

KNOWLEDGE DISCOVERY IN DATABASE USING INTENTION MINING

Ayesha Rashid, Umar Shoaib, Muhammad Shahzad Sarfraz

Department of Computer Science, Faculty of Computing and Information Technology, University of Gujrat, Pakistan
 ayesha.rashid@uog.edu.pk, umar.shoaib@uog.edu.pk, shahzad.sarfraz@uog.edu.pk

ABSTRACT –Text mining is a frame work to retrieve valuable knowledge from unstructured form of textual documents. Extracted knowledge presents in the user understandable form of facts and knowledge. Text mining further classified into information retrieval, NLP, statistics, web mining and Intention mining. An intention is a human mental state represents a current or future action. Intention mining is an up-and-coming research area of text mining describes what actually customer wants and what actions he can take in future. It explicitly finds what people want to happen not just what they like or dislike. This paper aims to take a deep insight on intention mining by categorizing user intentions like sale / purchase, wish, emotional, search and real-time intentions. It will help the upcoming researchers to take a view on research have done on intention mining and compare the adopted approaches to find the optimal method. User may express the intention implicitly or explicitly. Explicit intention is the direct explosion of user’s wishes which can easily detect from text documents. Implicit intentions communicated indirectly by user in perspective of other features of related object. Multiple classification, clustering, keyword based and machine learning techniques are used on different datasets to extract the user intentions. It is analyzed that now a days the most frequent used dataset for intention mining is micro blog tweets and frequent used techniques are support vector machine and Naïve Bayes with maximum accuracy rate.

Keywords: Data Mining, Intention Mining, Intentional Process, Text Mining, Knowledge Discovery

I. INTRODUCTION

In this era of digital world, millions of customers visit the social media to sale or purchase a product, to leave comments on political personalities, to share the ideas and feelings with friends and family [1]. In the result of such communication, a bulk of text documents generated consists of interested hidden knowledge [2]. Such hidden patterns and connections are extracted using Knowledge discovery technique from textual documents [1, 2]. Pattern’s clarity for humans, usefulness and novelty are the parameters used to judge the quality of extracted knowledge [3]. To extract the hidden facts, targeted dataset selected from problem domain and data are preprocessed by removing the irrelevant information like HTML and XML tags and DBMS schema constructed. Data mining model like summarization, classification, regression or clustering are applied on selected dataset to discover knowledge. Then results are interpreted and if needed go backward to the concerning steps, remove the irrelevant patterns and report the extracted knowledge [1,2]. Data preprocessing is most concerning to Text Mining [2] which converts the textual form of data into more suitable for data mining algorithms [4]. Text mining uses multiple disciplines: data mining, machine learning techniques and computational linguistics to form mining algorithms [5]. Some researchers define text mining as a tool to discover the new knowledge from a huge volume of natural language text using computational algorithms [6]. The most recent promising research area of text mining is intention mining used to detect the future intentions of users on the basis of their past activities [12]. Intention mining is an up-and-coming research area of text mining describes what actually customer wants and his possible activities in future. It explicitly defines the wish and want of user instead of liking and disliking [13]. In this paper the main emphasis is on the categorization of the intentions into different forms of human psychological state. To detect the intentions of people from textual documents, researchers have used a variety of datasets. The current digital world has facilitated the users to express their reviews and comment on current affairs to share ideas. Those textual

documents have been used to extract the useful knowledge in the form of user intentions. Datasets like product reviews, twitter blogs, web logs, surveys, Facebook comments have been used for intention modeling [14]. This paper made observations on the basis of current research that the most frequently used dataset to detect intentions of users is twitter blogs. Multiple classifications, clustering and machine learning techniques like naïve bayes, support vector machine, KNN algorithm have been used with good accuracy rate. But support vector machine and naïve Bayes classifier generated approximately 82% good results as compared to other techniques.

