

Available online at <http://www.mecs-press.net/ijeme>

Opinion Mining of Online Product Reviews from Traditional LDA Topic Clusters using Feature Ontology Tree and Sentiwordnet

D. Teja Santosh^a, K. Sudheer Babu^a, S.D.V. Prasad^a, A. Vivekananda^b

^a Assistant Professor, GITAM University, Rudraram, Telangana, India

^b Graduate Student, GITAM University, Rudraram, Telangana, India

Abstract

Online product reviews provide data about the user's perspective on the features that were experienced by them. Product features and corresponding opinions form a major part in analyzing the online product reviews. Extracting features from a huge number of reviews is classified into three major categories such as utilizing language rules, sequence labeling as well as the topic modeling. Latent Dirichlet Allocation (LDA) is one such topic model which clusters the document words into unsupervised learned topics using Dirichlet priors. The words so clustered are the features and opinion words in the product reviews domain. To identify appropriate product features from these clusters a hierarchical, domain independent Feature Ontology Tree (FOT) is applied to LDA clusters. The opinion bearing words of obtained product features are identified by utilizing the document indicators available from topic matrix of LDA. These indicators are useful to backtrack to the corresponding online review in which the product feature is present. The polarity of the opinion bearing word is calculated with the help of SentiWordNet. This improves the accuracy of the features using extracted LDA topic clusters and machine interpretation of polarity of opinion word is satisfactory.

Index Terms: Document indicator, Feature Ontology Tree, Latent Dirichlet Allocation, Opinion word, SentiWordNet.

© 2016 Published by MECS Publisher. Selection and/or peer review under responsibility of the Research Association of Modern Education and Computer Science.

1. Introduction

Current World Wide Web contains many web pages and websites that seek reviews from customers about their experience after using the product. The deluge of the reviews which are regularly fed into the site were not read by the customers completely to make decisions. To make it simple, various product feature specific summarization systems were provided as a constructive view of the reviews in NLP [1] research. Product features and corresponding opinion words were extracted and visualized with corresponding count to it. The

* Corresponding author:
E-mail address:

knowledge hidden in the online reviews be discovered using the concept of probability theory. The model learned for this kind of knowledge discovery is the probabilistic graphical model. This is a declarative representation of the particular task involving uncertainty on the final reasoned conclusions. This representation is useful for efficient training of the multifaceted distribution on the underlying data. Latent Dirichlet Allocation (LDA) developed by Blei [2] is an unsupervised learning model under topic modeling category which is a probabilistic graphical model used to generate word clusters under pre-specified topics from the collection of documents (reviews). The focal idea of a topic is to extract the thematic summary of the document collection. The thematic summary itself is a document which is used to discover the topic proportions from each document. LDA extends Probabilistic Latent Semantic Analysis (PLSA) [4] topic model by adding priors (Dirichlet Distribution) to the document collection. It uses the concept of Bayesian probability to perform information retrieval, document similarity and etc. These clusters contain words which are non features of the product. To uniquely identify the product features from these clusters, a domain independent Feature Ontology Tree (FOT) is engineered and used. FOT is an Web Ontology Language (OWL) Lite ontology in which product features were arranged in a hierarchy considering from parent feature and drill down to the most specific part of the object feature. The opinion bearing words of obtained product features are identified by utilizing the document indicators of LDA to backtrack to the corresponding online review where the product feature is present. The polarity of the opinion bearing word is calculated with the help of SentiWordNet.

