International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

Survey On Data Mining Approach Of Social Media Analytics For A Brand Building

Ayush Soni*
Department of Computer Science, Laxmi Narayan college of technology (RIT) , Indore

Guided by- Kushal Joshi*
Department of Computer Science, Laxmi Narayan college of technology (RIT), Indore

Abstract-The data mining is the technique which is utilizes the different kinds of algorithms and methods that consumes data and extract the application oriented patterns. These patterns are used for various task decision making, prediction, recognition and others. In this work the ability of data is mining is investigated in the domain of business brand building and the business intelligence. In this context the social media data is involved for analysis and finding the different aspect of business growth and understanding future consumer's expectation and interest. In this work the overview of proposed research direction and the objectives of the work are demonstrated.

Keywords: Data mining, business intelligence, brand analysis, marketing, digital evaluation

I. INTRODUCTION

The business is influenced with the various local and global factors among them the marketing and branding of products and business is a key factor. In order to market the brand name or product various different strategies are used such as television advertisement, print media, email, SMS and others. But now in these days the trends traditional advertisement techniques are updated manner. Additionally more and more marketing and advertisement agencies are working on digital marketing and strategies for promotions and product marketing. Among them the social media sites are played an important role to reach the people who are the target consumers. In social media a user can be involved directly and indirectly in business promotions. For example, when a post through the particular brand is seen by someone it shows the insight, if that post is attractive and a user like the post or if it is helpful for other one then user share the post that activity is termed as the user engagement. In addition of that different other factors are utilized for promotions and marketing. Such as cross marketing and affiliate marketing and others. But in order to know how the business is growing, what is the

user interest and how the target people are reacting on the products and brands the analysis of these kinds of data is required. In this proposed work a data mining model is presented that help to understand market, consumer's interest which helps to prepare the future business plans and decisions for improvements.

II. RELATED WORKS

B. Singh, H.K. Singh, [1] Web mining is used to understand customer behavior, evaluate the effectiveness of a particular Web site, and help quantify the success of a marketing campaign.Vijiyarani, S., and E [2] Web content mining has number of research issues since it can remove the data from the web indexes

Sillitti, Alberto, et al [3] introduced Web content mining has alternate ways to deal with mine the information. These are unstructured text information mining, structure mining, and semi-structure text mining and media information mining

Bai, X., [4] proposed All things considered, thorough, quantitative examinations via web-based networking media content, especially on electronic trade and data administration, stay rare

Sodhganga et al [5] This 'wise' utilization of innovation can be through information mining which draws upon broad work in territories, for example, measurements, machine learning, design acknowledgment, databases, and superior registering to find intriguing and already obscure data in datasets

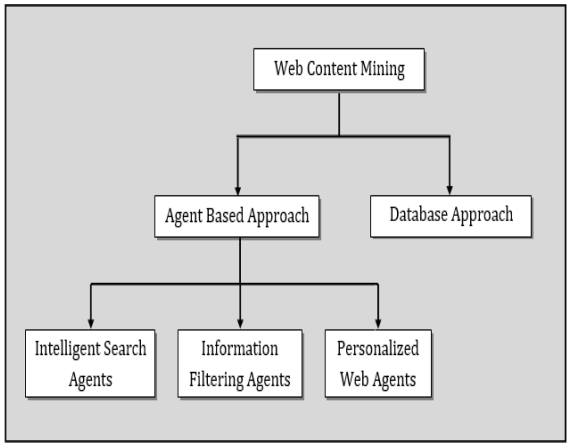
Jaideep Srivastava, Prasanna Desikan and Vipin Kumar, et al [6] proposed . Web mining it likewise assumes a vital part for E-Commerce and E-Service web webpage to comprehend their web destinations and administration are utilized and give better support of the two clients and clients. Hardly any applications are

Adnan Nadeem et al [7] Use of web mining is associate with the quick development of World Wide Web. Web mining turns into an exceptionally hot and mainstream point in web inquire about zone

