research topic, this particular design helps researcher to compare different types of variables at same time.

Different meanings of a research work are gained by a researcher with help of this research approach. Researcher can easily maintain proper methods to collect and analyse all gathered data by this research approach. Several types of research approaches are available such as: "inductive", "deductive" and "abductive". In this study, researcher uses an "inductive" research approach to generate conclusions of this study. Logical way of a research work is maintained by a researcher with help of this "inductive" research approach.

Every research work has aims and objectives and these are advantageous for a research work to maintain proper work strategy. "Inductive" research approach helps researcher to achieve several types of goals of this particular research work [4]. In order to gather several types of information related to a research work, research type helps researcher to investigate research problems and questions to gain new knowledge related to a research work.

Scientific advancement of a research work is maintained with help of this research type. In this study, researcher uses "qualitative" research types to gather clear ideas about this study. Logical and coherent way of research work is maintained by researcher with help of this particular research type [5]. Researcher collects secondary data related to this particular study. These individuals mainly gather existing data from online journals, books, websites and peer reviewed journals, which are published after 2019. This data collection process is advantageous for this particular study. Theoretical analysis of a research work is maintained properly with help of this methodological view of a research work.

RESULTS

A concept about social media data mining process

Social media is a useful tool to share and gather several types of information related to a study. In recent days, everyone uses social media platforms to communicate and maintain a strong relationship with each other. Several types of information are gathered by everyone by this social media platform. Up to date information is collected by a person by this social media platform. Aggregate data and uncover trends are known to everyone by this social media data mining process [6]. Wild language landscape related to this particular research work is maintained with help of this data mining process. Social media data mining process helps to spot trends of a business activity. This social media is effectively beneficial to gather information related to domestic demands of markets and customers. In case an organisation understands demand of local customers and markets, this should be beneficial for a firm to maintain their economical functionality.

Social media data mining process helps to spot trends and techniques of a business activity in an organised manner. Social listening plays a positive impact on market demand and customer demand of an organisation. Topics and keywords related to this study are maintained with help of this social media data mining process. This particular data mining process is helpful for a company to identify proper context related to work. More accurate prediction of work is gained with help of this social media data mining process. Traditional methods of work are maintained by a firm in global market to earn more profit [7]. These traditional methods are known to each and every customer. These individuals are attracted by this traditional method in a significant way. Relevant and accurate prediction of outcomes of an organisation is effectively beneficial to maintain their organisational culture. Every company always

tries to achieve aims and objectives to enhance engagement of employees. This social media data mining concept helps every firm to maintain work process in a simple way.

Social media data mining helps to watch several types of events unfold in real time. Another excellent reason to use this particular technique is offered with help of an event detection process. Emerging situations and context of a business activity are maintained properly by this social media data mining process [8]. Social media data mining process helps to maintain traditional sensor methods in an organised way. For this reason, work process is maintained properly by these concepts. This social media data mining process helps to provide relevant content related to this particular study. An annoying reality of an online presence is boosted with help of this social media data mining. Social media platforms are helpful for an organisation to interact with customers, employees and other stakeholders in a significant way. Spam activity of a company is mitigated with help of this social media data mining. Social media data mining process helps a company to maintain their work timing and process. Several types of micro trends are available in global market, every company has a responsibility to understand work process of those micro trends.

Social media heat map is leveraged with help of these social media data mining processes. A firm always tries to identify their target markets and customers by using social media platforms. Potential customers are gathered by a company to enhance their financial health. Data mining process helps a company to support research process related to work. Every firm always tries to enhance their economic growth. For this reason, a company produces several types of products to enhance engagement of customers. Every feature of a new product is known to every customer with help of this social media data mining process [9]. In case customers are easily attracted by these new products, this company can produce new products in a huge amount globally. For this reason, profitability of an organisation is also enhanced with help of this social media data mining process.

Brief idea about data mining techniques

Process of analysing dense volumes of data to identify accurate patterns. Data techniques are helpful for a company to discover modern trends of new markets. This particular process helps to gain insights related to usage of data. Decision making process of a person is enhanced by this social media data mining process. Large blocks of information are explored and analysed with help of these data mining processes [10]. Exploration process of data mining techniques is immensely helpful for an organisation to maintain their economic growth. Several types of data techniques are available in global market such as: data visualisation, clustering, data cleaning, association, classification, prediction, machine learning, neural networks, outlier detection and data warehousing. Visual presentation of data related to work is maintained properly with help of this clustering process of data mining. Buying trends and demographic sales of a company are represented in a graph

by this clustering process [11]. Processes of making groups of different data points based on their characteristics are referred to as clustering concepts.

