2018 5th International Conference on Behavioral, Economic, and Socio-Cultural Computing (BESC)

User Intention Mining in Bussiness Reviews: A Review

Anam Habib
Institute of Computing and Information Technology
Gomal Univesity
D.I.Khan,KP,Pakistan
anamhabib19@gmail.com

Furqan Khan Saddozai
Institute of Computing and Information Technology
Gomal University
D.I.Khan,KP, Pakistan
furqan.saddozai@hotmail.com

Anum Sattar
Institute of Computing and Information Technology
Gomal University
D.I.Khan,KP, Pakistan
anumsattar786@gmail.com

Aurangzeb Khan
Department of Computer Science, University of Science and Technology
Bannu, Pakistan
aurangzeb.ustb@gmail.com

Ibrahim A. Hameed
Faculty of Information Technology and Electrical Engineering
Norwegian University of Science and Technology(NTNU)
Alesund, Norway
ibib@ntnu.no

Fazal Masud Kundi
Institute of Computing and Information Technology
Gomal University
D.I.Khan,KP, Pakistan
fmkundi@gmail.com

Abstract*—In recent times, the rapid use of social media sites like Facebook and Twitter has produced oceans of textual data from different sources like discussion channels, review websites and blogs, containing user's emotions, opinions, sentiments and certainly the intentions. Intention mining is becoming an emerging area in social computing, playing a key role in understating the needs, desires and wants of users/customers from business reviews, posts, and comments. However, a brief work in the field of intention mining regarding business-related reviews is still needed, and there lack of systematic literature review. This work aims at presenting a brief review of the literature with respect to intention mining from business-related online reviews including various techniques and methods employed. The results presented, show that this review will provide a roadmap for future researchers working in the field of business-related intention mining.

Keywords—intention detection, business intelligence, explicit intention identification, social media big data, implicit intention identification, intention analysis, decision making.

I. INTRODUCTION

The rise of different social media websites such as Twitter and Facebook allows online users to share their daily activities, feelings, beliefs and mainly their wants, needs, and aims which in other words called intents regarding various entities [1]. Intention mining deals with the identification and analysis of intents underlying user-generated content. It is also called intent detection, intent classification, intent analysis and intent extraction. The psychology defines intent as a mental state, revealing a user's future desires, wants and plans [1].

The term Intention mining was first introduced by Khodabandelou [2] in her Ph.D. thesis. It is defined as the acquisition and analysis of intents expressed by customers in online posts and reviews on social network websites [3].

The intent is defined as the actual wishes, needs, desires, goal and future action of an individual [1].For Example. "*I am looking for a Canon DSLR."*

The reviews and posts that don't illustrate any wants, desires, needs and future action of an individual are known as non-intents. For Example. " *Actually if you borrow money from any banks at this time, you have to pay high loan interests rate"* [1].

The business organizations are showing a keen interest in analyzing their customer's intentions to improve the quality of their products and to sustain in the competing business environment [4]. This propels the need of developing intention mining-based Business Intelligence (B.I) systems to sustain in the current digital world of business competition.

Many studies have been carried out with respect to [5, 6, 7, 8] business intelligence systems. However, limited works are performed on intention mining-based B.I.

Nowadays, the online community is expressing intentions frequently regarding products, companies, brands etc. For example, in the sentence: "I want to by Toyota GLI latest model", a user is expressing explicit intent to buy a car of a particular brand.

In this article, we present a comprehensive review of some recent and related studies conducted on intention mining in B.I domain. The major inspiration behind this work is that the identification and classification of user intents in business are becoming the need of the day to devise enhanced business strategies for customer satisfaction.

In this study, we try to answer the following research questions.

- RQ1: What are the different techniques with respect to machine learning and deep learning for intent classification?
- RQ2: What are the open challenges identified by reviewing literature in business-related intention mining?
- RQ3: What is the trend of the research community for business and non-business intent mining with respect published articles?