III. WEB CONTENT MINING TECHNIQUES

The Internet has pulled in the consideration of research networks. Specifically, the critical part of breaking down internet based life and systems to propel our comprehension of data sharing, correspondence assessment arrangement, and dispersal has been perceived. All things considered, thorough, quantitative examinations via web-based networking media content, especially on electronic trade and data administration, stay rare Web content mining data may be composed or unstructured/semi sorted out notwithstanding the way that a considerable amount of web is unstructured. It is the path toward recouping the information from the web into more sorted out structures and

requesting the information to recuperate quickly or finding beneficial information from web substance or web records. Web content mining joins the web reports which may comprise of text, html, mixed media records i.e., pictures, sound, video and sound and so forth. The output mining contains the web list items. It might be a structure reports or unstructured records. Web content mining can be arranged in the accompanying system. This can be delineated utilizing figure.



IV. ISSUES ON WEB CONTENT MINING

Web content mining has number of research issues since it can remove the data from the web indexes.

- Data/Information Extraction center around extraction of sorted out data from web pages, for instance, things and recorded records.
- Web data coordination and composition coordinating. The web contains vast measure of information; every website acknowledges comparative data in an unexpected way. Comparable information revelation is a vital issue with loads of practical applications.
- Opinion extraction from online sources i.e. client ensures items, discussions, web journals and talk rooms. Mining assessments are of huge result for promoting insight and item benchmarking.

Naturally sectioning web pages and distinguishing clamor is an intriguing issue in web application. It couldn't have notices, route connections and copyrights takes note. Subsequently, separating the primary substance of the web page is critical issue in web application.

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

Figure -2

V. METHODOLOGY

The amount of main memory required to process the input data using the algorithm is known as memory consumption. Space Complexity of the system also termed as the Memory consumption in terms of algorithm performance. That can be calculated using the following formula: The performance of the implemented data mining approach of brand building is given using figure

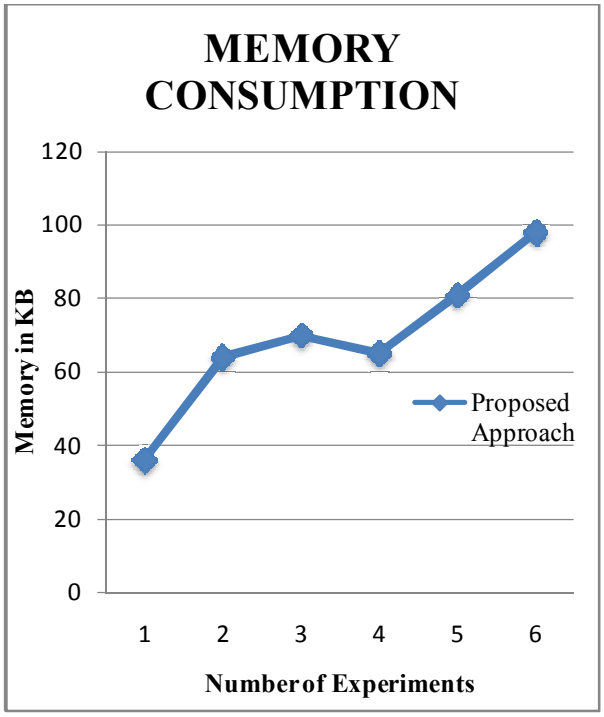


Figure -3 Space complexity

Memory Consumption

= Total Memory

– Free Memory

VI. SIMULATION RESULT

The amount of memory consumption depends on the amount of data reside in the main memory, therefore that affect the computational cost of an algorithm execution. data is numerically show by table 1. For clarification of the result, X axis of figure contains the different amount of code execution and the Y axis shows the respective

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

memory consumption during execution in terms of kilobytes (KB). According to the obtained results the performance of algorithm demonstrates similar behavior with input dataset. This consumed memory represents the required space by this algorithm process facebook dataset to produces efficient output.

