International Journal of Applied Engineering Research ISSN 0973-4562 Volume 12, Number 17 (2017) pp. 6670-6677

A Collaborative filtering based sentiment analyzer to evaluate textual user feedbacks /opinions

Alok Kumar
Assistant Professor, Department of Computer Science and Engineering,
University Institute of Engineering and Technology, Chhatrapati Shahu Ji Maharaj University, Kanpur-208024, Uttar Pradesh, India.

Orcid Id: 0000-0003-1051-0672

Renu Jain
Associate Professor, Department of Computer Science and Engineering,
University Institute of Engineering and Technology, Chhatrapati Shahu Ji Maharaj University, Kanpur-208024, Uttar Pradesh, India.

Orcid Id: 0000-0002-4077-3345

Abstract

We propose a system to automatically analyze the textual feedbacks of products or services. User-to-User and Aspect-to-Aspect based collaborative filtering is used to estimate the sentiment scores of those aspects of a feedback which are not cited in the feedback but are important aspects. The proposed system is able to perform a more reliable sentiment analysis of all necessary aspects of products/services/movies/books, etc., and on the basis of the reports generated by the system, two products or two services can be compared. It allows the users to make quick decisions about the product they want to procure. We have implemented the system to process the feedbacks of the product "laptop" but the system can be used for other domains easily.

Keywords: Sentiment Analysis, Collaborative Filtering, Feedback Analysis, Opinion Mining, Natural Language Processing.

INTRODUCTION

While buying a product or service, sometimes the process of decision making becomes very tiresome and confusing. User always looks for feedbacks so that he/she can get the real facts about the product which he/she wants to purchase. Mostly the final decision is influenced by the opinions of friends, co-workers and family members. But in today's scenario, it is very easy to examine the reviews of many users online and make a more reliable decision. Online reviews provide a wide platform to customers. However, it is not easy to extract the required information from textual feedbacks posted online by many users. The expectations of every user are different and everyone has its own writing style making it difficult for the customers to comprehend his/her sentiments. Therefore, analyzing feedbacks is not an easy task. To generate finite conclusions from a huge collection of textual data is a very cumbersome and tedious job even for human beings. On the basis of the sentiment analyzer results, customers can choose appropriate item or service that can fulfill their requirements and manufacturer can also use these feedbacks for the betterment of products and services.

Existing sentiment analyzers perform aspect based sentiment analysis on each sentence of the textual feedback. Sentences are classified as subjective and objective on the basis of their linguistic properties. For each subjective sentence, system identifies the aspects mentioned in the sentence with their sentiments; objective sentences are ignored by the system. After processing all the textual data of feedbacks, a summary of all cited aspects with their sentiments is generated as output to assist the customers/manufactures etc.

In this paper, we have incorporated collaborative filtering with the sentiment analyzer to get more consistent sentiment scores for each aspect and a better sentiment aggregation score. This also allows us to compare the feedbacks of two brands of the same product. We have divided our paper in five sections i.e. role of collaborative filtering, related works, architecture of proposed system and finally results and conclusions.

ROLE OF COLLABORATIVE FILTERING IN SENTIMENT ANALYZERS

Sentiment analyzer collects feedbacks from different platforms like social networking sites, blogs, etc., where users share their opinion in the form of running text about the products / services they have used. There is no specific format in which users share their opinion. Users always like to express their opinion with no restrictions and limitations and this leads to collection of unsystematic data. Despite having same features of the product, users give feedback as per their usefulness and utility of the product and the information acquired about that product. Due to this, it becomes difficult for the customer to compare two feedbacks. Different aspects can have different significance for different users, one aspect may be important for one user while same aspect may not be important for the other user. Generally, we human beings have the tendency to comment about only those features which have either very good or very bad performance and we forget to comment about the features having average performance. In addition to this, users tend to comment more about modern or newly publicized features by company instead of common but important features. Due to this, existing sentiment analyzers face following limitations.

