Aspect and Review Based Recommendation System

Mrs. Sujata Khedkar
Associate Professor
Department of Computer Engineering
V.E.S. Institute of Technology
Mumbai-400074
sujata.khedkar@ves.ac.in

Mugdha Govilkar
Computer Engineering
V.E.S. Institute of Technology
Mumbai-400074
mugdha.govilkar@ves.ac.in

Tejas Ingale
Computer Engineering
V.E.S. Institute of Technology
Mumbai-400074
tejas.ingale@ves.ac.in

Mihir Joshi
Computer Engineering
V.E.S. Institute of Technology
Mumbai-400074
mihir.joshi1@ves.ac.in

Abhishek Kuvar
Computer Engineering
V.E.S. Institute of Technology
Mumbai-400074
abhishek.kuvar@ves.ac.in

***Abstract*—In this paper, we are proposing an advanced recommendation system which performs natural language processing on the user reviews and gains a sentiment score for various aspects of the product/service. The sentiment score based on user reviews provides a more fine grained result for the aspects rather than coarse grained approaches like collaborative filtering which is purely based on users rating. Such classified recommendation can be used to recommend products/services to users based on their profiles and also gather the negative sentiments about certain aspects to report areas of improvement to the businesses selling the products/services.**

Keywords—Natural Language Processing, Sentiment analysis, Opinion parsing

I. Introduction

The Internet is booming and the state-of-the-art e-commerce has now spread over every field. This boost has a major contribution from the rise of smartphones and their availability for the masses. It has essentially brought the user close to the market because of the facility of providing instant reviews and ratings. There is a multitude of products and services available on various e-commerce websites. Choosing the best among the better products available on the Internet is observed to be an arduous task.

This was the dawn of the recommendation systems which give the personalised product recommendations to the users. Many approaches have been explored and implemented in order to make recommendations more accurate and relevant to the user preferences(eg: Collaborative Filtering). However most of these approach considers only the features of the product and not how those features are rated by the users. Natural Language Processing(NLP) can be used to perform sentiment analysis on user reviews to extract sentiment toward the individual feature(aspect).

II. Literature Review

The literature survey involved looking for relevant IEEE papers published by students or well renowned authors. These papers played a role in giving us a guideline for our approach toward the problem and an overall depth of the problem and its extent. It also helped us to get an idea about the extent to which the problem has been perused and dealt with. We were also able to define the scope up to which our proposed system will take care of the issue. The challenges faced by the authors of these papers gave insights and will help us to plan mitigation strategies if we face similar challenges in the implementation. The key words or phrases we used to find the most relevant research papers are Natural Language Processing, Sentiment analysis, Opinion parsing, Aspect Extraction, Sentiment Extraction.

In [1], state-of-the-art studies into two principal branches: review-based user profile building and review-based product profile building. In the user profile sub-branch, the reviews are not only used to create term-based profiles, but also to infer or enhance ratings. Opinions with a broad perspective can further be exploited to extract the weight/value preferences that users place on particular features. In [2], a novel approach to introduce aspect-based sentiment analysis into recommender systems is proposed.

The aspect of the product using the topic model is extracted and then the aspect-specific sentiment words are identified using the SentiWordNet (a sentiment lexicon). The use the result of sentiment analysis is then used to make user interests model and the product model. By comparing two models of each user-product pair, we obtain the similarity of the user's interest and the product.

The sentiment analysis system proposed in [6] performs two key functions, aspect extraction and aspect sentiment classification. Aspect extraction has the aim to get the sentiment targets on which some sentiments have been expressed. These targets are usually different aspects of entities (e.g., products or services), which are products in our context. Aspect sentiment classification classifies whether the sentiment expressed on an aspect is positive, neutral, or negative. The main advantage of this new model is the novel additional functionality of providing not only recommendations of items to users, but also recommendations of the most valuable aspects that may enhance user experiences with items.

III. Traditional and Existing System

Traditional Recommendation methods usually focus on utilizing product features obtained from structured behaviour information, which only contains coarse grained user interests. The sentiments in textual reviews are not considered.