2. Related Work

The topic model based data analysis has its research started with Brody and Elhadad [6] in which they were used LDA to mine the topics containing product features. To come to a decision on the polarity of each feature, the researchers started with a set of start opinion words and spread the polarities to further adjectives by using a label propagation algorithm. They have employed cluster evaluation scheme by calculating the difference between connectivity matrix and comparison matrix. The difference serves as a score to find the correct number of clusters but in the process they were not removed the stop words in the pre-processing phase. This has lead to unimportant cluster generation in the process. Instead of considering feature detection and opinion classification as two individual problems, the researchers Lin and He [7] and Jo and Oh [8] directly integrated the opinion categorization in the LDA model, so that it captures the opinion towards the features. The LDA parameter values of α and β that are considered in their work was not guaranteed the exact co-occurring words for a specific aspect in the cluster. Although these LDA-based approaches provide an elegant model of the problem, they generated topics that are not directly interpretable as features, and used manual annotation to achieve a clear output. The work of Zofia Stankiewicz & Satoshi Sekine [9] combined Ontology with LDA topic model to surf the similar products in online shopping website. They used ontology to map the product category with the topic name of the cluster. In their procedure they were not concentrated on identifying the exact product features from the obtained topic clusters using ontology as it identifies product features, product parts, part features as features. OWL-Lite ontology applied on the LDA generated topic clusters to identify and extract the appropriate product features so that precise review categorization is possible while searching on a particular feature.

The techniques to extract features from product reviews have gained considerable research in the literature. The work of Popescu and Etzioni [11] applied an unsupervised technique to mine product features and opinions from the reviews. They have introduced the system called OPINE to do the same. OPINE works on syntactic patterns towards identification of opinion phrases and their orientation. The research carried out by Carenini et al. [12] recommended a representation based on user defined knowledge to create classification of product features. The works of Ferreira et al. [13] and Yi et al. [14] put their effort on the problem of product features extraction through likelihood ratio test with lexical pattern. This approach produced best results. The researchers Ghobadi & Rahgozar [15] advised ontological approach to extract the products information. The ontology is based on product features. The proposal uses ontology for generating virtual catalogs by reasoning

the ontology on the product feature data.

Clearly, the taxonomy and the ontology of product features are domain dependent models for extracting product features. Hence, a need for the development of domain independent product feature Ontology is to be made for appropriate extraction of product features from the online reviews collection.

The organization of the paper is as follows: Terminology and Notation used in the paper is described in section 3, generation of feature and opinion topic clusters and generation of topic reviews are explained in section 4, engineering the FOT and extracting the appropriate product features using domain independent FOT on topic clusters is presented in section 5, identifying the opinion bearing word of the obtained feature by backtracking to reviews with the help of document indicators from the topic matrix and using Turney and Litmann word opposition sets and calculating its opinion orientation using SentiWordNet are discussed in section 6, a brief discussion of results is presented in section 7 and conclusion and further work is specified in section 8.

3. Terminology and Notation

The symbols and notations used in this paper are presented below.

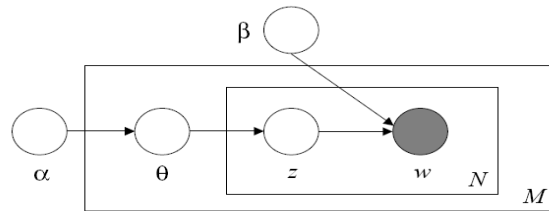


Fig.1. LDA Plate Model

A Document is a collection of topics. “A document exhibits multiple topics” [2] is the basic assumption of the LDA. A Topic is a distribution over words of a document. A topic contains collection of vocabulary of words. These words get regularly updated when topics were refined. The final solution of LDA is to generate a coherent topic. A Word is a term used to describe a thing or to express a concept. In LDA, a document contains words which are grouped under topics. Together dirichlet distribution and multinomial distribution are used in the generation of unknown parameters like topic distributions, topics and co-occurring words in the unsupervised learning process. Multinomial distribution is used to describe another distribution. It is a generalization of binomial distribution. It is the probability of observing each possible outcome c_i exactly X_i times in a sequence of yes/no trials. The formula for multinomial distribution is;

$$P(X_1 = x_1, \dots, X_k = x_k) = \frac{n!}{x_1! \dots x_k!} p_1^{x_1} \dots p_k^{x_k} \tag{1}$$

In our work, X_i 's are the individual word count across all documents and ‘n’ is the total word count of the document collection. The probability ‘ p_i ’ is the probability of the word fit in to a topic.

