Journal of Engineering Technology

AMERICAN SOCIETY FOR ENGINEERING EDUCATION

Vol. 6, Special Issue (Emerging Trends in Engineering Technology) Mar. 2018, PP. 97-111

RESEARCH ARTICLE -ENGINEERNG TECHNOLOGY

Machine learning in computational treatment of opinions towards better product recommendations – an ontology mining way: a survey

K.C. Ravi Kumar*[a], D. Teja Santosh[b], B. Vishnu Vardhan[c], P. Chiranjeevi[d]

[a] Research Scholar, Dept. of CSE, JNTUH, Kukatpally, Hyderabad, T.S., India.
[b] Assistant Professor, Dept. of CSE, GITAM, Hyderabad, T.S., India.
[c] Vice-Principal and Professor, Dept. of CSE, JNTUHCEM, Manthani, T.S., India.
[d] Assistant Professor, Dept. of CSE, ACE Engineering College, Hyderabad, T.S., India.

*[a] Corresponding author

Abstract:

The emerging field of opinion mining was investigated by Natural Language Processing (NLP) community for nearly two decades. The growth in this research field has materialized various sub areas wherein each sub area deals with a different level of research question. This survey is inclined to feature level opinion mining of online reviews, in which the main purpose is to identify and extract product features and opinions mentioned in them and to determine the sentiment of each feature. Special attention is paid towards the significance of ontology towards product recommendations as future research is focused on utilizing the constructs of the feature concepts and corresponding relationships from the ontology. The comparative results among different feature extraction techniques and sentiment classifiers are also provided in order to learn the best performer on the corpus. Finally, the gaps in the surveyed literature are preseneted and a direction for filling the gaps are suggested towards making better purchase decisions from E-commerce websites.

Keywords: opinion mining, product features, opinion orientation, sentiment, ontology, better product recommendations.

1. Introduction

Over the last two decades, the amount of web information has exploded in a rapid way in all formats of data. The web content is growing at a lightning fast speed. The need for sharing information among the web users is also increased. The term web 2.0 is coined by Tim O' Reilly [67] has this social web as a part of it. Popular web 2.0 online shopping websites like Amazon, Flipkart and etc., which utilize E-commerce Business-to-Consumer (B2C) business model for conducting online transactions are contributing to the content development over the web.

The size of the reviews database is getting scaled from time to time due to the E-commerce websites providing a facility to script consumer opinions. These reviews are regularly fed into the system and are not useful for certain cross section of people for finding the relevant sources of review information, and it seems to be a formidable task. This phenomenon paved the way for opinion mining. The available feature extraction approaches in the literature [6][37][68][70] were extracted either based on explicit product features or implicit features. Few works [34][69] concentrated on extracting both kinds of features.

The current opinion mining approaches are slowed down by major problems such as the nonexistence of connecting concepts of semantic associations in feature search process. The researchers [68] have expounded the need of ontology for the extraction of automatic product features as it is engineered by using product reviews domain which is taken into consideration. The obtained product features are determined for their sentiments with ontology analysis towards better product recommendations.

The objective of this work is to study the various opinion mining algorithms and techniques on product reviews available in current literature with their merits, demerits, and limitations and also to know about the various performance measures and provide the direction to fill the identified gaps in the literature.

2. Review based Opinion Mining

Bing Liu defined [1] opinion mining as “given a set of evaluative text documents D that contains opinions about an object, opinion mining aims to extract components and their corresponding attributes (features) of the product that were commented on in each document d ∈ D and to determine whether the comments are positive, negative or neutral”. Opinion mining on online product reviews is understood as the enhanced information retrieval as it involves learning the features and sentiments by the machine from the review sentences following the rules of the natural language [1].

2.1 Implications of online reviews in E-commerce field

Online reviews are essential for businesses because more customers rely on the opinions of others when making their purchase decisions. Reviews also help the business organizations to perfectly understand the views and opinions of the customers so that they can assess the customer mind and are successful in serving them. The web applications built on opinion oriented information access present two major concerns that are to be dealt with. The first is the privacy of the content gathered in order to know the preferences of the people. These preferences must be kept in mind by the web application developers as algorithms that are running in the background of web application personalize the content without the concern of the content holder. This is called as Filter Bubble [3]. Manipulation of online reviews is another dire issue which needs serious concern because mining for opinions from these kind of reviews often lead to biased results.

3. Review on various explored techniques on opinion mining

Many different algorithms on opinion mining of product reviews were developed in the past but it remains a complex and challenging task. It is very hard to achieve a generic opinion mining method that is universally applicable for a broad range of problem domains. A large number of books and journal publications are available on opinion mining. Exclusive book chapter and book [4, 10], surveys and review articles [5, 6] are available in the literature.

The standard taxonomy for feature level opinion mining [7] is illustrated in Figure 1 below.
This survey is initially aimed at exploring the impact of statistical methods that are used in natural language processing approaches, in which supervised approaches and in unsupervised learning approaches are to extract the product features and opinion words in order to classify the opinions. The purpose of the survey also encompasses the delineation of using ontology and mining it for automated and meaningful extraction of product features. Finally it is ended with studying the influence of

Vol. 6, Special Issue (Emerging Trends in Engineering Technology) Mar. 2018, PP. 97-111

sentiments in recommending the products to the customer and exploiting the constructs of ontology for providing better recommendations to the customer.

