International Journal of Scientific Research in Computer Science, Engineering and Information Technology
ISSN : 2456-3307 (www.ijsrcseit.com)

An Efficient Aspect-Based Opinion Mining on Smart Phone Reviews With LDA

P Rajanath Yadav[1], Dr. S. Ramakrishna[2]
[1]PG Scholar, Sri Venkateswara University, Tirupati, Andhra Pradesh, India
[2]Professor, Department of Computer Science, Sri Venkateswara University, Tirupati, Andhra Pradesh, India

Article Info
Volume 4, Issue 10
Page Number : 89-97
Publication Issue :
July-2020

Article History
Published : 20 July 2020

ABSTRACT

With the development of e-commerce stage, online shopping has become an easy and preferable mode of shopping. As one of the largest e-commerce stages worldwide, Amazon enjoy numerous user communities. Volumes of user-generated information of users' preferences and opinions towards items, for the most part for specific aspects of a ware, popped up every day. Albeit loaded with information, these texts are often unstructured information that requires a careful analysis for the two consumers and manufactures to extract meaningful and relevant information. Customary lexicon-based sentiment analysis considers extremity score of words however ignores the differences among aspects. Document level subject modeling help overcome these lacunae. Right now, guarantee that the aspects ought to likewise be weighted for featuring significance of different aspects appropriate to a space. Along these lines, manufacturers can understand what potential consumers may need as improvement in the expected items. To showcase our framework, more than 400,000 Amazon unlocked phone reviews were collected as preparing information. LDA models were used to cluster subject words with their corresponding likelihood values. Based on the machine learning framework results, a corpus of nearly 1,000 Amazon reviews of a new mobile phone mode, iPhone X, was tested utilizing this framework to perform subject labeling and sentiment analysis. Performance analysis was done utilizing Confuse Matrix and F-measure.

Keywords : Sentiment Analysis; LDA; iPhone X; E-Commerce, Opinion Mining, Information Systems, Information Retrieval, Retrieval Tasks and Goals, Sentiment Analysis.

I. INTRODUCTION

With the quick development of web 3.0, e-commerce emerged as a well known and preferable alternative to consumers. Off-hours purchases, item correlations, review by other users, are some of the benefits which made e-commerce successful. Among every one of these realities, corpus of user-generated reviews, which is generally found in any e-commerce stage requires an intensive analysis, as it isn't just used by other potential consumers to select an item, yet additionally by manufactures to refine their item based on users' information sources. These corpuses of user-generated reviews are for the most part plain texts that pose huge challenge to analyze. Doing this physically is a very onerous work, and, all things considered, use of computational and automated instruments has become imperative. Normal language processing (NLP) is very active and dynamic field and researchers contributing and developing state-of-the-craftsmanship sentiment analysis techniques to deal with this undertaking.

Sentiment analysis techniques can be divided into two general categories: lexicon-based methods and machine learning based methods. Lexicon-based method is easy to perform. Yet, for larger informational indexes having numerous aspects to analyze, lexicon-based methods are not sufficient [1, 2]. While machine learning methods use order to appoint appropriate extremity and require top notch information which is difficult to get or clean [3].

Customarily, lexicon-based sentiment analysis deals with the issue about negation handling and intensifier/diminishers to qualify and measure appropriate sentiment scores [4, 5, 6]. Researchers additionally bring up that the same word may have different sentiments under different context. For instance, the word 'quick' is associated with positive value when it will describe the speed of battery charging in the context of an advanced mobile phone, while a negative value ought to be assigned to 'quick' when it is associated to exhaustion speed of battery. With this challenge, researchers [7, 8] developed different context-based lexicons to increase the exactness of sentiment score assignment. Aminu et al. [5] developed a lexicon-based model SmartSA, which is a run of the mill aggregation of conventional technique, to calculate the sentiment score in sentence level based on neighborhood context and worldwide context. Neighborhood context deals with the negation handling and modifier issue while worldwide context refers to the extremity of the terms shows different sentiment extremity in different scenarios. However, this sort of technique just relegate value to contextual words yet not the point words, which indicate certain aspect of an item mentioned in a review or document. Weighted aspects, which reflect the greater part users' preferences for different attributes of an item, can help manufacturer to understand what potential consumers may need as improvement in the item.