The Dirichlet distribution [10] is an exponential family distribution over the simplex, i.e., positive vectors that add up to the value one. It is conjugate to the multinomial distribution. This means, given a multinomial observation, the posterior distribution of θ is a Dirichlet. The topic proportions are a K dimensional Dirichlet (allocation of topics in the reviews collection). The topics are a V dimensional Dirichlet (distribution of words in a topic collection). The formula of Dirichlet distribution is;

$$P(\theta | \alpha) = \frac{\Gamma(\sum_{i=1}^k \alpha_i)}{\prod_{i=1}^k \Gamma(\alpha_i)} \theta^{\alpha_1-1} \dots \theta^{\alpha_k-1} \quad (2)$$

The Dirichlet distribution is the “conjugate prior” to the multinomial distribution. This signifies that once the prior distribution of the multinomial parameters is Dirichlet then the posterior distribution is also a Dirichlet.

In computer science, ontology [3] represents formally the knowledge as set of concepts within a domain, and the relationships among individual concepts. It is used to engineer a domain and carry reasoning about concepts.

4. Product Reviews Clusters Generation using Latent Dirichlet Allocation Topic Model

LDA topic model clusters were generated using the proposed model specified in Fig. 2. The unstructured dataset of the product reviews are to be first preprocessed for removing stop words. After preprocessing, the obtained dataset is left with only the product features, sub features and corresponding opinion words.

A. Generation of Feature and Opinion Topic Matrix

The obtained dataset is given as an input to the LDA topic modeling tool. This generates the document as specified in the Fig. 2.

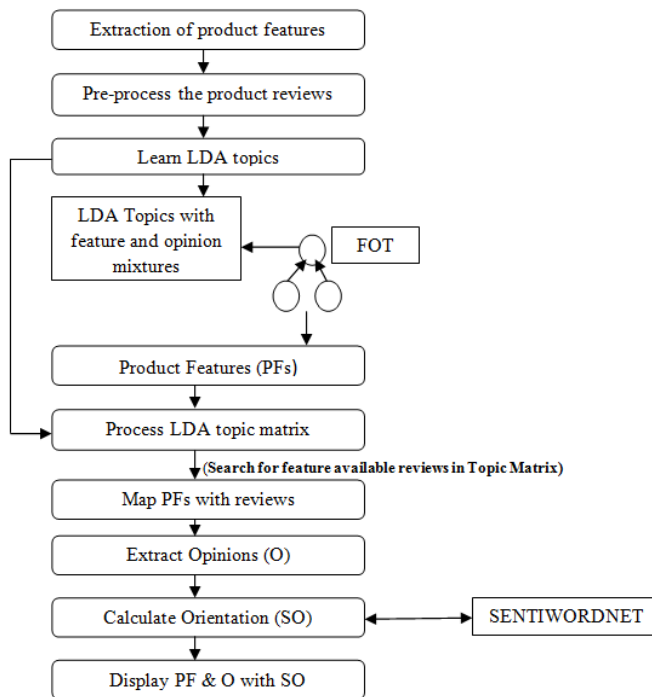


Fig.2. Proposed Model

LDA learns user specified topics by first randomly assigning vocabulary of words of the documents to the topics using dirichlet distribution. Then, it calculates the Term Frequency-Inverse Document Frequency (TF-IDF) value in the form of word probability to clearly assign the topics to the words so that LDA maximizes the generation of documents with near appropriate co-occurring words. These words with corresponding word position in a document and TF-IDF value are stored in a matrix called Topic Matrix. The pseudo code is explained in Fig. 3 which is presented below.

Step 1: LDA model learns ‘k’ input topics.
 Step 2: It randomly assigns topics to every word in the document. This assignment is based on the dirichlet distribution.
 Step 3: Is to refine the topic assignment, each word in the every document is checked with the current topic assignment and is updated based on the following two steps.

