Grouping of Aspects into Relevant Category Based on WordNet Definitions

Sheikh Muhammad Saqib[1], Shakeel Ahmad[2], Asif Hassan Syed[2], Tariq Naeem[1], Fahad Mazaed Alotaibi[2]

[1]Institute of Computing and Information Technology, Gomal University, Dera Ismail Khan, Pakistan.
[2]Faculty of Computing and Information Technology in Rabigh (FCITR) King Abdul Aziz University (KAU) Jeddah Saudi Arabia

Summary
Machine Translation, Information Retrieval and Knowledge Acquisition are the three main applications of sentiment analysis based on aspect word. But in some cases, there may be two aspect words having different spellings, yet belonging to the same aspect category, e.g. the words ‘costly’ and ‘expensive’ belong to the same category ‘price.’ Grouping such aspects into the same category is critical for sentiment analysis. Some have done this by using synonym, LDA (Latent Dirichlet Analysis) groups and Constrained-LDA. The WordNet dictionary is mostly used for such analysis. Although it contains a lot of synonyms, yet sometimes there are such aspects which are not synonymous, but still refer to the same broad category. For example, in the following two sentences: “this camera will not easily fit in a coat pocket” and “this camera is of large size;” aspects ‘fit’ and ‘size’ both are not synonyms yet belong to the same category. In this study, the proposed framework consists of four steps: DefString, Supporting-Words, SetWords and Intersection for Decision. The fourth stage will decide that the two aspects belong to the same category or not. After experimenting on a dataset comprising 252 pairs of same-meaning aspects and 62 pairs of different-meaning aspects, 91% accuracy and 86% F score were detected.

Key words:
Machine Translation, Information Retrieval, Sentiment Analysis, Aspect Extraction.

1. Introduction

In opinion mining field, when we want to analyze any documents e.g. summary analysis or sentiment analysis of document; besides all other efforts, it is of primary importance to find out whetherthe selected document belongs to the required aspects or not. Detection of aspect is also a challenging problem, but the grouping of aspect is another challenge.

A semi-supervised learning method was proposed to grouped aspect expressions into some user-specified aspect categories [1][2]. In such an analysis manual labelling is required for each category. Work based on Expectation-Maximization (EM) algorithm also required labeling and synonyms[3][4]. Labelling is an extra effort and synonym-problemsare discussed at the end of this section. A method called multilevel latent semantic association was presented. Here, all the words in aspect expressions (eachaspect expression can have more than one word) are grouped into a set of concepts/topics determined by LDA. Then, the detected topics can be combined with their surrounding words to obtain a particular category[5]. But most of the online product reviews have single lines, while the proposed work will determine topic/aspect using JJ or NN easily without LDA.Constrained-LDA uses a technique where two aspect expressions ai and aj share one or more words [6]. But there are also reviews which belong to the same aspect but do not share the same words.

Similarity path distance can also be useful for grouping aspect into category as used in concept similarity model [7] , but in WordNet dictionary score of word “car” and “auto” is 1.0, while some word-pairs not similar in sense so they have less or none score, yet they belong to same aspect category as shown in Table-1.

Works based on synonyms also follow a very good approach for grouping aspect but in dictionary word1 is the synonym of word2 while word2 might have no synonym as word1 e.g. the word ‘picture’ has a synonym, “movie;” while the word “movie” does not have the word “picture” among its synonyms. In Table-1, G1 means Grouping Through Synonym, G2 means Grouping Through Path Distance and G3 means Grouping Through Proposed Method.

Table 1: Problem Identification vis-à-vis Previous Works

S.No.	*Words*	*Path Distance*	*G1*	*G2*	*G3*
1	('picture', 'movie')	0.2	Yes	No	Yes
2	('car', 'auto')	1.0	Yes	Yes	Yes
3	('expensive', 'costly')	None	No	No	Yes

In Table-1:
-All those words which belong to S.No. 1 can be grouped through G1 but not through G2, while the proposed work (G3) will determine the group of those words.
-All those words which belong to S. No. 2 can be grouped through G1, G2 and also through the proposed work G3.
-All those words which belong to S. No. 3 cannot be grouped through G1 and G2, while the proposed work will

Manuscript received February 5, 2019
Manuscript revised February 20, 2019

be able to group them in the same category. This is the major contribution of the work.