For example, a mobile phone item, for example, iPhone X, has numerous attributes: show screen, camera quality, battery life, operating system and so on. Some consumers prefer to use their phone to watch videos. In like manner, they may be sensitive to the screen quality. Consequently, when they write their opinion, they tend to describe their feeling towards the screen show. Other consumers enjoy selfie and center more around camera quality.

Therefore, their reviews consistently mention the aspect camera. Suppose in a corpus, containing aggregated reviews pretty much a wide range of smartphones and their huge aspects, predominantly mention the set of words related to the aspect camera, then, the weight of the aspect camera ought to be higher than the other aspects. Researchers are utilizing different aspect - based modeling instruments to analyze aspect-based opinion mining, including Latent Dirichlet designation (LDA) [9] that can be used effectively for this purpose.

Right now, propose a cosmology framework which can naturally extract useful information from user-generated reviews to determine user's specific experience and opinion towards item's aspects utilizing LDA. The preparation dataset is collected by Rathan, M., et al [10]. The corpus of reviews of advanced cell iPhone X, one of the latest very good quality items from Apple Inc, from both Amazon and

Flipkart are collected to test this framework and measure the exactness of the proposed model. The python package, SentiWordNet [11], is used as lexicon.

II. RELATED WORK

So as to reduce manual remaining task at hand just as to acquire useful information from online reviews, researchers have developed different machine learning techniques to collate and extract huge information from corpus of online documents and calculate sentiment scores of opinions. However, robotization of such techniques is still not achieved. Human explanation is yet to be replaced by computer calculations; however, we have achieved certain advancements right now. Characteristic language processing (NLP) calculations have been developed for syntax-based analysis of sentences and documents. However, identifying context and subjects from the documents despite everything poses challenges to researchers.

Same word may have different meanings as different pieces of dependency: (like, camera). In Hridoy et al's. methodology, the information doesn't contain any of these three dependencies will be discarded. Numerous researchers [e.g., 10, 14] have pointed out the significance of performing spell checker on the pre-processing stage of online review analysis, since online texts contain bunches of slangs, abbreviation and misspelled words. For example, Mamgain et al. [15] applied a probabilistic model based on Bayes' theorem to correct the spell mistake which overlooked from [16]. There additionally exist available open-source library to perform such undertaking. Rathan et al. [10] invoke a java package named jSpellCorrect [17] to deal with the misspelled words. So also, the current paper will use the python version of jSpellCorrect, called autocorrect [18], to achieve spell correction.

Recently, a new generation of emoticons have additionally been well known to express one's emotions in the online texting area. In the year 2015, Oxford Dictionary recorded an emoji as the expression of year. Instead of composing a long descriptive sentence, people prefer utilizing several emojis to express their feeling. Novak et al. [19] developed a lexicon named Emoji Sentiment Ranking to record the sentiment value of 751 most frequently used emojis. This paper uses Unicode to detect emojis and this emoji sentiment lexicon to calculate the sentiment score.

Latent Dirichlet Allocation, which clusters subjects based on the words' co - occurrence, is widely used in theme modeling literarily. Buschken and Allenby [20] applied LDA model at the sentence-level in reviews to predict the customer evaluations. Calheiros et al. [21] use the beta value in LDA as the confidence to qualify the sentiment extremity for a set of hotel reviews. Jabr et al. [22] combine subject modeling with sentiment analysis to evaluate customer's extent of fulfillment with the different aspects of an item. What Jabr have done is like us, yet not center to the speech. For example, the word 'like' has a positive sentiment as a verb while a neutral sentiment when it is a combination. Sentence-level sentiment analysis technique, which identify the syntactic relations between the words in a sentence, is introduced to increase exactness of such models. SentiWordNet appoints different sentiment values to the word as different grammatical features. Moreover, Hridoy et al. [13] applied the NLP devices, Stanford dependency parser, provided by SNLP (Stanford NLP gathering) to discover the words' dependency in the information pre-processing phase. They search for three ordinary dependencies: nsubj, amod and dobj, which contain useful information to filter the information. The nsubj dependency is the relationship between things and verbs or adjectives. The review, "The phone is beautiful", will have nsubj dependency: (phone, beautiful). The amod dependency refers to adjectival modifier. For example, the sentence, "the awesome camera impressed me a great deal.", will be identified an amod dependency: (awesome, camera). Dobj stands for direct objective.

