International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878, Volume-8 Issue-3, September 2019

Ontology Reasoning Towards Sentimental Product Recommendations Explanations

Vidya Kamma , Teja Santosh Dandibhotla, Sridevi Gutta

Abstract— In the last two decades various organizations like service industries, study communities, academic world and public industries are working intensely on sentiment analysis, to extract and analyze public views. The reviews given on the social websites, commercial websites, etc. enable customer to share their point of view. Explainable Recommendation algorithms help the user by providing explainable recommendations, which improves user satisfaction. Recently, many researchers proposed explainable recommendations. In this survey Firstly, various opinion-mining approaches are explored. Secondly, we reviewed sentiment-based and ontology based recommendation systems. Finally, prospects for the research in opinion mining is discussed.

Keywords: Opinion mining, opinion orientation, ontology, product features, sentiment.

I. INTRODUCTION

The information sharing on the web is growing rapidly. To carry out online transactions social websites utilize B2C model. These websites provide large amount of information about products and services, which is leading to the content development and the customers, write their views on the products purchased from their websites. The customers can evaluate the product based on the opinions and can take the purchase decision. The online reviews have major impact on buyers. Facts and Opinions are two groups of textual content. Objective statements regarding the products and their properties are facts. Subjective statements that disclose the people's feelings about the entities and their properties. The business organizations adopted surveying and other mechanisms to know the feedback on their products and services. Reviews are the opinions on the products purchased by users. The reviews contain product feature information. To understand the user's perspective on the product, analysis on reviews are done by business organizations. The sentiment is calculated by analyzing the customer views from the product reviews. Positive, negative and neutral are the polarities used for calculating the sentiments of the product. Obviously, the decision of the customer is in the direction of positive polarity featured product. Online reviews are pre-processed prior to determine the polarity, tokenization, removal of stop words and POS-tagging are the done during pre-processing. Extracted features from above mentioned task are linked to form feature opinion pair. Various supervised machine learning algorithms or models like ontology are used to determine the polarity. In providing user recommendations, the current systems are lacking in accuracy and efficiency. To improve the above recommender systems can be provided with sentimental information. The recommendations can be improved by providing ontology, which helps to make accurate decisions by customers.

Manuscript published on 30 September 2019
* Correspondence Author
Vidya Kamma*, Research Scholar ,K L (Deemed to be University), Andhra Pradesh State,India and Assistant Professor,Neil Gogte Institute of Technology, Hyderabad, Telangana, India.
(Email: kammavidya@gmail.com)
Teja Santosh Dandibhotla, Associate Professor,Sreenidhi Institute of Science and Technology Yamnampet, Ghatkesar, Telangana , India.
(Email: tejasantoshd@gmail.com)
Sridevi Gutta , Professor,K L (Deemed to be University), Andhra Pradesh , India (Email: sridevi.gutta2012@gmail.com)

II. EXPLORED TECHNIQUES IN OPINION MINING

Mining opinions from product reviews is a challenging task. The algorithms developed in past work well on one domain may not well on a different domain, because the opinion words used changes depending on the context.

A. NLP Approaches

Frequency-based Approaches

It has seen that a limited set of words are used in reviews. These frequent single and compound nouns are considered as aspects. This approach finds frequent occurring nouns as aspects from the online reviews in a given domain. Features are often expressed by a verb, adjective, noun or adverb. Mostly people use frequent nouns as aspects, but not all frequent nouns are aspects. Frequent nouns (product features) are identified using frequency based methods. The research work developed Red Opal system to determine online products based on collected features. Another work used clustering to extract features from products. To compare two noun phrases they used dice analogy coefficient as a measure.The advantage of this approach is quite simple and effective method. The limitations are it needs manual tuning of number of parameters and miss low-frequency aspects. These frequency based approaches can be enhanced to solve some of these issues.