Following are key contributions of this work:

978-1-7281-0207-8/18/$31.00 ©2018 IEEE
DOI 10.1109/BESC.2018.00057

- An overview of different approaches and tasks required for the business-related user intention classification is presented.
- Different open problems and trends in business-related intent mining are investigated.
- Investigating future research directions in user intention mining.
- Presenting future research guidelines related to user intention mining in a business domain.

A. Related Reviews/Surveys

User intention mining is an emerging research field and a limited number of survey works are performed and there is a need for an up-to-date review, to understand the existing work performed in such field. Following are the two related studies carried out in this area.

Rashid et al. [9] conducted a survey for intent detection and classification. It is reported that supervised machine learning classifiers are frequently used on the microblogs (Twitter) dataset. Whereas a short review was carried out on intent detection form user reviews by [10].This review has different limitations, such as considering explicit intent and a limited discussion on consumption intent. However, the review presented in this article is different from the aforementioned two studies in terms of different parameters as shown in the classification diagram (Fig. 1).

Rest of the paper is organized as follows: (i) section II presents the review mechanism and conducted literature review, (ii) section III outlines open challenges in business-related intent mining, (iii) in section IV, recent trends in business-driven intent mining, are presented, and (iv) finally, conclusion and future directions are resented in the last section.

II. Review Methodology And Related Work

The review process starts by searching the related articles on the well-known digital libraries, such as Science Direct [11], IEEE Xplore [12], Springer Link [13], Google Scholar [14] and Emerald Insight [15]. In the next stage, inclusion and exclusion measure is used to filter the total downloaded articles as follows: (i) only English articles are considered; (ii) The peer-reviewed articles (journals and conferences) are included only, (iii) publications in the range 2010-2018 are included; (iv) Relevant non-English articles with are also considered. The exclusion criteria include: (i) articles not related to intent mining, (ii) and articles not in-line with the inclusion criteria.

The classification of user intention mining methods in the business domain is categorized into different classes as illustrated in Fig. 1.

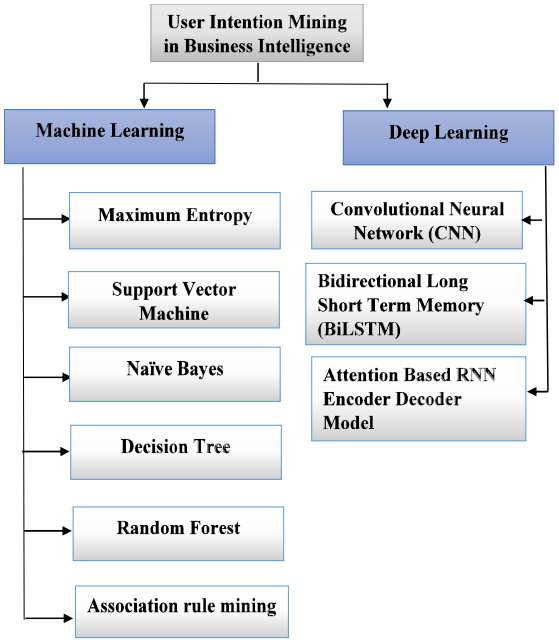

Fig. 1. Classification diagram of user intention mining methods in the business domain

A. Business-Related Intention Mining Using Machine Learning

In this section, relevant works performed on business-related intention mining using machine learning algorithms are presented.

Pérez-Vera et al. [4] conducted a study on intent mining of user's reviews about the Chilean electricity company by evaluating the performance of different classifiers using cross validation. It is observed support vector classifier performs better than the other methods

In their work on analyzing customer intentions expressed in commercial tweets, Hamroun et al. [16] performed an extraction of semantic patterns based on lexico semantic pattern based approach and ontologies. The experimental results show that the performance of CI-Pattern is better as compared to the baseline methods typically on Tweeter dataset.

Zhao et al. [17] proposed a product recommender system to detect purchase intents of customers on e-commerce sites for improving sales and customer experience. Moreover, in product recommendation task, user and product demographics, acquired from the tweets and online product comments, are matched and incorporated into the learning algorithm. The results show that the system performs efficiently.

