

# ASPECT CATEGORY POLARITY DETECTION USING MULTI CLASS SUPPORT VECTOR MACHINE WITH LEXICONS BASED FEATURES AND VECTOR BASED FEATURES

<sup>1</sup>Mr. Mannava Yesu Babu , <sup>2</sup>Dr. P. Vijaya Pal Reddy , <sup>3</sup>Dr. C. Shoba Bindu

<sup>1</sup>Research Scholar, Department of Computer Science and Engineering, JNTU Ananthapur, AP, India.

Email:[mannavababu@gmail.com](mailto:mannavababu@gmail.com)

<sup>2</sup>Head & Professor, Department of Computer Science and Engineering, Matrusri Engineering College, Saidabad, Hyderabad, Telangana, India.

Email:[drpvijayapalreddy@gmail.com](mailto:drpvijayapalreddy@gmail.com)

<sup>3</sup>Professor, Department of Computer Science and Engineering, JNTU College of Engineering, Anantapur, AP, India.

Email:[shobabindhu.cse@jntua.ac.in](mailto:shobabindhu.cse@jntua.ac.in)

Received: 21.03.2020

Revised: 22.04.2020

Accepted: 23.05.2020

## Abstract

Sentiment Analysis (SA) is the classification and interpretation of the opinions or emotions within the text data using text analysis methods. Aspect Level Sentiment Analysis (ALSA) is a type of SA technique which uses different types of algorithms to extract the aspects of the entities from the text data and determining the sentiment of the aspects. ALSA is divided into two subtasks such as Aspect Category Detection (ACD) and Aspect Category Polarity Detection (ACPD). ACD determines the aspect categories discussed in each sentence from a set of review sentences and a given predefined set of aspect categories. Aspect categories need not occur as terms in the sentences and are characteristically coarser than the aspect terms. For every review sentence the aspect categories are provided to identify the polarity of aspect category. The objective of ACPD is to identify the polarity such as negative, positive, neutral or conflict of every aspect category mentioned in each sentence. In this paper, the task aspect category polarity detection is addressed. The three types of features such as linguistic features, Lexicons features and Vector Space features are extracted from the reviews. The classifier is trained with these features. The learned model is used to detect the polarities of the categories that are present in the test review sentences. The combination of proposed features with the existing features improves the accuracy of the system. The accuracy of the proposed system has enhanced by 1.3 % compared with best performing system.

**Key words:** Aspect Category Polarity Detection, Vector based Features, Lexicons Based Features, Linguistic Features, Multi Class SVM.

© 2020 by Advance Scientific Research. This is an open-access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>)  
DOI: <http://dx.doi.org/10.31838/jcr.07.12.202>

## INTRODUCTION

Social media has become part of everyday life. It was actively used by millions of companies and Billions of people around the world. Giant websites like Facebook, Twitter, MySpace, Amazon, etc. allow an option of sharing real time information. Because of this, web is exponentially increasing with huge amount of personal, subjective and affective data. Intelligent and useful applications like recommendation systems and customer's reviews summarization can be developed by mining this huge volume of information which also brings interesting challenges [1]. Based on this information derived by mining data, companies uses it a lot to attract their customers by introducing a variety of marketing strategies. They keep on advertising their products by saying that they are cheap, easy to use, the best and long lasting. But as customer, we need to think whether these advertisements are really true or not. Yes, everything that we see is not true. Companies don't expose the limitations of their products, but definitely exaggerate the product's quality. In this situation we as a customer need to take a rational and wise decision on products which is the good between the varieties of products. This can be sometimes stressful for a customer. In order to avoid such situations, we need to trust in the opinions and experiences of others who have already purchased the same product or similar one. Many websites like Ciao , Epinions and Cnet gives a platform to the customers to share their experiences and opinions about products. It is very tedious and impossible task to conclude a final decision about any product from that large volume of reviews. If we think carefully about these user generated content on hand, specifically the reviews of products, the most important advantage is that we can derive information from these data using computer system.

The main aim of SA is to determine the overall polarity such as negative or positive of a given text spans or text [2] . However, the need of Aspect Based Sentiment Analysis (ABSA) became apparent for SA as a more fine-grained approach [3] . The aim of ABSA is to determine the aspects of entities from reviews and identify the expressed sentiment on aspects by the reviewers. In SA, ABSA is a fine grained level of sub task [3]. The aim of this task is determine the aspects like screen, battery, service, food, time-life, weight, size etc. of entities like restaurants, laptops, camera etc. identify the sentiment like negative, positive or neutral expressed towards every aspect. This task is composed by two basic steps like feature extraction and classification of features polarity.