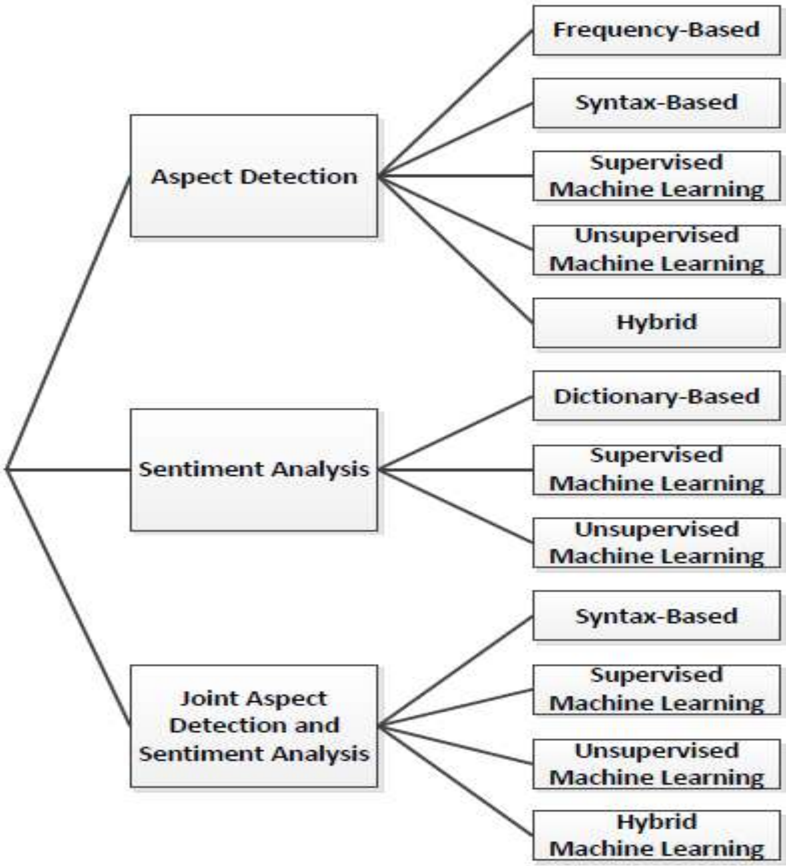
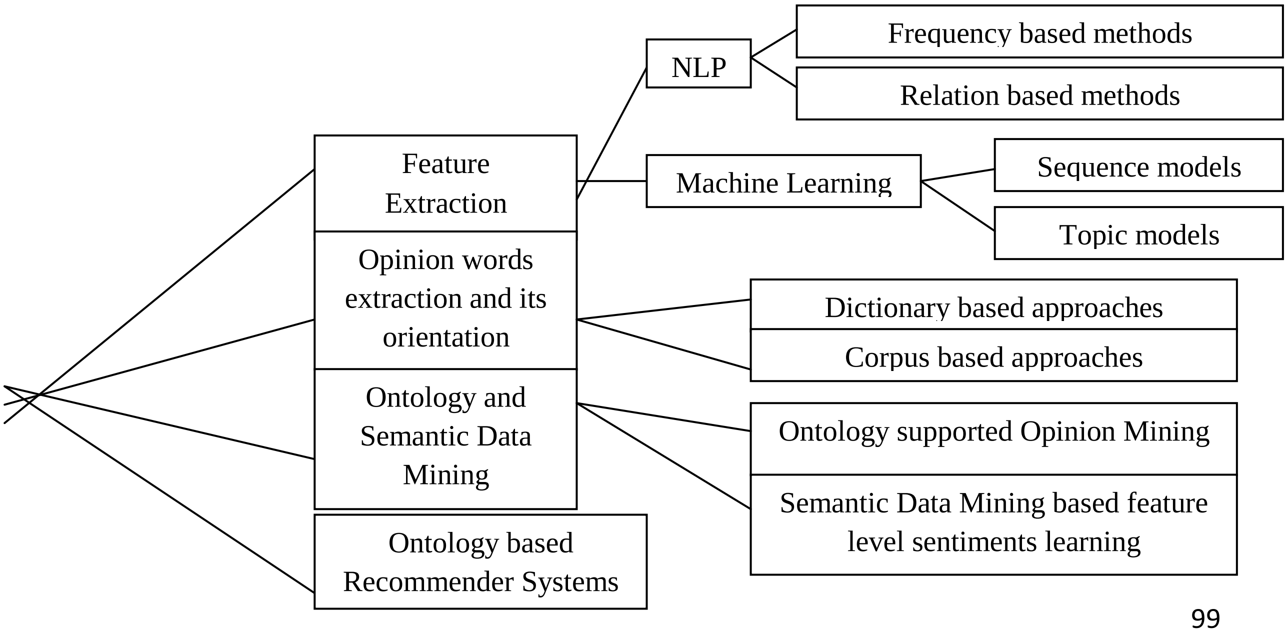


Figure. 1. Standard Taxonomy for feature level opinion mining[7]

The considered taxonomy for this survey is slightly updated according to the requirements of this thesis. The updated taxonomy for feature level opinion mining is illustrated in Figure 2 below.