2. Related Work

A word in a review can be extracted as aspect and another word in another review extracted as aspect, both different words may relate to same aspect i.e. both reviews will be put in same category. This is essential because the majority of research studies to date, have investigated opinion words or examined text mining through the creation of summaries of a given document based on aspect [8], [9]. By grouping aspect, a document or sentence can easily be understood by users as well as by intelligent machines. For example, human intelligence can automatically sense and detect the meaning of a word from examining a sentence and grouped the aspects. However, in the field of artificial intelligence, efforts are continuing to be made to understand a sentence from the correct dimension or aspect given that a different word may have same aspect.

How can a company improve the quality of a product, place, etc. from the huge amount of reviews? Many studies have been carried out with regard to sentiment analysis, which is about determining the sentiment orientation of a review or comment [10][11][12][13]. Sentiment orientation means that a positive opinion will be an exact positive, and a negative opinion will be an exact negative [14]. The view, assessment or feeling of a person towards a product [15], aspect [16], or service is known as a sentiment [17][18][19]. Such a feeling, which is either positive or negative, can be assigned a score. Most of the work in sentiment analysis is based on binary classification, which means that reviews or blogs are divided into “positive” and “negative” classes [20][21]. The classification of text sentiments can be done in two ways, i.e. through machine learning and score-based approaches [22][23]. Machine learning uses training data [24], while the other method uses several attributes of an entity to determine the scores. In the score-based approach, opinions can be oriented as positive or negative [25][26]. The work of [27] used a combined approach of SentiWordNet and lexical resources to determine the scores for slangs. A lexicon-based approach for extracting sentiment orientations of opinions has been used for scoring [28][29]. Studies by [30][31] used lists of positive and negative words to determine the polarity of a sentence by creating a training matrix and random forest classifier based on supervised learning. A sentiment analysis can be performed using different methods [32][33][34][35], with each method having an improved accuracy with respect to the previous one. Although a lot of work is involved in sentiment orientation [36] with the use of adjectives, frequent nouns and noun phrases, sentiment shifters, handling of ‘but’ clauses, decreased and increased quantity of an opinionated item; high, low, increased and decreased quantity of a positive or negative potential item; desirable or undesirable facts; deviations from the norm or a desired value range; and the production and consumption of resources and waste, etc., these are very important for determining the polarity of a document or sentence [37][38][29]. However, a large amount of online data is generated every day with unprecedented speed and size. Most of the available information on the Internet is in text and unstructured forms, i.e. online reviews, blogs, chats, and news. An aspect-based sentiment analysis, which can be carried out by using only particular aspects [29][39][40][41], requires less effort compared to a sentiment analysis of an object with respect to all aspects. Reviews are rated according to an object, so there should be a direct method to determine whether a review is positive or negative. LSI (Latent Semantic Indexing) is better for such a purpose [42][43]. LSI [44] has been used for the clustering of documents and for concept representations. An extended method based on LSI is able to filter unwanted emails in Chinese and English [45].To improve accuracy of sentiment analysis, lot of work has also been done on words sense disambiguation[46][47][48]. Machine learning approaches, also called corpus-based approaches, do not make use of any knowledge resources for disambiguation [49]. Most accurate WSD systems to date exploit supervised methods which automatically learn cues useful for disambiguation from manually sense-annotated data [50], [51].

Now to detect an entity from an opinion has been done using different methods. Author [52] has used two sets i.e. the set of seed entities Q and the set of candidate entities D to determine which of the entity in D belongs to C. For entity extraction [53][54] have used a method distribution similarity by comparing the similarity of the surround words of each candidate entity with those of the seed entities; and then ranking the candidate entities based on the similarity values. Topic modeling has also been used for entity extraction. Topic modeling is an unsupervised learning method that assumes each document consists of a mixture of topics and each topic is a probability distribution over words[55]. Latent Dirichlet Analysis LDA and Probabilistic Latent Semantic Analysis (PLSA) have been used for detection of topic from a document/documents [56]. To search a required document from a huge amount of published articles is very time consuming and laborious work, topic modeling offers a computational tool to find relevant topics by capturing meaningful structure among the collections of documents [57][58]. Grouping aspect into categories has also improved aspect based sentiment analysis [59].