With increasing volume of user generated textual data, there is also increasing the need of efficient techniques for analyzing it. According to Liu, since 2000 SA become more interesting and attracts the researchers to propose solutions [3]. In SA, many techniques from Natural Language Processing (NLP) are used to determine the polarity on an entity expressed in the text data. Further, many levels of granularity is identified in SA like Document level SA (DLSA), Sentence Level SA (SLSA) and Aspect Level SA (ALSA).

DLSA techniques determine the sentiment like negative, positive or neutral of whole document of an entity expressed in the document. SLSA consider the sentence of a document to determine the sentiment like negative, positive or neutral of each sentence. Some researchers worked on clause level SA but this is not still enough. ALSA is a more fine grained analysis wherein the aspects of the entities are extracted first, later it determines the sentiment of the aspects which

are identified in the first step. The task of ALSA contains several sub problems like sometimes some of the documents are a combination of multiple entities such as laptop, restaurant, printer, etc. An entity can be referred in many ways. And users use these different variations of same entity in many different ways in the form writings. Each entity is expressed with multiple aspects which are its attributes or parts, some of them can be a different entity like laptop screen, which was not taken by anyone in their works. Opinion can be defined by the quintuple  $(e_i, a_{ij}, s_{ijkl}, h_k, t_l)$  [3]. Where,  $e_i$  is the  $i^{th}$  entity,  $a_{ij}$  is the aspect  $a_j$  of entity  $e_i$ ,  $s_{ijkl}$  is the sentiment expressed by the opinion holder  $h_k$  on aspect  $a_j$  of entity  $e_i$  at time  $t_l$ ,  $h_k$  is the opinion holder of the review means who writes the review and  $t_l$  represents the time of the review. This definition doesn't consider the issue of when an aspect is defined with some other aspects which represents the hierarchy of aspects. In order to handle this issue, some researchers proposed a solution to avoid the loss of information by representing the aspects as a tree of aspect terms [4].

DLSA and SLSA are no longer give user's need of obtaining more precise and accurate information from a document. Aspect level sentiment analysis helps to users to gain more insights on the polarities of different aspects of the target entity. For sales of products and services, the key point is the user reviews. Before purchasing of a product or consuming of any service, online users largely depending on the online reviews which are ongoing trend. In SA or Opinion Mining (OM), the aim of aspect extraction is extract the features or aspects of an entity on which the sentiments are expressed by the reviewers [5]. An aspect is a component or an attribute of a product or entity that is discussed in a review. For an instance, if we look at one example of online review by one user on Dell laptop "Dell Laptop has very good battery life and click pads", in this sentence battery life and click pads are the aspect terms. Identifying the polarity (negative, positive or neutral) of the review is the main task of SA.

The sentiment polarity can be influenced by the aspect terms within a single domain. For example, the aspect "cheap" indicates positive polarity when discussing about food but this word becomes negative polarity when talking about ambience and decoration [6]. The main task of ABSA is identifying the possible aspects of entities and determining the corresponding sentiments to aspect terms which are discussed in a review document. Recently identifying aspects and sentiments simultaneously has become a huge interest. Hu proposed a method by using association mining which is based on information extraction (IE) which identifies most frequently occurring noun phrases [5]. Some other works include the methods by using unsupervised clustering technique [7], semantically motivated technique and subset of Wikipedia category hierarchy which is specified manually to define the aspect terms.

The ACPD determines the sentiment polarity of every aspect category. The aspect categories are provided for every review sentence. The aim is to identify the polarity such as negative, positive, neutral or conflict of every aspect category which are discussed in every sentence. The goal of ACPD is to identify the sentiment of an aspect category specified in a given sentence. This technique takes the input as the pair of aspect category and sentence and gives output as a sentiment label like negative, positive, conflict or neutral.