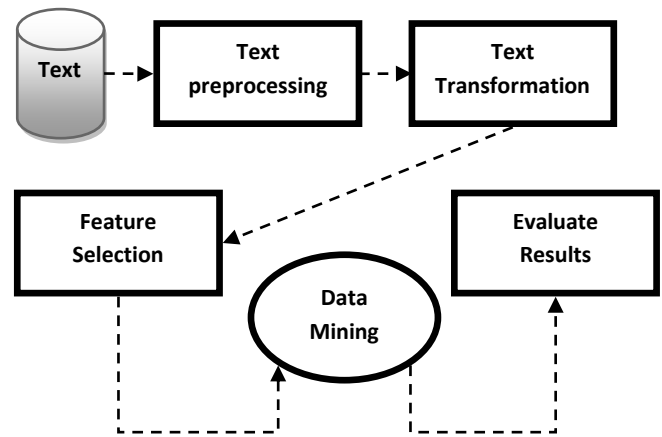


Figure 1: Text Mining Process

II. TEXT MINING

Text mining is an emerging research area of computer science [3]. It collaborates with natural language processing, data mining, web mining, educational mining and machine learning in the formation of information retrieval and knowledge management [3,4]. Estimation is that 85% of business, science, and medicine related documents are in unstructured form of text [5]. Text mining aims to extract the meaning, concepts, relationship of concepts, hidden patterns and text indexing from the targeted datasets and present them in the user understandable form of facts and knowledge [2]. Figure 1 defines the complete process of text mining. First of all the unstructured textual datasets are saved in database and all the unnecessary tags, sentences and special characters are removed. Break the paragraphs in the form of sentences to make symmetry of sentences. Then the unstructured data is transformed into structured form to separate the features along with their concerning parameters. Data mining techniques like classification, clustering, machine learning and natural language processing are applied on the selected attributes to extract the concerning knowledge. In the end, the results are evaluated in the form of accuracy or precision, recall to judge the methodology adopted to locate the knowledge [15].

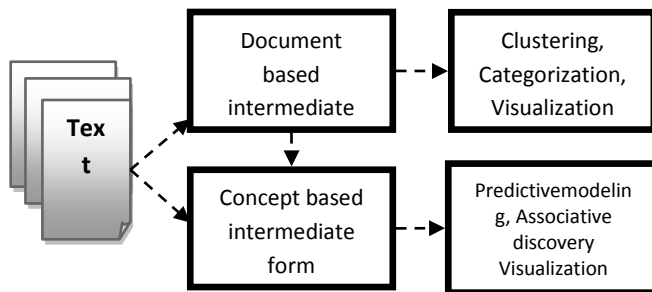


Figure 2: Text Mining Model [5]

After preprocessing the text is converted from the unstructured form of data into intermediate form (IF) as shown in figure 2. IF can be document based or knowledge based. Document based IF extract the knowledge or patterns across document used to organize the documents according to contents [7]. Concept based IF extract knowledge or patterns on the basis of objects used to extract the specific organization's behaviour knowledge. Figure 2 defines Complete Text mining model to process the text mining from selecting the targeted dataset to pattern discovery and interpretation of evaluated knowledge [6]. Facts are stored in relational database [8] for the analysis. Multiple tools and techniques are available in text mining to capture logic, fuzzy, unknown patterns and knowledge from text in efficient manner [9].

TEXT MINING TECHNIQUES

The textual datasets are analyzed using different text mining techniques like information retrieval, natural language processing, information extraction, summarization, clustering and visualization etc. Following section briefly described the text mining techniques

Information Retrieval (IR)

Information retrieval is the most promising research area to find the text document related to the requested question. For automatic retrieval of the related document from the huge collection of text documents, statistical methods and measures are used [11]. IR deals with the whole process of data retrieval to knowledge retrieval. These document retrieval systems are frequently used in search engine [12].

Natural language processing (NLP)

NLP is used for the better understanding of natural text, includes the reliable fast text processing techniques which range from the simple string manipulation to natural language processing [13].

Information Extraction (IE)

Information extraction aims to find the specific information from the entire document i.e. to find the answers of the asked questions from the related documents [12]. The extracted answers and patterns are stored in database as knowledge database for future use [14].

Computational Linguistic Research (CLR)

The most important research area of text mining is computational Linguistic Research, aims to automatically extract the meaning, facts and knowledge from text [6]. The basic idea is that if computer learn the rules of grammar and usage then it can easily derive the meaning of words being used in document [5]. Although it is a promising research area but the problems arise with long sentences [15].