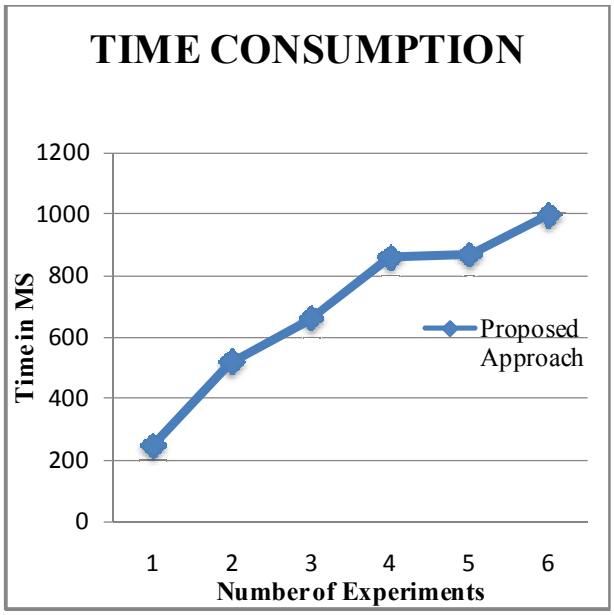
Table 1 Numerical Values of Space Complexity

TABLE I. Simulation Configuration

Number of experiment	Proposed Apporach
1	36
2	64
3	70
4	65
5	81
6	98

Time complexity: The amount of time required to calculate future trend of the social media analysis using dataset is known as the time consumption of the system. That can be computed using the following formula:

$$\text{Time Consumed} = \text{End Time} - \text{Start Time}$$



In this diagram the X axis contains the program execution of the system and the Y axis contains time consumed which is measures in milliseconds. Additionally, to demonstrate proposed and base approach we use blue and red line respectively. According to the evaluated performance of the proposed technique is process the future tend by their post types. For processing algorithm consume time which is illustrated in table 2 in numerically.

Table 4.2 Numerical Values of Time Complexity

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

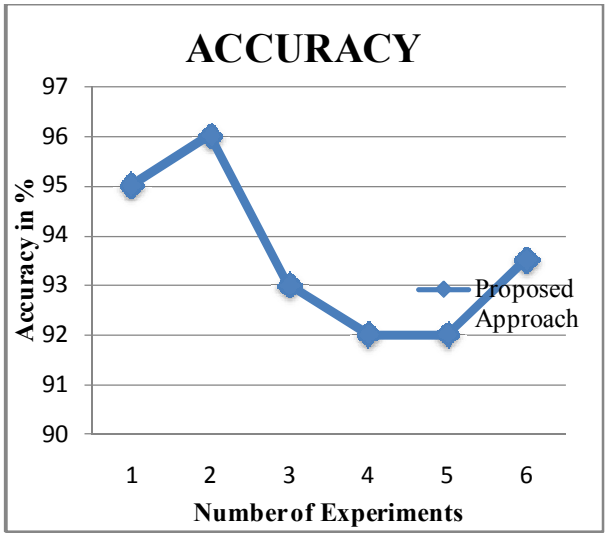
February 2021

Number of Experiments	Proposed Approach
1	248
2	521
3	663
4	861
5	870
6	998

Figure -4

Accuracy: The performance of the correctly classified patterns using trending future post is representing in terms of accuracy. The performance evaluation of proposed approach is evaluated using classification concept. The accuracy of the data mining approach can be evaluated using the following formula:

$$\text{Accuracy} = \frac{\text{Total correctly classified Patterns}}{\text{Total input Words to Patterns}} X100$$



The accuracy of the implemented proposed algorithm of predicting social media analysis on brand building is represented using table 4.3 and figure 4.3. The given figure 4.1 contains the accuracy of the implemented algorithms. The X axis of the diagram shows the different experiments and Y axis contains the obtained performance in terms of (%). To demonstrate the performance of the proposed technique is representing using blue line and traditional approach is representing orange line. This technique is evaluated on the basis of input facebook dataset.