3. Grouping of Aspect into Relevant Category Based on WordNet Definitions

In our previous work [58], we have determined how to extract aspect words from a review. Now these extracted words will be grouped with already detected categories. If category found, then review will be grouped in the same category; otherwise the review will be placed in a new group with its aspect name as shown in Figure-1.

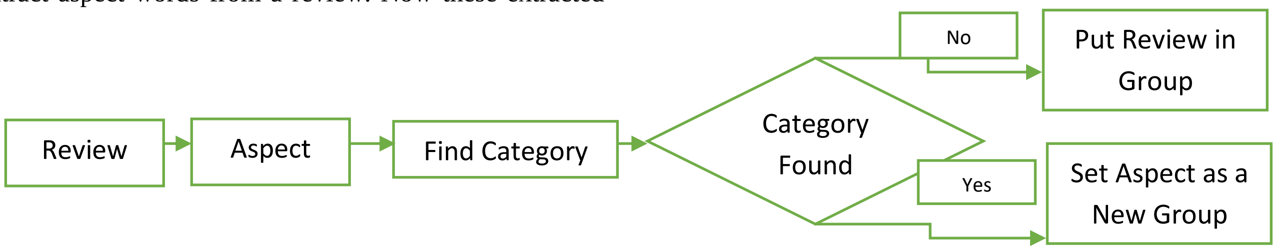

Fig. 1 Working Application of Proposed Work

Following is the work to determine the category to group the Aspect-1 and Aspect-2 as shown in Figure-2.

3.1. DefString

This is a string which is the combination of all definitions of Aspect from WordNet dictionary. Here DefString1 contains all the definitions of Aspect-1, while DefString2 contains all definitions of Aspect-2.

3.2. Supporting-Words

Tag the DefString, extract only Adjectives (JJ) and Nouns (NN) from Tagged string. In Figure-2, Supporting-Words1 are the collection of 'JJ' & 'NN' from DefString-1 and Supporting-Words2 are the collection of 'JJ' & 'NN' from DefString-2.

3.3. SetWords

Now convert Supporting-Words as a set, then we can take all operations of sets on it. Here SetWords1 is a set of Supporting-Words1 and SetWords2 is a set of Supporting-Words2.

3.4. Intersection for Decision

Now take intersection of these two sets SetWords1 and SetWords2. If resultant set is not empty, it means Aspect-1 and Aspect-2 both belong to the same category otherwise both belong to a different category.

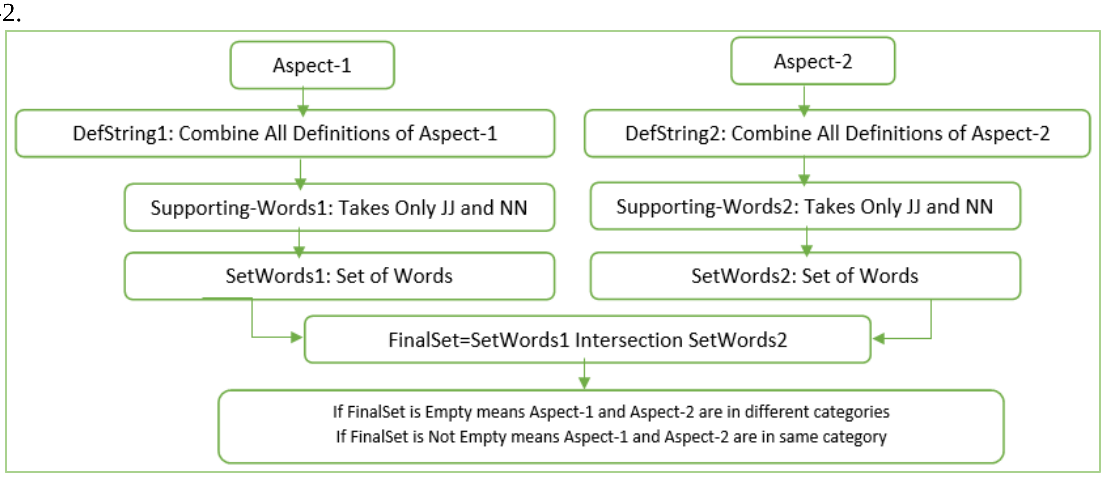

Fig. 2 Proposed Methodology

Here we are taking two pairs, first pair ("expensive", "costly") belongs to the same aspect, while the second pair ("beautiful", "large") belongs to different aspect. In Table-2, total elements from intersection is 2 (non-empty) it means first pair belongs to same aspect. In Table-3, total elements from intersection is 0 (empty) it means second pair belongs to a different aspect.