Vol. 6, Special Issue (Emerging Trends in Engineering Technology) Mar. 2018, PP. 97-111

Figure. 2. Considered Taxonomy for feature level opinion mining

3.1 Pre-processing of online reviews

Opinion mining of online reviews starts with the basic and important task of pre-processing. In the perspective of online reviews, pre-processing involves logical restructuring of the unstructured reviews. The important pieces of data available in online reviews are product features and opinions. In order to identify these processing tokens [8] initially the review is tokenized into individual words. Once the processing tokens were determined, stop words that are compiled from the reviews are applied to remove the semantically less significant words and frequently occurring articles that have no value in the feature search. The words after the application of stop list are given to the stemming algorithm to obtain the standard semantic representation of the word. Once the words are obtained after optional stemming operation, Parts-of-Speech (PoS) tagging is carried out. This is performed to identify the word uniquely at the time of extracting the product features and opinions. Stanford Log-Linear Maximum Entropy Part-of-Speech tagger framework [9] is utilized to carry out this task. The PoS tagged words are given as input to the process of feature extraction and opinion extraction. The research conducted by Liu [10] in 2007 specified that 60% to 70% of the product features are explicit nouns that are available from the reviews. Pang et al. [11] in 2002 achieved an accuracy of 82.8% in movie review domains using only adjectives in them. From these findings it is clear that explicit nouns and adjectives are useful in the feature extraction and opinion extraction.

3.2 Feature Extraction Approaches

Feature (aspect) level opinion mining aims at obtaining the features from the unstructured reviews and finding the opinion orientation of the feature. This analysis reveals the impressions of the users about the product whether they are enchanted by the product or otherwise. Based on the literature [6] available, feature extraction approaches are primarily divided into three ways. These are namely developing natural language rules, using sequence models and using topic models.

3.2.1 Developing natural language rules

Natural language rule based methods have an extensive history of practice in information extraction. The rules are based on background patterns which confine to various properties of one or more terms and their relations in the sentences. The natural language rules are further classified into two categories. These are specifically frequency based methods and relation based methods.

Frequency based methods

A feature is expressed either by a noun, adjective, verb or adverb. In reviews, people tend to talk about features which suggest that features are frequent nouns but all frequent nouns are not features. Hu and Liu used [12] Apriori algorithm to extract product features using nouns & noun phrases. Some of the infrequent features were also identified. Blair-Goldensohn et al. improved [13] the Hu and Liu [12] approach by considering those noun phrases that are in sentiment bearing sentences. WordNet is used [14] to build sentiment lexicon and orientation is based on Maximum Entropy [15] (ME) classifier. Popescu et al. developed [2] OPINE tool to extract common patterns to identify potential features from online reviews. The potential features are the explicit features identified by OPINE. OPINE uses information system named KnowItAll to find frequent noun phrases that contain the product features in them. The idea of mining frequent nouns and noun phrases as product features is simple and more effective. The limitations with frequency based techniques are those which produce too many non-features and avoid low frequency features which are the actual features.

Relation based methods

The relationship patterns among the product features and the corresponding sentiments are learned from the training data. The learned templates are applied on the test data to extract the product features from the reviews. Dependency relations are the grammatical relations in a sentence. These relations are used to relate the features and the opinions in a review sentence. Zhuang et al. worked [16] on the idea of dependency relation to extract the features from online movie reviews. The typed dependencies used for feature extraction in this work are specified by Marie-Catherine and Christopher Manning [17]. The unigram words in the review sentences are merely identified. Wu et al. improved [18] the work of Zhuang et al. [16] wherein the phrases in the review sentences were related instead of considering the relations to the word level. This method never considered the association between the features and opinion words. Wang and Wang proposed [34] an algorithm to identify and to extract the product features and opinion words in the simultaneous manner. They used seed opinion words to bootstrap the feature extraction process from the collection of reviews. The idea of modifying the orientation of the product features using opinion words was not concentrated in that work. Qiu et al. suggested [21] an algorithm which is the improvement over Wang and Wang [34] approach. The suggested algorithm is termed as Double Propagation. It works on the piece of information that opinion words are used to modify the orientation of the product feature. The ability to find low frequency product features that are the actual features is the potential advantage of the relation based methods. The major drawback of these methods is that they produce many non-features which would match up with learned patterns.

3.2.2 Sequence models

Sequence models were extensively used in information extraction tasks. The feature extraction task is deemed to be a sequence labeling task because the features and opinion words are often dependent on each other and occur in a sequence in the review. There are two sequence models namely Hidden Markov Model (HMM) [22] and Conditional Random Fields (CRF) [23]. Jin et al. extended [24] basic HMM. They added Part of Speech (PoS) data to the observations to extract the hidden word as a component or a function or an entity which is understood as the product feature. Low precision was achieved due to the need for more ground truth data for training. Jakob and Gurevych utilized [25] the concept of linear chain CRF and extracted features from sentences. They considered the following things for identifying the hidden state: the word or the token, PoS tag, short dependency path and the word distance from Inside-Outside-Beginning (IOB) labeling scheme. This approach is not worked well for long distance dependencies having conjunctions in the review sentences. Fangtao et al. proposed [26] Skip-Tree CRF approach to overcome the problem faced in Jakob and Gurevych [25] work. This method skips the conjunctions using the conjunction structure and syntax tree structure information from training sequence of data to extract product features. The strength of the supervised learning techniques is that they overcome the frequency based limitations by learning the model parameters from the training data. The only limitation in these approaches is that they require huge manually labeled (ground truth) data for training.

