PREDICTING THE FUTURE TRENDS BY MINING THE SOCIAL WEB: A SURVEY

Sevinç Ömer[1], Güzel Serdar M.[2] & İman Askerbeyli[3]
[1]Computer Engineering Ondokuz Mayıs University Samsun, Turkey
[2,3]Computer Engineering Ankara University Ankara, Turkey

ABSTRACT

Social network is an important part of people's daily life, which essentially can be considered as a black box, providing great insight into the current trends of people. It has been proved that data obtained from Social Network platforms, including Facebook, Twitter, Instagram, Baidu etc., can be employed in a variety of different fields. Technically, social networks can be shown as nodes and edges, inspired from graph theory, this theory play an important role to express the structure and relations of the network. Big data mining, natural language processing (NLP), text mining and machine learning techniques are mainly utilized for processing and analysing the data. The analysis of social web can reveal the real trends of the people on any subjects, can detect terror activities before attacks, can help to understand political tendencies, cultural or global believes etc. This study is a review of social web mining that aims to examine the literature of social web mining and help to understand the techniques. However this study also proposes some novel approaches by taking into accounts data of the images on the web which will enrich the handled information while mining the social web. The study also points to the semantic web for reaching more clear inferences from social networks than at present.

Keywords: *Social web mining, big data, text mining, graph theory, natural language processing, semantic web*.

1. INTRODUCTION

Personalization [1] which are interpreting, sharing of feelings, opinions, and thoughts lies [2] in the philosophy of social Networks. Nowadays very huge data flows over social networks. A critical manner is required for first acquiring these data and then processing them to get valuable information. However, it is a fact that social networks are portrayed on a graphical theorem, in which individuals are nodes in social networks and edges that indicate relationships between them. Accordingly, social networks can be analyzed on graph theory by determining the nodes and edges for target groups to identify the structure of them as subgroups. On the other hand, the fact that big data arising from the flow of information between these nodes can be processed with data mining techniques to reveal different patterns, rules and trends [3]. Under the concepts of opinion mining, sentimental analysis and clustering very important inferences can be made [4]. For instance, clustering on a graphical model can reveal links, nodes, and subscripts of a community. Besides, the information in the data obtained from the shares can be analyzed and interpreted and new deductions can be made. According to recent statistical studies, the number of social network users is expressed in billions and increasing day by day, as seen in Fig.1 [5]. In the aforementioned image, the data obtained from 2016 is illustrated, detailing total number of social network users. For instance, in 2012 alone, the content produced by social networks and web based applications is more than content of all written documentations during previous 6000 years until then [6]. Analyzes obtained using social network mining will shed light on future health, education, trade, politics and minds. For example, any terrorist organization or attack can easily be detected even when the initiative is still in the preparation phase. It can be quickly determined in which regions a disease is spreading or political tendency in a country can be deduced that the thoughts about a political issues are positive or not. Regarding with increasing use of mobile devices and rapid social sharing and the ability to make inferences in the analysis of these data prove that the future of the social network mining is so brilliant.