Duan et al. [18] exploited user's intents from the question and answer forum of baby products. A collocation extraction method is applied using intent keywords for the extraction of intents associated with a product. The system has certain malfunctions with respect to the extraction and identification of intents and the integration of method with the recommender system is also not performed.

Hollerit et al. [19] identified commercial intents from Twitter using an automatic method. The method covers n-gram features and part-of-speech labels to perform commercial intents detection. Improved results in terms of precision and recall are achieved with respect to baseline methods.

In their work on detecting customer's wishes from the product review, Ramanand et.al [20] proposed a rule-based technique. However, their system considered only product reviews with emphasis on explicit intents. Furthermore, and lack of sufficient training data was another issue. Moreover, in terms of precision and recall, the system attains a score of about 80% and 50% respectively.

Identification of product-based business wishes is carried out by [21], by proposing graph driven technique. The experimental results show that in the case of movies and sports domains, the system achieved promising results with an accuracy of 81% and 52% respectively in each domain.

The work performed by [22] aims to detect and classify user intent classes using a semi-supervised approach. Using the bootstrap technique, a test collection is built and an intent graph driven optimization approach is implemented on a test collection. According to the experimental results, the performance of the proposed approach is effective for intent classification related to tweets.

Gupta et al. [23] worked on the identification of buy wishes using a supervised learning approach (SVM), preceded by a feature selection module. Satisfactory results are obtained from the two datasets, namely: Quora and Yahoo Answers.

In their work on automatic recognition and classification of user reviews into intent and non-intent tags, Goldberg et al. [24] used intent detector features trained on wish corpora. Different evaluation metrics like precision, recall, and f-measure are used to measure the effectiveness of the system in business (product) domain [25].

B. Business-related Intention Mining Using Deep Learning

Deep learning is a subset of machine learning methods. It consists of different neural network models, providing an assistance in classification and recognition tasks [26].Deep learning is an emerging approach to business-related intention mining, providing information at a deeper level [27]. In the section, a review of studies pertaining to deep learning with business-related intention mining is presented.

Ding et al. [28] performed identification of implicit intentions underlying social media text using a convolutional neural network-based deep learning technique. The experimental results show that the proposed model outperformed the baseline approaches. Furthermore, the intention words are validated as the best indicator of a user's demands. However, the recommendation of products is required to enhance system performance.

In their work on implicit intent identification, Li et al. [29] proposed a CNN classifier based on RNN driven encoder decoder system. Parallel corpora are used to conduct experiments and the results obtained are promising with respect to comparing methods.

A hybrid approach including Conditional Random Fields (CRF) and Bi-directional Long Short Term Memory (Bi-LSTM) is exploited by [30,31] to determine user intents expressed in social media posts. The experimental results show that the hybrid model achieves high accuracy than the baseline methods.

Table I. presents a review summary of the selected studies in a tabular form.