Figure 3 depicts the Text mining as an interdisciplinary method used in different fields like machine learning, web mining information retrieval, information extraction, statistics, computational linguistic and data mining [5, 6].

TEXT MINING APPLICATIONS

The main applications of text mining are enriching content, systematic literature review, discovery, and computational linguistic research [6]. Publishers are more ambitious to make their contents (text) more machines accessible [7]. Text mining can be used to find the fixed vocabulary words from text and extract that in which parts of the contents that words appears and how many times [5]. Those facts can be used for further analysis and to discover the hidden patterns [8]. Text mining can also be used to conduct the literature review of related fields [9]. Such type of text mining helps the scientists, engineers and other users by creating the summary of all the work which has been done in their related fields [7, 8]. It helps the researchers to keep up date with the current research areas of their field and reduce the risks to miss current knowledge [10]. Utopia Docs [12] is the tool which helps the scientists to get the most recent literature review in the related field.

III. INTENTION MINING

In cognitive science "An intention is a thought of a person before taking an action" [10]. Intention mining is an prevailing research area describes what customer wants the system to do. It explicitly finds what people wish to happen not just what they like or dislikes [11]. Intentions or suggestions are depicted by the sentences written by users to express their views. For this purpose Hidden Markov Models are most emerging to find

the human intentions from text [14]. Viterbi Algorithm and Baum-Welch Algorithm (BWA) are most commonly used intention mining algorithms [15]. Yet there is no intention mining tool developed to discover the aim behind activities [14, 15]. Social media data widely expresses the user’s expressions, desires, likes and dislikes. Researchers are ambitious to expose the user’s intentions from the social conversation platforms like twitter, Facebook and chat rooms [16]. Figure 4 defines the complete process of intention mining. Crawl the dataset containing textual data and preprocess it by removing the unwanted sentences, characters and tags. Transform the data in required form to apply the proposed approaches of data mining in supervised or unsupervised format. Now the dataset is ready to prepare the intention map, pseudo map or to be divided in intention categories [17, 18].

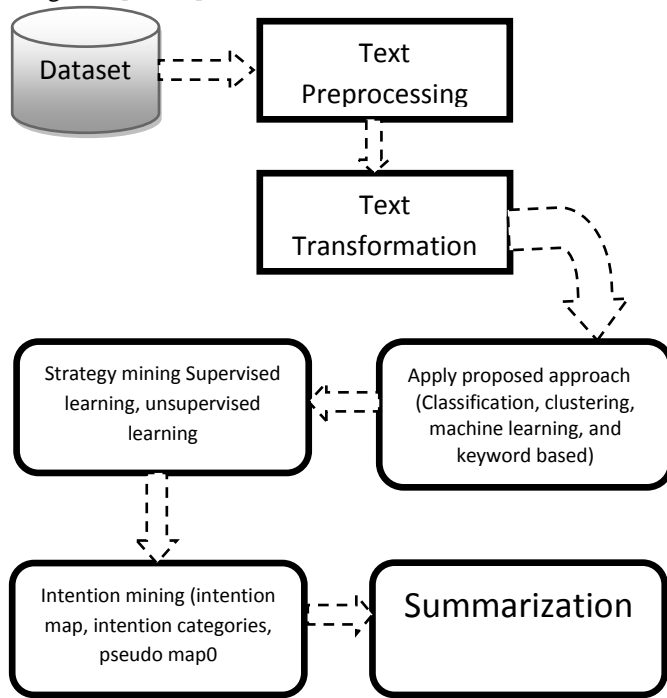


Figure 3: Intention Mining Process

IV. CATEGORIES OF INTENTION

Intention can be expressed in multiple ways. Users feel free to share their ideas with other people on web, they used free writing style format to express their feelings. In such scenario, it becomes difficult and ambiguous to categories the intentions from free format text. In this paper, intentions are categories in five groups according to the nature of written sentences. The five categories are described as following