In this paper, section 2 explains various ABSA approaches for aspect term identification, detection of aspect term polarity, ACD and ACPD. The proposed system is described in section 3. The various kinds of features such as linguistic features

and lexicons features are presented. The features generated from vector space model are proposed in this paper. In section 4, the restaurant dataset description, the evaluation measure and experimental evaluations with the combination of three feature sets are performed. The comparison of the accuracy level of the proposed system with the baseline system developed by the organizers and the top scoring system is provided. The observations from the obtained results are concluded and possible interventions for future extension are presented in section 5.

## LITERATURE SURVEY

In [5], they found that many of the existing works for aspect detection are based on information extraction to identify the noun phrases which are most frequent. This approach is very helpful in case of single noun to identify the aspects. But this approach cannot determine the aspect terms which are of low frequent and noun phrases like different dish names of dosa, uttapam, biryani, etc. for the aspect category, "food", which is the biggest disadvantage of this approach. The proposed work of such problem involves rule-based, semantic hierarchy or mixing of both [7]. For aspect detection, many recent approaches used topic modeling technique of Latent Dirichlet Allocation (LDA) [8]. But there is one problem with LDA approach for the aspect detection task is LDA identifies the global topics in the data rather than capturing the local aspects associated with the entity. Later this approach was modified in Aspect and Sentiment Unification Model (ASUM) and Sentence LDA (SLDA). In Das [9], they identified opinion targets and topics by focusing on text spans. Snyder discussed the problem of determining the categories related to multiple aspect terms which are appeared in the text [10]. Something can be looked in a restaurant review there will be many categories like service, ambience, food etc which we call them as review or aspect categories in our task.

Good Grief decoding algorithm outperforms over the famous PRank algorithm on a dataset collected from reviews of restaurants [11]. Restaurant reviews collected from City search New York was classified by Ganu [12] into six classes such as Service, Food, Ambience, Price, Miscellaneous and Anecdotes. Support Vector Machine classifiers have been used for experimenting and for identification of sentiment associated with each category. Finally, they experimented with the regression based model which uses MATLAB regression function of `mvregress` to assign rating 1 to 5 to each review. We need a prior annotated sentiment lexicon for finding the polarity or sentiment of the aspect term and it's associated aspect category. In different English languages, many researchers conducted experiments on developing emotional corpora like WordNet Affect [13], SentiWordNet [14] etc.

SentiWordNet is one of the most popular and widely used sentiment lexicon among all freely accessible sentiment lexicons because the number of citations or the number of referred researchers are more for this sentiment lexicon compared with other resources. It was used in many applications like SA and emotion analysis. On the automated polarity identification or opinion detection from reviews [5], several works have been performed. Yu [15] concentrated on characterizing the facts and opinions in a common way, without analyzing what the opinion is about or who is the opinion holder. Naive Bayes classifier was used for finding the sentiment or polarity of the fact. Customers reviews summarization and identification of the polarity of review was successfully done by Hu [5]. They have obtained good accuracy for determining the reviews sentiment.

For a machine learning based sentiment classifier, Pang [16] has proved that Part Of Speech (POS) tags, adjectives, unigrams and bigrams are important features. And then later, they also identified verbs and adjectives as equally important features. Meena [17] performed SLSA based on clauses of a sentence by using rules. But in this case they can't consider blindly the verbs and adjectives as features because they may be related to various aspects. If we take an example, the sentence "The pizza is the best if you like thin crusted pizza" shows positive sentiment towards 'pizza' because of the 'best' adjective word. However, the 'like' word act as a sentiment verb for the term 'thin crusted pizza'. Therefore, only those verbs and adjectives which are having relationship with target aspect are considered in the indication of their polarity. Wilson [18] proved that the words which are sharing relations of dependency with aspect terms are more useful to specify the sentiments associated with those terms.

The co-relation among topic and polarity in tweets was showed by Saif [19]. They asserted that especially in case of positive sentiments, several people tend to express same type of polarities on same topics. The baseline approach for this task also associates aspect terms with their polarity. Therefore, in this work, we also considered the aspect term as a probable feature.

The utilization of sentiment resource was proven to be very useful to train, build and evaluate the systems which are proposed by the researchers for sentiment analysis. Several approaches have been presented in order to build sentiment resource. In one of the first works presented by Hatzivassiloglou [20], they proposed an approach for detecting the polarity by considering if the adjectives are linked with copulative or adversative conjunctions. In Turney [21], the authors exposed a method which uses point-wise mutual information (PMI) for measuring a set of negative and positive paradigm words for inferring the word semantic orientation. Hu[5] suggested a technique, to expand the lexicon by using antonym and synonym relations which are provided in WordNet [22].