TABLE I. AN OVERVIEW OF ADOPTED STUDIES FOR USER INTENTON MINING

Study	Objective	Methods	Dataset and Results	Limitations And Future Work
Pérez-Vera et.al (2017) [4]	To categorize Intent of social media data.	• Feature Engineering(Binary,TF-IDF,TF-RFL) • Supervised Classification techniques(NN,NB,SVM,DT)	Twitter Dataset SVM performs better in case of TF-IDF representation. DT performs better in case of TF-RFL computations	In limitations, it is observed that a complete automatic answering process of user queries is needed to achieve favorable outcomes In future, the focus is to enhance representation and algorithms for effective results.
Hamroun et.al (2015) [16]	Analyzing customer intention made in commercial tweets.	• Semantic Patterns • Ontology .	Epinions.com , Mouthshut.com , TREC Dataset Performance of CI-Pattern is improved as compared to the baseline methods typically on tweeter dataset.	Limited set of features In future, a set of features needs to be upgraded
Duan et.al (2014) [18]	To exploit question and answer forum in order to derive user's intents.	o Unsupervised Technique o Pattern based approach to obtain candidate product o Intent- relevant product detection	Baby Tree, Taobao Wenda, Sina Baby & Child Care Q&A. The proposed method achieve 20% upgradation with respect to the baseline approach. Morover, the average weighted R-Precision of the proposed method is 63.3%.	Identification Error and Extraction Error The future works focuses on: Inclusion of online product search engine to eliminate the noise. Merging of intention detection with intention related product recommendation. Inclusion of filter process. Utilizing purchase data made from intention related products.
Holleri et.al (2013) [19]	To detect commercial intents made in twitter.	o Supervised technique(Linear SVM,NN,Linear LR) o Feature Selection(Parts of speech n-grams and Words)	Twitter website Precision: 57% Recall: 77%	Lack of attribute types. The future aims include: Creation of link between buyers and seller The calculation of inter-rater agreement will be performed to verify commercial intent definition. Further types of attributed like temporal or retweet information is required to be explored. Automatize the potential commercial intent detection.
Ramnand et.al (2010)[20]	Detection and classification of wishes.	o Unsupervised Technique Suggestion Wishes: o Rule based Approach Buy Wishes: o Rule based Approach	Suggestion wishes: Weblogs, MouthSut.com , Epinions.com ,Customer Surveys Product reviews, ,Goldberg et.al Wish Corpus, Purchasing Wishes**:** Yahoo! Answers ,Alibaba.com, Precision: Canon Camera achieve high precision among other datasets that is 97.50% Recall: ipod achieve high precision among other datasets that is 68.51%	Exploration of Implicit wishes is not presented It is required to perform corpora upgradation Future aims are: Inspecting Tense and Mood of a sentence It is needed to analyze advance features about parts of speech and semantic roles. Expansion of wish detectors. Robust quality training data deficiency.
Wang et.al (2015) [22]	To perform user intent extraction made in tweets. To detect and classify user intent.	o Semi supervised graph based technique o Intent Graph(Optimization model)	Kwak et al. 2010 twitter dataset. Satisfactory outcomes as compared to the state of art approaches	A set of seed record is limited Future Aim: A set of seed record needs to be enlarged.

Gupta et.al (2014) [23]	Detection and classification of purchase intents.	○ Feature Engineering(Phrase and Word) • Purchase Action Word And Purchase Object ○ Supervised technique (SVM)	Quora Dataset: 15,000 Yahoo! Answers Dataset:7,000 Results on Quora dataset: Delta-TFIDF:AUC=0.79 Delta-TFIDF+ PA-features+ PO-features: AUC=0.86 Inclusion of All features: AUC=0.93 Results on Yahoo Answers: Delta-TFIDF:AUC=0.77 Delta-TFIDF + PA-features + PO-features: AUC=0.83 Inclusion of All features: AUC=0.89	Limited usage of social media certain characteristics Existence of class imbalance In future, the focus will be not only on getting an insight about other certain features of social network sites like friend network but also on the exploitation of these features.
Goldberg et.al (2009)[24]	To perform detection and categorization of social media text into wishful and non-wishful classes automatically. Building of wish detector.	○ Supervised technique(Linear SVM) ○ Template Recognition(bipartite graph)	Product Domain Dataset: Reviews from Amazon.com Cnet.com Politics Domain Dataset: poltics.com Results of Politics domain: Area Under Curve(AUC) Template detector:0.73±0.03 Words + Template detector:0.80±0.03 Results of Product domain: Area under curve(AUC) Template detector:0.47±0.16 Words + Template detector:0.56±0.16	Language of wishes is only English, other languages like Portuguese, Spanish, Chinese, and French are not used. Future work aims: Developing additional features Exploring semi supervised technique.
Ding et.al (2015)[27]	Recognition of Implicit intention exists in social media big data.	○ Consumption Intention Mining Model ○ Convolutional Neural Network	Sina Weibo According to results the proposed model outperforms as compared to the baseline approaches Furthermore, the intention words are validated as a best indicator of user's demands.	Explicit intention detection is not performed. In future the aim is to develop a product recommender system.
Luong et.al (2017)[29]	Extracting information about intents made in online social media textual content.	○ Supervised technique ○ Machine Learning Model (CRF) ○ Deep Learning Model (Bi-LSTM) ○ Feature templates(n-gram , look-up dictionary , regular expression)	Some well-known sites: diendanthammy.net , muaban.net , webtretho.com , vatigia.com Accuracy Results: Performance of LSTM-CRF(char+pre+drop)for Cosmetic and Beauty Domain: Accuracy:91.17% Performance of LSTM-CRF(char) for Real-Estate Domain: Accuracy:91.37%	Limited dataset size Only Text related to Vietnamese language is utilized Future focus is on: Increment in Dataset size Utilize more language for intention extraction task. Explore more feature templates.