SALES / PURCHASE INTENTIONS

The Consumption Intention Mining Model developed [9] to mine the user consumption intentions from social media to indicate the implicit or explicit user intentions towards sale or purchase of product. They use the SinaWeibo twitter data consist on 76 million sentences and extract 245 intention words for baby and kid corpus and 78 intention words for movie reviews.[10] Proposed way to categories the users comments into ten intentional categories. They used a user comments corpus and apply simple machine learning techniques to automate the classification of intentions. Commercial intentions detects from tweets to link up buyers

and sellers for industry benefits. Twapper-keeper used to extract the tweets containing any commercial keyword like piece, discount etc. 16 keywords set was created and per keyword 100 tweets annotated then using Weka tool preprocess the data and apply n-gram part of speech tagging. They achieved up to 57.1% precision and 77.4% recall using Naïve Bayes classifier for classification task [14].

Learning algorithms are used to develop a set of New Year’s Day wishes from a corpus of wishes for the domain of product and politics [21]. Models to improve the detection of Suggest and purchase wishes are studied in [30]. A new rule based approach introduced in [17] to categorize sale and purchase intentions. Rules are tested on multiple corpus like product reviews, customer surveys and comments from consumer forums. The wish to suggest and the wish to purchase using variants of Class Sequential Rules (CSRs) studied in [31]. Search intents are attempted to differentiate in [32] by using interaction features such as mouse movements or scrolling behavior. Commercial intents are detected in queries partly served as a pre-processing step to analyze the correlation of click-through behavior and rank/location of sponsored links or ads [33].

WISH INTENTIONS

Wish described what user actually wants instead of his liking or disliking. A keyword strategy formed to find the wish sentences from product reviews extracted from amazon.com [16]. Extracted wish sentences formulated in keyword set denoted as K and remaining as candidate wish sentences denoted as C, then C sentences judge manually whether a sentence is wish or not denoted as S. Sequential pattern mined from S set of sentences using sequential pattern mining then used patterns as features to train a classifier which detect wish sentences [11]. A rule based method used in [15] to extract wish sentences from industrial product reviews which help the product managers to take insight in consumers mind. Rules were based on moral verb (would, could, should etc), auxiliary verb and positive opinion. They detect purchasing and suggestion wishes on the basis of rules constructed with the combination of keywords.[17] Showed that search query logs represented a viable, yet largely untapped, source for acquiring knowledge about human goals.

Four types of wish detectors proposed in [18] to take insight of world’s wants and desires. They analyzed 80,000 English wish sentences of New Year. First two wish detectors found the top ten wish templates like “I wish to ___”, “I want ___”, “I desire ___” etc and use them as features in product reviews. Strict and non-strict matching were made using extracted feature templates. Third wish detector is based on SVM and fourth wish detector is based on word + template using technique of support vector machine.[19] Proposed an ontological based work to find the human goals from web and develop a knowledge base. Social psychological theoretical framework classified the sentences into 135 intention categories to find the descriptive wish phrases. Wish phrases associated with extracted actions to formulate the knowledge base using Yahoo BOSS API. They segmented the product reviews into set of sentences, each sentence is entered in knowledge base and using default Lucene similarity measure, the most similar entry is saved into intent profile. Results

show the proposed method is quick and effective for intention EXTRACTION.

EMOTIONAL INTENTIONS

Emotions of instant messaging using emotion annotation and feature engineering are mined in [12]. They worked on six types of emotions Sad, Disagree, fear, joy, anger, surprise and natural. They used a set of classifiers like Naïve Bayes, J48, Decision Tree, IBK and SVM to classify the emotions. Results reveal that their proposed system outperforms the results as compared to existing systems. An adaptation modal proposed to find the intentional factors towards the airline e-ticketing by conducting a questionnaire survey. "Saving, Effort, Expectancy and Facilitating Conditions" -in sequence- have significant positive effects on airlines passenger's intentions to use online e-ticketing systems in Thailand [16].

SEARCH INTENTIONS

A method proposed to mine the user search intention and build an intention map according to needed information. From the query search log document the intention features (nouns) were extracted and using SOM clustering algorithm user intentions was classified. Then collaborative filtering, query search and query clustering was applied to represent the intention map [20]. User's intentions are extended hierarchically in [19] through a previous vector model, sorted the information into positive or negative documents with SVM and suggested a way to recommend documents proper to a user's intentions [10]. However, the model cannot consider pattern learning of early queries and may affect badly a search function without a user's participation because it expands a query and reflects its pattern though user feedback [21].