Several groups refer to several types of characteristics related to data. Data miners help to divide data into subsets with help of clustering data mining techniques. Several types of data clustering methods are available such as: model-based method, partitioning, density based, grid based, hierarchical method. The types of data are used to make different types of models in workplace to identify target markets and customers. Model-based method mainly refers to creating several models for each data cluster. Involvement process of data into a group is referred to as the partitioning method [12]. Specific cluster is used for this particular study to fulfil criteria of each individual cluster. Data points belong to just one cluster or group for this specific method of data cluster. Data points are collected and analysed with help of this density-based method. Density based method helps a company to enhance machine learning process. Several types of data related to work are gathered in workplace to maintain a smooth process [13]. Those data are divided into some cells on a grid. An individual cell is clustered by this grid-based method.

Overall database process is managed by this particular cluster process. Grid based method helps to maintain fast processing time of work within a company. Data points are referred to as a single cluster globally. Hierarchical method of data clustering helps to make a group of similar data points [14]. These data points are analysed separately from each other by this hierarchical method. This method helps a company to identify necessary characteristics of products, by which customers are easily attracted by this company. Sales of graphs of an organisation is also maintained by this clustering data mining process. In order to find correlations among data points in a data set, association rules are immensely helpful for an organisation. Data miners use association to discover interesting relationships among variables in database. Marketing strategy and research is maintained by a firm with help of rules and regulations related to association [15]. In this process two types of methods are available such as: multidimensional association single dimensional association.

Single dimensional association involves looking for one instance of a data point. In case a company focuses on only one product, this firm may not be able to earn high profit from global market. Single dimensional association creates a negative impact on economic growth of a company in global market [16]. Company management team has a responsibility to focus on every product and service, which are provided by a firm. Multidimensional association helps an organisation to manage all products and services. Several types of data points are available within a data set. All data points are critically evaluated with help of this multidimensional association. Economic stability of an organisation is managed with help of this association data cluster.

Usage of software's and techniques for social media data mining

Data mining applications are immensely useful for each and every company in global market. Several types of data mining areas are available such as: retail industry, financial data analysis, biological data analysis, telecommunication industry, intrusion industry and other scientific applications [17]. High quality data analysis is done in banking and financial industry with help of this financial data analysis in an organised manner. Financial data analysis process helps to maintain a systematic data analysis framework. Multidimensional data analysis and data mining processes are maintained with help of this design and construction of data warehouses. Prediction related to loan payment is managed by a company by this particular analysis. Financial data analysis helps to maintain customer credit policy analysis. Target markets and customers are maintained properly by a company to enhance their economic growth. Financial condition of an organisation is managed by this financial data analysis. In recent days, everyone uses social media on a daily basis to be up to date on a daily basis. Social media analytics market value is expected to grow 8.5 billion in the year of 2022 [18]. Usage of social media is increasing on a daily basis. This value will increase 26 billion dollars within the year of 2028. This creates a positive impact on economic growth of an organisation.
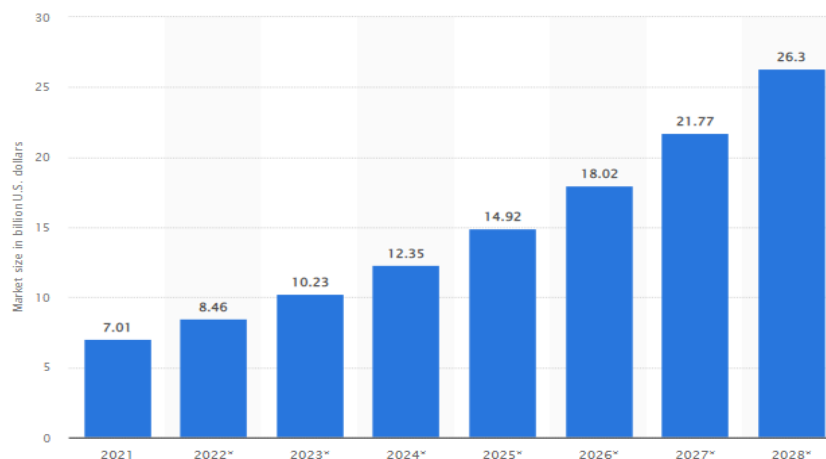

Figure 1: Social media data analytics market size

Great application of retail industry is also managed with help of this data mining process. This particular industry collects large amounts of data from sales, goods transportation, customer purchasing history and services with help of this data mining process. Quality and quantity of data are measured by a company to maintain their work process and strategy [19]. Every firm always wants to expand their business in different countries to enhance their financial condition. Availability and popularity of a firm is enhanced with help of this data mining process. Data mining process of retail industry helps to identify patterns and trends of loyal customers. In order to improve quality of customer services, retention and satisfaction of customers are maintained properly by a company with help of this data mining process. Telecommunication industry helps to provide various types of services such as: fax, pager, cellular phone and internet messenger. New computer and communication technology are also enhanced with help of modern technology and software's.