In Cruz [23], a sentiment resource is created based on a graph. This is developed by observing the conjunctive expressions among adjectives pairs in corpus of reviews. In [24], a PageRank algorithm was adapted by using graphs with negative and positive edges in order to acquire the words semantic orientation. There are more number of existing proposals are available for resources construction, but the obtained results are far from what we expected. They observed that the polarity of a word majorly dependent on that domain in ABSA and the publicly available sentiment resources such as SentiWordNet [14], General Inquirer, HowNet and WordNet-Affect do not capture this dependency. The human annotators are not in a position to create a specific sentiment resource whenever a new product is launched into the market. Because of this the proposed methods need to create their own sentiment resources which become a challenging task to the researchers.

## SYSTEM DESCRIPTION

The aim of our system is to determine the expressed sentiment towards a specified aspect category in a given sentence. As the aspect category terms not present directly in the sentence, it is required to consider the the sentence for each category and find the category sentiment label as negative, positive, neutral or conflict. To address this multi-class classification problem, a linear SVM classifier is used to train them with various features extracted from the reviews presented in training dataset. To classify the targeted category sentiment with all other categories, the classifier is trained with whole dataset to learn the common sentiment features as well sentiment features specific to aspect from the training samples pertaining to a specific aspect category. The steps involved in category polarity detection are presented in Figure 1. The features are extracted from the review sentences of train dataset. The best combination of the features to increase the accuracy of the system is trained as a part of future engineering. The classifier is trained with these classifiers to learn the classification model. The test review sentence with already identified category label is used to find the sentiment of the aspect category using the learned system.

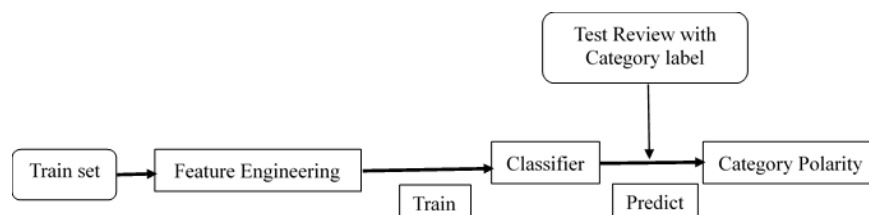


Fig. 1. Aspect Category Polarity Detection Model

Sentences are tokenized and tokens are represented with part-of-speech tags. Then, the following group of features is used to represent the every sentence as a vector of features.

### Features

From the Yelp restaurant reviews corpus, the reviews start rated with three and above are considered as positive and the remaining are considered as negative reviews. We automatically developed a sentiment lexicon for restaurants domain. We computed a sentiment score for every term  $w$  in the corpus using the point wise mutual information (PMI) which computes the differences among negative and positive scores. The presence of negative word before the term, negation the score of a term. The manually formed sentiment lexicons such as SentiWordNet, NRC Emotion Lexicon,

Sentiment140 lexicons, Bing Liu's Lexicon and MPQA Subjectivity Lexicon are used to label the sentiment with the word present in the sentence.

### Linguistic Features

#### Word N-grams

The counts of every word unigram and word bigram words in the sentence.

#### TF-IDF

Term Frequency and Inverse Document Frequency of all the terms in the sentence to identify the importance of a term to a category at that instance.

### Lexicon Features

#### Frequency count

count of words with negative and positive scores.

#### Polarity score

The overall score of a sentence (sum of scores of all words in the sentence).

#### Word2Aspect

The association of each word in a sentence with one or more categories. The level of association is calculated using PMI.

#### Word Cluster

The occurrence of a cluster in the sentence formed on Yelp Restaurant dataset using LDA.

### Proposed Features

#### Word Vector Features

We used skip-gram model of Word2Vec to compute the vector representations of words using a Yelp dataset of restaurant reviews with 300-dimensional vector space. The following features are derived from the word vectors.

#### Dependency Vector Average (DVA)

is attained by averaging the dependency words (DW) vector representations for each term in the dependency tree of the input sentence.

#### Topic Vector Average (TVA)

The LDA could generate the document distribution among predefined topics. The average of the computed vector representations trained on different topics.