Step 3.1: If the probability of topic of a word is exists in the given document (Term Frequency), then $P(\text{topic} | \text{document})$
 Step 3.2: If the probability of a word is exists across the document collection given the current topic assignment (Inverse Document Frequency), then $P(\text{word} | \text{topic})$

Step 4: The weighing conclusion of assignment of a topic to word is the probability of the word. A topic matrix is generated with word and probability values.

Fig.3. Feature and Opinion Topic Matrix

B. Probabilistic Document Generation – Topic Clusters

LDA learns the number of words in a generated document using Poisson distribution. When the size of the vocabulary is huge, Poisson distribution changes to Multinomial and in some cases this transforms to the dirichlet distribution. Then, LDA generates prior topic proportions in every document using dirichlet distribution. For every word in the reviews collection, a single topic is chosen at random using multinomial distribution applied on the generated topic proportions. Now, from the chosen topic, a word is chosen using multinomial distribution to generate the document. This is explained in Fig. 4 as a process to ensure the document generation using the words from the topic matrix which is presented below.

Step 1: Choose $N \sim \text{Poisson}()$, where N represents sequence of words in a document.
 Step 2: Choose $\theta_d \sim \text{Dir}(\alpha)$.
 Step 3: For each of the N words W_n :

Step 3.1: Choose a topic $z_{d,n} \sim \text{multi}(\theta_d)$.

$P(z_{d,n} | \theta_d)$

Step 3.2: Choose a word $w_{d,n} \sim \text{multi}(\beta_{z_{d,n}})$, a multinomial probability conditioned on the topic.

$P(w_{d,n} | \beta_{z_{d,n}})$

Fig.4. Probabilistic Document Generation using words from Topic Matrix

5. Identification of Appropriate Product Features from the LDA Clusters Using Domain Independent FOT Tree

The topic clusters generated from LDA overlap as same words come from different documents with different probabilities. Some of the feature words and opinion words were available in these clusters and contain the words which are non features of the product when comparative reviews hold on other product (its features) data in them. To identify the appropriate product features from these clusters, domain independent FOT tree is applied on the clusters. The formal definition of FOT is specified below.

Definition: FOT is an abbreviation for Feature Ontology Tree that is a tree like ontology structure $T(v, T')$. v is the root node of T which represents a feature of a given product. T' is a set of subtrees. Each element of T' is also a FOT $T''(v1, T''')$ which represents a sub feature of its parent feature node.

Feature Ontology Tree (FOT) is a domain independent OWL-lite Ontology that specifies product features and their sub features as features in a hierarchy. The excerpt of FOT is illustrated in Fig. 5.

The steps in extracting exact product features starts with mapping the OWL class with the corresponding topics by considering the OWL classes Feature and ObjectPartFeature to map with the LDA generated topic clusters and then the ontology instances were compared with the topic words in the clusters. The words are extracted from the cluster when a match is found and storing the obtained product features.