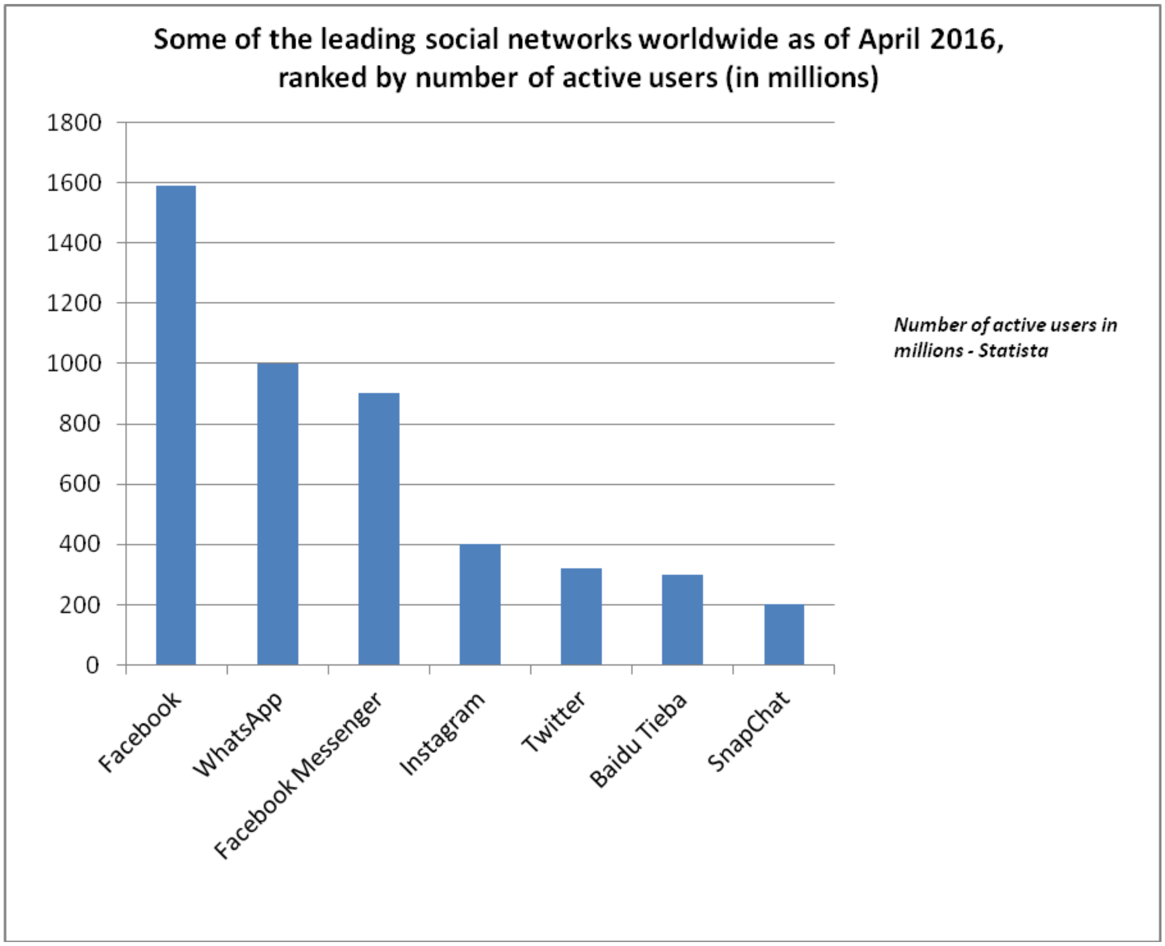

Figure.1 Active User Numbers of Leading Social Networks by April 2016

1.1. Social Network Mining Components

The process of obtaining the data from the sources is the first step of the data mining section of the social network mining problem. Afterwards, the data is pre-processed and cleared so to create corresponding document word matrices. Analyzes of document similarity are made using those matrices [7]. At this point, workspaces such as text mining, machine learning, etc. are introduced which are able to analyze the current situation or create prospective predictions [8]. On the other hand, in the study of nodes and edges in social networks, it is tried to find people with the most effective publishing power in a social network by means of centrality measurement by going over the theory of graphs [9]. Social networks are presented as nodes and edges by the theorem of graphs, and then the associated subdivisions or communities are tried to be found by using the clustering methods [10]. Hierarchical clustering can follow a strategy either from top to bottom or from top to bottom. On the other hand, the data on the web increase drastically over a short period of time. As well as data is diversified. Consequently this data is called big data for the definition. At this point, researches regarding the big data field also play an important role in social networking [11]. However for some analysis works to be done as understanding the emotional situations of people's comment about any subject, the natural language processing is introduced which provides that the human languages are understood and interpreted by the computers [12]. However, studies such as analyzing texts with text mining, passing frequency values through statistical studies, or doing similarity analyzes through document matrices can be performed, thus all these constituting the steps of the data mining [13]. Therefore, all of these work field titles that is considered in this study can be included in the context of the study and which are also critical for social networking.

1.2. Big Data

It's the bigger thing computers can not handle. This data is constantly growing. A better definition of this is the "5V" definition, namely, volume, velocity, variety, veracity and value [14]. When it comes to big data, it is actually the greatest volume as the first dimension to the mind. The second 'V' is Velocity. For example, a large number of new messages are written every day on facebook, and a certain processing speed must be caught to process them. It is also possible to mention the variety of the data which is the third of the 5V concept. Briefly, different sources come from different contents with different characteristics. Therefore, big data has volume, speed and resource diversity [15]. Therefore, in big data studies, we try to solve the problems of volume, diversity, speed and reliability expressed by 4 V, and valuable information that can be obtained, that is to say the value that will result in 5 V, is produced [16]. The daily Wiki data that can be expressed at Terabytes level can serve as a good example for big data [17].