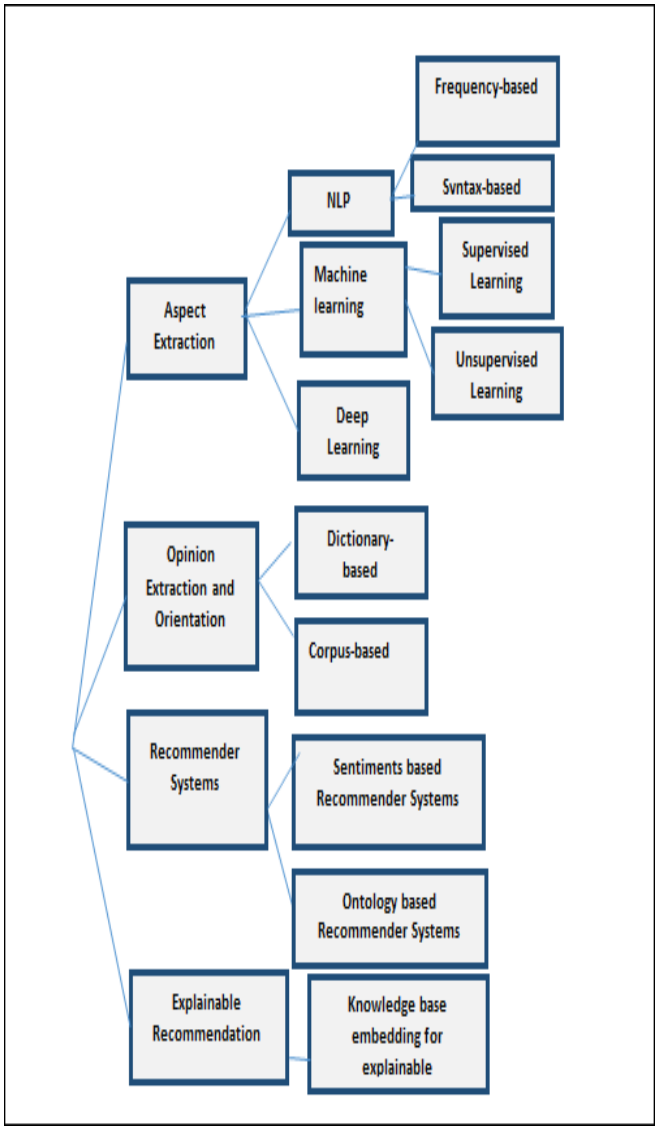

Fig-1:Considered Taxonomy

Syntax-based Approaches

Syntactical relationship are used to find aspects based on syntactic patterns. This method finds extraction rules between aspect and opinion words by using syntactic patterns and grammatical relation. Zhuang et al. used dependency parser for aspect extraction from movie reviews. Wu et al. expanded Burges, C. J. work by classifying the features and sentiments using support vector machine (SVM) classifier. This process never measured the relationship between opinions and features. Cruz, I., Gelbukh, A. F., Sidorov, G. suggested Opinion Observer to differentiate consumer opinions of various products. Qiu et al. developed Double Propagation algorithm. Double propagation algorithm identifies the opinion words and product features from the given input and used to modify the product feature orientation. A tree kernel function for finding syntactic patterns was proposed. Chiranjeevi et al. (2018) used distant supervision technique to improve the precision in the number of extracted aspects.The limitation with this approach chances to produce various non-aspects that match syntactic pattern.

B. Aspect Extraction using Machine learning approaches

Supervised and Unsupervised learning Approaches. Supervised approaches uses frequency based approach to produce features that are more considerable. Jakob, N., Gurevych, I. solved labeling problem by linear chain CRF. Fangtao et al. proposed the skip-tree CRF approach to overcome the problem faced by Jakob, N., Gurevych, I. This method skips using the combination structure and syntactic tree structure information from the training sequence of data to collect the product features. The strength of supervised learning methods is to overcome frequency-based restrictions by learning sample parameters from training data. The limitation is it requires manually labeled data for training. Therefore, accuracy of learned model depends on what aspects are labeled. Unsupervised learning approach solves problem of domain dependency and reduce need for labeling training data in sentiment classification. This approach usually needs a huge amount of trained data to successfully perform. Most common method used was LDA Blei et al. extracted aspect words from product reviews using a MaxEnt-LDA (Maximum Entropy and LDA combination). Lu et al. implemented structured PLSA and unstructured PLSA for feature detection on user reviews and found that structured PLSA has extracted good number of features. Titov et al. altered model of Blei et al. and named as MG-LDA (Multi-grain Latent Dirichlet Allocation). They considered particular local and global topics for feature extraction. The limitation of the model is lack of association among topics and features. Titov et al. enhanced their MG-LDA model and named as MAS (Multi-Aspect Sentiment) (MAS) for extracting features. Lin et al. proposed (JAS) Joint Aspect/Sentiment Model. The limitation of this approach is it requires large amount of data.

C. Deep Learning Approaches for Aspect Extraction

Deep learning provides learning models to automatically extract aspect from a sentence without human participation. Katiyar, A., Cardie, C. studied usage of relationships connect entities and deep bidirectional LSTMs for combined extraction of opinion entities. Wang et al. proposed a model combining CRF, RNN for extracting aspects and expressions. This model double propagate data among aspects and opinion terms. Li, X., Lam, W. improved model proposed in Katiyar, A., Cardie, C. for sentiment and aspect extraction. This model consists three LSTMS in which two used for sentiment and aspect extraction and one LSTM for sentiment polarity identification. Zhang et al. modified CRF model using neural network. This modified model substitutes the original features with word embeddings, neural layer between input node and output node. Poria et al. developed deep seven layer CNN by integrating linguistic patterns. He et al. proposed a model for non-supervised aspect extraction. Main objective is to focus more on aspect-related words. The limitation of this approach is supervised and reinforcement learning need to be reinforced by huge amount of data.