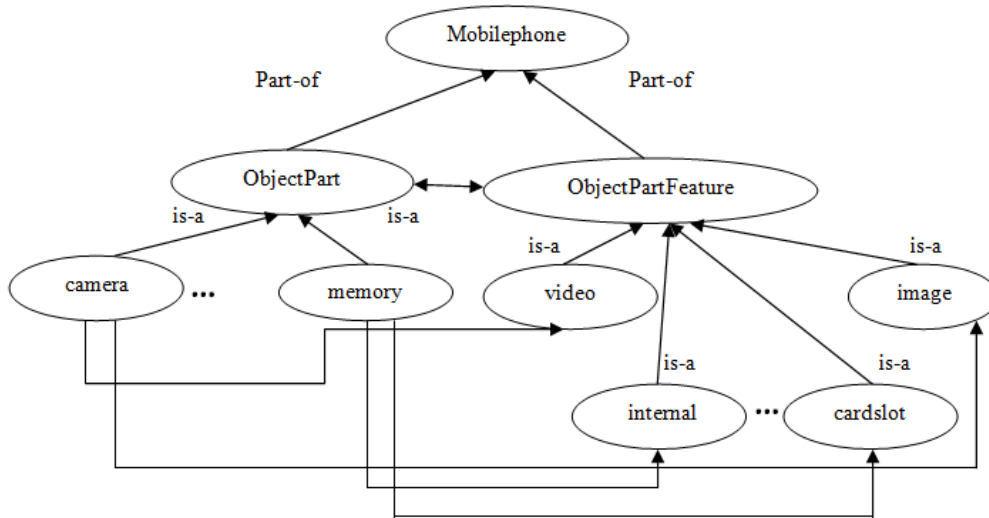


Fig.5. Excerpt of Feature Ontology Tree

6. Feature Specific Opinion Word Identification and Orientation

The work of the researchers in [21] is extended in this section. The identified product features from the LDA topic clusters by applying FOT tree on them are used in searching the associated review documents in which the feature was commented on. The opinion words of the features are identified and orientation of these words is calculated. This is carried out by the following steps.

A. Relevant review documents search in Topic Matrix

The task of searching the appropriate review documents for the identified product features is important as

these documents contain the reviews written on these features. The review documents of the identified features are obtained from the topic matrix (β) (contains document indicator, chosen word, TF-IDF value) generated earlier for learning the pre specified topics. This task is possible as the words are chosen from the multinomial distribution conditioned on the dirichlet distributed topics in the review document for topic matrix generation. This results in the reviews having the features upon which it is commented on and as one of the multiple features commented on in the review. The excerpt of the topic matrix is illustrated in Fig. 6.

Topic 1	Topic 2
$W_{1,3}$ price 0.60	$W_{3,1}$ good 0.40
$W_{1,2}$ bluetooth 0.45	$W_{2,6}$ fast 0.08
$W_{2,2}$ price 0.45	$W_{2,4}$ cheap 0.08
$W_{3,6}$ price 0.40	$W_{1,4}$ awesome 0.05
...	...
...	...

Fig.6. Excerpt of the Topic Matrix

B. Extraction of feature surrounding adjective

The subjective word (opinion word) written alongside the product features are the adjectives in most of the written reviews. Pang et al. [20] achieved 83% accuracy in identifying positive and negative opinion words by working with only adjectives in their research. In this work also the researchers concentrated on the adjectives surrounding the identified features. The adjective enclosing the feature from the explored review documents are extracted by simply learning the adjectives tagged by the Stanford Log Linear Part of Speech Tagger [22] on the specified review. The extracted adjectives are combined with the corresponding features for further analysis.

C. Identification of adjective as opinion word

The identification of adjective as opinion word is carried out by considering the positive and negative seed sets as specified by Turney and Littmann [17]. These adjectives are compared with the considered seed sets. The matched adjectives are considered as opinion words and are stored separately with corresponding features and document indicator.

D. Calculation of opinion orientation of the opinion word

The identified opinion words are analyzed for the opinion orientation of the feature. This is carried out with the help of SentiWordNet [18] which is a lexical resource for calculating sentiment of a review, a statement or a word respectively. The learned opinion orientations are counted upon the need for learning the number of positive and negative views on a particular feature.

7. Results and Discussion

The precision for the appropriate product features were identified from the above discussion which is calculated by considering the manual features done by Hu & Liu [5]. The number of features from the considered cellular phone product is 67. We have considered only one product type for the analysis as the FOT tree is developed for a class of mobile phones of different manufacturers. The product reviews considered for extraction is Nokia 6610 cellular phone. The number of reviews extracted from Amazon.com website is 5000.