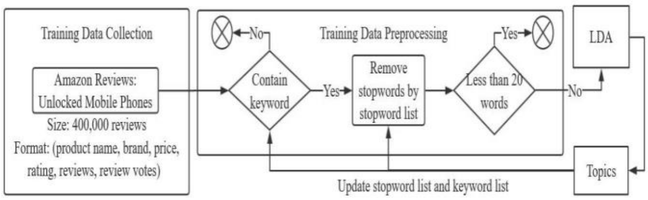

Fig 1. Topic Modeling Framework

Phones, is created by Rathan et al. [10] which contains more than 400 thousand reviews of unlocked mobile phones sold on Amazon.com. To perform LDA, there are two significant issues required to be discussed: information pre-processing and the choice of theme number K.

Fig. 1 reveals the fundamental steps of information pre-processing. Right off the bat, a keyword list was utilized for screening out uproarious reviews. The keyword list was initialized as the related terms in 100 most frequent terms appearing in the corpus by human interpretation. 13 terms were selected: "screen", "size" and "contact" for aspect show screen; "battery", "charge", "charger", "hours", "power" and "life" for aspect battery life; "camera", "light", "pictures" and "photographs" for aspect camera quality. The reviews which didn't contain any word in the keyword list were discarded. Right now, will be exposed. After that, for the reserved reviews, potentially problematic images, for example, "- ", "/", etc. and stop words, for example, "an", "an", etc. were removed. The first stop word list is provided by NLTK (Natural Language Toolkit) [23] in python. Since measurable models, for example, LDA treat texts as packs of word, short texts, which need enough content, short text will confuse the LDA model [24]. Empirical, the threshold of least text length was set to 20 to filter out short texts. The remaining reviews were passed to the LDA model to generate first-round subjects.

The subject number K is a significant pre-decided parameter which directly influence the order quality. Moro et al. [25] set K to half of the terms considered. The proposed framework followed Moro's methodology; hence six points were modeled. Gensim [26] is a freely available python library for discovering the subjects of reviews applying LDA model. Table 1 is a piece of the first-round result generated by Gensim LDA. Keyword rundown and stop word list was updated based on this table. Some significant terms which belong to these three subjects are discovered. For example, "front" for camera; "fingerprint", "waterproof", "dropped", "plastic" and "glass" for screen. These terms are added to keyword list. Likewise, some high - frequency yet meaningless words, for example, "phone", "phones" and "applications" are added to the stop word list. After refreshing the keyword list and the stop word show, some useful reviews were recalled, and some useless information were filtered out. The refined information was sent to Gensim LDA to discover better result. Table 2 is a piece of the second result.

The second result is more interpretable. Point 1 includes 'blackberry', 'keyboard', 'qwerty', etc. These words are attempting to describe the composing issue. Since some of the Blackberry model have physical keyboard, numerous reviews may mention it to examine the composing issue. Accordingly, the term 'blackberry' has the highest weight. The representative expression of point 2 is 'screen' whose likelihood value reached 0.054. Clearly, this theme can be represented by 'screen'. Point 3 mentions 'camera', 'pictures', 'light' which describe the attribute 'camera'. The highest weighted word in subject 4 is 'battery' and this gathering of point contains 'charge', 'life'. Therefore, this point describes the attribute 'battery'. The representative expression of theme 5 is likewise screen, however the associated likelihood value is low. Consequently, this subject isn't run of the mill and was discarded. Point 6, which mentions 'call', 'sim', 'wifi', describe the connection issue. For these themes, just subject 2 as screen, point 3 as camera and theme 4 as battery were utilized in the sentiment analysis section. Moreover, some words are not related to their assigned theme. For example, in subject 4, the words 'seller', 'received' and 'purchase' will talk about the delivery issue. The reason why these words come together with the word 'battery' is that people tend to describe battery with some time issue, and people consistently mention 'time' together with delivery. These words were physically screened out by human interpretation. Table 3 is a piece of the last subject weight matrix.

Table 1: First-round Result of LDA

Topic 1	Topic 2	Topic 3	Topic 4	Topic 5	Topic 6
0.073 screen	0.042 phone	0.041 phone	0.113 phone	0.084 phone	0.038 phone
0.047 phone	0.017 screen	0.015 nokia	0.035 battery	0.028 sim	0.032 camera
0.013 protector	0.017 apps	0.013 wifi	0.017 time	0.018 card	0.015 sony
0.011 glass	0.008 time	0.008 windows	0.011 bought	0.012 service	0.014 quality