REAL TIME INTENTION

Real time intentions detected from twitter data by implementing a heuristic based ensemble approach on feature knowledge base. The knowledge base developed by using wide range of linguistic and statistical techniques. Content indicative and usage indicative words used to detect real time intention from the topic of Movie, Sports, Music, Food and Dance and developed four feature extractors and on the basis of features Real time intentions and Non-Intention sentences were classified using Ensemble classifier [22]. People used micro-blogs to search or share information, post their activities to stay connect with people of their interest. They

used the two-level framework to find the intentions of users to use micro-blogs. The main user intentions detecting on micro-blogs are daily chatter, conversation, sharing information and reporting news [25]. A classification model was developed to make prediction which user will switch from traditional telephone service to internet telephony service on the basis of user intentions. They used logistic regression, classification methods and neural network techniques to develop the method which find the user switching intentions. Proposed method suggests some customer relationship management strategies that can be helpful to find the management strategies [27].

Twitter data gathered over ten cities and make analysis to express how user show interest in real time micro-blogs and how text mining techniques can be applied to find the user intentions from their tweets. They extracted the meaningful keywords and find the association between these words to find the intention sentences. Then classification techniques are used to classify tweets according to user intention [28]. Development Emails Content Analyzer classified the written documents of email according to defined six categories ("feature request, opinion asking, problem discovery, solution proposal, information seeking and information giving"). They proposed "Development Emails Content Analyzer" to classify the email contents according to intent by manually reading and observing the nature of each sentence [29].

V. CLASSIFICATION OF ARTICLES BY DATASET TYPE

Dataset is gathered in different forms from real time observations of surroundings. Multiple methods are used to gather data like warehouses, surveys, interviews, social Blogs etc. Table 1 categories the research articles used in this paper on the basis of dataset used for intention mining. Products reviews, social Blogs, Web history Political reviews, Questionnaire surveys, Twitter data and Chat logs are major data sets used to detect intention mining. This is the era of tweets and Blog to share informations and updates quickly with friends and family. Following Table clearly shows that the most frequently used dataset for intention mining is tweets blog.

Table 1: Classification of Articles by Dataset Type

Dataset	Research paper	Quantity
Tweets	[9],[14],[22], [25], [28], [29], [23]	7
Facebook comments	[10]	1
Product reviews	[11], [15], [18], [19], [30]	5
Questionnaire Survey	[16], [27], [24]	3
Web Search Log	[21], [31], [32], [17], [19], [20]	4
Political reviews	[18]	1
Chat logs	[12]	1

Figure 5 reflects the classification of articles by quantity of used dataset as follows:

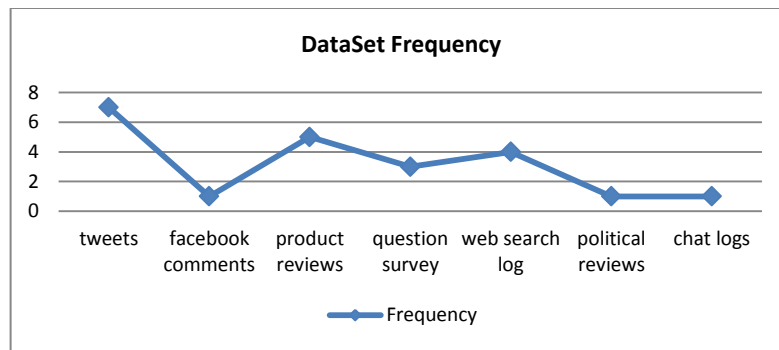


Figure 4: classification of articles by dataset

CLASSIFICATION OF ARTICLES BY TECHNIQUES

Multiple data mining techniques have been used to detect human behavior in form of intentions from textual documents. Researchers frequently used Classification, rule based, clustering, neural network and machine learning techniques to achieve the required results. In this paper, it is analyzed numerously used tools and techniques to depict intentions from

textual document. Such information is represents in table 2 and table 3.

Table 2 describes the techniques used in intention mining with respect of their evaluated results by using parameters precision and recall. It clearly defines that Support Vector Machine is most promising technique used in intention detection with high precision 0.92 and recall 0.99.