1. **Two brands of a product cannot be compared**: If two users give their opinion about two brands of same product but they do not mention all the features in their feedback then two feedbacks cannot be compared. This means: the sentiment analyzer system cannot prepare the comparison matrix of two brands of a product or we can say that two brands of a product cannot be compared.
2. **Two feedbacks cannot be assimilated**: Let us take an example: suppose one user comments on the 'touch screen', 'battery life' and the 'screen size' of mobile phone and other user comments on the 'operating system', 'weight' and the 'battery life' of mobile phone. In this case, if combined feedback of each feature is evaluated on the basis of all feedbacks then no data is entered for a non-cited feature. This way the calculated overall feedback of each feature fails to take into account the opinion of all users and may not represent the genuine sentiment value. The reason is that many users do not like to give their opinions about common features or about those features which have mediocre performance.
3. **Unbiased aspect sentiment aggregation:** It is a known fact that few customers just want to know the overall rating of a product without going in detail about each feature. Sentiment aggregation gives us overall picture of the performance of a product or a service. Mostly sentiment analyzers evaluate sentiment aggregation on the basis of the actual information available for an aspect in the feedback. In analyzing the overall performance of a product or a service, feedback on every aspect is equally important but as stated above every user does not cite each feature and this leads to dissimilar data. Due to this heterogeneity in data, results may not be unbiased and may not provide a realistic sentiment score.

To overcome the above mentioned problems, we have proposed an improved sentiment analyzer system which uses collaborative filtering to estimate the sentiments of missing features in a user feedback. This system either directly extracts sentiments of a user about each important aspect of a product from its feedback or estimates it using collaborative filtering if not mentioned in the feedback. It uses User-to-User and Aspect-to-Aspect collaborative filtering to estimate sentiment of missing features. The proposed system is more flexible and versatile than the existing systems. The proposed system performs fine grained analysis on the collected feedbacks and produces more consistent results.

RELATED WORK

There are many methods proposed by the researchers to perform aspect based sentiment analysis of textual opinion documents. Broadly proposed methods can be divided in three categories: Rule based approaches, Supervised approaches and Unsupervised approaches. Few important researches are given below.

Hu and Liu [1] proposed an association rule mining based method to extract aspects in opinion documents. In this method authors first determine list of possible nouns in opinion documents with intuition that aspects are nouns. Only those nouns are considered valid aspects which are frequently occurred in opinion document. Chin-Sheng Yang et al [2] proposed a rule based system to perform aspect based sentiment analysis on customers' reviews. Proposed system learns rules to extract aspect of items, and then opinions are extracted from sentences and then in the last phase opinion orientation is identified.

Syntactic relations help in identifications of aspects. Syntactic relations between aspects and sentiment words can be captured by using dependency parser. Identified relations are treated as templates; new aspects can be extracted with the help of syntactic relation templates. Zhuang et al[3] and Qiu et al[4] proposed a method based on syntactic structure to extract aspects and sentiments pairs in reviews. A. Kumar and R. Jain [5] built a sentiment analyzer to process textual feedbacks of teachers. This system analyzes teachers' feedback in three steps: identify important aspect of teachers, perform subjectivity analysis of sentences, compute sentiment polarity of identified aspects in subjective sentences and finally calculate overall performance of each teacher on each aspect on a predefined scale.

Jin et al[6] proposed a lexicalized HMM based system to identify aspect terms and sentiment terms in review sentences. Authors consider aspect and sentiment word identification problem as a sequence labeling problem. Some linguistic features like part of speech tagging (POS tagging) and lexical patterns are incorporated in lexicalized HMM model. To enhance performance of HMM based method, Jakob and Gurevych [7] proposed a conditional random field (CRF) based method. Sequence independent and long distance linguistic functions are used to extract opinion targets (aspects) from subjective sentences.

Lin and He[8] proposed a LDA based joint topic-sentiment model to extract aspects and sentiment words in reviews. The system extracts opinion words and sentiment words simultaneously. Prior linguistic language knowledge is incorporated to enhance system performance. A clustering based method was proposed by Masashi Hadano et al [9] to identify aspects in subjective sentences. Initially all sentences are divided in some clusters with the assumption that sentences related to same aspects must be in the same cluster. All sentences in a cluster are labeled with same aspect term associated with sentence in centre of cluster and support Vector Machine learning method is used to classify sentiment sentences. Rishabh Soni, K. James Mathai[10] suggested a method to perform sentiment analysis of tweets on products. Tweets on some products are divided into K clustered using K-means clustering method. Highly dominating terms in each cluster are considered as aspects of product and Classification and Regression Trees (CART) are used to identify the sentiment orientation of tweets.

Item based and user based recommender system was proposed by Yao, G. and Cai, L. [11] using user to user and item to item collaborative filtering. In this method, system first identifies similar items and then identified similar items are recommended to selected users. Recommended users are selected on the basis of sentiment rating given by the users for different items. Song Jie Gong [12] proposed a collaborative filtering recommendation algorithm based on user clustering and item clustering. Initially similar items are grouped into clusters and then similarity of any item is computed with respect to cluster head item instead of all items to reduce computation time.