IV. Proposed System

The goals of our new proposed system includes two perspectives viz., customers and businesses. The customer perspective includes recommendation of products based on customer interests, preference based product rating, notification related to new developments based on user interests and user or product profile building. The other perspective includes overall sentiment towards product in market, suggested areas of improvement for product, flexible report generation templates and current trends in the market.

Advancement that we are trying in our system is to provide accuracy by comparing to existing purely feature based recommendation systems,, robustness by developing the ability to handle large datasets with minimal slowdown or crash, usability by making user interface minimalistic with well spread-out features having responsive layout and scalability by improving the ability to extend the system to a distributed environment.

Our new proposed system mostly divided into two main components viz., Data Preprocessor and Opinion Parser.

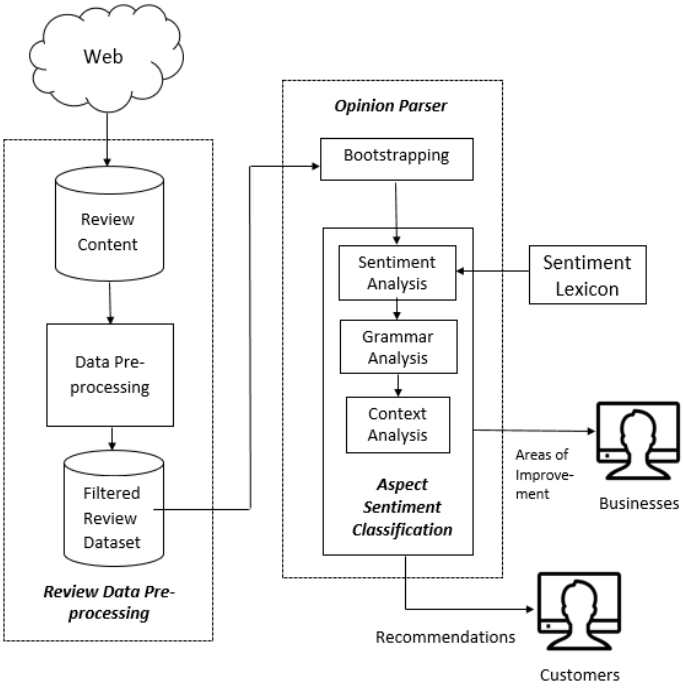

Fig 1: Block Diagram of Aspect and Review Based Recommendation System

In data preprocessing, we check for opinion spamming by parsing content of all reviews for the same product, for similar selective products. If any reviews are found to have exactly matching content, we mark all such reviews as irrelevant. Further we calculate the weight of

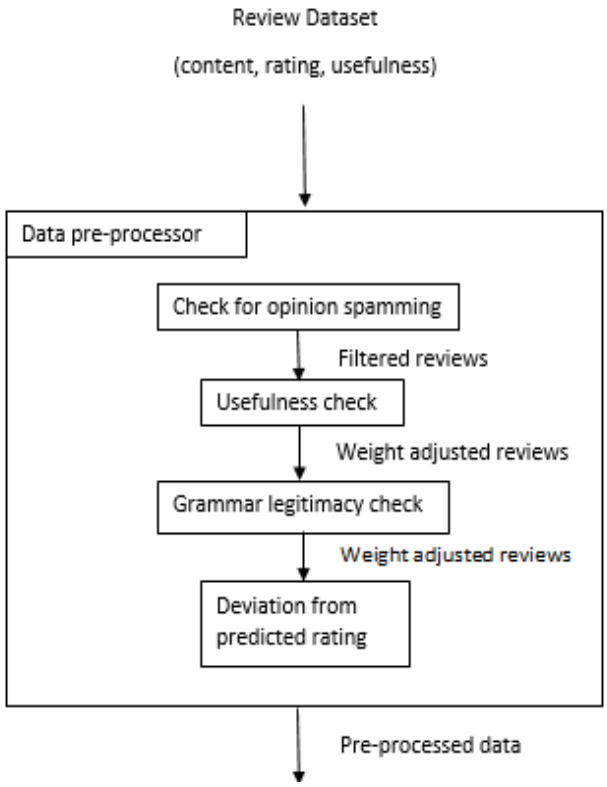

Fig 2: Data Preprocessor

usefulness according to the ratio of the number of people who have found the review useful to the number of people who have viewed the review.