The number of topics to learn by LDA is considered as 50. The concentration parameters α and β are set as $50/k$ and $200/W_n$ (W_n is the total number of words in the reviews collection) respectively. These values of α and β are suggested by Griffiths & Steyvers [19] for better topic clusters generation. The FOT technique is implemented using Protégé [16] Ontology development environment. The domain independent FOT is applied on the generated topic clusters and the number of identified product features is 64. These product features are searched in the topic matrix for clear identification of its presence in the reviews. SentiWordNet is applied on these reviews to learn the opinion orientation of the extracted feature. Carenini. G et al., [12] pioneered a superior unsupervised method for feature mining that uses the categorization of the product features. The outcomes of this experiment are higher than the existing unsupervised technique presented in [5] which has precision in extracted product features of 76.1%. Nevertheless, the pre-knowledge base mechanism makes the approach domain dependent. We have compared our FOT results with both OPINE [11] and Likelihood Ratio Test (LRT) [13] techniques. The precision of extracted features of the product using FOT outperforms over the other two techniques. The comparison we have done is shown in Table 1.

Table 1 Comparison of Information Retrieval measures of Extracted Product Features using OPINE, LRT+dBNP-L and FOT Approaches

S.No.	Features Extraction Techniques	Precision	Recall	F-measure
1.	LRT+dBNP-L	0.742	0.197	0.31
2.	OPINE	0.79	0.76	0.77
3.	FOT	0.955	0.853	0.90

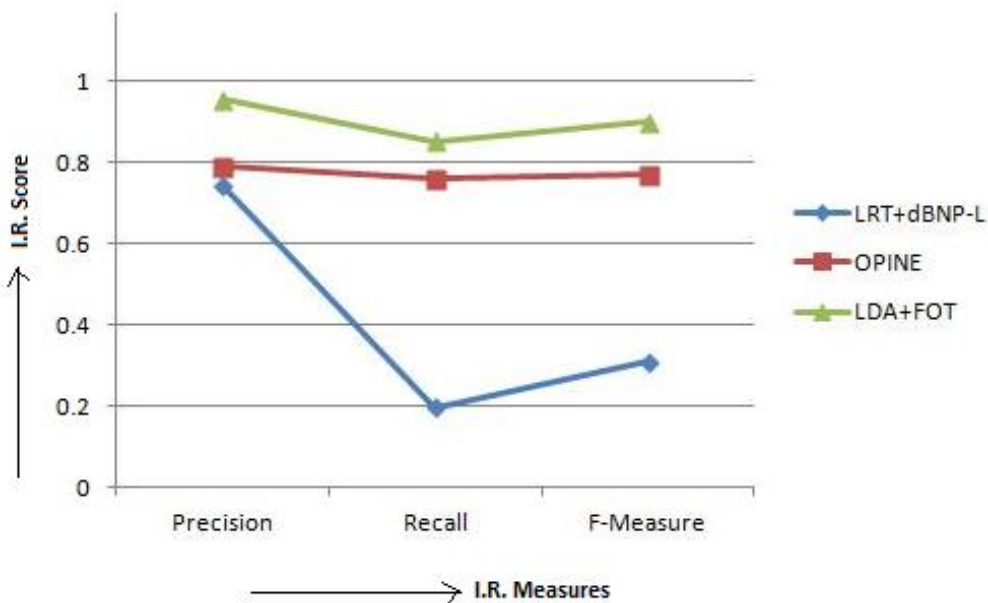


Fig.7. Line Graph

The recall using FOT for product features extraction is 10% less w.r.to recall when compared with OPINE

[11] as 10 product features which are retrieved as relevant features found to be the plural form of the node in the FOT tree. The table below is the population of extracted features and the identified polarities of the extracted features (both positive and negative orientations).