Table 2 : Second-round Result of LDA

Topic 1	Topic 2	Topic 3	Topic 4	Topic 5	Topic 6
0.045 blackberry	0.054 screen	0.034 camera	0.033 battery	0.014 screen	0.016 sim
0.037 keyboard	0.012 button	0.027 battery	0.017 screen	0.014 android	0.015 card
0.014 touch	0.011 protector	0.019 screen	0.012 life	0.011 gb	0.009 nokia
0.013 bb	0.01 touch	0.018 love	0.011 amazon	0.01 apps	0.009 apps

Table 3 : Topic-Weight Matrix

Topic Screen	Topic Camera	Topic Battery
screen 0.054	camera 0.034	battery 0.033
protector 0.011	quality 0.015	time 0.012
touch 0.01	fast 0.014	charge 0.011
cover 0.009	size 0.009	charger 0.01

III. TEST DATA COLLECTION AND PRE-PROCESSING

The test information is the iPhone X reviews collected from Amazon.com and Amazon.in. Table 4 contains the detailed information of these information.

Figure 2 is the framework of test information pre-processing. Some reviews may not merely describe one aspect. For example, the review, "The camera is awesome, yet the battery is very weak.", mentions two aspects: camera and battery. To evade some significant themes being covered up, each review was part into sentences by accentuation and combination. Comparative as the preprocessing part of point modeling, the sentences which don't mention any aspect were discarded based on the keyword list. Each remaining sentence was passed to two destinations: subject labeling and sentiment score labeling.

Table 4. Test Data Acquisition*

Website	Time	# of reviews
Amazon.com	November 2, 2017 – November 14, 2018	333
Amazon.in	November 30, 2017 – November 15, 2018	693

*The data was collected from these websites on November 21, 2018

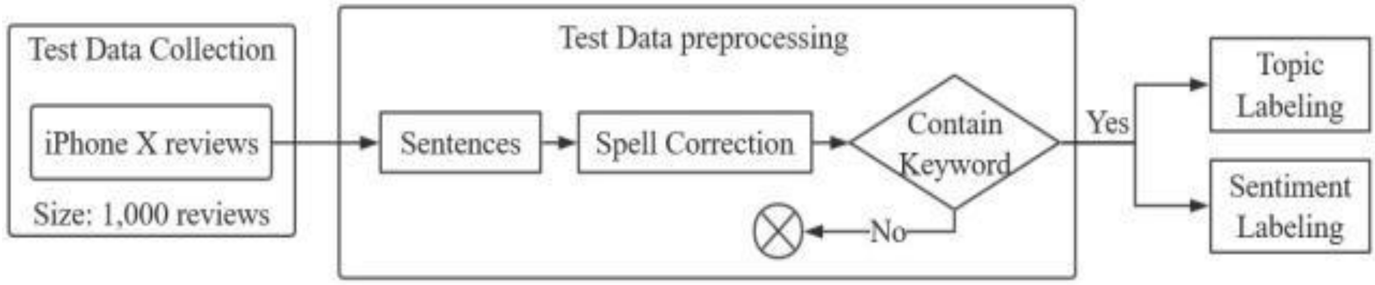

Fig 2. Test Data Pre-processing

3.1 Sentiment Labeling

Figure 3 provide the overall framework of sentiment labeling process. Right off the bat, emoji was detected by Unicode. Based on the lexicon provided by Novak el al. [19], the emoji sentiment score of a sentence was calculated by the whole of all the emoji sentiment values in that sentence. Then, context-based sentiment analysis was performed based on the area specific word lexicon. Next, based on the Stanford NLP instrument: Stanford POS tagger, each word in the sentences was labeled with its grammatical form. Based on this label, SentiWordNet, the celebrated lexicon, was utilized. In the event that a word was detected by both space specific lexicon and SentiWordNet, just the sentiment score provided by area specific lexicon were considered. The last sentiment weight of a sentence was calculated by the aggregate of all these three values: sentiment score of emoji, investigation based on area specific word lexicon and the examination based on SentiWordNet. Eventually, the score of a sentence in detailed aspect is the duplication of the subject weight and the sentiment weight.

IV. RESULT AND EVALUATION

The trained model was applied on the user-generated online reviews of one specific mobile phone: iPhone X. Altogether, there were nearly 1,000 reviews collected by a crawler program and after parting these reviews into sentences based on combination and accentuation, 7200 sentences were extracted. After filtering out the useless information with a keyword list, there were approximately 500 sentences left to be labeled. We applied the methodology described in section 3.3 to label the information. To evaluate the performance of our model, these sentences were physically labeled with sentiment extremity and aspect. Table 5 refers to the performance evaluation of the sentiment grouping. The F-measure of positive sentiment achieved 70% while negative sentiment is marginally lower. This seems to be due to the linguistic and typographical errors in the online texts which will mislead the Stanford POS tagger. With an inappropriate grammatical feature, SentiWordNet can't provide the correct sentiment value.