3.2.3 Topic models

Vol. 6, Special Issue (Emerging Trends in Engineering Technology) Mar. 2018, PP. 97-111

Supervised learning requires manually annotated huge size corpus for learning the model from the algorithm. In unsupervised learning, no prior learning takes place on the corpus. The pre-processed corpus is provided directly as an input to the algorithm and it generates the clusters. Topic Modeling is an unsupervised learning method that generates topic clusters containing the mixture of words in the documents.

There are two most important and basic topic models: Probabilistic Latent Semantic Analysis [27] (PLSA) and Latent Dirichlet Allocation [28] (LDA). PLSA and LDA both use bag-of-words notation of documents. LDA generates the topics better than PLSA. This is because the LDA describes about generating topic allocation for the unseen document. Lu et al. proposed [29] a method for feature detection and grouping in short comments using PLSA Algorithm. In this work, the researchers implemented both unstructured PLSA and structured PLSA on the phrases in the reviews. They found that a good number of extracted features were extracted from structured PLSA. Blei et al. proposed [28] a model to generate latent topics from the underlying document collection (LDA). The problem with LDA is that it cannot generate local topics. Titov et al. modified [30] the already proposed model of Blei [28] and extracted the features by sampling the words which are specific to global and local topics. The researchers named the modified model as Multi-grain Latent Dirichlet Allocation (MG-LDA). This model lacks the correspondence between features and topics. Titov et al. extended [31] their MG-LDA model by constructing more coherent model which is named as MultiAspect Sentiment (MAS) for feature extraction. Owing to low feature rating value, MAS could not extract some local features. Lin et al. identified [32] features by considering feature distributions for each sentiment word in their Joint Sentiment/Topic (JST) model. This model never worked for feature specific sentiment words. Yohan et al. extended [33] JST and identified the product features of related sentiment word. The researchers named the model as Aspect and Sentiment Unification Model (ASUM). Topic models require manually labeled data for training. These models perform both feature extraction and clustering at the same time in an unsupervised manner.

Least percentage of online reviews were expressed in terms of the product feature in the indirect manner. These features are called as implicit features. The feature indicators which are present in these reviews help to identify the implied product feature. Wang and Wang identified [34] implicit features from online reviews by using the property of context association between product features and opinions. They used a statistical measure named as Revised Mutual Information (RMI) and deduced the implicit features by learning the property. Qi et al. utilized [35] COP-KMeans clustering algorithm to cluster and link the compatible product feature and opinion words. The implicit features were identified from the clusters by finding out the unlinked feature words in the product feature cluster. Ivan et al. extracted [36] implicit product features by identifying the implicit feature indicators using CRF classifier first and then extracted the implicit feature using SenticNet [37] knowledge base service.

3.3 Opinion word extraction approaches

Extracting features from the online reviews is the first part of the opinion mining task. The second part is to extract the opinion words associated with the extracted features. The task of opinion extraction is performed by Yi et al. [38]. They collected opinionated words from one of the various sentiment lexicons namely General Inquirer (GI) [20, 39], Dictionary of Affect of Language (DAL) [40], and WordNet [16]. They refined the opinion patterns obtained from training dataset by learning dependencies among the words in a sentence. The obtained opinion words were added to the sentiment lexicon. They applied the expanded sentiment lexicon to the extracted subjective sentences to extract the actual opinion words.

3.4 Opinion word orientation approaches

The third part is to analyze the orientation of the extracted opinion word. This helps in identifying the number of positive and negative opinions for the features. Opinion orientation or sentiment classification at feature (word) level were carried out in two ways namely dictionary based models and corpus based models.

3.4.1 Dictionary based models

Minqing Hu and Bing Liu exploited [41] WordNet bipolar adjective structure to find out the orientation of the adjectives from the online reviews. Esuli and Sebastiani used [42] the glossary definitions to learn the opinion bearing word orientation. Esuli and Sebastiani developed [43] a lexical resource for opinion mining called SentiWordNet 1.0. Baccianella et al. improved [44] the opinion mining lexical resource and called it as SentiWordNet 3.0. The disadvantage with dictionary based models is that these dictionaries are not able to find opinion words with domain and context specific orientations.

3.4.2 Corpus based models

Hatzivassiloglou and Weibe analyzed [45] the conjunctions between adjectives using a log linear classifier and created adjective clusters of similar orientation for learning opinion word orientations. Turney classified [46] the sentiment of the reviews using the Semantic Orientation (SO) of the phrase using Pointwise Mutual Information (PMI) [47]. Turney and Littmann classified [48] the word level sentiments by calculating the SO of the word using PMI. The disadvantage with corpus based models is that it is hard to prepare a huge corpus to cover all English words.