Table 2: Example based on First Pair (Same Aspect)

Processes	***Aspect-1 (expensive)***	***Aspect-2 (costly)***
DefString	high in price or charging high prices	entailing great loss or sacrifice having a high price
Supporting-Words	[(u'high', 'JJ'), (u'price', 'NN'), (u'high', 'JJ')]	[(u'great', 'JJ'), (u'loss', 'NN'), (u'sacrifice', 'NN'), (u'high', 'JJ'), (u'price', 'NN')]
SetWords	Set {'high', 'price', 'high'}	Set {'great', 'loss', 'sacrifice', 'high', 'price'}
Intersection	Set {'high', 'price'}	
Decision	Total Element =2, it means "expensive" and "costly" belongs to same category	

Table 3: Example based on Second Pair (Different Aspect)

Processes	***Aspect-1 (beautiful)***	***Aspect-2 (large)***
DefString	delighting senses exciting intellectual emotional admiration (of weather) highly enjoyable	garment size large person above average size number quantity magnitude extent fairly large important effect;
Supporting-Words	[(u'intellectual', 'JJ'), (u'emotional', 'JJ'), (u'admiration', 'NN'), (u'enjoyable', 'JJ')]	[(u'garment', 'NN'), (u'size', 'NN'), (u'large', 'JJ'), (u'person', 'NN'), (u'average', 'JJ'), (u'size', 'NN'), (u'number', 'NN'), (u'quantity', 'NN'), (u'extent', 'JJ'), (u'large', 'JJ'), (u'important', 'JJ'), (u'effect;')]
SetWords	Set{'intellectal', 'emotional', 'admiration', 'enjoyable'}	Set {'garment', 'size', 'large', 'person', 'average', 'size', 'nmber', 'qantity', 'extent', 'large', 'important', 'effect;'}
Intersection	Set { }	
Decision	Total Element =0, it means "beautiful"and "large"belongs to different category	

4. Results and Discussions

A complex example from [14] has been solved easily from proposed method i.e. "This camera will not easily fit in a coat pocket" and "this camera is of large size" here fit and size both are not synonyms but belongs to same aspect category as depicted in S. No. 1 of Table-4. To check the accuracy of proposed model, we have collected 252 synonyms (same meaning) pair of words and 62 words which have not shared same meaning[60][61]. Some of the samples are shown in Table-4. In this table, G1 means Grouping Through Synonym, G2 means Grouping Through Path Distance and G3 means Grouping Through Proposed method. All pairs are detected correctly by the proposed method G3, i.e. ('beautiful', 'large') and ('chair', 'car') both belong to different aspect so proposed method decided correctly.

Table 4: Sample of Results from 314 collected pairs

S.No	***Words***	***Path Distance***	***Intersection***	***Total Words***	***G1***	***G2***	***G3***
1	('fit', 'size')	0.1	set([u'something', u'size'])	2	No	No	Yes
2	('expensive', 'costly')	None	set([u'high', u'price'])	2	No	No	Yes
3	('car', 'auto')	1.0	set([u'engine', u'combustion', u'wheels;', u'internal', u'motor', u'vehicle'])	6	Yes	Yes	Yes
4	('chair', 'car')	0.1	set([])	0	No	No	No
5	('smart', 'clever')	None	set([u'quick'])	1	No	No	Yes
6	('interesting', 'fascinating')	0.1428571	set([u'attention', u'interest'])	2	No	No	Yes
7	('administer', 'manage')	0.1666666	set([u'charge', u'direct'])	2	No	No	Yes
8	('native', 'local')	0.1	set([u'characteristic', u'particular'])	2	No	No	Yes
9	('object', 'thing')	0.125	set([u'entity'])	1	No	No	Yes
10	('picture', 'movie')	0.2	set([u'sound', u'story', u'form', u'entertainment', u'continuous', u'sequence', u'illusion', u'movement'])	8	Yes	No	Yes
11	('beautiful', 'large')	None	set([])	0	No	No	No

4.1 Statistical Measures

A confusion matrix [62] is formed from the four outcomes produced as a result of binary classification. A binary classifier predicts all data instances of a test dataset as either positive or negative.Here we consider classification as Yes and No i.e. 'Yes' means two words belong to same aspect and No means two words belong to different aspects. This classification (or prediction) produces four outcomes: –True-Yes (TY): correct same aspect prediction, False-Yes (FY): incorrect different aspect prediction, True-No (TN): correct different aspect prediction, False-No (FN): incorrect same aspect prediction. Various measures can be derived from a confusion matrix as shown in Table-6.