Retrieval Number C6852098319/2019©BEIESP
DOI: 10.35940/ijrte.C6852.098319
Journal Website: www.ijrte.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

III. Opinion Extraction and Orientation

After opinion extraction next task is finding orientation of extracted opinion word. Nasukawa, T., Yi, J. collected sentiment words, refined the sentiment patterns by learning dependencies among the words and added to the sentiment lexicon to extract real opinion words. Opinionated words are collected from GI (General Inquirer) Stone et.al.,WordNet Fellbaum, C. and DAL (Dictionary of Affect of Language) Kim et al. expanded sentiment lexicon by collecting the synonyms of sentiment words from WordNet and applied to subjective sentences extracted to obtain actual opinion words.

A. Dictionary-based Approaches

Hu, M., Liu, B. generated the word lists by collecting the information manually and further extended the list by adding antonyms synonyms of collected words. Antonyms synonyms are taken from WordNet. Esuli et al. developed SentiWordNet 1.0 resource for opinion mining. Kennedy et al. used dictionary called General Inquirer (GI) for sentiment class identification. Baccianella et al. developed SentiWordNet 3.0 which is extension of the resource developed in Esuli et al. The versions differ in the versions and the algorithms used. The limitation is unable to find opinion words with domain specific orientations.

B. Corpus-based Approaches

Corpus-based approaches to identify opinion words depend on syntactic patterns. Hatzivassiloglou, V., McKeown, K. R. Considered opinion adjectives, opinion words and opinion orientations. Some words like BUT in reviews changes the opinion orientation. So, orientation of adjective is identified and graph is generated by linking adjectives. To, produce negative and positive words clustering is performed on the graph. Jiao, J., Zhou, Y. Lafferty, J., McCallum, A., Pereira, F. C. Presented a statistical method called (CRF) Conditional Random Fields to extract opinion phrases and distinguish sentiment polarities. The disadvantage of corpus based approach difficult to build large corpus. This approach is efficient in finding orientations.

C. From Opinion to Sentiment for Opinionated Recommendations and explanations

To help users take accurate decisions and increase their satisfaction automatically generated recommendations should be explained. To achieve this, the detected opinions are to be analyzed further for orientations. Ravikumar et al. (2017) proposed an opinion orientation approach that modifies distance formula upon opinionated adjectives. Ravikumar et al. (2108) improved the accuracy of the number of opinion orientations by learning ontology from the product reviews for detecting maximum number of opinions. The quantitative orientations of the opinions are transformed into settled opinion called Sentiment. McAuley, J., Leskovec, J. combined topic models with rating prediction to provide text-based explanations. The limitation with this approach is opinions on specific topic or feature is missing which leads wrong explanations. Ren et al. mapped aspect-level comments into numeric ratings by ignoring the detailed reason of opinions. Baltrunas et al. considered holistic view of an item recommendation and developed factorization based recommendation algorithms.

IV. RECOMMENDER SYSTEMS

The recommender systems (RS) use information filtering technology to suggest user's items of preferences based on item and user attributes, users preferences. Collaborative filtering (CF) algorithms recommend highly rated items to other users with same user preferences and item attributes. Resnick et al. used RS to find articles which are mostly liked by the users developed a recommender system called GroupLens. Stavrianou, A., Brun, C. employed an application to recommend products. Content based filtering uses attributes derived from documents. The limitation is that there is no variety in the recommendations. Lang, K. used words as features and developed system called NewsWeeder. Zhou, J., Luo, T., Cheng, F. developed a recommender system which recommends the products to the users based on the user shopping history. Knowledge based filtering based on preferences of user knowledge based recommender system provide the suggestions. Kolodner, J. used case based reasoning system to recommend the restaurants. Burke, R. D., Hammond, K. J., Yound, B. C. developed FindMe for recommending the products. Holland, S., Ester, M., Kieling, W. used log data for mining product preferences.

A. Sentiment Based Recommender Systems

The current research is concentrating on sentiment based product recommendations. The features and opinions extracted from online reviews are useful in recommendations. Chen, G., Chen, L. implemented an interface and compared multiple products by considering similarity using common sentiment features. Gurini et al. suggested a novel approach by considering both user sentiments and interests for friend recommendations in Twitter. They named the approach based Sentiment Volume-Objectivity (SVO). Li, X., Wang, H., Yan, X. developed a recommender system to recommend products based on sentiment strength and expressions. Recently, Teja Santosh et al. (2018) used case-based reasoning for providing sentiments utilized intentions based product recommendations.