Table 2. Extracted Features and Opinion Words from Topics using FOT and SentiWordNet

Extracted Features	ringer, nokia, roaming, calling, apps, email, fm, backlight, menu, earpiece, games, infrared, gsm, speaker, hardware, network, vibrate, bluetooth, color, picture, appearance, voice, button, screen, phone, voice dialing, signal, graphics, camera, keys, headphones, battery, jack, light, sound, layout, stereo, upgrading, calendar, modem, tone, volume, speakerphone, messenger, resolution, radio, gprs, wallpaper, music, size, keypad, plan, ringtone, zip, mms, browser, outlook, laptop, weather, scroll, weight, software, port
Opinion Words	Positive words: {loud, best, glad, highly, fully, excellent, convenient, super easy, accessible, comfortable, favourite, great, bright, nice, good, amazing, better, well, more, powerful, cute, glad, awesome, decent, cool, superb, perfect, small, compact} Negative words: {too quickly, disappointment, only, lacks, not crisp, not clear, small, tiny, sticky, not have, distorted, funky, subtle, hard, useless, slow}

8. Conclusion and Future Work

The extraction of appropriate product features using domain independent FOT on LDA topic clusters is carried out successfully. Opinion orientation of the identified features was carried out successfully using SentiWordNet. The LDA parameters ‘ α ’ and ‘ β ’ are clearly specified in topic clusters generation process. The obtained features and their opinions help the new customer to read feature specific reviews.

In the future, the traditional LDA model is modified to identify the opinion bearing words and classifies the orientations of their opinions of the extracted features and can compare with the sentiwordnet calculated opinion orientations.

References

- [1] Dandibhotla, T., & Bulusu, D., Obtaining Feature- and Opinion-Based Linked Instance RDF Data from Unstructured Reviews using Ontology-Based Machine Learning, *International Journal Of Technology*, 6(2), 198-206, 2015. doi:10.14716/ijtech.v6i2.555.
- [2] David M. Blei, Andrew Y. Ng and Michael I. Jordan, Latent Dirichlet Allocation, *Journal of Machine Learning Research*, Issue 3, pages 993-1022, January 2003.
- [3] Ian Horrocks, *Ontologies and the Semantic Web*, ACM 2009.
- [4] Hofmann, T. (1999). Probabilistic latent semantic indexing. *Research and Development in Information Retrieval*, pages 50–57.
- [5] Hu, M., Liu, B., 2004. Mining and Summarizing Customer Reviews. In: *Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, August 22–25, 2004, Seattle, WA, USA.
- [6] S. Brody, N. Elhadad, An unsupervised aspect-opinion model for online reviews, in: *HLT-NAACL, The Association for Computational Linguistics*, 2010, pp. 804–812.
- [7] C. Lin, Y. He, Joint opinion/topic model for opinion analysis, in: *Proceedings of the 18th ACM Conference on Information and Knowledge Management (CIKM’11)*, 2009.