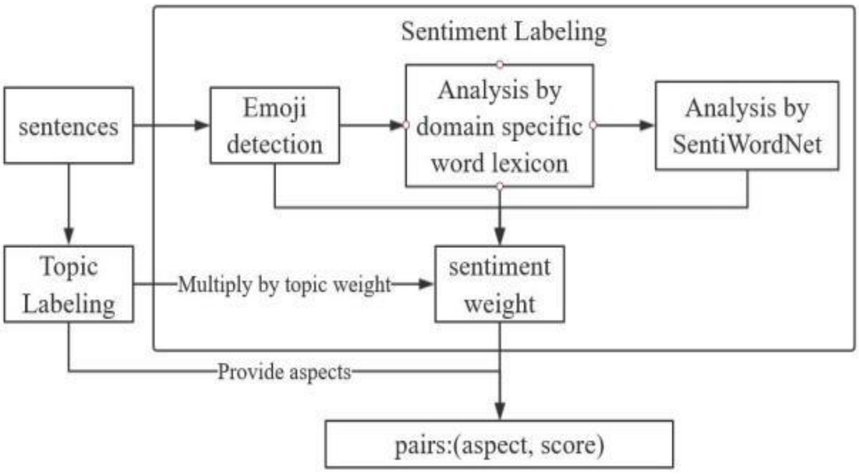

Fig 3. Sentiment Labeling Framework

Table 5 : Result Evaluation: Sentiment Classification

Confuse Matrix				Result		
	Pos	Neu	Neg	Precise	Recall	F-measure
Pos	185	28	49	0.706107	0.700758	0.703422
Neu	50	41	27	0.347458	0.465909	0.398058
Neg	29	19	69	0.589744	0.475862	0.526718

Table 6 : Result Evaluation: Topic Classification

Confuse Matrix					Result		
	None	Screen	Camera	Battery	Precise	Recall	F-measure
None	19	29	27	19	0.202	0.292	0.239
Screen	27	159	1	3	0.837	0.787	0.811
Camera	14	9	104	9	0.765	0.787	0.776
Battery	5	5	0	79	0.888	0.718	0.794

Table 6 provides the result evaluation of topic classification. “None” means that the sentence doesn’t describe any one of these three subjects. Since the information was right off the bat filtered with a keyword list, the low F-measure is acceptable. The exactness of aspect screen was measured at 81%, aspect camera was measured at 77%, and aspect battery was measured at 79%. Moreover, during the physically labeling process, a minor improvement of the overall precision was found. Consider this sentence, "great showcase likewise camera.", based on our methodology, this sentence scored 0.006 on theme screen, and scored 0.034 on subject camera, and point camera, which was assigned with higher score, was chosen to be the subject of that sentence. However, the subject screen was likewise mentioned and had the same sentiment weight as point camera. Therefore, it is better to permit a sentence to contain more than one aspect.

V. CONCLUSION

The presented paper introduced a cosmology framework that is qualified for online review sentiment analysis on a feature level. Further, the significance of weighted aspects of items in mining online reviews was claimed. We considered the likelihood value of a word in a point in LDA as the weight of the word in that subject and modified the sentiment score by the theme weights. Currently, just three aspects were considered: show screen, battery life and camera quality. However, a mobile phone item contains more attributes: ergonomics, memory, operating system, In the future, these attributes will be included.

V. REFERENCES

[1]. Poria, S., Chaturvedi, I., Cambria, E., and Bisio, F. (2016, July). Sentic LDA: Improving on LDA with semantic similarity for aspect-based sentiment analysis. In 2016 international joint conference on neural networks (IJCNN) (pp. 4465-4473). IEEE.

[2]. Ruder, S., Ghaffari, P., and Breslin, J. G. (2016). A hierarchical model of reviews for aspect-based sentiment analysis. arXiv preprint arXiv:1609.02745.

[3]. Hutto, C. J., and Gilbert, E. (2014, May). Vader: A parsimonious rule-based model for sentiment analysis of social media text. In Eighth international AAAI conference on weblogs and social media.