3.5 Ontology supported feature level sentiment classification – A semantic data mining approach

Sentiment Classification is a text categorization problem. Text Categorization as given by Fabrizio Sebastiani [50] is the task of assigning a Boolean value to each pair $<d_j,c_i> \in D * C$, where D is a domain of documents and C is a set of predefined categories. A value of 'T' assigned to $<d_j,c_i>$ indicates a decision to a file d_j under c_i , and 'F' if not. The review documents are to be categorized for learning the orientation of the opinion. Ontology based feature level opinion mining allows the machine to automatically identify the product features and sentiment words. However, the purpose of ontology in opinion mining is beyond this level of convention requiring it to categorize the reviews based on the sentiments of the extracted features. This is towards developing the better recommendation systems. Machine learning algorithms are to be improved in this direction. Ontology contains abundance of knowledge about a particular domain with the instance data. The mined knowledge from the improved machine learning algorithms extracts set of rules that are useful for learning the target concepts. This is called Semantic Data Mining [51]. Semantic Data Mining helps in target concept learning

Colace .F et al. provided [52] an approach to automatically extract positive and negative sentiments with ontological filtering. The word synsets were not considered in the ontology development. Alberto Salguero and Macarena Espinilla applied [53] an ontology based Description Logic (DL) class expression learner on the review documents to classify their sentiments. The work never concentrated on the feature level sentiment classification.

3.6 Recommender Systems in E-Commerce

The recommender systems (RS) are the information filtering systems which deal with the large amount of information that is dynamically generated based on users preferences, interests and observed behaviours. These traditional recommender systems fall into three categories. They are; collaborative filtering based RS, content based RS and knowledge based RS.

The collaborative recommender systems are the most popular and widely implemented systems. It identifies the users who are similar with the user from whom recommendations are to be provided. Resnick et al. developed [54] a system called GroupLens to help people to find articles they are most interested in. Anna Stavrianou and Caroline Brun developed [55] an application to recommend products based on the opinions and suggestions written in the online product reviews. The content based recommender systems learns the user profile based on the product feature where the user has targeted. Lang developed [56] a system called NewsWeeder which uses the words of the text as the features. Jia Zhou and Tiejian Luo developed [57] a content based recommender system that views customer shopping history to recommend the similar products based on the similarity between the product features. The knowledge based recommender systems provide the entity suggestions based on the deductions from users needs and preferences. The user profile is also required to provide good product recommendations to the user. Case based reasoning (CBR) is a kind of knowledge based recommender system. Kolodner used [58] CBR to recommend the restaurants based on the user’s choice of features. Robin Burke used [59] the FindMe system to recommend the online products.

3.6.1 Sentiments based product recommendations

Sentiment based product recommendations have gained research importance in the recent times. Li Chen and Feng Wang proposed [60] a novel explanation interface that fuses the feature sentiment information into the recommendation content. They also provided the support for multiple products comparison with respect to similarity using the common feature sentiments. Gurini et al. proposed [61] friends recommendation technique in Twitter using a novel weighting function which is called Sentiment-Volume-Objectivity (SVO) that considers both the user interests and sentiments. Finally, Dong et al. developed [63] a product recommendation strategy that combines both similarity and sentiments to suggest products.

3.6.2 Ontologies utilization for better product recommendations

The utilization of ontologies for better product recommendations is an emerging research area. Uzun and Christian developed [64] a semantic extension to FOKUS recommender system. This extension is capable of integrating contextual and semantic information in the recommendations. Hadi and Mohammadali introduced [65] a semantic recommendation procedure using ontology on online products based on the usage patterns of the customers.

4. Evaluation techniques in opinion mining

Evaluating the review at feature level by a machine is a difficult task because of the diverse natures of written reviews in the review sites. In this section the metrics used to evaluate the performance of opinion mining systems and classifiers are discussed.

4.1 Measures in Opinion Mining

Precision

Precision is defined as the fraction of number of retrieved items which are relevant to the query to the total number of retrieved items.

$$\text{Precision} = \frac{Number\_\mathrm{Re}trieved\_\mathrm{Re}levant}{Number\_Total\_\mathrm{Re}trieved}$$

Recall

Recall is defined as the fraction of number of retrieved items which are relevant to the query to the number of relevant items.

$$\text{Recall} = \frac{Number\_\mathrm{Re}trieved\_\mathrm{Re}levant}{Number\_Possible\_\mathrm{Re}levant}$$

F1-score

The definition of F1-score is the harmonic mean of Precision and Recall.

$$\text{F1-score} = \frac{2 * \mathrm{Pr}ecision * \mathrm{Re}call}{\mathrm{Pr}ecison + \mathrm{Re}call}$$

Table 1 and Table 2 tabulate the results of feature extraction techniques and sentiment classification models.