After analysis of proposed work on said words, obtained results are shown in Table-5. In Table-5, G1 means Grouping Through Synonym, G2 means Grouping Through Path Distance and G3 means Grouping Through Proposed Method.

Table 5: Detected Parameters for Confusion Matrix

Methods	***Outcomes***	***Score***	***Outcomes***	***Score***
G1	True-Yes	0	False-No	252
	True-No	61	False-Yes	0
G2	True-Yes	39	False-No	213
	True-No	61	False-Yes	0
G3	True-Yes	204	False-No	48
	True-No	43	False-Yes	18

Applying above values in online confusion matrix, and calculated results shown in Table-6.

Table-6: Statistical Measures

Measure	***Derivations***	***Values***
Sensitivity or True Yes Rate (TYR)	TYR = TY / (TY + FN)	0.8095
Precision or Yes Predictive Value (YPV)	YPV = TY / (TY + FY)	0.9189
Accuracy	ACC = (TY + TN) / (Y + N)	0.7891
F1 Score	F1 = 2TY / (2TY + FY + FN)	0.8608
False Yes Rate	FYR = FP / (FP + TN)	0.2951
False No Rate	FNR = FN / (FN + TP)	0.1905

Table-6 is depicting the results of grouping of detected aspects of product reviews and obtained improved precision (91%), accuracy (78%) and F1-score (86%) when compared to the previous methods from Table-5.

5. Conclusion

Searching of documents is not only a part of search engine, but also very important for sentiment analysis, text summarization, etc. Aspect based sentiment analysis also requires those documents related to same aspects. How these documents are identified, there is need of some words which can determine those documents which are related to these aspects. Instead of comparing these aspects with all the documents, just compare these aspects with the aspects of an individual document. If there is a list of aspect words, then there is need to group the aspects, if same aspects are defined in different words. In this study, grouping aspects into same category is investigated. Based on the experimental results through confusion matrix, it was found that results generated from the proposed method show a significant improvement from existing texts related not only to precision and accuracy, but also to FYR (False Yes Rate) and FNR (False No Rate) which are 29% and 19% respectively. Grouping aspect is not only fruitful for searching, information retrieval, clustering; but also helpful for sentiment analysis, natural language processing and word sense disambiguation.

Conflict of Interest: Sheikh Muhammad Saqib, Shakeel Ahmad, Asif Hassan Syed, Tariq Naeem and Fahad Mazaed Alotaibi declare that they have no conflict of interest.

References

[1] S. Zhou, Q. Chen, and X. Wang, "Active Deep Networks for Semi-Supervised Sentiment Classification," Technology, no. August, pp. 1515–1523, 2010.

[2] A. Mukherjee and B. Liu, "Aspect extraction through semi-supervised modeling," ACL '12 Proc. 50th Annu. Meet. Assoc. Comput. Linguist. Long Pap. - Vol. 1, no. July, pp. 339–348, 2012.

[3] T. Wang and G. Hirst, "Refining the Notions of Depth and Density in WordNet-based Semantic Similarity Measures," Emnlp, pp. 1003–1011, 2011.

[4] J. J. Jiang and D. W. Conrath, "Semantic Similarity Based on Corpus Statistics and Lexical Taxonomy," 1997.

[5] H. Guo, H. Zhu, Z. Guo, X. Zhang, and Z. Su, "Product feature categorization with multilevel latent semantic association," Proceeding 18th ACM Conf. Inf. Knowl. Manag. - CIKM '09, p. 1087, 2009.

[6] Z. Zhai, B. Liu, H. Xu, and P. Jia, "Clustering product features for opinion mining," Proc. fourth ACM Int. Conf. Web search data Min. - WSDM '11, p. 347, 2011.

[7] S. Mohammad and G. Hirst, "Distributional measures of concept-distance: A task-oriented evaluation," Proc. Conf. Empir. Methods Nat. Lang. Process., pp. 35–43, 2006.