PROPOSED SYSTEM
(System architecture and implementation details)

We propose a system to perform aspect based sentiment analysis on textual feedbacks integrated with collaborative filtering. Our system first extracts all the important aspects from the collection of feedbacks known as essential aspects of a product and then after identifying essential aspects, the system evaluates each feedback to determine sentiment scores of all cited and non cited aspects. Sentiment score of cited essential aspects is computed with the help of contextual words present in the sentences of the feedback. To estimate sentiment score of not cited aspects, we have used User-to-User and Aspect-to-Aspect and Hybrid (combination of User-to-User and Aspect-to-Aspect) collaborative filtering.

There are overall five main modules of the proposed system namely ***Feedback Collection and Cleaning***, ***Aspect Identification***, ***Aspect Sentiment Score Generator***, ***Aspect Sentiment Score Aggregator*** *and* ***Sentiment Report Generator***. The block diagram of the system is given below in figure 1. Details of each module are explained in subsequent sub sections. We have implemented the system for the product 'Laptop' and the data has been taken from a public repository 'SemEval-2014'[13].The collected feedback data contains total 3048 review sentences.

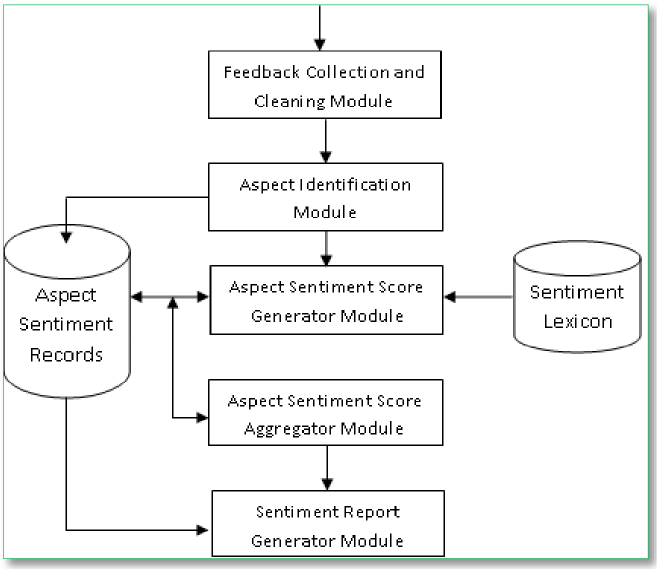

Figure 1: Block Diagram of Proposed System

Feedback Collection and Cleaning Module

Feedback collection module collects all feedbacks of a specified product or a service from all the identified sources and Feedback Cleaning module is used to perform pre-processing of the collected text. Processing of the text data involves removal of unwanted characters, removal of stop words, spelling correction of misspelled words and text normalization. To implement this module, we have used Natural Language Tool Kit (NLTK) libraries [14].

Aspect Identification Module

This module extracts the set of essential aspects for a particular product or a service. Input to this module is the output of first module i.e. pre-processed feedbacks and syntactic information of related words. This module works in two phases, in the first phase, system generates the list of all possible aspects (words / phrases) and in the second phase irrelevant words/ phrases are removed to get the more refined list of possible aspects. The authors have applied the semi supervised approach used in their paper [5] to get the essential features in this implementation.

Aspect Sentiment Score Generator

This module assigns sentiment scores to all the essential aspects identified by the aspect identification module. We have divided this module into two sub modules: Cited Aspect Sentiment Scorer and Non-Cited Aspect Sentiment Scorer. It takes as input –the complete list of essential aspects, textual feedback user wise and sentiment lexicon. The list of aspects is generated on the basis of all the feedbacks gathered from all the users. But, the feedback received from a single user may not have all the aspects present in the list. Therefore, for each user feedback, Cited Aspect Sentiment Scorer computes sentiment scores of those aspects which are cited in the feedback and Non-Cited Aspect Sentiment Scorer computes sentiment scores of those aspects which are not cited in that feedback. The functionalities of Cited Aspect Sentiment Scorer and Non-Cited Aspect Sentiment Scorer modules are explained in the following sub sections.

