Plagiarism Scan Report

Summary	
Report Genrated Date	21 Apr, 2018
Plagiarism Status	**95% Unique**
Total Words	991
Total Characters	6432
Any Ignore Url Used	

Content Checked For Plagiarism:

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, There are three steps involved:

A. check for opinion spamming
Parse the content of all reviews for the same product.

Parse the content of reviews for similar selective products.
If any reviews are found to have exactly matching content, mark all such reviews as irrelevant.
Set their weightage for the recommendation system to 0.
Blacklist the spam reviewers.
Example: If the review has the following content: “This phone is really good. Please buy this phone! It’s has the best touch screen and the camera is awesome!”, this content is compared with reviews on the same product and with the content of reviews on similar products. If the content with other reviews matches, all such reviews are marked irrelevant and their weightage is set to 0, also, the users who wrote these reviews will be added to a blacklist and other reviews by these users will be marked unimportant.

Fig 2: Data Preprocessor

B. Helpfulness check
Obtain the number of people who have viewed the review, call it ‘v’.
Obtain the number of people who have found the review useful, call it ‘u’.
Calculate the ratio of u to v.
Approximate the weight according to this ratio to one decimal place.
Example: If the review has a high votes-to-views ratio, say a review has been marked helpful by 200 users and not helpful by 150 users, the ratio is 1.33. This number is compared to a threshold number, 0.6, which, if the number fails to exceed, the review is discarded.

C. Deviation from predicted rating
The adjectives extracted from the review are scored on a scale of 1 to 5.
An overall expected rating associated with the review is calculated.
This calculated rating is 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.
Example: If a review has all positive sentiments, like “This is the best phone in this price range. The screen is very smooth and responsive. The clarity and vividity of the camera is amazing and the processor can handle heavy apps.”, and the entered rating of the phone is 2.5/5, the rating is not in sync with the review content. Such reviews are discarded.

The preprocessed data from data preprocessor is fed as input to the opinion parser. Opinion Parser is a heart of the system. It takes preprocessed data from data preprocessor and produces output which contain recommendations (for users) and overall product sentiment(for businesses). This process takes place in two steps:

Fig 3:Opinion Parser

A. 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.
Initially bootstrapper uses initially provided Opinion Lexicon to identify opinion words in the reviews. It then extracts corresponding aspects and forms pairs.
Known Opinion lexicon and extracted opinion words and target (aspects) are then used together to further extract opinion words and targets.
Subtasks included in this process are:
A. extracting targets using opinion words
B. extracting targets using extracted targets
C. extracting opinion words using extracted targets