[8] and J. R. Blair-Goldensohn, Sasha, Kerry Hannan, Ryan McDonald, Tyler Neylon, George A. Reis, "Building a Sentiment Summarizer for Local Service Reviews," Proc. WWW-2008 Work. NLP Inf. Explos. Era., 2008.

[9] K. Lerman, K. Lerman, R. McDonald, and R. McDonald, "Contrastive summarization: an experiment with consumer reviews," in Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics, Companion Volume: Short Papers, 2009, no. June, pp. 113–116.

[10] T. Chen, R. Xu, Y. He, and X. Wang, "Improving sentiment analysis via sentence type classification using BiLSTM-CRF and CNN," Expert Syst. Appl., vol. 72, pp. 221–230, 2017.

[11] F. Wu, Y. Huang, and Z. Yuan, “Domain-specific sentiment classification via fusing sentiment knowledge from multiple sources,” Inf. Fusion, vol. 35, pp. 26–37, 2017.
[12] wang maoquan, C. Shiyun, X. yufei, and Z. lu, “EICA at SemEval-2017 Task 4: A Simple Convolutional Neural Network for Topic-based Sentiment Classification,” Proc. 11th Int. Work. Semant. Eval., pp. 737–740, 2017.
[13] D. Bollegala, T. Mu, and J. Y. Goulermas, “Cross-Domain Sentiment Classification Using Sentiment Sensitive Embeddings,” in IEEE Transactions on Knowledge and Data Engineering, 2016, vol. 28, no. 2, pp. 398–410.
[14] B. Liu, Sentiment Analysis and Opinion Mining, vol. 5, no. 1. 2012.
[15] J. Jin, P. Ji, and R. Gu, “Identifying comparative customer requirements from product online reviews for competitor analysis,” Eng. Appl. Artif. Intell., vol. 49, pp. 61–73, 2016.
[16] L. Shu, H. Xu, and B. Liu, “Lifelong Learning CRF for Supervised Aspect Extraction,” in Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics, 2017, pp. 148–154.
[17] F. e-M. K. Khan, B.B. Baharudin, A. Khan, “Mining opinion from text documents,” Adv. Res. Comput. Commun. Eng., vol. 3, no. 7, pp. 217–222.
[18] B. Pang and L. Lee, “Opinion mining and sentiment analysis,” Found. Trends Inf. Retr., vol. 2, no. 1–2, pp. 1–135, 2008.
[19] M. Z. Asghar, A. Khan, S. Ahmad, and F. M. Kundi, “A Review of Feature Extraction in Sentiment Analysis,” J. Basic. Appl. Sci. Res, vol. 4, no. 3, pp. 181–186, 2014.
[20] S. Wang, D. Li, Y. Wei, and H. Li, “A Feature Selection Method Based on Fisher ’s Discriminant Ratio for Text Sentiment Classification Suge,” Springer-Verlag Berlin Heidelb., pp. 88–97, 2009.
[21] A. Esmin and S. Matwin, “Hierarchical Classification Approach to Emotion Recognition in Twitter Hierarchical Classification Approach to Emotion Recognition in Twitter,” 2016, no. March.
[22] W. Y. and L. H. A. Wang S, Li D, Song X, “A feature selection method based on improved fisher’s discriminant ratio for text sentiment classification,” Expert Syst. Appl., vol. 38, no. 7, p. 8696–8702., 2011.
[23] L. C. H. and C. H. Chen LS, “A neural network based approach for sentiment classification in the blogosphere.,” J. Informetr., vol. 5, no. 2, pp. 313–322, 2011.
[24] T. Brychcín, M. Konkol, and J. Steinberger, “Machine Learning Approach to Aspect-Based Sentiment Analysis,” in Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014), 2014, pp. 817–822.
[25] F. M. Kundi, A. Khan, S. Ahmad, and M. Z. Asghar, “Lexicon-Based Sentiment Analysis in the Social Web,” J. Basic. Appl. Sci. Res, vol. 4, no. 6, pp. 238–248, 2014.
[26] S. Muhammad and F. Masud, “MMO: Multiply-Minus-One Rule for Detecting & Ranking Positive and Negative Opinion,” Int. J. Adv. Comput. Sci. Appl., vol. 7, no. 5, pp. 122–127, 2016.
