Addressing Cold-Start Product Recommendation Using Microblogging Reviews

Nilesh Kumbhar[1], Krushna Belerao[2]
[1, 2] Department of Computer Engineering
[1, 2] Trinity College of Engineering and Research, Pune India

Abstract-Collaborative filtering is most common technique used in recommender systems. It uses explicit information for making recommendation like product ratings, user's transactional information and previous purchase records. The problem associated with collaborative filtering is cold-start because of new product launch in market or new user with recently registered over E-Commerce. In these days, many E-Commerce allows user to sign in their website through social media account. In such cases E-Commerce fails to recommend items to these users because E-Commerce doesn't have any previous information about them. This is also called as cross website cold-start product recommendation problem. To address this issue we proposed new framework that incorporate reviews of products from social media-Twitter and uses user-item feature based matrix factorization technique for recommending more relevant list of products to user. We also considered the audio reviews about products for getting more accuracy while predicting items. With the help of proposed system accuracy of recommender system in cold-start situation is increased.

Keywords- Collaborative Filtering, E-Commerce, Microblogging, Recommender Systems, Sentiment Analysis, Social Networks, Speech Recognition

I. INTRODUCTION

Recommender systems are software and tools which are used to suggest more relevant list of items that are similar to the item being search by the user on e-commerce website [1]. Most common methods used in recommender systems are content-based method and collaborative filtering. Content-based systems uses the features of items for example in case of movies one can extract features like movie is animated or not, is action movie? Or is a kid's movie, which actors are there and so on.

Collaborative filtering method uses the explicit information like user's ratings about products, previous product purchase history, location travelled by targeted user. Because of this, it leads two well known problem data sparsity and cold-start. Data sparsity is because of most of the users rate only few features of product, hence there is lack of

sufficient information about product's pros and cons. Cold-start situation is due to new product launched in market which is not yet rated by anyone. Another case of cold start problem is new user to e-commerce website and hence e-commerce fails to recommend items to such users. Nowadays, many e-commerce websites allow user to login to their website through social media account and lead to cold-start situation because the e-commerce doesn't have any previous purchase record or interest of user. These users are completely new for e-commerce. This situation also known as cross website cold start product recommendation and we made attempt to reduce this problem using microblogging reviews collecting from social media Twitter. We propose a system model that incorporates social information in recommender system and accuracy is improved using text as well as audio reviews from microblogging platform. This is first work that considers audio reviews in recommending the list of similar product.

II. LITERATURE SURVEY

Hao Ma et al. proposed a system that addresses the data sparsity and poor prediction accuracy problem in recommender systems by adding social media information like social network and social tags. User-Item rating matrix used with social contextual information using probabilistic matrix factorization [2].

Recommender systems are personalized, non-personalized and may be based on user demographics, content-based and collaborative filtering. These systems make use of social trust for alleviating cold-start situation, user's product ratings, location travelled by user, implicit information like user's web browsers history, product mostly liked on social media for improving functioning of recommender systems [3].

R. Sandanalakshmi et al. proposed speech to text converter for mobile application. Conversion involves two main steps, first is front end analysis and second is pattern recognition. Features from speech are extracted using Mel Frequency Cepstral Coefficient (MFCC) and generalized regression neural network used as a speech recognizer [4].

Bernd Hollerit et al. given the relationship between products purchased from e-commerce and their acceptance by users in terms of tweet on social media twitter. Undesired advertisements can be avoided from showing to users on twitter. Hence buyers can handle much more relevant advertisement on twitter [5].

Lakshmish Kaushik et al. made attempt to extract sentiment from natural audio stream. Audio is considered from YouTube videos. Maximum Entropy and Part of Speech techniques used for text based sentiment detection. Audio to text conversion carried out using Mel Frequency Cepstral Coefficient method [6].

Sayed Ali et al. introduced hybrid polarity detection model for sentiment analysis and text summarization of twitter data. Twitter post also known as tweets are rich in pros and cons information of product purchased from e-commerce websites that can be used in recommender system [7].

Chinsha T et al. addressed problem of manually extracting the sentiment or opinion of reviews of products. Opinion mining is carried out at three different level namely document level, sentence level and aspect level. SentiWordNet and Aspect table is used for the sentiment classification [8].