Issues and mitigation process related to social media data mining process

Security and social challenges are faced by an organisation in global market to maintain data mining processes in a significant way. Critical innovations for researchers and businesses in numerous spaces are managed with help of these data mining concepts [20]. A new database of a company is maintained properly with help of this particular data mining process. Impressive security of data is managed by a firm with help of these dynamic techniques in an organised manner. Private information's of about people is gathered by an organisation to maintain a strong security base and social challenges. Secrete data related to customers of a company is managed properly. Relevant and authentic tools are used by a company to maintain their security in a significant way. Process of obtaining huge volumes of data related to this study is managed with help of this noisy and incomplete data [21]. This process mainly refers to noisy, incompleteness and heterogeneity in global market. Several types of data are available in global market, and the company management team has a responsibility to identify their relevant and authentic data related to work process.

True data is normally put away at different stages of data mining process. Several types of data are collected from individual systems, internet and data databases. This data process is effectively beneficial for a company to maintain their financial growth. Complex data creates a negative impact on performance of an organisation. Natural language text, spatial data, temporal data and time series are included in this complex data analysis method [22]. Various data and necessary information related to data mining process are collected by a company management team to maintain their process of work. Presentation of data is a vital segment of a company to manage data mining process. In case a company cannot be able to maintain performance globally, speed and efficiency of work of this company is hampered.

Improvement of mining algorithms plays an essential role to maintain organisational performance. Incorporation of background knowledge is not gathered properly with help of this data mining process. Data visualisation process is also hampered due to lack of knowledge and skills related to data mining process.

DISCUSSION

Social media data mining process refers to a framework by which a company can easily communicate with their employees, customers and other stakeholders. Personalised interactions among stakeholders and employees are enhanced with help of this social media data mining process. Company management team always tries to fulfil demand of markets and customers to enhance their performance. This particular data mining process helps to support several types of research work in an organised manner. A large volume of public data is collected by a company to maintain quality and quantity of products. This process is immensely beneficial to maintain security of data related to public data globally. Technological, healthcare and social science related data magnitude are applied with help of this social media data mining process [23]. Trends of a business activity are maintained properly by this social media data mining process. Accuracy and authenticity of work is managed by an organisation with help of this data mining process. Social media platforms are used by everyone on a daily basis. In case a company provides their advertisements related to their products in social media platforms.

Customers can easily know about internal and external characteristics of a product with help of this advertisement policy. Every company has an aim and objective to enhance their financial growth. Hence social media data mining process helps an organisation for maintaining economic condition. Prediction of outcomes of an organisation is managed by this particular method of data mining. This process of social media data mining helps to maintain a traditional method of work, by which every employee can easily provide their better performance in workplace. A firm can provide better quality products and services in global market to attract more customers. Profitability and productivity of this company are increased by this process. Behaviour of a customer plays an important role to fulfil demand of customers and markets. A firm always tries to predict market and customer segmentation. Social media data mining process helps to predict those characteristics.

Several types of data mining techniques are mentioned in this study. Data cleaning process is included in this data mining technique. Data cleaning process involves to organise data, eliminate duplicate data and filling in null values. Useful and relevant information related to work is gathered by a company with help of this data cleaning process [24]. This process follows several steps such as: verifying data, converting data type, removing unnecessary data, eliminating errors and completing missing values. Each and every data point is gathered and verified by an organisation with help of

this data mining process. Uniform data sets must be managed by a company globally to maintain their work process. Irrelevant data are eradicated by a firm to prepare a fresh data set in an organised way. Company management team has a duty to fulfil missing values for all relevant and authentic data.

Data visualisation process is referred to as a data mining process within a company. Several types of methods are available in this data visualisation such as: maps, heat maps, histograms, comparison chats, network diagram, density plots and scatter plots. These visualisation processes are immensely helpful to understand visualisation concepts of data. This creates a positive impact on financial and organisational culture of an organisation. Classification and machine learning process are included in this data mining framework. Logistic regression, decision trees, naive bayes and support vector machines are used as methods of data mining classification. Semi supervised, supervised, unsupervised learning is referred to as machine learning processes. A business activity is enhanced with help of this machine learning process.

CONCLUSION

Social media data mining process helps a company to maintain their performance globally. For this reason, a company can easily enhance their rate of production and profitability with help of this process. Social media platforms are immensely helpful for each and every form to identify their target market and customers. In case a firm produces better quality products in market, this company can easily enhance their brand value and brand equity in an organised manner. Frequency, correlation and cross turbulence is managed by an organisation with help of this social media data mining process. Social media platforms are used by each and every customer on a daily basis. For this reason, a company can easily provide their discounts, offers related advertisements on this particular platform. Cross sectional" research design is followed by researcher in this study. Secondary data are collected by researcher with help of qualitative data analysis.