Table 1. Comparative results among different feature extraction techniques. Bold face indicates the best performer on the data

Feature Extraction Technique	Results reported by	No. of reviews	Precision (in %)	Recall (in %)
Association Rule Mining	Hu and Liu [14]	4254	72	80
Static and Dynamic Aspects learning	Blair-Goldensohn et al. [15]	3492	85	66
Dependency Grammar Graph	Zhuang et al. [16]	1000	48.3	58.5
Phrase Dependency Tree	Wu et al. [18]	4254	66.5	65.75
Double Propagation	Qiu et al. [21]	4334	88	**83**
CRF	Jacob and Gurevych [25]	2578	64	37.75

MG-LDA	Ivan and Titov [30]	10000	**90**	61.66
Works on Implicit features extraction				
COP-KMeans	Qi et al. [35]	18867	78.9	90.73
Rules based on Dependency Grammar & SenticNet	Cambria et al. [37]	4254	**93.25**	**94.15**

Table 2. Sentiment Evaluation metrics applied on various classifiers. Bold face indicates the best performer on the data

Classifier/System	Type	Results reported by	Evaluation Metric and value in %
Log linear classifier	Regression	Vasileios et al. [45]	Accuracy(Yes label) – 82.05
SO-PMI	Cluster Accuracy	Turney and Littman [48]	Accuracy of sentiment classes – 71
SVMLIGHT	NB, SVM	Esuli and Sebastiani [42]	Accuracy – **88.05**

The published experimental outcomes that are listed in Table 1 and Table 2 allow attempting some considerations on the performance of the various feature extraction techniques and the sentiment classifiers discussed. Careful interpretation of these approaches with respect to the results can help in drawing interesting conclusions. By comparing the explicit feature extraction techniques on the basis of the test carried out on the same collection (sub section 4.1), the best performer on the data is found to be the Double Propagation approach [21] in terms of recall. Ivan and Titov suggested [30] that the LDA variant is found to be the best performer on the data in terms of precision. By comparing the techniques used in implicit feature extraction the best performer on the data is found to be Cambria et al. [37] rule based approach. Ivan and Titov [30] and Esuli and Sebastiani used [42] the SVMLIGHT system to learn SVM [19] classifier to classify the sentiments of the reviews. Esuli and Sebastiani considered [42] Kamps [66] WordNet lexical relations dependent data to estimate the accuracy of the SVM model. Esuli and Sebastiani [42] sentiment classifier works better than all others.

The present works are aimed at extracting explicit and implicit product features from online reviews, the extraction of product features by improving the traditional topic model, the learning of semantic rules by mining the ontology and catering the sentiments in the products recommendation process. There is a need for intelligent discovery of sentiments of the product features by mining the constructs of the ontology towards improving product recommendations. This helps the users to make a faster and better decision making while purchase.

5. Conclusions

The speed at which the reviews are increasing rapidly in the review databases has called for a need to analyze and summarize them for effective opinion visualization and retrieval. Although the research started in opinion mining about twenty years ago, this field is still observed to be in growing stage. This survey emphasized various feature extraction approaches using both statistical and machine learning methods. Also, opinion word extraction and orientation of the extracted opinion words are reviewed. Ontology support for opinion mining has been studied. The recent works in learning the polarity of the opinion by mining the ontology is also emphasized. The influence of sentiments in recommending the products using recommender systems is also studied. The suggestions that opinion mining research have provided to the field of E-commerce are highlighted. For evaluating opinion mining, commonly used measures like precision, recall and F1- score are frequently used in the research literature. The comparative results in learning the best performer on the data are also presented. Finally, the problems were identified by highlighting the gaps identified in the literature.

References

[1] Liu, Bing. "Opinion mining." Encyclopedia of Database Systems. Springer US, 2009. 1986-1990.

[2] Popescu, Ana-Maria, Bao Nguyen, and Oren Etzioni. "OPINE: Extracting product features and opinions from reviews." Proceedings of HLT/EMNLP on interactive demonstrations. Association for Computational Linguistics, 2005.

[3] Pariser, Eli. The filter bubble: What the Internet is hiding from you. Penguin UK, 2011.

[4] Taylor, Edison Marrese, et al. "Web opinion mining and sentimental analysis." Advanced Techniques in Web Intelligence-2. Springe Berlin Heidelberg, 2013. 105-126.

[5] Pang, Bo, and Lillian Lee. "Opinion Mining and Sentiment Analysis." Foundations and Trends® in Information Retrieval 2.1–2 (2008): 1-135.

[6] Zhang, Lei, and Bing Liu. "Aspect and entity extraction for opinion mining." Data mining and knowledge discovery for big data. Springer Berlin Heidelberg, 2014. 1-40.

[7] Schouten, Kim, and Flavius Frasincar. "Survey on aspect-level sentiment analysis." IEEE Transactions on Knowledge and Data Engineering 28.3 (2016): 813-830.

[8] Kowalski, Gerald. "Information retrieval systems: theory and implementation." Computers and Mathematics with Applications 5.35 (1998): 133.

[9] 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.

[10] Liu, Bing. Web data mining: exploring hyperlinks, contents, and usage data. Springer Science & Business Media, 2007.

[11] 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.

[12] Hu, Minqing, and Bing Liu. "Mining opinion features in customer reviews." AAAI. Vol. 4. No. 4. 2004.

[13] Blair-Goldensohn, S., K. Hannan, R. McDonald, T. Neylon, G. A. Reis, and J. Reyna. Building a sentiment summarizer for local service reviews. In Proceedings of International Conference on World Wide Web Workshop of NLPIX (WWW-NLPIX-2008), 2008.

[14] Christiane Fellbaum (1998, ed.) WordNet: An Electronic Lexical Database. Cambridge, MA: MIT Press.