Mayy Habayeb et al. introduced location based recommender system model that combines user's web interaction and historical trajectory information of user's smart phone. Location information like latitude, longitude and timestamp is used for recommending list of more relevant products to the user's current location. Open street map is used for the same [10].

Sneha Pasarate et al. gave comparison among different sentiment analysis techniques. Total weighted score computing method, Neutral or polar or irrelevant classifier method, weighting and aggregation scheme, intrinsic or extrinsic or domain relevance approach are some common methods used for sentiment analysis [11].

Zhipeng Gao et al. addressed cold-start product recommendation using limited resource table method. This method is based on limited resource. As more number of product launching day by day it is difficult to recommend more relevant items that the user interested. Hence attempt was made to improve the recommender system [12].

III. MATHEMATICAL MODEL

Proposed system can be viewed as 7-tuples as

$S = \{N, M, L, O, P, F, G\}$

Where,

N = Number of users
M = Product ID
L = Tweets
O = Output
P = Processing
F = Case of failure
G = Case of success

A. Input

Inputs to the systems are user, product Id and reviews in the form of text and audio.

B. Output

List of more relevant items with respect to item that is being searched by the user.

C. Intermediate State

It involve following processes

- User registration
- User login| logout
- Product searching
- Features extraction
- Reviews collection
- Audio to text conversion
- Sentiment classification on reviews
- Features mapping
- Relevant products list generation
- Recommendation

D. Case of failure

User's opinion not agrees on any group of peoples. Hence irrelevant recommendation to the user.

E. Case of success

More relevant products are generated.

IV. AUDIO TO TEXT CONVERSION

Audio to text conversion process is divided into two phases namely features extraction from audio stream and speech recognition.
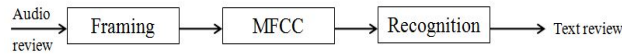
Figure 1. Audio to text review conversion

Audio stream is divided into small frames of size 25-45 millisecond. Fast Fourier Transform is performed followed by logarithmic compression [4]. Then MFCC coefficients are formulated using inverse fast Fourier transform of Mel logs. When particular frequency is matched the word gets identified.

V. SYSTEM ARCHITECTURE

Following Figure 2, show proposed system architecture. Users with social media account or E-commerce account are allowed to login into system. User has options like search items, logout. When user search a required item, task of recommender system is to show more relevant list of items similar to item he is being searching. Generation of list of more relevant items involves following processing. Product features extraction, in this, features of products is listed out and the items that are having these features are listed. Reviews of listed items from social media-Twitter are retrieved. Reviews involve text as well as audio reviews. Audio to text conversion, in this, audio reviews are processed first and then converted to text reviews. The sentiment analysis is carried out on those reviews. Sentiment analysis gives positive or negative or neutral nature of the reviews and is used for determining list of more relevant items to targeted user. Feature mapping, at this level the items having more similar product feature are formulated.
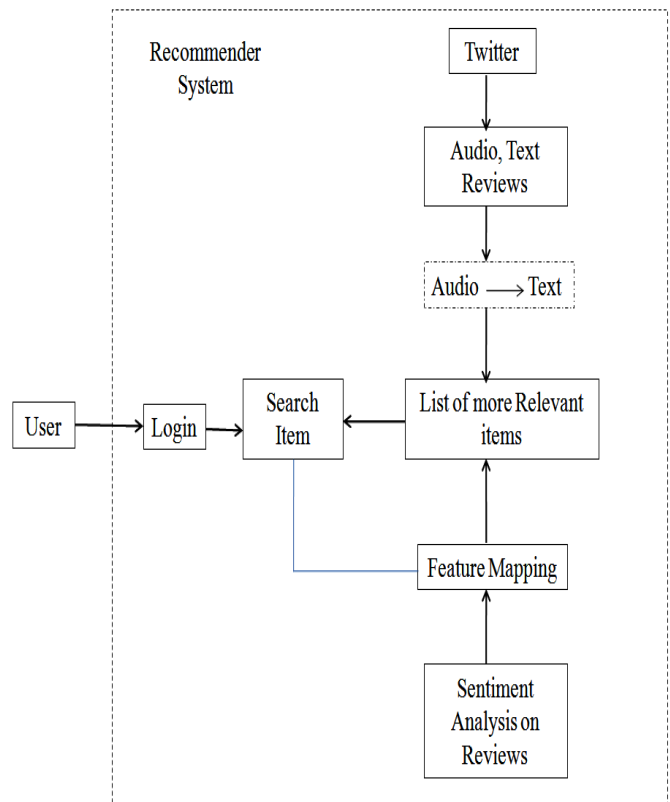