"Inductive" research approach is maintained by a researcher to gather several types of data related to this study. Theoretical analysis of a study is also managed by this particular approach in global market. In this study, social media data mining process of a research work is critically evaluated. Several types of data mining techniques are also mentioned here. Model based, grid based, density based and hierarchical methods are critically evaluated in this study. This creates a positive impact on economic growth and financial condition of an organisation.

REFERENCE

[1] Fedushko, Solomia, Taras Ustyianovych, and Michal Gregus. "Real-time high-load infrastructure transaction status output prediction using operational intelligence and big data technologies." *Electronics* 9.4 (2020): 668.

[2] Aljawarneh, Shadi, et al. "Particularities of data mining in medicine: lessons learned from patient medical time series data analysis." EURASIP Journal on Wireless Communications and Networking 2019.1 (2019): 1-29.

[3] Rashid, Yasir, et al. "Case study method: A step-by-step guide for business researchers." International journal of qualitative methods 18 (2019): 1609406919862424.

[4] Patel, Mimansha, and Nitin Patel. "Exploring Research Methodology." International Journal of Research and Review 6.3 (2019): 48-55.

[5] Ronkainen, Noora J., and Gareth Wiltshire. "Rethinking validity in qualitative sport and exercise psychology research: A realist perspective." International Journal of Sport and Exercise Psychology 19.1 (2021): 13-28.

[6] Belcastro, Loris, et al. "ParSoDA: high-level parallel programming for social data mining." Social Network Analysis and Mining 9.1 (2019): 1-19.

[7] Tien, Nguyen Hoang. "Customization and Standardization of the Business Strategy of Foreign Enterprises in Vietnam–the McDonald's Case and the Fast Food Sector."." International journal of research in marketing management and sales 1.2 (2019): 44-50.

[8] Sivarajah, Uthayasankar, et al. "Role of big data and social media analytics for business to business sustainability: A participatory web context." Industrial Marketing Management 86 (2020): 163-179.

[9] Abkenar, Sepideh Bazzaz, et al. "Big data analytics meets social media: A systematic review of techniques, open issues, and future directions." Telematics and Informatics 57 (2021): 101517.

[10] Ageed, Zainab Salih, et al. "Comprehensive survey of big data mining approaches in cloud systems." Qubahan Academic Journal 1.2 (2021): 29-38.

[11] Yoseph, Fahed, et al. "The impact of big data market segmentation using data mining and clustering techniques." Journal of Intelligent & Fuzzy Systems 38.5 (2020): 6159-6173.

[12] Schleder, Gabriel R., et al. "From DFT to machine learning: recent approaches to materials science–a review." Journal of Physics: Materials 2.3 (2019): 032001.

[13] Kaasinen, Eija, et al. "Empowering and engaging industrial workers with Operator 4.0 solutions." Computers & Industrial Engineering 139 (2020): 105678.

[14] He, Li, Bruno Agard, and Martin Trépanier. "A classification of public transit users with smart card data based on time series distance metrics and a hierarchical clustering method." Transportmetrica A: Transport Science 16.1 (2020): 56-75.

[15] Donthu, Naveen, and Anders Gustafsson. "Effects of COVID-19 on business and research." Journal of business research 117 (2020): 284-289.

[16] Streimikiene, Dalia, et al. "Sustainable tourism development and competitiveness: The systematic literature review." Sustainable development 29.1 (2021): 259-271.

[17] Priyanka, E. B., and S. Thangavel. "Influence of Internet of Things (IoT) In Association of Data Mining Towards the Development Smart Cities-A Review Analysis." Journal of Engineering Science & Technology Review 13.4 (2020).

[18] Statista Research Department.Social media analytics market size worldwide from 2021 to 2028. *Statista.* (2023) https://www.statista.com/statistics/1330450/global-social-media-analytics-market-size/. Accessed 9th January, 2023.

[19] Abbas, Jawad. "Impact of total quality management on corporate green performance through the mediating role of

corporate social responsibility." Journal of Cleaner Production 242 (2020): 118458.

[20] Williamson, Ben, Felicitas Macgilchrist, and John Potter. "Covid-19 controversies and critical research in digital education." Learning, Media and Technology 46.2 (2021): 117-127.

[21] Li, Chunquan, Yaqiong Chen, and Yuling Shang. "A review of industrial big data for decision making in intelligent manufacturing." Engineering Science and Technology, an International Journal 29 (2022): 101021.

[22] Dagaeva, Maria, et al. "Big spatio-temporal data mining for emergency management information systems." IET Intelligent Transport Systems 13.11 (2019): 1649-1657.

[23] Yang, Jin, et al. "Brief introduction of medical database and data mining technology in big data era." *Journal of Evidence-Based Medicine* 13.1 (2020): 57-69.

[24] Pejic-Bach, Mirjana, et al. "Text mining of industry 4.0 job advertisements." *International journal of information management* 50 (2020): 416-431.