Cited aspect sentiment scorer

Cited aspect sentiment scorer processes each user's feedback one by one. From each feedback, system first removes all the objective sentences and then looks for the aspects cited by that user from the list of essential aspects. Sentiment score of each cited aspect is evaluated on the basis of linguistic properties of the contextual words present in different sentences in which cited aspect is present. Sentiment values are classified as average, above average, below average and a sentiment score '0','+1'or '-1'is assigned to each cited aspect. Before the final assignment of sentiment score, the scores are passed to sentiment shifter function which may invert the sentiment score on the basis of the semantic properties of sentiment shifter words. Overall sentiment score of each cited aspect is computed by aggregating sentiment scores of all the sentences in a feedback in which that aspect was present and the value is stored in User Aspect Sentiment vector. Sentiment scores of all non cited aspects are recorded as undefined. Flow chart of Cited Aspect Sentiment Scorer is shown in figure 2.

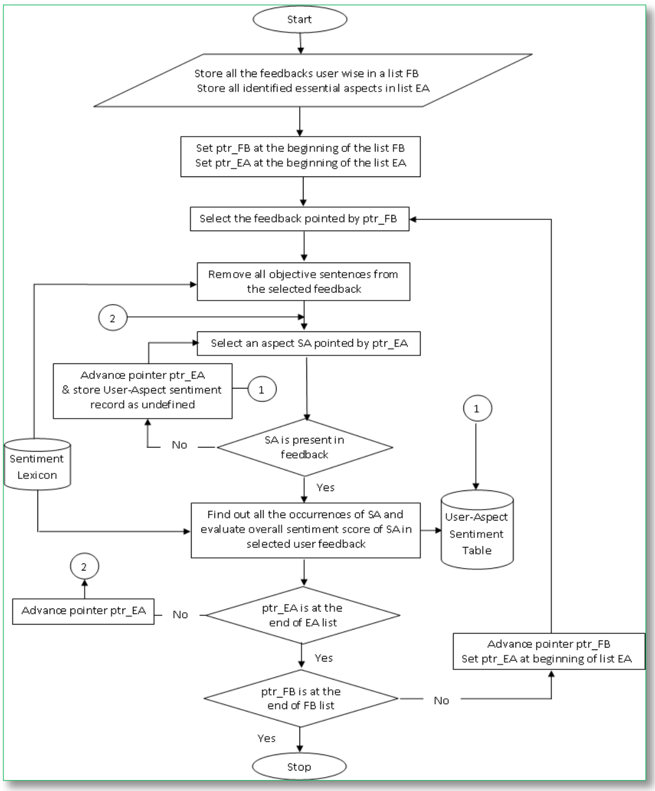

Figure 2: Flow chart of Cited Aspect Sentiment Scorer

The output of Cited Aspect Sentiment Scorer is a table where aspects/features represent columns, users represent rows and the values of the table represent the sentiment scores. A value representing undefined ('?') is stored in the table for those features which were not cited in the feedback. A sample output table generated from Cited Aspect Sentiment Scorer will be as same as shown below in table 1.

Table 1: Output of Cited Aspect Sentiment Scorer

	Aspect-1	**Aspect-2**	**Aspect-3**	**….**	**Aspect-m**
User-1	**+1**	**-1**	**+1**	**….**	**?**
User-2	**+1**	**+1**	**?**	**….**	**-1**
User-3	**-1**	**-1**	**+1**	**….**	**+1**
User-4	**?**	**+1**	**?**	**….**	**+1**
….	**…..**	**…..**	**…..**	**…..**	**…..**
User-n	**+1**	**?**	**-1**	**…..**	**-1**

Non-Cited Aspect Sentiment Scorer

This module estimates sentiment scores of aspects which are not cited by a user in his/her feedback. Input to this module is the User Aspect Sentiment table generated by Cited Aspect Sentiment Scorer and output of this module is a new User Aspect Sentiment table having no 'undefined' entries. The 'undefined' entries are substituted by estimated sentiment scores calculated using collaborative filtering. Sentiment scores of non cited aspects are computed by implementing collaborative filtering using three different methods.

In the first method, sentiment score of missing aspect is computed with the help of sentiment scores given by similar users (similar users are identified on the basis of their feedback for other aspects/features).

Second method is based on aspect similarity. On the basis of some defined criteria, similar aspects are chosen and on the basis of sentiment scores of similar aspects, sentiment score of undefined aspect is computed.

In the third method, we compute sentiment scores of non-cited aspects by taking average of the computed sentiment scores from method 1 and method 2. Formal algorithm for calculating the sentiments scores of missing features using User-to-User collaborative filtering is given below.