[27] M. Z. A. Fazal Masud Kundi, Shakeel Ahmad, Aurangzeb Khan, “Detection and Scoring of Internet Slangs for Sentiment Analysis Using SentiWordNet Fazal,” Life Sci. J., vol. 11, no. 1, pp. 66–72, 2014.
[28] J. Fiaidhi, O. Mohammed, S. Mohammed, S. Fong, and T. Kim, “Opinion Mining over Twitterspace : Classifying Tweets Programmatically using the R Approach,” Digit. Inf. Manag. (ICDIM), Seventh Int. Conf. on. IEEE, pp. 313–319, 2012.
[29] A. Jeyapriya and C. S. K. Selvi, “Extracting aspects and mining opinions in product reviews using supervised learning algorithm,” in 2nd International Conference on Electronics and Communication Systems, ICECS 2015, 2015, pp. 548–552.
[30] D. K. Kirange, R. R. Deshmukh, and M. D. K. Kirange, “Aspect Based Sentiment analysis SemEval-2014 Task 4,” Asian J. Comput. Sci. {&} Inf. Technol., vol. 4, no. 8, pp. 72–75, Aug. 2014.
[31] Deepak Kumar Gupta and Asif Ekbal, “Supervised Machine Learning for Aspect based Sentiment Analysis,” in Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014), 2014, pp. 319–323.
[32] T. Shaikh and D. Deshpande, “A Review on Opinion Mining and Sentiment Analysis,” in IJCA Proceedings on National Conference on Recent Trends in Computer Science and Information Technology, 2016, no. 2, pp. 6–9.
[33] A. Alghunaim, M. Mohtarami, S. Cyphers, and J. Glass, “A Vector Space Approach for Aspect Based Sentiment Analysis,” Proc. NAACL-HLT 2015, pp. 116–122, 2015.
[34] M. Cuadros, S. Sebastian, G. Rigau, E. H. Unibertsitatea, and S. Sebastian, “V3: Unsupervised Aspect Based Sentiment Analysis for SemEval-2015 Task 12,” no. SemEval, pp. 714–718, 2015.
[35] S. Rosenthal, N. Farra, and P. Nakov, “SemEval-2017 Task 4 : Sentiment Analysis in Twitter,” Proc. 11th Int. Work. Semant. Eval., vol. 3, no. 4, pp. 502–518, 2017.
[36] V. S. Rajput, “An Overview of Use of Natural Language Processing in Sentiment Analysis based on User Opinions,” Int. J. Adv. Res. Comput. Sci. Softw. Eng. An, vol. 6, no. 4, pp. 594–598, 2016.
[37] S. M. Street and B. Liu, “Identifying Noun Product Features that Imply Opinions,” Proc. 49th Annu. Meet. Assoc. Comput. Linguist. Hum. Lang. Technol., no. Issue: 2009, pp. 575–580, 2011.
[38] S. S. Htay and K. T. Lynn, “Extracting product features and opinion words using pattern knowledge in customer reviews,” Sci. World J., 2013.
[39] M. Gamon, A. Aue, S. Corston-Oliver, and E. Ringger, “Pulse: Mining Customer Opinions from Free Text,” in Proceedings of the 6th international conference on Advances in Intelligent Data Analysis, 2005, pp. 121–132.
[40] L. Zhuang, F. Jing, and X.-Y. Zhu, “Movie review mining and summarization,” in Proceedings of the 15th ACM international conference on Information and knowledge management - CIKM ’06, 2006, pp. 43–50.
[41] S. Gojali and M. L. Khodra, “Aspect based sentiment analysis for review rating prediction,” in 4th IGNITE Conference and 2016 International Conference on Advanced Informatics: Concepts, Theory and Application, ICAICTA 2016, 2016.
[42] A. O. Shakeel Ahmad, Sheikh Muhammad Saqib, “LSI Based Search Technique: Using Extracted Keywords and Key Sentences,” VAWKUM Trans. Comput. Sci., vol. 14, no. 2, pp. 1–8, 2017.
[43] S. M. Saqib, K. Mahmood, and T. Naeem, “Comparison of LSI algorithms without and with pre-processing : using text