#### Bag of Tags (BoT)

Bag of syntactic dependency tags of children, siblings and parent of each term from the sentence parse tree.

### RESULTS AND DISCUSSIONS

We trained a four-way SVM classifier using software of LibSVM for each restaurant domain using Weka's SMO implementation. J48 is a type of Decision Tree (DT) algorithm which builds on the concept of Information entropy. The algorithm chooses the most informative feature using information gain. Naive Bayes (NB) classifier is a linear classifier based on Bayes rule. IBk is an implementation of K-Nearest Neighbor (KNN) classifier. It computes the distance among the feature vectors of the training examples and the unseen examples. The label of the unseen example is based on the k nearest neighbors training examples.

#### Dataset Description

The details of the reviews dataset of restaurant domain are displayed in Table 1. The training data of restaurants contain 3041 English sentences which are annotated with aspect terms occurred in the sentences, polarities of aspect terms, coarse aspect categories and their sentiments. The test data of restaurants contain 800 English sentences which are also annotated in the same manner specified in the training data. The organizers provided the dataset in an XML format for the ABSATask.

Table 1. Reviews in Restaurant Train and Test Datasets

Domain	Train	Test	Total
Restaurant	3041	800	3841

The aspect categories of the reviews presented in the train dataset and test dataset with its polarity are presented in the Table 2. From the table, it is identified that the most of reviews contains FOOD as a category label. The number of categories labeled with positive polarity is more compared with other sentiment polarities. The conflict polarity is

labeled with few numbers of categories. This imbalance may affect the performance of the classifier to learning rules. The similar observation is made for test set reviews also. Most of the reviews in test set contains FOOD category and the more number of categories are labeled with positive sentiment.

Table 2. Aspect category with its Polarity distribution in Train and Test datasets

Category	Positive		Negative		Conflict		Neutral		Total	
	Train	Test	Train	Test	Train	Test	Train	Test	Train	Test
FOOD	867	302	209	69	66	16	90	31	1232	418
PRICE	179	51	115	28	17	3	10	1	321	83
SERVICE	324	101	218	63	35	5	20	3	597	172
AMBIENCE	263	76	98	21	47	13	23	8	431	118
MISCELLANEOUS	546	127	199	41	30	15	357	51	1132	234
Total	2179	657	839	159	163	52	500	94	3713	1025

### Evaluation Measures

We computed the accuracy of each system to evaluate aspect category polarity detection. Accuracy is defined as the ratio

among the count of aspect category polarity labels correctly predicted and the total count of aspect category annotations.

$$\text{Accuracy} = \frac{\text{Number of correctly predicted Aspect Category Polarity labels}}{\text{Total Number of Aspect Category annotations}}$$

### Empirical Evaluations and Analysis

We trained a SVM using the LibSVM software to address this multi-class classification problem. The extracted features were used to train four different classifiers namely NB, SVM (LibSVM), Decision Tree (J48) classifier and K-Nearest Neighbor (IBk) classifier. The Weka 2.7 implementation of J48, IBk and Naive Bayes were used with 5 as the k value.

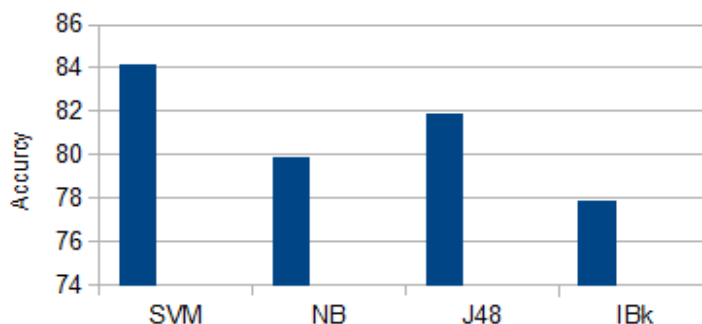
Semantic features and manually crafted rules were used to identify sentiment scores for the aspect terms. The proposed system performance is measured with various classifiers by inputting the linguistic features, lexicon based features in combination with vector space model features on the restaurant dataset. The accuracy of the proposed system with the combination various features are presented in Table 3.