Data preprocessing also includes grammar legitimacy check based on the content of review. The sentence is segregated into different parts of speech such as nouns/pronouns, verbs, adjectives, etc. Syntactical rules are applied to the sentence and the order of these words is checked. If there is an unusually high number of consecutive adjectives (sentiments), the review is considered to be fickle and the weight of this review falls considerably.

Last part of data preprocessing help to provide deviation from predicted rating with the extraction of adjectives from the review which are scored on a scale of 1 to 5. An overall expected rating associated with the review is calculated. This calculated rating is then compared to the actual rating. The weight of the review is adjusted on the basis of the deviation of the calculated rating from the actual rating.

The preprocessed data from data preprocessor is fed as input to the opinion parser, which is the most important component of system provides output which contain recommendations for users and overall product sentiment (for business). Opinion parsing includes bootstrapping process for aspect extraction. Bootstrapping is an opinion words and opinion target (aspect) extraction process. In this process a set of opinion words like *"good","bad", "amazing",* called as Opinion lexicon, is given as an input to the bootstrapper. This Opinion lexicon is used by the bootstrapper to identify opinion words from the reviews.

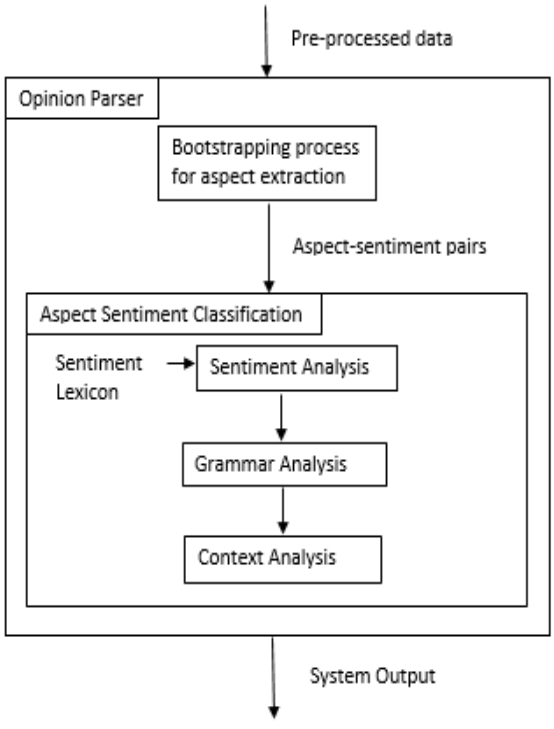

Fig 3:Opinion Parser

It then extracts corresponding aspects and forms <aspect, sentiment> pairs. Known Opinion Lexicon and extracted opinion words and target (aspect) are then used together to further extract opinion words and targets. Subtasks included in this process are extracting targets using opinion words, extracted targets and extracting opinion words using extracted targets and using both the given and the extracted opinion words. This process goes on till no opinion words and target are left to be extracted.