B. Ontology Based Recommender Systems

The use of ontologies for improved product recommendations is a developing area. Santosh and Vishnu (2018) learned sentiments based semantic web rules on the top of their ontology for providing improved product recommendations to the e-commerce customers.

V. EXPLAINABLE RECOMMENDATION & RESULTS

To improve the efficiency, user satisfaction and effectiveness of recommendation systems explainable recommendations are provided. These systems provide the users with explainable recommendations.

A. Knowledge based Embedding for Recommendation Explanation

Huang et al. for (RNN) sequential recommendations combined Recurrent Neural Network with (KVMN)

Key-Value Memory Network .Further the model is enhanced by providing value level interpretability. Catherine et al. proposed a method which uses Personalized Page Rank approach to rank knowledge graph entities and items. This method produced explainable recommendations. Ai, Q. et al. constructed user item knowledge graph includes item, entity user relations. They adopted knowledge graph embedding to obtain relation, entity and user embedding's and provided recommendations for user.

VI. PROSPECTS FOR RESEARCH IN OPINION MINING

The present work is aimed at extracting product aspects from online reviews, determining valence of extracted opinions of aspects, reasoning ontology for providing product recommendations using the content-based knowledge graph. It is observed that i) There is no linguistic and clustering based combined approach for extracting maximum number of product aspects. ii) There is a semantic gap existing between textual reviews and actual opinionated knowledge in order to identify opinions of the extracted aspects. iii) There is a need to express learned ontology to intelligently mine explanations for the statistical recommendations. The imminent offerings to the above-specified shortcomings in opinion mining research are (a) to provide a linguistic and clustering based approach for aspects extraction. b) To reduce semantic gap using ontology learning for opinions analysis of extracted aspects and c) To construct expressive rules on top of learned ontology in order to explain the sentiment and similarity based product recommendations for explaining them.

VII. CONCLUSION

As the review databases are growing rapidly there is a need for summarizing and analyzing these reviews for effective information retrieval and efficient decision making. As the work in opinion mining started out about two decades ago, still it is determined in growing level. This paper highlighted various statistical and machine learning feature extraction approaches. The opinion word extraction and orientation of opinions is reviewed. Also, Ontology support for opinion mining is emphasized. Recommender systems for recommending the products are also emphasized. Also, Knowledge based embedding for explainable recommendation is studied. For evaluating recommender systems metrics from machine learning and information retrieval perspectives are emphasized. Prospects for the research in opinion mining are also included.

REFERENCES

1. Kumar, KC Ravi, D. Teja Santosh, and B. Vishnu Vardhan. "Determining the semantic orientation of opinion words using typed dependencies for opinion word senses and SentiWordNet scores from online product." (2017).
2. KC, Ravi KUMAR, SANTOSH DANDIBHOTLA Teja, and Vishnu VARDHAN BULUSU. "Learned ontology guided opinions analysis of extracted aspects from online product reviews." Journal Europeen des Systemes Automatises 51.1-3 (2018): 25.
3. Pandi, Chiranjeevi, Teja Santosh Dandibhotla, and Vishnu Vardhan Bulusu. "Corpus Linguistic Rules Based Review Sentence Selection for Opinion Targets Extraction and Opinion Orientation: A Distant Supervision Approach."
4. Santosh, Dandibhotla Teja, and Bulusu Vishnu Vardhan. "Sentiment-Based Semantic Rule Learning for Improved Product Recommendations." Machine Learning: Advanced Techniques and Emerging Applications (2018): 185.
5. Santosh, D. Teja, KC Ravi Kumar, and P. Chiranjeevi. "Feature level intentions based product recommendations with case-based reasoning." (2018).

AUTHORS PROFILE

Vidya Kamma is presently working as Assistant Professor in CSE at Neil Gogte Institute of Technology, Hyderabad. She is the research scholar in the faculty of Computer Science and Engineering K L (Deemed to be University). Her areas of interest are Data Mining, Machine Learning, Natural Language Processing.

Teja Santosh Dandibhotla Associate Professor, Sreenidhi Institute of Science and Technology , Ghatkesar, Hyderabad. He is awarded with Ph.D. in the faculty of Computer Science and Engineering under JNTU Kakinada. His areas of interest are data mining, machine learning, natural language processing. He was the Additional Reviewer for 2015 International Conference on Advances in Computing, Communications and Informatics (ICACCI) in Kerala, India. He is the member of professional bodies of international repute namely IEEE, ACM, IAENG. He is one among the 16 research scholars of Dr. B. Vishnu Vardhan.

Sridevi Gutta Professor K L (Deemed to be University).She is Awarded with Ph.D from Acharya Nagarjuna Univerisity. Her areas of interests are Software Engineering and Data Mining .She is the member of various professional bodies.