**Table 3. Results for the aspect category Polarity detection on Restaurant dataset**

Classifier	Features	Accuracy
Support Vector Machine	Linguistic Features	73.48
	+ Lexicon Features	82.73
	+ Word Vector Features	84.21
Decision Tree	Linguistic Features	72.51
	+ Lexicon Features	79.26
	+ Word Vector Features	81.92
Naive Bayes	Linguistic Features	71.42
	+ Lexicon Features	78.13
	+ Word Vector Features	79.89
K-Nearest Neighbor	Linguistic Features	70.63
	+ Lexicon Features	76.48
	+ Word Vector Features	77.93

From the results, it is identified that the multi class SVM has achieved 7.9 % of more Accuracy using linguistic features compared with baseline system. With the combination of lexicon features the proposed system has increased its accuracy by 17.1%. The influence of knowledge gained from external sources has influenced much in detecting correct polarities for the aspect categories. The combination of set of proposed features, vector space features has influenced the performance of the system. The accuracy has increased by 18.5 % compared with the baseline system. The improvement

in the system performance with the combination vector based features performance has increased by 1.3% compared with the top scoring system participated in SemEval 2014 conference. The empirical evaluations are performed with three other classifiers such as K-Nearest Neighbor classifier, NB and DT classifier. The performance of DT classifier is comparable with SVM classifier. The comparison of the four classifier performance with combination of linguistic, lexicon and vector based features are presented in the Figure 2.



**Fig. 2. Classifier performance on Restaurant dataset**

The proposed system accuracy is compared with the baseline system and best performing system in the conference as presented in table 4. The baseline approach provided by the SemEval 2014 organizers for aspect category polarity detection is allocate aspect category c to a test sentence s based on the majority polarity label that aspect category had in the k most similar training instances to test sentence from the same domain. This system only considers the training instances which are having aspect category label c.

In SemEval 2014 conference, NRC-Canada is the best scoring system for the sub task of ACPD. The system used a SVM classifier. NRC-Canada depends on SVM classifier with features mostly based on parse trees, n-grams and several sentiment lexica such as SentiWordnet, MPQA, and Opinion Lexicon of Bing Liu's. NRC-Canada system also used two predefined sentiment lexica for restaurants, obtained from YELP and Amazon data, respectively. From the table, it is observed that the proposed system with the addition of vector based features to linguistic features and lexicon



features, its accuracy has improved by 18.6 % compared with baseline system. When the proposed system compared with NRC-Canada system, it is increased by 1.3 %. The NRC-Canada system has used number of features to label the

categories with their polarities. As compared with NRC-Canada system, a very less features are used for sentence vector construction to train by the classifier.

**Table 4. Comparison of Proposed method with State-of-the-art methods**

Aspect Category Polarity Detection Approach	Accuracy
Baseline System	65.66
NRC-Canada	82.92
Proposed System	84.21

## CONCLUSIONS AND FUTURE SCOPE

In this work, the task of aspect category polarity detection on restaurant dataset is addressed. The impact of linguistic features, lexicon features and vector based features are explored in polarity detection. Lexicon features are highly influencing for polarity detection. The vector based features are useful in increasing the accuracy of the system. The semantic lexical sources constructed using Yelp dataset and the manually encoded semantic lexicons enhance the system performance effectively. The vector space model captures the semantic and sentiment relations exists among the terms present in the reviews. The features developed using word vector space representations effectively captures the semantic and sentiment relations among the words.

The distributions of polarities of categories are very imbalanced with most sentences on negative or positive but with very less number of sentences on conflict. This leads to very low performance in the classification for the conflict class. This leaves us the future work to increase the imbalanced datasets classification performance. Deep learning techniques need to explore with more number of features to observe the impact on various tasks on Aspect based Sentiment Analysis.