[4]. Chatzakou, D., and Vakali, A. (2015). Harvesting opinions and emotions from social media textual resources. IEEE Internet Computing, 19(4), 46-50.

[5]. Muhammad, A., Wiratunga, N., and Lothian, R. (2016). Contextual sentiment analysis for social media genres. Knowledge-Based Systems, 108, 92-101.

[6]. Song, Y., Wang, H., Wang, Z., Li, H., and Chen, W. (2011, June). Short text conceptualization using a probabilistic knowledgebase. In Twenty-Second International Joint Conference on Artificial Intelligence.

[7]. Labille, K., Gauch, S., and Alfarhood, S. (2017, August). Creating domain-specific sentiment lexicons via text mining. In Proc. Workshop Issues Sentiment Discovery Opinion Mining (WISDOM).

[8]. ark, S., Lee, W., and Moon, I. C. (2015). Efficient extraction of domain specific sentiment lexicon with active learning. Pattern Recognition Letters, 56, 38-44. Blei, David M., Andrew Y. Ng, and Michael I. Jordan. Latent dirichlet allocation. Journal of machine Learning research 3: 993-1022.

[9]. Rathan, M., et al. Consumer insight mining: Aspect based Twitter opinion mining of mobile phone reviews. Applied Soft Computing 68: 765-773.

[10]. Baccianella, S., Esuli, A., and Sebastiani, F. (2010, May). Sentiwordnet 3.0: an enhanced lexical resource for sentiment analysis and opinion mining. In Lrec 10, 2200-2204.

[11]. Ali, F., Kwak, D., Khan, P., El-Sappagh, S., Ali, A., Ullah, S., ... and Kwak, K. S. (2019). Transportation sentiment analysis using word embedding and ontology-based topic modeling. Knowledge-Based Systems.

[12]. Hridoy, Syed Akib Anwar, et al. Localized twitter opinion mining using sentiment analysis. Decision Analytics 2.1 (2015): 8.

[13]. Lu, C. J., Aronson, A. R., Shooshan, S. E., and Demner-Fushman, D. (2019). Spell checker for consumer language (CSpell). Journal of the American Medical Informatics Association, 26(3), 211-218.

[14]. Mamgain, N., Mehta, E., Mittal, A., and Bhatt, G. (2016, March). Sentiment analysis of top colleges in India using Twitter data. In 2016 International Conference on Computational Techniques in Information and Communication Technologies (ICCTICT) (pp. 525-530). IEEE.

[15]. Segaran, T., and Hammerbacher, J. (2009). Beautiful data: the stories behind elegant data solutions. " O'Reilly Media, Inc.".

[16]. http://developer.gauner.org/jspellcorrect/ Access date: April 30, 2019

[17].https://pypi.org/project/autocorrect/ Access date: April 30, 2019

[18].Novak, P. K., Smailović, J., Sluban, B., and Mozetič, I. (2015). Sentiment of emojis. PloS one, 10(12), e0144296.

[19].Büschken, J., and Allenby, G. M. (2016). Sentence-based text analysis for customer reviews. Marketing Science, 35(6), 953-975.

[20].Calheiros, A. C., Moro, S., and Rita, P. (2017). Sentiment classification of consumer-generated online reviews using topic modeling. Journal of Hospitality Marketing & Management, 26(7), 675-693.

[21].Jabr, W., Cheng, Y., Zhao, K., and Srivastava, S. (2018). What Are They Saying? A Methodology for Extracting Information from Online Reviews.

[22].https://www.nltk.org/ Access date: April 30, 2019

[23].Song, Y., Wang, H., Wang, Z., Li, H., and Chen, W. (2011, June). Short text conceptualization using a probabilistic knowledgebase. In Twenty-Second International Joint Conference on Artificial Intelligence.

[24].Moro, S., Cortez, P., and Rita, P. (2015). Business intelligence in banking: A literature analysis from 2002 to 2013 using text mining and latent Dirichlet allocation. Expert Systems with Applications, 42(3), 1314-1324.

[25].https://radimrehurek.com/gensim/ Access date: April 30, 2019

Author

P. Rajanath Yadav, received Bachelor of Science degree from SRI KRISHNADEVARAYA UNIVERSITY, ANANTHAPUR in the year of 2014-2017 Pursuing Master of Computer Applications from Sri Venkateswara University, Tirupati in the year of 2018-2020. Research interest in the field of Data Mining.