Input : User Aspect Sentiment table with 'undefined entries'
Output: User Aspect Sentiment table without 'undefined' entries.

Step 1: *Read the table row and column wise and find an entry marked as 'undefined'.*
Step2: *Store the row and column location of that entry (Assume it is at (I, J) location in the table).*
Step 3: *Identify all the users similar to User I on the basis of the sentiment scores assigned for other aspects and who have cited aspect J.*
Step 4: *Store the locations of closest users which fall within a certain threshold.*
Step5. *Evaluate the estimated value of (I,J) entry by taking the weighted average of sentiment scores given by the closest users for the J^{th} aspect.*
Step6: *Repeat steps 1 to 5 till all the 'undefined' entries of the table are processed.*

With the help of collaborative filtering we estimate the sentiment value of a feature or aspect (say 'J') about which a user has not stated his/her opinion but it is an essential aspect of the product. In User-to-User collaborative filtering, we first find those users who have cited the aspect 'J' and then calculate similarity extent value for those. Degree of similarity between two users is computed with the help of cosine similarity. In estimation of cosine similarity, each user is represented as a vector and angle between two vectors is computed. The angle between two user vectors demonstrates the similarity extent between the two users.

The cosine similarity between two users 'U_1' and 'U_2' is evaluated as:

$$\mathrm{Sim}(U_1,U_2)=\frac{U_1.U_2}{|U_1|\,|U_2|}=\frac{\sum_{i=1}^{m} S_{1i}.S_{2i}}{\sqrt{\sum_{i=1}^{m}(S_{1i})^2}\sqrt{\sum_{i=1}^{m}(S_{2i})^2}} \qquad \text{----(1)}$$

In equation 1, U_1 and U_2 are the vectors for user-1 and user-2 and S_{1i} and S_{2i} are the sentiment scores of i^{th} aspect with respect to user-1 and user-2 respectively. After identifying K similar users, system estimates sentiment score of a particular non cited aspect by arithmetic mean and weighted arithmetic mean of sentiment scores of K similar users.

Hence, using arithmetic mean, the sentiment value of j^{th} aspect with respect to i^{th} user is computed as:

$$S_{ij}= \frac{1}{K}\left(\sum_{x\in n} S_{xj} \right) \qquad \text{-----(2)}$$

Weighted arithmetic mean sentiment score of j^{th} aspect with respect to i^{th} user is computed as:

$$S_{ij}=\sum_{x\in n}(\mathrm{Sim}(i,x) * S_{xj})/\sum_{x\in n}\mathrm{Sim}(i,x) \qquad \text{----(3)}$$

In equation 3, S_{ij} is sentiment score of j^{th} aspect in feedback of i^{th} user, $\mathrm{Sim}(i,x)$ is similarity between i^{th} and x^{th} users/feedbacks of a product.

Similarly Aspect-to-Aspect collaborative filtering is used to calculate the estimated value of sentiment of missing features. We have also computed sentiment scores of not cited aspects using combination of User-to-User and Aspect-to-Aspect collaborative filtering. Results obtained from three different methods are discussed in Results and Conclusion section.

Aspect Sentiment Score Aggregator

This module evaluates absolute sentiment score and relative sentiment score about an aspect/ feature of a product. Absolute sentiment score of an aspect of a product represents overall sentiment orientation of that aspect. It is evaluated by taking arithmetic mean of all sentiment scores of that particular aspect.

$$\text{Abs\_senti (product, aspect)} = \frac{\sum_{U\in Users} Senti\_score(U,product,aspects)}{Number\ of\ feedbacks} \qquad \text{---- (4)}$$

Here, Abs_senti (product, aspects) represents sentiment score of a particular aspect of a particular product in a feedback given by the user '*U*'.

Relative sentiment score (Rel_senti (product, aspect)) signifies the overall sentiment orientation of an aspect. It is calculated by taking the logarithmic ratio of positive orientation and negative orientation. Positive orientation takes into account positive and neutral sentiments and negative orientation takes into account negative and neutral sentiments.

$$\text{Rel\_senti (product, aspect)} = \log\left(\frac{\text{POS (product,aspect)+NEUT (product,aspect)}}{\text{NEG (product,aspect)+NEUT (product,aspect)}}\right) \qquad \text{---(5)}$$

Where, POS (product, aspect), NEG (product, aspect) and NEUT (product, aspect) represent the number of users having positive, negative and neutral sentiments for particular aspect of a product respectively. Absolute and relative sentiment aggregations of essential aspects are shown in figure 3 and figure 4 respectively.