## REFERENCES

- Ivlitskaya, I.L., Korokin, M.V., Loktionov, A.L. Pharmacological efficiency of statins and L-norvalin at an endotoxin-induced endothelial dysfunction (2016) Research result: pharmacology and clinical pharmacology, 2 (2), pp. 25-35.
- Jiang, L., Cheng, Y., Yang, L. *et al.* A trust-based collaborative filtering algorithm for E-commerce recommendation system. *J Ambient Intell Human Comput* 10, 3023-3034 (2019). <https://doi.org/10.1007/s12652-018-0928-7>
- Turney, Peter. (2002). Thumbs Up or Thumbs Down? {S}emantic Orientation Applied to Unsupervised Classification of Reviews. *Computing Research Repository - CORR.* 417-424. 10.3115/1073083.1073153.
- Bing Liu. 2012. Sentiment Analysis and Opinion Mining. *Synthesis Lectures on Human Language Technologies*, 5(1):1-167.
- Kim, Jianwen Zhang, Zheng Chen, Alice Oh and Shixia Liu (2013). A Hierarchical AspectSentiment Model for Online Reviews
- M. Hu and B. Liu. 2004. Mining and summarizing customer reviews. In *Proceedings of the 10th KDD*, pages 168-177.
- S. Brody and N. Elhadad. 2010. An unsupervised aspect-sentiment model for online reviews. In *Proceedings of NAACL*, pages 804-812, Los Angeles, CA.
- Popescu and Oren Etzionir. 2005. Extracting product features and opinions from reviews. In *Proceedings of the Conference on HLT/EMNLP*, pages 339-346.
- S. Brody and N. Elhadad. 2010. An unsupervised aspect-sentiment model for online reviews. In *Proceedings of NAACL*, pages 804-812, Los Angeles, CA.
- Dipankar Das and Sivaji Bandyopadhyay. 2010. Extracting emotion topics from blog sentences: use of voting from multi-engine supervised classifiers. In *Proceedings of the 2nd international workshop on Search and mining user-generated contents*, pp. 119-126
- Snyder and Regina Barzilay. 2007. Multiple Aspect Ranking Using the Good Grief Algorithm. In *Proceedings of the Human Language Technologies: The Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL-HLT 2007)*, pp. 300-307
- Koby Crammer and Yoram Singer. 2001. Pranking with ranking. In *NIPS*, vol. 14, pp. 641-647.
- Ganu Gayatree, Noemie Elhadad, and Amelie Marian. 2009. Beyond the stars: Improving rating predictions using review text content. In *Proceedings of the 12th International Workshop on the Web and Databases*, Providence, Rhode Island.
- Strapparava, and Alessandro Valitutti. 2004. WordNet Affect: an Affective Extension of WordNet. In *LREC*, vol. 4, pp. 1083-1086.
- Baccianella, Andrea Esuli, and Fabrizio Sebastiani. 2010. SentiWordNet 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining. In *LREC*, vol. 10, pp. 2200-2204.
- Yu and Vasileios Hatzivassiloglou. 2003. Towards answering opinion questions: Separating facts from opinions and identifying the polarity of opinion sentences. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP-2013)*, pp. 129-136.
- Bo Pang, Lillian Lee, and Shivakumar Vaithyanathan. 2002. Thumbs up?: sentiment classification using machine learning techniques. In *Proceedings of the ACL-02 conference on Empirical methods in natural language processing - Volume 10, EMNLP '02*, pages 79-86, Stroudsburg, PA, USA. Association for Computational Linguistics
- Meena and Prabhakar T.V. 2007. Sentence level sentiment analysis in the presence of conjuncts using linguistic analysis. In *ECIR*, volume 4425 of *Lecture Notes in Computer Science*. Springer.
- Wilson Theresa, Janyce Wiebe, and Paul Hoffmann. 2009. Recognizing contextual polarity: An exploration of features for phrase-level sentiment analysis. *Computational Linguistics*, pages 399-433
- Hassan Saif, Yulan He, and Harith Alani. 2012. Semantic sentiment analysis of twitter. In *International Semantic Web Conference (1)*, volume 7649 of *Lecture Notes in Computer Science*, pages 508-524. Springer.
- Hatzivassiloglou and Kathleen McKeown. 1997. Predicting the semantic orientation of adjectives. In

- Proceedings of the Joint ACL/EACL Conference, pages 174–181
22. Peter Turney and Michael Lederman Littman. 2003. Measuring praise and criticism: Inference of semantic orientation from association. *ACM Transactions on Information Systems (TOIS)*, 21(4):315–346
23. Christiane Fellbaum, editor. 1998. *WordNet: An Electronic Lexical Database*. MIT Press
24. Cruz, José Antonio Troyano, Francisco Javier Ortega, and Carlos García Vallejo. 2009. Inducción de un lexicon de opinión orientado al dominio. *Procesamiento del Lenguaje Natural*, 43:5–12.
25. Page, Sergey Brin, Rajeev Motwani, and Terry Winograd. 1999. The PageRank Citation Ranking: Bringing Order to the Web.