Table 2: Techniques Used in Intention Mining With Precision & Recall

Research Articles	Techniques	Precision	Recall
11	SVM using Naïve Bayesian (NB) classifier	0.73	0.78
14	Naïve Bayes, J48, Decision Tree, IBk and SVM	0.74	0.90
15	Rule based analysis	0.69	0.72
17	Support Vector Machine and Naïve Bayes Classifier	0.77	0.63
20	TF/IDF weight, SOM algorithm	Mean: 5.56	
21	Support Vector Machine	F Score 21% improve then Google	
30	Algorithm based on class sequential rule, Naïve Bayes Classifier	0.47	0.96
31	SVM Classifier	0.92	0.99

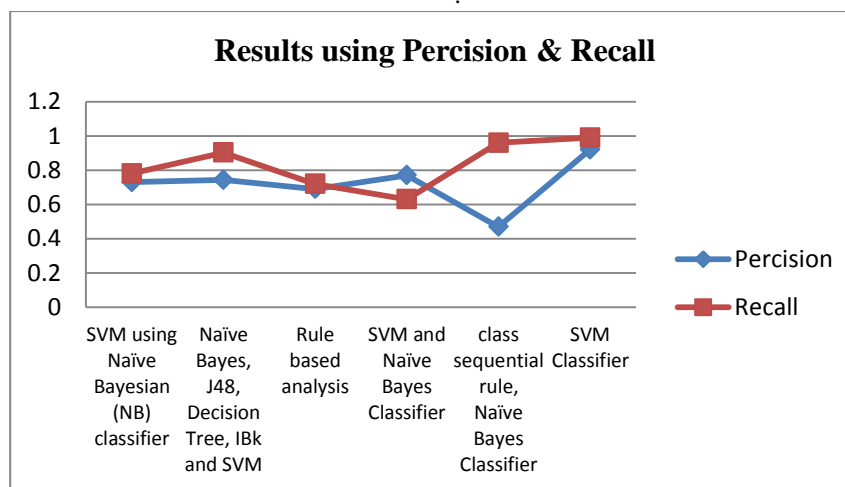


Figure 5: Intention Mining Process

Figure 6 represents the used techniques and generated results in term of precision and recall

Baradwaj, B. K. and Pal, S, Ding, X. , Liu, T., Duan, J .Nie,J.Y *et al* used accuracy as evaluated parameter of used techniques.

Table 3 depicts the accuracy along with used method to detect

different types of intention from textual documents. It is clearly represents that the support vector machine is more robust to mine the human behavior. SVM achieve 72 % of accuracy rate in intention modeling.

Table 3: Techniques Used in Intention Mining With Accuracy

Research Articles	Techniques	Accuracy
9	CIMM based on neural network	71.51
10	Naive bayesian (NB) classifiers, maximum entropy (ME) classifiers, and support vector machine (SVM) classifiers	66.97
16	Analysis using SPSS	69.5
18	Support Vector Machine	72
19	Lucene similarity measure	59.5
22	Keyword based intention extraction	70.0
25	Keyword based intention extraction	69.0
28	Keyword based intention extraction	65.5