document based search," Accent. Trans. Inf. Secur., vol. 1, no. 4, pp. 44–51, 2016.
[44] A. Huang, D. Milne, E. Frank, and I. H. Witten, "Clustering Documents using a Wikipedia-based Concept Representation," in PAKDD 2009: Advances in Knowledge Discovery and Data Mining, 2009, pp. 628–636.
[45] Q. YANG, "SUPPORT VECTOR MACHINE FOR CUSTOMIZED EMAIL FILTERING BASED ON IMPROVING LATENT SEMANTIC INDEXING," in , Proceedings of the Fourth International Conference on Machine Learning and Cybernetics, Guangzhou, 2005, pp. 18–21.
[46] S. Streat, "10 English Words with the most multiple meanings – Part 1," Succeed in English with Business, 2014. .
[47] A. Rios, L. Mascarell, and R. Sennrich, "Improving Word Sense Disambiguation in Neural Machine Translation with Sense Embeddings," Proc. Conf. Mach. Transl. (WMT, vol. 1, pp. 11–19, 2017.
[48] C. Engineering, "A Survey on Word Sense Disambiguation Used In NLP," Int. J. Innov. Res. Comput. Commun. Eng., vol. 5, no. 3, pp. 5116–5117, 2017.
[49] J. C.-C. and R. N. D. Alessandro Raganato, "Word Sense Disambiguation: A Unified Evaluation Framework and Empirical Comparison," Proc. 15th Conf. Eur. Chapter Assoc. Comput. Linguist., vol. 1, pp. 99–110, 2017.
[50] T. Wang, W. Li, F. Liu, and J. Hua, "Engineering Applications of Artificial Intelligence Sprinkled semantic diffusion kernel for word sense disambiguation," Eng. Appl. Artif. Intell., vol. 64, no. May, pp. 43–51, 2017.
[51] A. Wawer and J. Kazimierza, "Supervised and Unsupervised Word Sense Disambiguation on Word Embedding Vectors of Unambiguous Synonyms," 120 Proc. 1st Work. Sense, Concept Entity Represent. their Appl., pp. 120–125, 2017.
[52] B. Li, L. Zhou, S. Feng, and K. Wong, "A Unified Graph Model for Sentence-based Opinion Retrieval," Acl, no. July, pp. 1367–1375, 2010.
[53] P. Pantel, E. Crestan, A. Borkovsky, A.-M. Popescu, and V. Vyas, "Web-scale distributional similarity and entity set expansion," Proc. 2009 Conf. Empir. Methods Nat. Lang. Process. Vol. 2 EMNLP 09, vol. 2, no. August, pp. 938–947, 2009.
[54] L. Lee, "Measures of distributional similarity.," Proc. Annu. Meet. Assoc. Comput. Linguist. (ACL-1999)., 1999.
[55] B. Liu, "Sentiment Analysis and Opinion Mining," Synth. Lect. Hum. Lang. Technol., vol. 5, no. 1, pp. 1–167, 2012.
[56] D. Andrzejewski and X. Zhu, "Latent Dirichlet Allocation with topic-in-set knowledge," Proc. NAACL HLT 2009 Work. Semi-Supervised Learn. Nat. Lang. Process. - SemiSupLearn '09, pp. 43–48, 2009.
[57] W. Z. Shi-Heng Wang, Yijun Ding, Weizhong Zhao, Yung-Hsiang Huang, Roger Perkins, "Text mining for identifying topics in the literatures about adolescent substance use and depression.," BMC Public Heal. 2016; 16 279. Publ. online 2016 Mar 19., 2016.
[58] S. M. Saqib, T. Naeem, and K. Mahmood, "Extraction of key / title / aspect words from document using wordnet," Accent. Trans. Inf. Secur., vol. 1, no. 3, pp. 36–43, 2016.
[59] A. Garcıa-Pablos, M. Cuadros, and S. Gaines, "V3: Unsupervised Generation of Domain Aspect Terms for Aspect Based Sentiment Analysis," SemEval 2014, no. SemEval, pp. 833–837, 2014.
[60] "https://7esl.com/synonyms/," 252 synonyms, Visit On Date: 20-01-2019. .
[61] "http://www.studypart.com/pair-of-words/index.php?id=&page=4," 52 Different Words Pairs, Visit On Date: 20-01-2019. .
[62] S. Muthulakshmi, "Comparative Study on Classification Meta Algorithms," Ijircce.Com, 2013. [Online]. Available: https://www.rroij.com/open-access/comparative-study-on-classification-metaalgorithms.php?aid=43760.