International Journal of Applied Engineering Research ISSN 0973-4562 Volume 12, Number 17 (2017) pp. 6670-6677

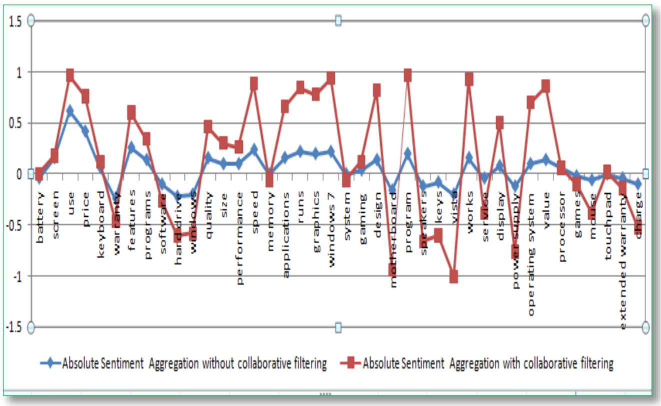

Figure 3: Absolute sentiment aggregation with and without collaborative filtering

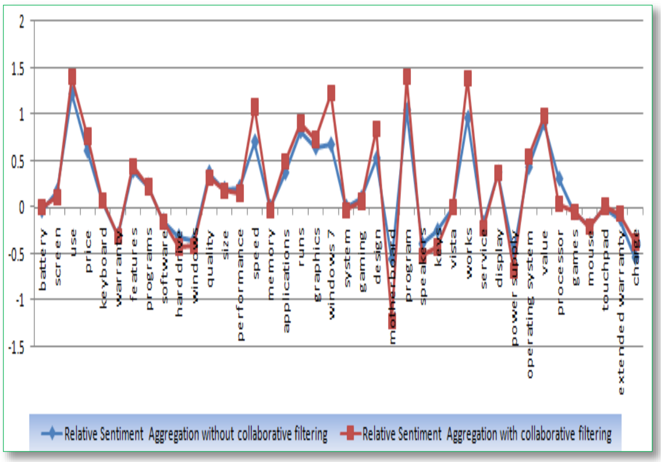

Figure 4: Relative sentiment aggregation with and without collaborative filtering

Sentiment Report Generator

This module is responsible for generating different types of graphical reports, displaying variations among users and their sentiments. With this module, we can compare two products with respect to selected aspects. This module is also helpful in identifying the strong and weak aspects of any product. Strong and weak aspects of any products are those aspects which are mostly and least commented by the users. Some reports generated by this module are discussed in next section.

RESULTS AND CONCLUSION

With our proposed system, we can perform fine grained analysis on textual feedbacks of products or services. It can determine users' likes and dislikes on any aspect of the product. System is able to discover weaknesses and strengths of product or service. We have implemented the system to analyze textual feedbacks of laptops. Feedback of Laptop is taken from public repository ('SemEval-2014' [13]). We extracted the essential aspects on the basis of frequency and linguistic information of words. We have considered only those aspects essential which are mentioned in ten percent feedbacks at least, forty such essential aspects are identified by our system. With the help of results obtained as shown in figure 5, we are able to categorize the aspects as highly important, important and average aspects.

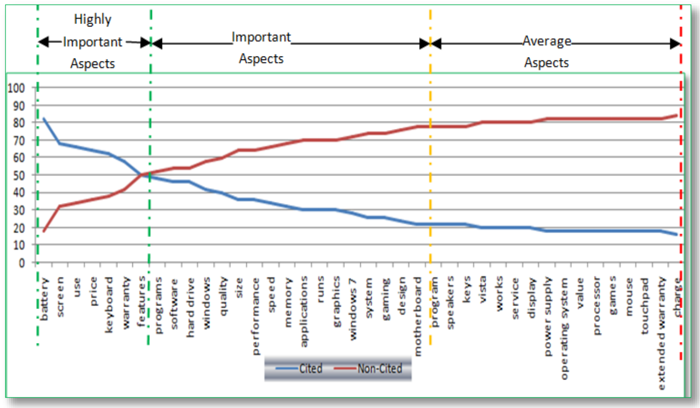

Figure 5: Cited versus Non-Cited Aspects (Categorization of Aspects)

Essential aspects may have positive, negative and neutral sentiment. Percentage of positive, negative and neutral sentiments of all essential aspects without using collaborative filtering are shown in figure 6. We have used user-to-user and aspect-to-aspect collaborative filtering to determine sentiment score of non cited aspects in each feedback. Percentage of positive, negative and neutral sentiments of all essential aspects after incorporating collaborative filtering are shown in figure 7. The comparison of two figures shows that we are able to get more accurate sentiment score of each aspect after taking into account the sentiment of all the users.