III. Open Challenges in Business-related Intention mining

Following are some of the open challenges in business-related intention mining identified from the literature.

A. Less focus on implicit intents

The detection of implicit intents is a demanding task due to the indirect expression of such intentions. Deep Learning approaches are recommended for handling implicit intentions [29]. Feature engineering for explicit intent extraction is based on manual tagging, which is not recommended for implicit intent identification due to degraded performance [1].

B. Applying Deep Learning techniques

To identify intents from business reviews, machine learning algorithms are widely used with less focus on applying deep learning methods. Therefore, there is a need to investigate deep learning techniques for intent classification from online business reviews [29].

IV. Trends in Business-related Intention Mining

In the process of user intention detection and classification, machine learning and deep learning are the frequently used methods (Fig. 2). However, deep learning techniques need further attention.

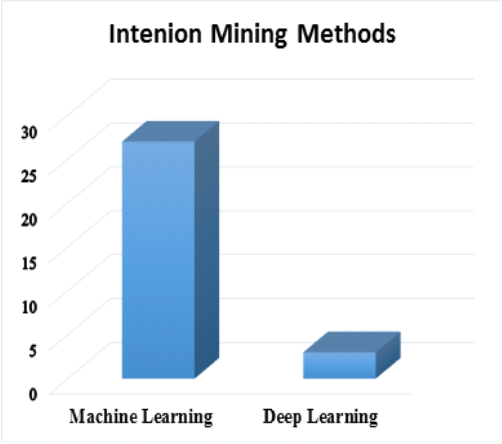

Fig. 2. Intention mining methods

It is observed that a lot of work has been carried out in the identification of non-business intents. While in recent time, identifying intent in business domain is an emerging research area. Fig. 3 provides a graphical representation of a number of research articles published with respect to business vs non-business intents. The diagram clearly illustrates the fact that business-related intents need to be explored.

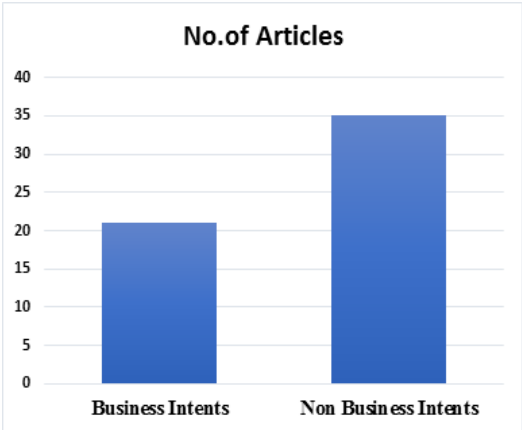

Fig. 3. Number of published articles with respect to business vs nonbusiness intents

Fig. 4 describes a list of research articles published with respect to explicit business intent and also it is noted that a large number of studies have been conducted for explicit business intent identification.