[15] Berger, Adam L., Vincent J. Della Pietra, and Stephen A. Della Pietra. "A maximum entropy approach to natural language processing." Computational linguistics 22.1 (1996): 39-71.

[16] Zhuang, Li, Feng Jing, and Xiao-Yan Zhu. "Movie review mining and summarization." Proceedings of the 15th ACM international conference on Information and knowledge management. ACM, 2006.

[17] De Marneffe, Marie-Catherine, and Christopher D. Manning. Stanford typed dependencies manual. Technical report, Stanford University, 2008.

[18] Wu, Y., Q. Zhang, X. Huang, and L. Wu. Phrase dependency parsing for opinion mining. In Proceedings of Conference on Empirical Methods in Natural Language Processing (EMNLP-2009), 2009.

[19] Burges, Christopher JC. "A tutorial on support vector machines for pattern recognition." Data mining and knowledge discovery 2.2 (1998): 121-167.

[20] Kennedy, Alistair, and Diana Inkpen. "Sentiment classification of movie reviews using contextual valence shifters." Computational intelligence 22.2 (2006): 110-125.

[21] Qiu, Guang, et al. "Opinion word expansion and target extraction through double propagation." Computational linguistics 37.1 (2011): 9-27.

[22] Rabiner, Lawrence, and B. Juang. "An introduction to hidden Markov models." ieee assp magazine 3.1 (1986): 4-16.

[23] Lafferty, John, Andrew McCallum, and Fernando Pereira. "Conditional random fields: Probabilistic models for segmenting and labeling sequence data." Proceedings of the eighteenth international conference on machine learning, ICML. Vol. 1. 2001.

[24] Wei Jin and Hung Hay Ho, "A Novel Lexicalized HMM-based Learning Framework for Web Opinion Mining", Proceedings of the 26th International Conference on Machine Learning, Montreal, Canada, 2009.

[25] Jakob, Niklas, and Iryna Gurevych. "Extracting opinion targets in a single-and cross-domain setting with conditional random fields." Proceedings of the 2010 conference on empirical methods in natural language processing. Association for Computational Linguistics, 2010.

[26] Li, Fangtao, et al. "Structure-aware review mining and summarization."Proceedings of the 23rd international conference on computational linguistics. Association for Computational Linguistics, 2010.

[27] Hofmann, Thomas. "Probabilistic latent semantic analysis." Proceedings of the Fifteenth conference on Uncertainty in artificial intelligence. Morgan Kaufmann Publishers Inc., 1999.

[28] Blei, David M., Andrew Y. Ng, and Michael I. Jordan. "Latent dirichlet allocation." Journal of machine Learning research 3.Jan (2003): 993-1022.

[29] Lu, Zhai and Sundaresan,"Rated feature summarization of short comments", Proceedings of International Conference on World Wide Web, 2009.

[30] Titov, Ivan, and Ryan McDonald. "Modeling online reviews with multi-grain topic models." Proceedings of the 17th international conference on World Wide Web. ACM, 2008.

[31] Titov, Ivan, and Ryan T. McDonald. "A Joint Model of Text and Aspect Ratings for Sentiment Summarization." ACL. Vol. 8. 2008.

[32] Lin, Chenghua, and Yulan He. "Joint sentiment/topic model for sentiment analysis." Proceedings of the 18th ACM conference on Information and knowledge management. ACM, 2009.

[33] Jo, Yohan, and Alice H. Oh. "Aspect and sentiment unification model for online review analysis." Proceedings of the fourth ACM international conference on Web search and data mining. ACM, 2011.

[34] Wang, Bo, and Houfeng Wang. "Bootstrapping Both Product Features and Opinion Words from Chinese Customer Reviews with Cross-Inducing." IJCNLP. Vol. 8. 2008.
[35] Su, Qi, et al. "Hidden sentiment association in chinese web opinion mining."Proceedings of the 17th international conference on World Wide Web. ACM, 2008.
[36] Cruz, I. V. A. N., Alexander Gelbukh, and Grigori Sidorov. "Implicit aspect indicator extraction for aspect based opinion mining." International Journal of Computational Linguistics and Applications 5.2 (2014): 135-152.
[37] Poria, Soujanya, et al. "A rule-based approach to aspect extraction from product reviews." Proceedings of the second workshop on natural language processing for social media (SocialNLP). 2014.
[38] Nasukawa, Tetsuya, and Jeonghee Yi. "Sentiment analysis: Capturing favorability using natural language processing." Proceedings of the 2nd international conference on Knowledge capture. ACM, 2003.
[39] Stone, Philip J., et al. "The general inquirer: A computer system for content analysis and retrieval based on the sentence as a unit of information." Behavioral Science 7.4 (1962): 484-498.
[40] Whissell, Cynthia et al. "Whissell's Dictionary of Affect in Language Technical Manual and User's Guide Dictionary Construction Word Selection." (2000).
[41] Hu, Minqing, and Bing Liu. "Mining and summarizing customer reviews." Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2004.
[42] Esuli, Andrea, and Fabrizio Sebastiani. "Determining the semantic orientation of terms through gloss classification." Proceedings of the 14th ACM international conference on Information and knowledge management. ACM, 2005.
[43] Esuli, Andrea, and Fabrizio Sebastiani. "Sentiwordnet: A publicly available lexical resource for opinion mining." Proceedings of LREC. Vol. 6. 2006.
[44] Baccianella, Stefano, Andrea Esuli, and Fabrizio Sebastiani. "SentiWordNet 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining." LREC. Vol. 10. 2010.
[45] Vasileios Hatzivassiloglou and Kathleen R. McKeown, "Predicting the Semantic Orientation of Adjectives", Proceedings of the 35th annual meeting of ACL, 1997.
[46] Turney, Peter D. "Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews." Proceedings of the 40th annual meeting on association for computational linguistics. Association for Computational Linguistics, 2002.
[47] Kim, Soo-Min, and Eduard Hovy. "Identifying and analyzing judgment opinions." Proceedings of the main conference on Human Language Technology Conference of the North American Chapter of the Association of Computational Linguistics. Association for Computational Linguistics, 2006.
[48] Turney, Peter D., and Michael L. Littman. "Measuring praise and criticism: Inference of semantic orientation from association." ACM Transactions on Information Systems (TOIS) 21.4 (2003): 315-346.
[49] Peñalver-Martínez, Isidro, Rafael Valencia-García, and Francisco García-Sánchez. "Ontology-guided approach to feature-based opinion mining." International Conference on Application of Natural Language to Information Systems. Springer Berlin Heidelberg, 2011.
[50] Sebastiani, Fabrizio. "Machine learning in automated text categorization."ACM computing surveys (CSUR) 34.1 (2002): 1-47.