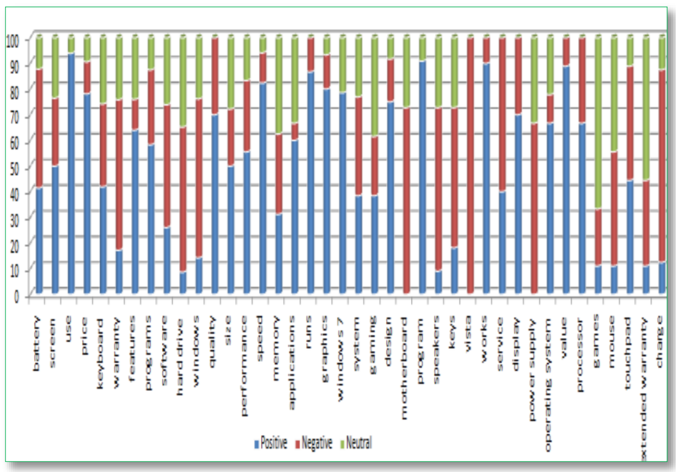

Figure 6: Sentiment orientation of all essential aspects without collabrative filtering

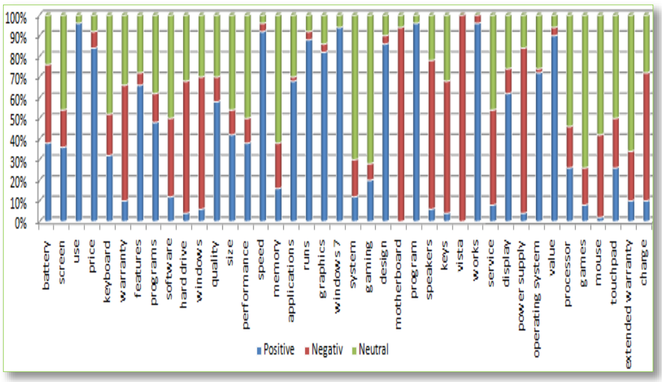

Figure 7: Sentiment orientation of all essential aspects after incorporting collabrative filtering

We have found that sentiment analysis of feedbacks using collaborative filtering allows us to compare one product /service with other products/services in the same domain. Such comparisons help the customers to take quick decision while selecting a product or a service as per their requirements. Sentiment comparison of two laptops i.e. ‘Product 1’ and ‘product 2’ with respect to some essential aspects is shown in figure 8. From figure 8, we can conclude that product 1 is better than product 2 in respect of ‘battery’, ‘screen’, ‘warranty’ and ‘speed’ while product 2 is better in ‘price’, ‘keyboard’, ‘memory’ and ‘software’. Such results will be helpful for new customers to select appropriate product as per their requirements as well as for product manufacturers to enhance quality of the products.

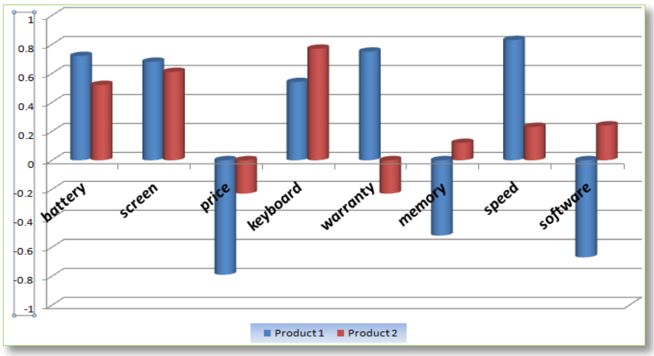

Figure 8: Sentiment compression of two products with respect to some essential aspects

A test data of nominal size was taken to measure the accuracy of the implemented system. The tables 2, 3 and 4 show the accuracy for different methods. The first column of a table represents the number of similar users or aspects (K) taken into account to evaluate the value of sentiment scores of non-cited aspects. The second column shows the accuracy if simple average is taken and column 3 shows the accuracy if weighted average is taken.

It can be observed from the data that initially as the value of K increases, accuracy increases but after an optimum value of K =20, the accuracy starts decreasing. It is also observed that the accuracy of the hybrid collaborative filtering (combination of user-to-user and aspect-to-aspect) is better.