1.3. Natural Language Processing

Natural language processing (NLP) is an area of research that allows computers to analyze and understand natural language texts or human spoken language. The aim of researchers in this area is to develop tools and techniques that enable computers to understand and manipulate natural languages, such as understanding and using natural languages, so that they can fulfil their duties [18]. Natural language processing involves computer science and disciplines such as informatics, mathematics, electrical and electronics engineering, artificial intelligence, robotics and psychology [19]. Besides, it involves Applications comprising computer translation, natural language text processing and summarization, user interfaces, CLIR, speech recognition, artificial intelligence and expert systems [20]. Another very important area of work is to be able to take full advantage of the resources available through the Internet and digital libraries using multilingual language processing and the CLIR [21]. Natural language processing is used to capture the meaning of texts, such as whether the texts obtained in web mining are negative or positive or the summary of it in short to be able get foremost meaning, and therefore they can establish a significant step in the process of the social network mining.

1.4. Text Mining

Text mining is a popular data mining operation and considered as one of the important pillars of the social network mining. It employs text data as the only source and aims to obtain structured data via text. Classification of texts, segmentation, subject extraction from texts, production of class particles, emotional analysis, text summarization, entity relationship modelling can all be considered in this working field. In order to reach these goals, several different methodologies are employed such as information gaining, syllable analysis, word frequency distribution, pattern recognition, labelling, information extraction, data mining and even visualization within the scope of text mining studies [22].

1.5. Graph Theory

Graph Theory is mainly used to show nodes and nodes groups in a network and details relationships between them. In terms of social sciences, these demonstration influencers and followers are given names. In other words, in a social network, other sources have this effect while others have a resource efficiency dissemination feature. Especially in very big databases, the use of the graph theory has an even bigger proposition [23]. Although the diagram theory is used to visually show a social network in books and other sources, it is possible to operate directly on the data without visualizing the problem. Although the methods used are graphical theorem methods and problems can be solved through matrices, tables or equations [24]. The centre of centrality measurement allows us to find the user with effective emission power. This approach is based on a simple computation based on the simple theorem, the parametric centre analysis methods [25] are applied by simple matrix operations through the creation of the α-centrality metric or connection matrix which based on a simple operation as counting the number of attenuated paths. These methods do not need to visualize while being implemented [4].

1.6. Clustering

Each node represents a person as social networks are designed as graphs, Subgroups, in which more than one node come together, can be thought of as a cluster. At this point, there are different algorithms for how many clusters there are. The simplest and most basic of these is the k-means algorithm. In the K-means algorithm, cluster centres are determined, out-of-centre samples are classified according to their distances, new centres are obtained according to the classification made, and this process continues until the centres become fixed [26]. The division of the society into sub-communities can be dealt with in several different ways. For instance, the Girvan-Newman algorithm divides social constructs into sub-social constructs according to their intensities [27]. On the other hand, a different method of community extraction in a very detailed approach is discussed in [28]. Hierarchical clustering actually unifies the power of independent groups by combining many other methods and divides the network into subgroups.

Here, it is expected that how many segments are to be included or some parameters in the algorithm, for example threshold values, should be defined. As it has been previously mentioned, one of the simplest and primitive algorithms used for segmentation is the k-means algorithm. The algorithm forms the cores of segments according to a value of k average of segments. For this algorithm, the number k must be given as the default. Figure.3 shows the example of clustering by using K-means algorithm in which the clustering process is performed by giving the number of cluster information.