Table 4.3 Numerical Values of Accuracy

Number of Experiments	Proposed Approach
1	95

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

2	96
3	93
4	92
5	92
6	93.5

According to the obtained results the performance of the proposed model provides more accurately recognized trending post. Additionally, the performance can be varying if we change the attribute of the input dataset.

Error rate: The amount of data of misclassified patterns during classification of algorithms is known as error rate of the system. This can also be computed using the following formula.

$$\text{Error Rate \%} = \frac{\text{Total Misclassified Patterns}}{\text{Total Input Patterns}} X100$$

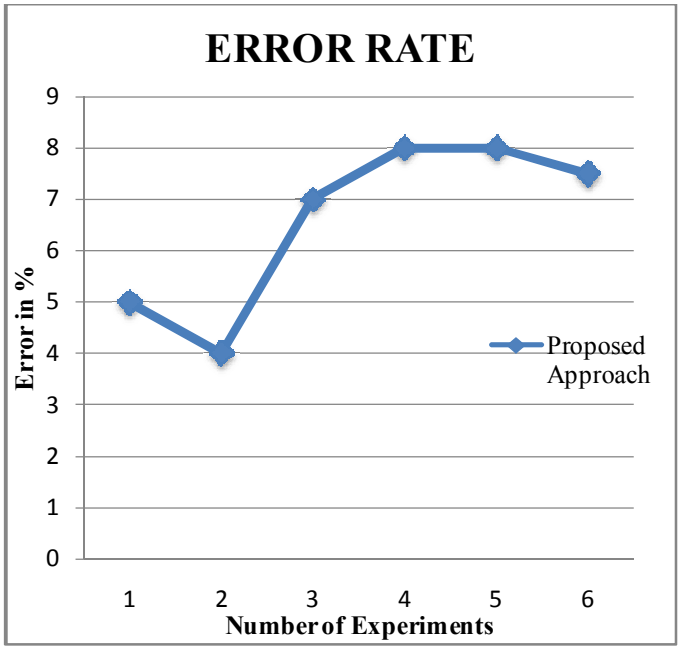


Figure -7

The figure 4.4 and table 4.4 shows the error rate of implemented algorithm of data mining. In order to show the performance of the system, X axis contains the experiments and the Y axis shows the performance in terms of error rate in percentage (%). The performance of the proposed data mining approach is given using the blue line. The performance of the proposed post type classification is effective and efficient during different execution and reducing with the amount of data increases of future trend post. Thus the presented classifier C 4.5 is more efficient and accurate for classifying future trend post

Table 4.4 Numerical Values of Error Rate

Number of	Proposed Approach

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

Experiments	
1	5
2	4
3	7
4	8
5	8
6	7.5

SYSTEM DOMAIN

Implementation of the required system utilizes software and hardware for successfully implementation is listed in this section.

(A) Tools-

User Interface Design (UI Design)-Net Beans IDE 8.0.2

(B) Technology/Framework-

Framework-JDK 1.8.0

(C) Hardware Specifications-

25 GB hard disk space

Minimum 2 GB RAM

Intel P4 Processor

(D) Software Specifications-

Operating System (Windows XP and above)

APPLICATION DOMAIN

The data mining is always help to understand the data patterns and the future and recent aspects for a target application.Therefore after successful implementation of the proposed social media analytic technique the following promising outcomes are expected.

1. Trends of business brand
2. How the demand of their products are increasing and decreasing
3. What the current user aspect and interest for their products and activities
4. Flow and influence of the users for the particular band.

Conclusion

The proposed technique helpful for measuring the impact of social media post impact and identifying the possible future popularity growth of any kind of content marketing. Therefore, the proposed model is acceptable for real world use with different applications. The main aim of the proposed

work is to analyzing the social media post user engagement data for predicting the new post trends for any brand or product. The implementation of the required data mining technique is performed successfully. In near future the following extensions are proposed for work.