Table 2: Accuracy of User-to-User collaborative Filtering

User-to-User	**Average Accuracy**	
	Average of K Similar Sentiments (Method-1)	**Weighted average of K Similar Sentiments (Method-2)**
K=5	56	58.43
K=10	67.56	71.78
K=15	73.24	74.15
K=20	85.43	89.57
K=25	82.87	84.56
K=30	82.58	84.78
K=35	69.78	72.21

Table 3: Accuracy of Aspect-to-Aspect collaborative Filtering

Aspect-to-Aspect	**Average Accuracy**	
	Average of K Sentiments (Method-3)	**Weighted average of K Similar Sentiments (Method-4)**
K=5	56.87	59.32
K=10	72.46	73.58
K=15	79.34	82.68
K=20	86.9	88.38
K=25	84.56	86.96
K=30	85.38	84.23
K=35	81.67	82.73

Table 4: Accuracy of Hybrid collaborative Filtering

Hybrid	**Average Accuracy**	
	Average of K Sentiments (Method-5)	**Weighted average of K Similar Sentiments (Method-6)**
K=5	54.43	57.72
K=10	74.13	75.27
K=15	81.54	83.32
K=20	89.92	91.16
K=25	88.79	88.99
K=30	86.48	87.69
K=35	79.87	81.92

REFERENCES

[1] Hu, Minqing and Bing Liu. Mining and summarizing customer reviews. In Proceedings of ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD-2004). 2004.

[2] Chin-Sheng Yang and Hsiao-Ping Shih. A Rule-Based Approach for Effective Sentiment Analysis, Pacific Asia Conference on Information Systems - PACIS 2012 Proceedings.

[3] Zhuang, L., F. Jing, and X. Zhu. Movie review mining and summarization. In Proceedings of ACM International Conference on Information and Knowledge Management (CIKM-2006), 2006.

[4] Qiu, Guang, Bing Liu, Jiajun Bu, and Chun Chen. Expanding domain sentiment lexicon through double propagation. in Proceedings of International Joint Conference on Artificial Intelligence (IJCAI-2009). 2009.

[5] Alok Kumar, Renu Jain, Sentiment Analysis and Feedback Evaluation**,** IEEE 3rd International Conference on MOOCs, Innovation and Technology in Education (MITE), 2015.

[6] Jin, Wei and Hung Hay Ho. A novel lexicalized HMM-based learning framework for web opinion mining. in Proceedings of International Conference on Machine Learning (ICML-2009). 2009.

[7] Jakob, Niklas and Iryna Gurevych. Extracting Opinion Targets in a Single and Cross-Domain Setting with Conditional Random Fields. In Proceedings of Conference on Empirical Methods in Natural Language Processing (EMNLP-2010). 2010.

[8] Lin, C. and Y. He. Joint sentiment/topic model for sentiment analysis. In Proceedings of ACM International Conference on Information and Knowledge Management (CIKM-2009), 2009.

[9] Masashi Hadano, Kazutaka Shimada, Tsutomu Endo, Aspect Identification of Sentiment Sentences Using A Clustering Algorithm, Procedia - Social and Behavioral Sciences, Volume 27, 2011, Pages 22-31, ISSN 1877-0428.

[10] Rishabh Soni , K. James Mathai. Effective Sentiment Analysis of a Launched Product using Clustering and Decision Trees, International Journal of Innovative Research in Computer and Communication Engineering, Vol. 4, Issue 1, January 2016

[11] Yao, G., Cai, L. User-Based and Item-Based Collaborative Filtering Recommendation Algorithms Design. University of California, San Diego.

[12] Song Jie Gong. A Collaborative Filtering Recommendation Algorithm Based on User Clustering and Item Clustering, Journal of software, Vol 5, No. 7, July 2010

[13] SemEval2014:

a. http://alt.qcri.org/semeval2014/task4/index.php?id=data-and-tools or http://metashare.ilsp.gr:8080/repository/browse/semeval-2014-absa-laptop-reviews-train-data

[14] Natural Language Tool Kit (NLTK): http://www.nltk.org

[15] Liu, Bing. Sentiment Analysis and Subjectivity, in Handbook of Natural Language Processing, Second Edition, N. Indurkhya and F.J. Damerau, Editors. 2010.

[16] Bing Liu. Sentiment Analysis and Opinion Mining, Morgan & Claypool Publishers, May 2012.