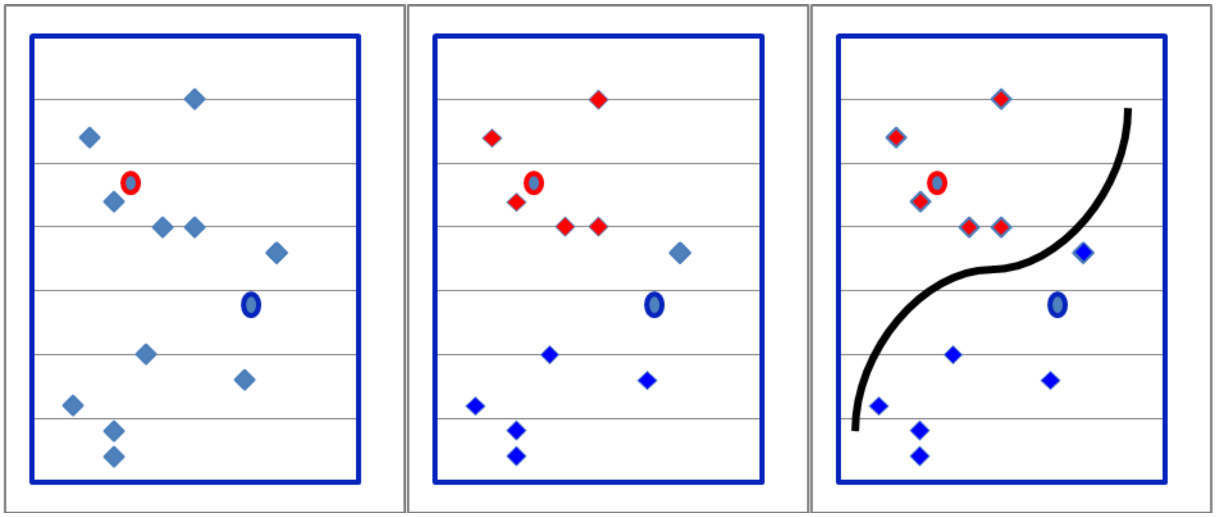

Figure.3 K-means Clustering Example Demonstration

2. PREVIOUS STUDIES

When the social network mining studies are evaluated based on their contents, It is revealed that the data obtained from the web is cleared and the text mining passes through the steps such as training the machine so as to create new knowledge or prediction using the corresponding analyzes results . On the other hand, by the theory of graphs, it is tried to determine the communities related to each other by accepting the relations between the individuals which are nodes and the relations between them are edges in the social network. Here, clustering is performed and meaningful segments are formed on the graph. K-means is one of the most well-known segmentation algorithms and the work done with this algorithm [29] is available. One of the biggest problems in the cluster is how many clusters will be determined. In this context, for example [30], the connections are primarily defined in an empty network with a hierarchical segmentation approach, which can be used in previous studies when segmented in the graph theory, and the initial value is assigned as 0. These associations are then weighted according to the degree of relationship. The highest weighted individual is the centre, since the process from highest to weakest weight is performed.

As the type of relationship weakens, the likelihood of error increases so that it is not true that the individual with a link belongs to that group. Alternatively, the Girvan-Newman algorithm [27] makes the exact opposite of the hierarchical algorithm's segmentation. Instead of starting with centralized individuals, they start with individuals who have remained in the vicinity and have low weighted connections. This continues to work by squeezing the remaining individuals. Again, contrary to the hierarchical segmentation, instead of adding a new individual to the system at every step from an empty network, Girvan Newman takes an individual at every step in his algorithm. There are also other clustering methods and studies. For instance, in addition to traditional methods such as graph partitioning, vertex similarity, hierarchical clustering, partitioned clustering or spectral clustering, the new approaches as overlapping communities, statistical methods, or capture of varying communities methods can be implemented too [28]. On the other hand, the Girvan-Newman algorithm proposes a new method to calculate vertex betweenness, which is considered as a critical innovation. The concept of middle individual has been studied for centrality for a long time [24]. The middle individual value for any individual is the number of the shortest paths that pass through corresponding individual. Let’s assume that n nodes is defined in a network, while middle individuality is calculated, the number of edges connecting the shortest path in Girvan-Newman algorithm is calculated, when nodes are normally counted on the shortest route from one node to another (See Figure.4)

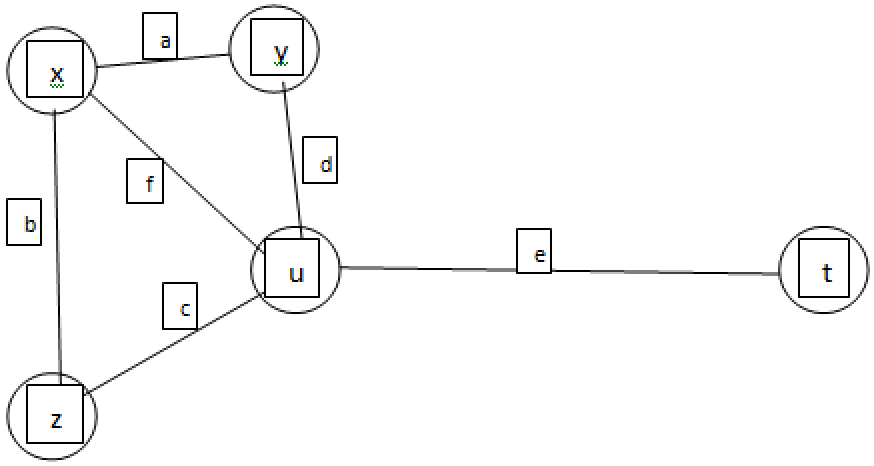

Figure - 4 Example Nodes and Paths Between them (Relations)