1. Improving the decision making rules for achieving high accurate future trends of the posts.
2. Involving the text based analysis for estimating the user's orientation for any kind of post.
3. Implementation of hierarchical clustering technique for estimating the reason of high and low performance posts.

REFERENCES

[1] Moro, Sérgio, Paulo Rita, and Bernardo Vala. "Predicting social media performance metrics and evaluation of the impact on brand building: A data mining approach." Journal of Business Research 69, no. 9 (2016): pp. 3341-3351.

[2] Berry Michael W., (2004), "Automatic Discovery of Similar Words", in "Survey of Text Mining: Clustering, Classification and Retrieval", Springer Verlag, New York, LLC, pp. 24-43.

[3] Tan, Ah-Hwee. "Text mining: The state of the art and the challenges." In Proceedings of the PAKDD 1999 Workshop on Knowledge Discovery from Advanced Databases, Volume 8, pp. 65-70. 1999.

[4] Kumar, Lokesh, and Parul Kalra Bhatia. "Text Mining: Concepts, Process and Applications." Journal of Global Research in Computer Science 4.3 (2013): pp. 36-39.

[5] Stavrianou, Anna, Periklis Andritsos, and Nicolas Nicoloyannis, "Overview and semantic issues of text mining", ACM Sigmod Record 36.3 (2007): 23-34.

[6] S. Vijayarani Ms. J. I lamathi and Ms. Nithya, "Preprocessing Techniques for Text Mining: An Overview", International Journal of Computer Science & Communication Networks, Volume 5(1), pp. 7-16

[7] Vishal Gupta, "A Survey of Text Mining Techniques and Applications", Journal of Emerging Technologies in Web Intelligence, Volume 1, No. 1, August 2009.

[8] B. Singh, H.K. Singh, "Web data Mining Research", in the Proceedings of 2010 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), pp. 1-10, Dec. 2010.

[9] Shen zihao, Wang Hui, "Research on E-commerce Application Based on Web Mining", 2010, IEEE.

[10] Li Haigang and Yin wanling, "Study of Application of Web Mining Techniques in E- Business

[11] Sebastiani, F., "Machine learning in automated text categorization." ACM Computing Surveys, Vol. 34, No. 1, pp. 1–47, 2002

[12] Sonali Vijay Gaikwad and Archana Chaugule, "Text Mining Methods and Techniques", International Journal of

Computer Applications (IJCA), Volume 85 – No 17, January 2014.

[13] Jaideep Srivastava, Prasanna Desikan and Vipin Kumar, "Chapter 21- Web Mining — Concepts, Applications, and Research Directions", http://dmr.cs.umn.edu/Papers/P2004_4.pdf

[14] S. Vidya, K. Banumathy, "Web Mining- Concepts and Application", International Journal of Computer Science and Information Technologies, Volume 6, Issue 4, 2015, pp. 3266-3268.

[15] "Chapter 1: Introduction", Sodhganga, available online at: http://shodhganga.inflibnet.ac.in/bitstream/10603/11075/5/05_chapter1.pdf

[16] "Web Content Mining: Tutorial given at WWW-2005 and WISE-2005", available online at: https://www.cs.uic.edu/~liub/WebContentMining.html

[17] Faustina Johnson and Faustina Johnson, "Web Content Mining Techniques: A Survey", International Journal of Computer Applications (IJCA) Volume 47– No.11, June 2012

[18] Bai, X., "Predicting Consumer Sentiments from Online Text," Decision Support Systems, Vol. 50, No. 4, pp. 732-742, 2011.

[19] Sillitti, Alberto, et al. "News miner: A tool for information retrieval", 7th International Conference on Intelligent Engineering Systems, 2003.

[20] Vijiyarani, S., and E. Suganya, "Research issues in web mining", International Journal of Computer-Aided Technologies (IJCAx), Volume 2, PP. 55-64, 2015.