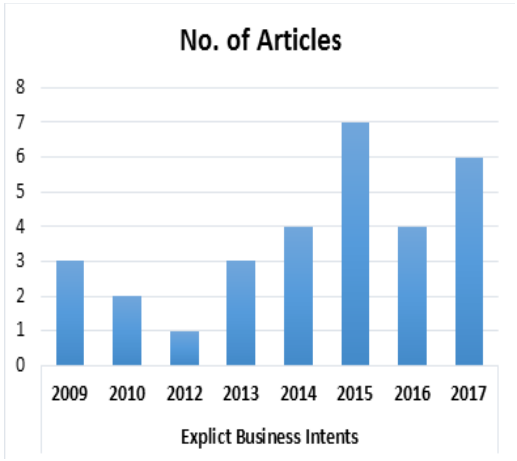

Fig. 4. Number of published articles with respect to explicit business intent

Fig. 5 depicts the number of published papers regarding business-related implicit intents. Furthermore, it is required to conduct more studies with respect to implicit business intents.

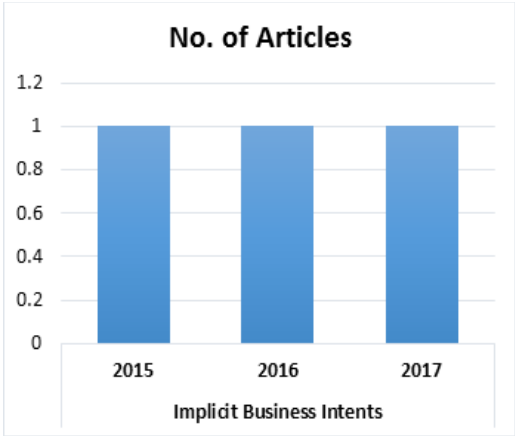

Fig. 5. Number of published articles with respect to implicit business intents

V. Conclusion and Future Work

The user intention mining with respect to business perspective is an important and challenging task due to the varying nature of customer-generated text data. The purpose of this review is to present a brief review of studies pertaining to user intention mining with emphasis on discussing different machine learning and deep learning techniques. Following are the possible future guidelines: (i). To investigate deep learning models for implicit intent classification, (ii).To provide a smart dashboard to the management of business companies with respect to user wishes for their manufactured products and (iii).enhanced recommender system needs to be developed using deep learning models on the basis of user intention for improved customer satisfaction.

REFERENCES

[1] Luong, T. L., Tran, T. H., Truong, Q. T., Phi, T. T., & Phan, X. H. (2016, March). Learning to filter user explicit intents in online vietnamese social media texts. In *Asian Conference on Intelligent Information and Database Systems* (pp. 13-24). Springer, Berlin, Heidelberg.

[2] Khodabandelou, G.; Hug, C.; Deneckere, R.; Salinesi, C. (2013). "Supervised intentional process models discovery using Hidden Markov models". IEEE Seventh International Conference on Research Challenges in Information Science (RCIS)

[3] https://www.lexalytics.com/technology/intentions

[4] Pérez-Vera, S., Alfaro, R., & Allende-Cid, H. (2017, July). Intent classification of social media texts with machine learning for customer service improvement. In International Conference on Social Computing and Social Media (pp. 258-274). Springer, Cham.

[5] S. Chaturvedi, V. Mishra and N. Mishra, "Sentiment analysis using machine learning for business intelligence," 2017 IEEE International Conference on Power, Control, Signals and Instrumentation Engineering (ICPCSI), Chennai, India, 2017, pp. 2162-2166.

[6] D. Delen, G. Moscato and I. L. Toma, "The impact of real-time business intelligence and advanced analytics on the behaviour of business decision makers," 2018 International Conference on Information Management and Processing (ICIMP), London, 2018, pp. 49-53.

[7] Khan, A., Asghar, M. Z., Ahmad, H., Kundi, F. M., & Ismail, S. (2017). A rule-based sentiment classification framework for health reviews on mobile social media. *Journal of Medical Imaging and Health Informatics*, *7*(6), 1445-1453.

[8] Q. Zhou, B. Xia, W. Xue, C. Zeng, R. Han and T. Li, "An Advanced Inventory Data Mining System for Business Intelligence," 2017 IEEE Third International Conference on Big Data Computing Service and Applications (BigDataService), San Francisco, CA, 2017, pp. 210-217.

[9] Rashid, A., Shoaib, U., & ShahzadSarfraz, M. KNOWLEDGE DISCOVERY IN DATABASE USING INTENTION MINING

[10] Khan, F., Borah, S., & Pradhan, A. Mining Consumption Intent from Social Data: A Survey.