In the Girvan-Newman algorithm, the median individual evaluation varies slightly and is calculated by the number of the shortest paths that passes over for any connection that connects two individuals with the concept of edge-to-edge. x-y: {a} x-z: {b} x-u: {f} x-t: {f, e} y-z: {a, b} or {d, c} y-u: {d} y-t: {d, e} z-u: {c} z-t: {c, e} u-t: {e} lists the edges showing the shortest path between nodes . Then the algorithm counts that how many times each edge has passed; a: 2 b: 1 or 2 c: 2 or 3 d: 2 or 3 e: 4. As a result, the busiest link is located at "e side". If more than one connection has the same number of shortest paths, these connections can be related to each other to achieve a longer connection [27]. The link with the highest medium connection value is removed and thus the segmentation is achieved. Then the intermediate connection values are recalculated for the remainder, resulting in a top-down tree. This algorithm produces better results in segmentation. The approach called HITS and HUBS is another algorithm of graph theory that solves the connections between web pages. It has been widely used in search algorithms, including wiki, blog and social network analysis applications [6].

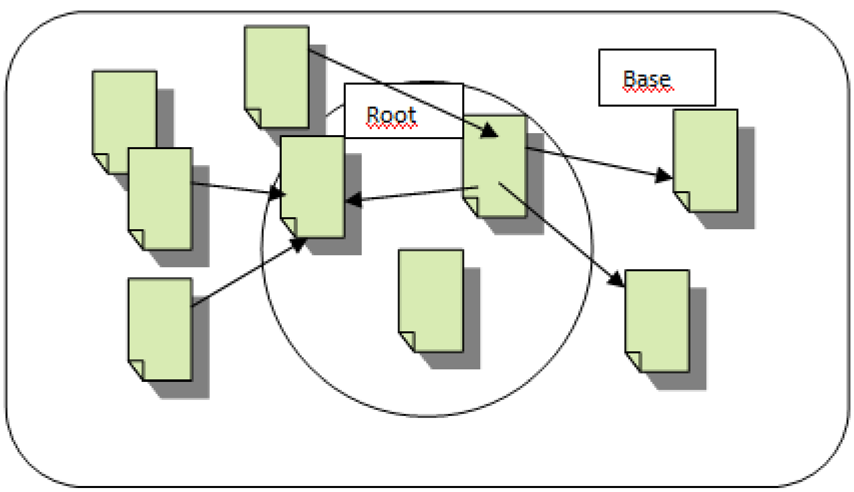

Figure-5 is the HITS and HUBS Algortiması Sample Scheme.