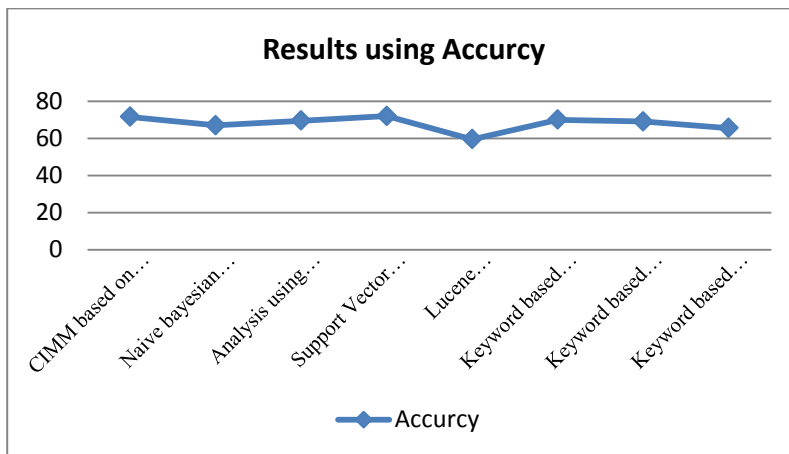


Figure 6: Intention Mining Process

CONCLUSION

This paper described the review of intention mining in the perspective of knowledge discovery. Intention mining is a promising research area of text mining used to predict the current and future actions of user. This article categorized the human intentions into five categories. Multiple social blog datasets are used to detect the intentions and it is analyzed that most recurrently used dataset is tweets blogs. Classification, clustering, neural network and keyword based techniques are used in said field and achieve high accuracy performance but Support Vector Machine has the maximum successful rate with 72 % accuracy, 0.99 recall and 0.92 precision with the respect of accuracy to predict user intentions.

REFERENCES

1. Basharat, S and Farooq. A, “Opinion Mining: Issues and Challenges (A survey)” International Journal of Computer Applications (0975 – 8887) **Vol. 49**– No.9, July(2012)
2. Baradwaj, R. J. and Pal. S “Mining Educational Data to Analyze Students Performance”

3. Prof. Maimon. O and Dr. Rokach. L“Data Mining and knowledge discovery Handbook” Maimon, Oded; Rokach, Lior (Eds.) 2nd ed, XX, 1285p. 40 illus. (. 2010)
4. Shelke, N. M., Deshpand, S. and Thakre, V. “Survey of Techniques for Opinion Mining” International Journal of Computer Applications (0975 – 8887) **Vol. 57**– No.13, November (2012)
5. Clark, J. and Loosdrech and Netherlands “Text Mining and Scholarly Publishing” Research Consortium (2013)
6. Nilanjan B., Dipanjan C., Anupam J., Sumit M., AngshuR.,andBalaraman Ra. “Mining of Sentence Level Opinion Using Supervised Term Weighted Approach of Naïve Bayesian Algorithm”, IJCTA | MAY-JUNE (2012)
7. Ayesha R. ,Naveed, A. and Muhammad S. “A Survey Paper: Areas, Techniques and Challenges of Opinion Mining” IJCSI International Journal of Computer Science Issues, Vol. 10, Issue 6, No 2, November 2013 ISSN (Print): 1694-0814 | ISSN (Online): 1694-0784
8. Ayesha R. ,Naveed, A. “Feature Based Opinion Mining of Online Free Format Customer Reviews Using Frequency Distribution and Bayesian Statistics” Networked Computing and