- [8] Jo, Yohan, and Alice H. Oh., Aspect and opinion unification model for online review analysis. *Proceedings of the fourth ACM international conference on Web search and data mining*. ACM, 2011.
- [9] Zofia Stankiewicz and Satoshi Sekine, SurfShop: Combining a Product Ontology With Topic Model Results for Online Window Shopping, NAACL-HLT 2012. Montreal, Canada, pp 13-16.
- [10] BA Frigyik, A. Kapila, and M.R. Gupta. Introduction to the dirichlet distribution and related processes. Department of Electrical Engineering, University of Washington, UWEETR-2010-0006, 2010.
- [11] Popescu A. and Etzioni O., Extracting Product Features and Opinions from Reviews, in Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing, USA, pp. 339-346, 2005.
- [12] Carenini G., Ng R., and Zwart E., Extracting Knowledge from Evaluative Text, in Proceedings of the 3rd International Conference on Knowledge Capture, USA, pp. 11-18, 2005.
- [13] Ferreira L., Jakob N., and Gurevych I., A Comparative Study of Feature Extraction Algorithms in Customer Reviews, in Proceedings of IEEE International Conference on Semantic Computing, Santa Clara, California, USA, pp. 144-151, 2008.
- [14] Yi J., Nasukawa T., Bunesco R., and Niblack W., Sentiment Analyzer: Extracting Sentiments about a Given Topic Using Natural Language Processing Techniques, in Proceedings of the 3rd IEEE International Conference on Data Mining, Washington, USA, pp. 427-434, 2003.
- [15] Ghobadi A. and Rahgozar M., An Ontology based Semantic Extraction Approach for B2C eCommerce, The International Arab Journal of Information Technology, vol. 8, no. 2, pp. 163-170, 2011.
- [16] Alani, Harith, et al., Using protege for automatic ontology instantiation, 2004.
- [17] Turney, P. D., & Littman, M. L., Measuring praise and criticism: Inference of semantic orientation from association. *ACM Transactions on Information Systems*, 21(4), 315–346, 2003.
- [18] Santosh, D. Teja, B. Vishnu Vardhan, and D. Ramesh. "Extracting Product Features from Reviews Using Feature Ontology Tree Applied on LDA Topic Clusters." *Advanced Computing (IACC)*, 2016 IEEE 6th International Conference on. IEEE, 2016.
- [19] Griffiths TL, Steyvers M (2004). "Finding Scientific Topics." *Proceedings of the National Academy of Sciences of the United States of America*, 101, 5228–5235.
- [20] B. Pang, L. Lee, and S. Vaithyanathan, "Thumbs up? Sentiment classification using machine learning techniques," in *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 79–86, 2002.
- [21] D. Teja Santosh, B. Vishnu Vardhan "Extracting product features from reviews using Feature Ontology Tree applied on LDA topic clusters." In *Proceedings of the Advance Computing Conference (IACC)*, IEEE, pp 163-168, 2016.
- [22] Kristina Toutanova, Dan Klein, Christopher Manning, and Yoram Singer. 2003. Feature-Rich Part-of-Speech Tagging with a Cyclic Dependency Network. In *Proceedings of HLT-NAACL 2003*, pp. 252-259.

Authors' Profiles



D. Teja Santosh is presently working as Assistant Professor in CSE at GITAM University Hyderabad. He is the research scholar in the faculty of Computer Science and Engineering, JNTUK, Kakinada. His areas of interest are data mining, machine learning, natural language processing. He was the Additional Reviewer for 2015 International Conference on Advances in Computing, Communications and Informatics (ICACCI) in Kerala, India. He is the member of professional bodies of international repute namely IEEE, ACM, IAENG. He is one among the 16 research scholars of Dr. B. Vishnu Vardhan.



K. Sudheer Babu is presently working as Assistant Professor in CSE at GITAM University Hyderabad. His areas of interest are data mining, machine learning, natural language processing. He is the member of professional bodies of national and international repute namely ISTE, IAENG. He has published many research papers in the elite areas of Computer Science and Engineering. In 2014 he has published an IEEE international conference paper at Kuala Lumpur, Malaysia.



S.D.Vara Prasad is presently working as Assistant Professor in the department of CSE, School of Technology, GITAM University, Hyderabad. He has 10 years of teaching experience at UG and PG Level. His areas of interests are Web & Data Mining, Information Security. He has published many research papers in the elite areas of Computer Science and Engineering. He is presently pursuing his Ph.D from Acharya Nagarjuna University, Guntur, Andhra Pradesh.



A. Vivekananda is the graduate student from the department of Computer Science and Engineering, School of Technology, GITAM University, Hyderabad. His areas of interests are Web & Data Mining, Natural Language Processing.

How to cite this paper: D. Teja Santosh, K. Sudheer Babu, S.D.V. Prasad, A. Vivekananda, "Opinion Mining of Online Product Reviews from Traditional LDA Topic Clusters using Feature Ontology Tree and Sentiwordnet", International Journal of Education and Management Engineering(IJEME), Vol.6, No.6, pp.34-44, 2016.DOI: 10.5815/ijeme.2016.06.04