This algorithm which is another graph theory algorithm, designed based on the principle that find outs the links between web pages. The algorithm is executed over two sets, one of which is the root and the other is the base set. While searching web documents in the root cluster, the documents that carry the best results are kept in the base set, and documents referring to the documents in the root set are linked. Then, a two-step calculation is performed in order to calculate the number of web documents under influence, that is, the number of web documents under the influence, and in the second step, the effect values of the documents affecting the other web documents are summed to obtain HITS values. These are mostly works on the side of the graph theory where the nodes are evaluating the web pages, while there are studies that analyze the contents obtained on the side of the tree and make predictions about the findings and forward. Examples include text mining, intellectual mining [31], text summarization, social network mining (Liu, D and Friends 2016), opinion extraction, and relationship rule extraction [32]. For example, it is the idea mining work to show semantically the idea of social media sharing. Attempts are performed to determine the formation of ideas on a community. The idea is that it is personal, not the truth, but the ideological case is checked. In general, it is determined whether there is a positive opinion about the subject within the context of emotional analysis [33]. While e-mail analysis and simple text mining problems can be categorized, there is now a sophisticated classification of where different opinions differ, or where the seller is, and to what extent the margin is affected by sales [34]. The frequency based idea mining is done through frequency of words such as number of words, name, adjective, adverb, or verb while ideas are derived from the texts. It is based on the recognition that ideas are expressed in terms of the relevant words. In a survey conducted in 2011, it was found that a significant proportion of ideas (60-70%) were based on the words in the text [35]. It has also been found that frequently recurring names in opinion-based mining reflect the views of the people. Another approach is to search for specific templates in the text to achieve certain results. For example, it is possible to obtain a view from the values coming to A by finding patterns such as "great A", "with A", "consist A" and then it is possible to make idea mining with polarity analysis [36]. The frequency-based idea mining can also be achieved by labelling words according to their linguistic. The concept of POS-Tagging is to label words in a text according to their properties such as name, adjective, and even number adjectival, special name. On the other hand, the model-based idea mining methods are applied to untagged texts that are previously labelled and thus formed from texts known to contain an idea, and try to come up with ideas from unconventional texts [37]. In another study it is developed a new hidden Markofian model with a subset approach and created a statistical model based on changing weights based on their relevance to the position of view and emotion level values [38]. An important part of the newly emerging opinion extraction work is a hybrid model arising from the use of view-based, frequency-based, or model-based approaches and graph theory. It includes approaches in which users having the same idea are grouped together in the same cluster and social network. Clustering could be done according to age groups or sex by means of intellectual mining methods in studies using a method that can also be evaluated for peer-review studies [39]. However, because of the unpredictable change of the web environment, the necessity of studying dynamic methods has been revealed. In order to overcome these problems the use of web 3.0 semantic networks come along with new web technologies that data analysis and inference will gain a very different dimension [40]. Because the semantic webs express the concepts (words) according to their contents on the web, the texts, videos, paragraphs, texts, block HTML tags and web content, the data analysis has gained a different dimension as it defines the relations of the connections between the entities [41]. Extracting useful information from the web is the most significant issue of concern for the realization of semantic web. This may be achieved by several ways among which Web Usage Mining, Web Scrapping and Semantic Annotation plays an important role. Web mining enables to find out the relevant results from the web and is used to extract meaningful information from the discovery patterns kept back in the servers [42]. Recent studies show that advanced mining, opinion, emotion analysis as well as data mining studies are carried out on semantic web and the conversion of exponentially increasing data into meaningful information in social network mining is being continued [43].

3. OVERALL CONCLUSIONS

As a review article some specific points of web mining subjects are evaluated which focused on making improvements in the web mining research area. Some valuable researches are considered such as social web mining, terror activity dedection and health protection depending on core methods in web mining which are machine learning, information retrieval, text mining, natural proceessing language and graph theory. It is pointed that if image data as well as semantic data on the web is used to produce hidden information beyond them, big contributions will be made [44]. Ragarding to all these, rapidly growing popularity of social networks and data sharing have led to attempts to make further inferences, especially to develop new methods for computer scientists. By applying big data mining and statistical methods, data pre-processing and machine learning to the data coming from social networks from the web environment, it is possible to make valuable information available to the public so that forward-looking predictions can be presented [45]. Rapidly growing popularity of social networks and data sharing have led to attempts to make further inferences, especially to develop new methods for computer scientists. By applying big data mining and statistical methods, data pre-processing and machine learning to the data coming from social networks from the web environment, it is possible to make valuable information available to the public so that forward-looking predictions can be presented [46]. The necessity of processing and monitoring the data on the social network effectively seems to be undeniable. Although data mining techniques include computer science as a field of statistics and application, it is thought that the techniques used on social networks in the future will be fed from different disciplines such as behavioural sciences, social sciences or business. If you imagine a web environment that runs on semantic networks with web 3.0 in the coming days, it is not hard to imagine that the data mining will go far beyond the understanding of today's data through meaningful data to be obtained from it [47]. It is understandable how important these studies are when we consider important information that can be obtained from education, politics, economy, security and health issues. Knowledge is power. Table-1 shows an overview and the most outstanding publications referenced in this survey.

Table- 1 Summary of milestones in web data mining.

Authors	*Category*	*Strategy*	*Outstanding Methods*
[1],[9],[10],[23],[27],[28], [32],[35],[40]	Social Web Mining	User – Groups Detection	Graph Theory, Clustering
[8], [9],[13],[22],[31],[34]	Text Mining	Sentiment Analysis, Summarization	Classification, Frequency Analyze
[12],[18],[19],[20],[21],[35]	Natural Language Processing	Get the main idea of the article.	Information retrieval, Artificial Intelligence
[1],[6],[32],[34]	Data Mining	Make future predictions, show unknown relations of the data	Information retrieval, Machine learning
[3],[11],[13],[14],[15],[16],[25], [45],[46]	Big Data Mining	Make future predictions, show unknown relations of the streaming huge data	Information retrieval, Machine learning, Parallel programming, Map and reduce
[7],[40],[42],[43],[44],[47]	Semantic Web Mining	Get meaningful inferences over annotated data	Semantic Web, SparQL queries

4. REFERENCES

[1]. Asur, S., Huberman, B. A.,. Predicting the future with social media. In Web Intelligence and Intelligent Agent Technology (WI-IAT), IEEE/WIC/ACM International Conference on Vol. 1, pp. 492-499 (2010).