- Advanced Information Management (NCM), 2010 Sixth International Conference on 16-18 Aug. 2010.
9. Baradwaj, B. K. and Pal, S. "Mining Educational Data to Analyze Students Performance" (IJACSA) International Journal of Advanced Computer Science and Applications, **Vol. 2**, No. 6, (2011)
 10. Ding, X. , Liu, T., Duan, J. Nie, J.Y. "Mining User Consumption Intention from Social Media Using Domain Adaptive Convolutional Neural Network," AAAI'15 Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence Pages 2389-2395 (2015)
 11. Suja ,C. arlos,C. adhulikaM. andYalamanchi,"Intention Analysis for Sales, Marketing and Customer Service", Proceedings of COLING 2012: Demonstration Papers, pages 33–40, COLING 2012, Mumbai, December (2012).
 12. WU. X., HE, Z. ,"Identifying Wish Sentence in Product Reviews", Journal of Computational Information Systems **Vol. 7** issue 5 1607-1613(2011)
 13. Shashank and Pushpak, B."Emotion Analysis of Internet Chat", Proceedings of ICON-2008: 6th International Conference on Natural Language Processing (2008).
 14. Sudhir,R. "A Survey on X'ing and Intelligent Systems" www.iiste.org ISSN 2222-1719 (Paper) ISSN 2222-2863 (Online) **Vol 2**, No.6,(2011)
 15. Hollerit,B. Kröll, M. Strohmaier,M. "Towards Linking Buyers and Sellers: Detecting Commercial Intent on Twitter", WWW 2013 Companion, May 13–17, 2013, Rio de Janeiro, Brazil. ACM 978-1-4503-2038-2/13/05 (2013).
 16. Ramanand, J. ,Bhavsar,K. andPedanekar. N."Wishful thinking: finding suggestions and 'buy' wishes from product reviews", In CAAGET '10 Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text, Stroudsburg, PA (2010).
 17. Premchaiswadi,W. andPorouhan,P. "Factors affecting the Passengers' Intention toward "Airline Electronic Ticketing" In Thailand ICT and Knowledge Engineering (ICT & Knowledge Engineering), 2011 9th International Conference on 12-13 Jan. 2012 pages 177 – 186 (2011).
 18. Strohmaier,M. Kröll,M. "Acquiring knowledge about human goals from Search Query Logs", Information Processing & Management **Vol. 48**, Issue 1,Pages 63–82 January (2012).
 19. Andrew,B. , Goldberg, N. Fillmore, D.Andrzejewski, Z. Xu, B. Gibson andZhu,X. "May All Your Wishes Come True: A Study of Wishes and How to Recognize Them", The 2009 Annual Conference of the North American Chapter of the ACL, pages 263–271 (2009)
 20. Kröll,M. andStrohmaier,M. "Analyzing human intentions in natural language text Proceedings of the fifth international conference on Knowledge", capture Pages 197-198 ACM NewYork (2009).
 21. Park, K., Lee, T. ,Jung, S. and Nam, S. "Extracting Search Intentions from Web Search Logs"Information Technology Convergence and Services (ITCS), 2010 2nd International Conference on 11-13 pages 1 – 6Aug. (2010)
 22. Park, G.,JinGiChae, D. H. Lee and Lee, S. "User Intention based Personalized Search: HPS(Hierarchical Phrase Serch)", the WSEAS International Conference on Applied Computing Conference, pp.205–210 (2008).
 23. N. Banerjee, D.Chakraborty, A. Joshi, S. Mittal, An. Rai, B.Ravindran "Towards Analyzing Micro-Blogs for Detection and Classification of Real-Time Intentions"(2012)
 24. Godbole,S.,andBhattacharya.I." Building Re-usable Dictionary Repositories for Real-world Text Mining". In Proc. of CIKM In(2010)
 25. Peng.W., and Park, D. "Generate Adjective Sentiment Dictionary for Social Media Sentiment Analysis Using Constrained Nonnegative Matrix Factorization". In Proceedings of ICWSM In (2011)
 26. Java,A.,Song,X. Finin,T. andTseng,B. "Why We Twitter: Understanding Microblogging Usage and Communities", Proceedings of the Joint 9th WEBKDD and 1st SNA-KDD Workshop (2009).
 27. Shrihari,C. R and Amish D. Article: "A Review on Knowledge Discovery using Text Classification Techniques in Text Mining". *International Journal of Computer Applications* **111(6)**:12-15, February (2015).
 28. Ho Ha S. , Yang J. "Classification of switching intentions toward internet telephony services: a quantitative analysis"Information Technology and Management **Vol. 14**, Issue 2, pp 91-104June 2013.
 29. Nilanjan B., Dipanjan C., Anupam J., Sumit M., AngshuR.,andBalaraman Ra. "User interests in social media sites: an exploration with micro-blogs"CIKM '09 Proceedings of the 18th ACM conference on Information and knowledge management Pages 1823-1826 ACM New York, NY, USA (2009)
 30. Goldberg,A. B., Fillmore,N., Andrzejewski,D.,Xu, Z., Gibson, B., and Zhu, X. "May all your wishes come true: a study of wishes and how to recognize them?" InNAACL'09Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics, pages 263–271, Stroudsburg, PA. Association for Consumer Research (2009).
 31. Wu, X. and He, Z. "Identifying wish sentence in product reviews". Journal of Computational Information Systems, **Vol.7(5)**:1607–161(2011).
 32. Guo, Q. and Agichtein, E. "Ready to buy or just browsing?: Detecting web searcher goals from interaction data" In Proceedings of the International Conference on Research and Development in Information Retrieval (2010)
 33. Ashkan, A. and Clarke, C. "Term-based commercial intent analysis". In Proc. of the International Conference on Research and Development in Information Retrieval (2009).