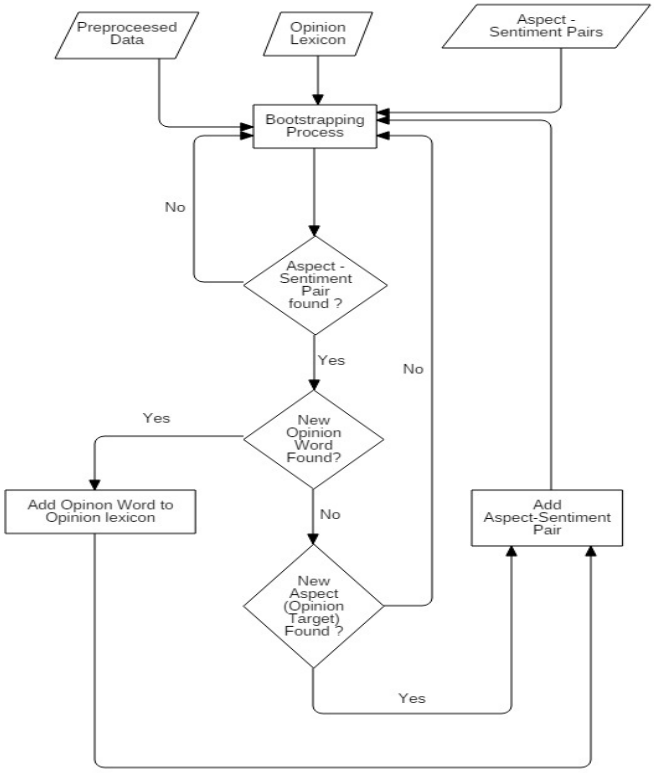

Fig 4: Bootstrapping Using Double Propagation

Other part of opinion parser is sentiment analysis, a block of opinion parser assigns sentiment weights to the aspects. It uses Sentiment Lexicon to identify sentiment weights to the aspects. It uses Sentiment Lexicon to identify sentiment intensity associated with the particular opinion word. Sentiment lexicon is a database of lexical units of a language along with their sentiment orientations. Strengths of these sentiments are averaged to get the overall sentiment for the given aspect

V. Conclusion

In particular, we are aiming to measure how much the ratings have changed for those users who "follow" our recommendations on the test set by mentioning the recommended aspect in the review. In addition to the average rating on the test set, we can compare our results with three strong baseline approaches. These baselines basically indicate the strengths and weaknesses of the establishments based on their user reviews: Trending

Similarly, those users who will not follow our negative recommendations (and experience the negative aspect of an item against our advice) might give higher/lower ratings to the items than the average rating of the items given by all users in the application and those users who did not follow the recommendations provided with the baseline approach. The results are likely to show that managers who "followed" our positive recommendations, obtained higher or lower ratings for the user experiences than the managers who followed recommendations provided with the baseline approaches. The results shall hold not only for the restaurants but also across hotels and beauty domains.

The state-of-the-art approach to enhance the functionality of recommender systems by recommending not only the product itself but also some positive aspects of the product to further enhance user experiences with the it. Recommendations of a set of valuable aspects are expected to work well as those users who follow our recommendations must rate their experiences significantly higher than those who follow the baseline recommendations.

The way in which users of the system behave is fundamental in judging the performance of the system. Considering the user profiles and product profiles for recommendation would really assist retaining the aspects of the user-product pairs and being highly relevant in making recommendations and provide even better insights to businesses with the help of product-user association.

VI. References

[1] Li Chen, Guanliang Chen and Feng Wang, "Recommender systems based on user reviews: the state of the art".

[2] Yu Zhang, RuiFang Liu, AoDong, "A Novel Approach to Recommender System Based on Aspect-level Sentiment Analysis".

[3] Deepali Virmani, Vikrant Malhotra and Ridhi Tyagi, "Aspect Based Sentiment Analysis to Extract Meticulous Opinion Value".

[4] Niharika Purbey, Aishwarya Iyer, "Categorization and Analysis of Yelp restaurant reviews".

[5] Warih Maharani, Dwi H. Widyantoro, Masayu Leylia Khodra, "Syntactic pattern based aspect extraction".

[6] Konstantin Bauman, Bing Liu et al, "Aspect Based Recommendations: Recommending Items with the Most Valuable Aspects Based on User Reviews".

[7] S.K. Bharti, B. Vachha, R.K. Pradhan, K.S. Babu, S.K. Jena, "Sarcastic sentiment detection in tweets streamed in real time: a big data approach".

[8] Vaishakh Suresh, Sayeda Roohi et al, "Aspect-Based Opinion Mining and Recommendation System for Restaurant Reviews".

[9] Sapna Negi, Paul Buitelaar, "Towards the Extraction of Customer-to-Customer Suggestions from Reviews"

[10] Toqir Ahmad Rana, Yu-N Cheah , "Hybrid Rule-Based Approach for Aspect Extraction and Categorization from Customer Reviews"