[2]. Pang, B., & Lee, L. Opinion mining and sentiment analysis. Foundations and trends in information retrieval, 2(1-2), 1-135 (2008).

[3]. Provost, Foster, and Tom Fawcett. "Data science and its relationship to big data and data-driven decision making." Big Data 1.151-59 (2013).

[4]. Seker S.E. Çizge Teorisi (Graph Theory), YBS Ansiklopedi, v.2, is.2, pp. 17-29 (2015).

[5]. Anonymous. Web Site: http://www.smartinsights.com/social-media- marketing/social-media-strategy/new- global-social- media-research/ Access Date : 11.03.2017 (2008)

[6]. Seker S.E. Sosyal Ağlarda Akan Veri Madenciliği, YBS Ansiklopedi, v.1, is.1, pp. 21- 26 (2014).

[7]. Sevinc, O., Huang, L., Loughzang, L., & Kilic, E. Geospatial Information Retrieval Base on Query Expansion and Semantic Indexing. Journal of Engineering and Fundamentals, 2(2), 51-68 (2015).

[8]. Alexander Pak, Patrick Paroubek. Twitter as a Corpus for Sentiment Analysis and Opinion Mining. In LREc (Vol. 10) (2010).

[9]. Borgatti, S. P., & Everett, M. G.. A graph-theoretic perspective on centrality. Social networks, 28(4), 466-484 (2006).

[10]. Wu, Z., & Leahy, R. An optimal graph theoretic approach to data clustering: Theory and its application to image segmentation. IEEE transactions on pattern analysis and machine intelligence, 15(11), 1101-1113 (1993).

[11]. Zikopoulos, P., & Eaton, C. Understanding big data: Analytics for enterprise class hadoop and streaming data. McGraw-Hill Osborne Media (2011).

[12]. Manning, C. D., Surdeanu, M., Bauer, J., Finkel, J. R., Bethard, S., & McClosky, D. The stanford corenlp natural language processing toolkit. In ACL (System Demonstrations) (pp. 55-60) (2014, June)..

[13]. Wu, X., Zhu, X., Wu, G. Q., & Ding, W. Data mining with big data. ieee transactions on knowledge and data engineering, 26(1), 97-107 (2014).

[14]. Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Byers, A. H. Big data: The next frontier for innovation, competition, and productivity (2011).

[15]. Seker, S. E. Büyük Veri ve Büyük Veri Yaşam Döngüleri YBS Ansiklopedi, v.2, is.3,pp.1-8 (2015)

[16]. Mayer-Schönberger, V., & Cukier, K.. Big data: A revolution that will transform how we live, work, and think. Houghton Mifflin Harcourt (2013).

[17]. Anonymous. Web Site: http://www.smartinsights.com/social-media- marketing/social-media-strategy/new- global-social- media-research/ Access Date: 22.11.2016 (2007)