[51] Liu, Haishan. "Towards semantic data mining." 9th International Semantic Web Conference (ISWC2010). 2010.
[52] Colace, Francesco, et al. "Ontological Filtering for Sentiment Analysis."DMS. 2012.
[53] Salguero, Alberto, and Macarena Espinilla. "Description Logic Class Expression Learning Applied to Sentiment Analysis." Sentiment Analysis and Ontology Engineering. Springer International Publishing, 2016. 93-111.
[54] Resnick, Paul, et al. "GroupLens: an open architecture for collaborative filtering of netnews." Proceedings of the 1994 ACM conference on Computer supported cooperative work. ACM, 1994.
[55] Stavrianou, Anna, and Caroline Brun. "Expert Recommendations Based on Opinion Mining of User-Generated Product Reviews." Computational Intelligence 31.1 (2015): 165-183.
[56] Lang, Ken. "Newsweeder: Learning to filter netnews." Proceedings of the 12th international conference on machine learning. 1995.
[57] Zhou, Jia, Tiejian Luo, and Fuxing Cheng. "Modeling learners and contents in academic-oriented recommendation framework." Dependable, Autonomic and Secure Computing (DASC), 2011 IEEE Ninth International Conference on. IEEE, 2011.
[58] Kolodner, Janet. Case-based reasoning. Morgan Kaufmann, 2014.
[59] Burke, Robin D., Kristian J. Hammond, and B. C. Yound. "The FindMe approach to assisted browsing." IEEE Expert 12.4 (1997): 32-40.
[60] Chen, Guanliang, and Li Chen. "Recommendation based on contextual opinions." International Conference on User Modeling, Adaptation, and Personalization. Springer International Publishing, 2014.
[61] Gurini, Davide Feltoni, et al. "Analysis of sentiment communities in online networks." Proceedings of the International Workshop on Social Personalization & Search, co-located at SIGIR. 2015.
[62] Li, Xiu, Huimin Wang, and Xinwei Yan. "Accurate recommendation based on opinion mining." Genetic and Evolutionary Computing. Springer International Publishing, 2015. 399-408.
[63] Dong, Ruihai, et al. "Combining similarity and sentiment in opinion mining for product recommendation." Journal of Intelligent Information Systems 46.2 (2016): 285-312.
[64] Uzun, Abdulbaki, Christian Räck, and Competence Center FAME. "Exploiting Ontologies for better Recommendations." GI Jahrestagung (1). 2010.
[65] Su, Qi, et al. "Hidden sentiment association in chinese web opinion mining."Proceedings of the 17th international conference on World Wide Web. ACM, 2008.
[66] Kamps, Jaap, et al. "Using wordnet to measure semantic orientations of adjectives." (2004): 1115-1118.
[67] What is Web 2.0, “http:// www.oreillynet.com/ pub/a/oreilly/ tim/news / 2005 / 09/30/what-is-web-20.html”.
[68] Penalver-Martinez, Isidro, et al. "Feature-based opinion mining through ontologies." Expert Systems with Applications 41.13 (2014): 5995-6008.
[69] Kumar, KC Ravi, D. Teja Santosh, and B. Vishnu Vardhan. "Extracting Opinion Targets from Product Reviews using Comprehensive Feature Extraction Model in Opinion Mining." Indian Journal of Science and Technology 10.21 (2017).

Vol. 6, Special Issue (Emerging Trends in Engineering Technology) Mar. 2018, PP. 97-111

[70] Yamada, Atsuhiro, Sho Hashimoto, and Noriko Nagata. "A Text Mining Approach for Automatic Modeling of Kansei Evaluation from Review Texts." International Conference on Kansei Engineering & Emotion Research. Springer, Singapore, 2018.