[11] http://www.sciencedirect.com/science/search

[12] http://ieeexplore.ieee.org/Xplore/home.jsp

[13] https://link.springer.com/

[14] https://scholar.google.com/

[15] http://www.emeraldinsight.com/

[16] Hamroun, M., Gouider, M. S., & Said, L. B. (2015). Customer intentions analysis of twitter based on semantic patterns. In The 11th International Conference on Semantics, Knowledge and Grids (pp. 2-6).

[17] Zhao, X. W., Guo, Y., He, Y., Jiang, H., Wu, Y., & Li, X. (2014, August). We know what you want to buy: a demographic-based system for product recommendation on microblogs. In Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 1935-1944). ACM.

[18] Duan, J., Ding, X., & Liu, T. (2014, November). Mining intention-related products on online q&a community. In Chinese National Conference on Social Media Processing (pp. 13-24). Springer Berlin Heidelberg.

[19] Hollerit, B., Kröll, M., & Strohmaier, M. (2013, May). Towards linking buyers and sellers: detecting commercial intent on twitter. In Proceedings of the 22nd International Conference on World Wide Web (pp. 629-632). ACM.

[20] Ramanand, J., Bhavsar, K., & Pedanekar, N. (2010, June). Wishful thinking: finding suggestions and'buy'wishes from product reviews. In Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text (pp. 54-61). Association for Computational Linguistics.

[21] Wang, J., Zhao, W. X., Wei, H., Yan, H., & Li, X. (2013, October). Mining New Business Opportunities: Identifying Trend related Products by Leveraging Commercial Intents from Microblogs. In EMNLP (pp. 1337-1347).

[22] Wang, J., Cong, G., Zhao, W. X., & Li, X. (2015, January). Mining User Intents in Twitter: A Semi-Supervised Approach to Inferring Intent Categories for Tweets. In AAAI (pp. 318-324)

[23] Gupta, V., Varshney, D., Jhamtani, H., Kedia, D., & Karwa, S. (2014, May). Identifying Purchase Intent from Social Posts. In ICWSM.

[24] Goldberg, A. B., Fillmore, N., Andrzejewski, D., Xu, Z., Gibson, B., & Zhu, X. (2009, May).May all your wishes come true: A study of wishes and how to recognize them. In Proceedings ofvHuman Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics (pp. 263-271). Association for Computational Linguistics.

[25] Asghar, M. Z., Khan, A., Khan, F., & Kundi, F. M. (2018). RIFT: A Rule Induction Framework for Twitter Sentiment Analysis. *Arabian Journal for Science and Engineering*, *43*(2), 857-877.

[26] http://adilmoujahid.com/posts/2016/06/introduction-deep-learning-python-caffe/

[27] Park, D. H., Fang, Y., Liu, M., & Zhai, C. (2016, October). Mobile App Retrieval for Social Media Users via Inference of Implicit Intent in Social Media Text. In Proceedings of the 25th ACM International on Conference on Information and Knowledge Management (pp. 959-968). ACM.

[28] Ding, X., Liu, T., Duan, J., & Nie, J. Y. (2015, January). Mining User Consumption Intention from Social Media Using Domain Adaptive Convolutional Neural Network. In AAAI (Vol. 15, pp. 2389-2395).

[29] Li, C., Du, Y., & Wang, S. (2017, August). Mining Implicit Intention Using Attention-Based RNN Encoder-Decoder Model. In International Conference on Intelligent Computing(pp. 413-424). Springer, Cham.

[30] Luong, T. L., Cao, M. S., Le, D. T., & Phan, X. H. Intent Extraction from Social Media Texts Using Sequential Segmentation and Deep Learning Models.

[31] Asghar, M. Z., Kundi, F. M., Ahmad, S., Khan, A., & Khan, F. (2018). T-SAF: Twitter sentiment analysis framework using a hybrid classification scheme. *Expert Systems*, *35*(1), e12233.