[18]. Hirschberg, J., & Manning, C. D. Advances in natural language processing. Science, 349(6245), 261-266 (2015).
[19]. Lehnert, W. G., & Ringle, M. H. Strategies for natural language processing. Psychology Press (2014).
[20]. Nadkarni, P. M., Ohno-Machado, L., & Chapman, W. W. Natural language processing: an introduction. Journal of the American Medical Informatics Association, 18(5), 544-551 (2011).
[21]. Chowdhury, G. G. Natural language processing. Annual review of information science and technology, 37(1), 51-89 (2003).
[22]. Seker, S, E, Mert Cihan, Al-Naami K, Ozalp N, Ayan U, Correlation between the Economy News and Stock Market in Turkey., International Journal of Business Intelligence and Review (IJBIR), vol. 4, is. 4, pp. 1-21, (2013)
[23]. Scott, J. Social network analysis: developments, advances, and prospects. Social network analysis and mining, 1(1), 21-26 (2011).
[24]. Van Steen, M. Graph theory and complex networks. An introduction, 144 (2010).
[25]. Ghosh, R., Lerman, K.. Parameterized centrality metric for network analysis. Physical Review E, 83(6), 066118 (2011).
[26]. Tsuda, K., & Kudo, T.. Clustering graphs by weighted substructure mining. In Proceedings of the 23rd international conference on Machine learning (pp. 953-960). ACM (2006, June)
[27]. Girvan, M., & Newman, M. E. Community structure in social and biological networks. Proceedings of the national academy of sciences, 99(12), 7821-7826 (2002).
[28]. Fortunato, S. Community detection in graphs. Physics reports, 486(3), 75-174 (2010).
[29]. Saeed, F., Salim, N., Abdo, A., & Hamza, H. Using graph-based consensus clustering for combining K-means clustering of heterogeneous chemical structures. Journal of Cheminformatics, 5(1) (2013).
[30]. Nascimento, M. C., & Carvalho, A. C. A graph clustering algorithm based on a clustering coefficient for weighted graphs. Journal of the Brazilian Computer Society, 17(1), 19-29 (2011).
[31]. Dey, L., & Haque, S. M. Opinion mining from noisy text data. International Journal on Document Analysis and Recognition (IJDAR), 12(3), 205-226 (2009).
[32]. Adnan, M., Nagi, M., Kianmehr, K., Tahboub, R., Ridley, M., & Rokne, J. Promoting where, when and what? An analysis of web logs by integrating data mining and social network techniques to guide ecommerce business promotions. Social Network Analysis and Mining, 1(3), 173-185 (2011).
[33]. Pang, B., & Lee, L. Opinion mining and sentiment analysis. Foundations and trends in information retrieval, 2(1-2), 1-135(2008).
[34]. Hung, S. Y., Yen, D. C., & Wang, H. Y. Applying data mining to telecom churn management. Expert Systems with Applications, 31(3), 515-524 (2006).
[35]. Liu, B. Chapter 11, Opinion Mining and Sentiment Analysis. B. Liu içinde, Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data (s. 459-526). Springer (2011).
[36]. Popescu, A.-M., & Etzioni, O. Extracting product features and opinions from reviews. Proceeding HLT Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing , s. 339-346 (2005).
[37]. Seker, S. E. Müşteri Kayıp Analizi (Customer Churn Analysis) YBS Ansiklopedi, v.3, is.1, pp. 26-29 (2015).
[38]. Lakkaraju, Himabindu, et al. "Exploiting coherence for the simultaneous discovery of latent facets and associated sentiments." Proceedings of the 2011 SIAM international conference on data mining. Society for Industrial and Applied Mathematics, (2011).
[39]. Jackson, M. O.. Social and economic networks . Princeton University Press (2010).
[40]. Chelmis, C., & Prasanna, V. K. 2011. Social networking analysis: A state of the art and the effect of semantics. In Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE Third International Conference on Social Computing (SocialCom), IEEE Third International Conference on (pp. 531-536). IEEE (2011).
[41]. Sevinc O., Kılıc E. XBRL Bilanço ve Gelir Tablosu Analizi ve Semantik Web'e Uyarlanması, XVI. Türkiye'de İnternet Konferansı, pp 105-114 (2011).
[42]. Malik, S. K., & Rizvi, S. A. M. Information extraction using web usage mining, web scrapping and semantic annotation. In Computational Intelligence and Communication Networks (CICN), International Conference on (pp. 465-469). IEEE (2011, October).
[43]. Madhu, G., Govardhan, D. A., & Rajinikanth, D. T. Intelligent semantic web search engines: A brief survey. arXiv preprint arXiv:1102.0831 (2011).
[44]. Velásquez, J. D., Dujovne, L. E., & L'Huillier, G. Extracting significant website key objects: A semantic web mining approach. Engineering Applications of Artificial Intelligence, 24(8), 1532-1541(2011).
[45]. McAfee, A., Brynjolfsson, E., Davenport, T. H., Patil, D. J., & Barton, D. Big data. The management revolution. Harvard Bus Rev, 90(10), 61-67 (2012).
[46]. Boyd, D., & Crawford, K.. Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. Information, communication & society, 15(5), 662-679 (2012).
[47]. Rettinger, A., Lösch, U., Tresp, V., d'Amato, C., & Fanizzi, N. Mining the semantic web. Data Mining and Knowledge Discovery, 24(3), 613-662 (2012).