Figure 2. Proposed system architecture

VI. COMPARISON

Following table shows comparison between existing systems and our proposed system.

Table 1: Existing systems Vs Proposed system

Sr. No.	Existing System	Proposed System
1.	Uses product's ratings	Microblogging information – Reviews
2.	Consider only text reviews for sentimental classification	Consider both text as well as audio reviews about product
3.	Focus on increasing accuracy	Focus on increasing accuracy
4.	Cold-start situation is not considered	Cold-start situation is addressed
5.	May include locations travelled by users	User's location trajectory is not considered

VII. CONCLUSION

A novel problem of cross website Cold-Start product recommendation occurs when user is allowed to sign in E-

Commerce websites with Social Media account and search an item on E-Commerce is addressed. By considering text as well as audio reviews of product for recommendation accuracy of recommender system is improved. Hence, incorporating social media information the cross website Cold-Start product recommendation is alleviated. Future scope may involve improving speed of recommender system, in proposed system focus is on improving the accuracy of product recommendation is made.

ACKNOWLEDGMENT

REFERENCES

[1] Ricci, Francesco, Lior Rokach, and Bracha Shapira. Introduction to recommender systems handbook. springer US, 2011.

[2] Ma, Hao, et al. "Improving recommender systems by incorporating social contextual information." ACM Transactions on Information Systems (TOIS) 29.2 (2011): 9.

[3] Bobadilla, Jesús, et al. "Recommender systems survey." Knowledge-based systems 46 (2013): 109-132.

[4] Sandanalakshmi, R., et al. "Speaker Independent Continuous Speech to Text Converter for Mobile Application." arXiv preprint arXiv:1307.5736 (2013).

[5] Hollerit, Bernd, Mark Kröll, and Markus Strohmaier. "Towards linking buyers and sellers: detecting commercial intent on twitter." Proceedings of the 22nd International Conference on World Wide Web. ACM, 2013.

[6] Kaushik, Lakshmish, Abhijeet Sangwan, and John HL Hansen. "Sentiment extraction from natural audio streams." Acoustics, speech and signal processing (icassp), 2013 ieee international conference on. IEEE, 2013.

[7] Bahrainian, Seyed-Ali, and Andreas Dengel. "Sentiment analysis and summarization of twitter data." Computational Science and Engineering (CSE), 2013 IEEE 16th International Conference on. IEEE, 2013.

[8] Chinsha, T. C., and Shibily Joseph. "A syntactic approach for aspect based opinion mining." Semantic Computing (ICSC), 2015 IEEE International Conference on. IEEE, 2015.

[9] Zhao, Wayne Xin, et al. "Connecting social media to e-commerce: cold-start product recommendation using microblogging information." IEEE Transactions on Knowledge and Data Engineering 28.5 (2016): 1147-1159.

[10] Habayeb, Mayy, et al. "A novel point of interest (POI) location based recommender system utilizing user location and web interactions." Big Data Computing Service and Applications (BigDataService), 2016 IEEE Second International Conference on. IEEE, 2016.

[11] Pasarate Sneha, and Rajashree Shedge, "Comparative study of feature extraction techniques used in sentiment analysis." Innovation and Challenges in Cyber Security (ICICCS-INBUSH), 2016 International Conference on. IEEE, 2016.

[12] Gao, Zhipeng, Zhixing Li, and Kun Niu. "Solutions for problems of existing E-Commerce recommendation system." Consumer Electronics-Taiwan (ICCE-TW), 2016 IEEE International Conference on. IEEE, 2016.