IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Volume-6, Issue-1 E-ISSN: 2347-2693

Feature level intentions based product recommendations with case-based reasoning

D. Teja Santosh* , K.C. Ravi Kumar, P. Chiranjeevi

[1*]CSE, GITAM (Deemed to be University), Rudraram(V), India
[2]CSE, Sridevi Womens Engg. College, Gopanapally, Hyd., India
[3]CSE, ACE Engineering. College, Ankushapur, Hyd, India

**Corresponding Author: tejasantoshd@gmail.com*

Available online at: www.ijcseonline.org

Received: 06/Jan/2018, Revised: 10/Jan/2018, Accepted: 25/Jan/2018, Published: 31/Jan/2018

Abstract: Crucial data like product features and opinions are mined from online reviews. The obtained opinions are further analyzed for orientations. These orientations that are positive, negative or neutral are counted to determine the sentiment of the feature. The product recommendations performed by using the sentiments lead to a problem called "customer churn". This is due to the tide of sentiment change. The reviewer intention on the product feature is important in finalizing the recommended list of product cases. In order to carry out this, the statistical intentions are calculated. The product cases are generated for a product by using these calculated intentions. The statistical intentions of the common features are stabilized to uncover the finalized features at the time of product similarity calculation. This "intent-to-opine" way of product recommendations addresses the problem of customer churn in the long run.

Keywords: Customer Reviews, Product Features, Sentiments, Customer Churn, Intentions, Product Recommendations, Case Based Reasoning

1. INTRODUCTION

Recommendation services have long been an important module of e-commerce platforms providing automated product suggestions that match the learned preferences of customers [1]. The product cases, which are automatically mined from product reviews, are used in the recommendation that emphasises similarity and sentiment. These recommendations depend on the sentiments of the static features which shift its value as and when the size of the reviews database scales up drastically [3] and the comparative product features are better than the preferred one [2]. The second reason leads to a problem called "customer churn." This rate of decrease in the number of customers of the product affects the productivity of the good leading to downfall of the product brand in the market. The reviewer intention on the product feature is also important in addition to sentiment in finalizing the recommended list of products.

Intention Analysis is the process of identifying intentions from text. The types of intentions are: the intention to *purchase*, *sell*, *complain*, *accuse*, *inquire*, *opine*, *advocate* or to *quit*, in incoming customer reviews [4]. Among these intentions, opine intention category is further subdivided into two sub types. They are the intention to "praise" and the intention to "criticize". These sub types provide the evaluative character of a word [5]. In the context of online reviews, the praise and criticize intentions are determined statistically by counting the number of positive opinions and the number of negative opinions on product features from the reviews. Customer Satisfaction (CSAT) [6] is the metric which is used to determine the satisfactory value of the customer on the product brand. The praise and criticize intention counts are used in CSAT and the obtained value is interpreted with the corresponding sentiment value. The CSAT and the sentiment are calculated upon the k-common features among the product cases and the query case. This "intent-to-opine" way of product recommendations addresses the problem of customer churn in the long run.

The paper is organized as follows: The related works are critically reviewed in Section 2, the proposed method is explained in Section 3, results and discussions are explained in Section 4, and finally, conclusion and future work is specified in Section 5.

2. RELATED WORK

The recommender systems (RS) are the information filtering systems which deal with the large amount of information that is dynamically generated based on users preferences, interests and observed behaviours. These traditional recommender systems fall into three categories. They are; collaborative filtering based RS, content based RS and knowledge based RS. The collaborative recommender systems are the most popular and widely implemented systems. These systems aggregate ratings from the set of users on the item and recommend it. It also identifies the users who are similar with the user from whom recommendations are to be provided. Resnick et al. developed [7] a system called GroupLens to help people to find articles they are most interested in. Anna Stavrianou and Caroline Brun developed [8] an application to recommend products based on the opinions and suggestions written in the online product reviews. The content based recommender systems learns the user profile based on the product features the user has targeted. Lang developed [9] a system called NewsWeeder which uses the words of the text as the features. Jia Zhou and Tiejian Luo developed [10] a content based recommender system that views customer shopping history to recommend the similar products based on the similarity between the product features. The knowledge based recommender systems provide the entity suggestions based on the deductions from users needs and preferences. These systems have the knowledge about how a particular product meets the customer requirement based on the factual data. The user profile is also required to provide good product recommendations to the user. Case based reasoning (CBR) is a kind of knowledge based recommender system. Kolodner used [11] CBR to recommend the restaurants based on the user's choice of features. Robin Burke used [12] the FindMe system to recommend the online products. Stefan et al. worked [13] on user log data to mine the product preferences based on the like or dislike information available in the log.

Sentiment based product recommendations have gained research importance in the recent times. The knowledge discovered in terms of product features and opinions from online product reviews among the category of products are useful to the customer in personalized recommendations. These feature level sentiments are aggregated to form the product sentiment. Li Chen and Feng Wang proposed [14] a novel explanation interface that fuses the feature sentiment information into the recommendation content. They also provided the support for multiple products comparison with respect to similarity using the common feature sentiments. Gurini et al. proposed [15] friends recommendation technique in Twitter using a novel weighting function which is called Sentiment-Volume-Objectivity (SVO) that considers both the user interests and sentiments. Xiu et al. proposed [16] a recommender system that recognizes the sentiment expressions from the reviews, quantified with the sentiment strength and appropriately recommend products according to customer needs. Recently, Dong et al. developed [17] a product recommendation strategy that combines both similarity and sentiments to suggest products.

In understanding the above body of literature, certain shortcomings were identified: The knowledge based recommender systems has the drawback of knowledge acquisition that comes in three forms namely catalog knowledge, functional knowledge and the user knowledge. The sentiment based product recommendations suffer with the problem of aggregating the opinion orientations which are susceptible to change over the timeline. The statistical intention when used in conjunction with sentiment provides better recommendations using product cases for the given query case. It also addresses the problem of customer churn in the long run.

3. RECOMMENDING PRODUCTS USING FEATURE LEVEL INTENTIONS

The principal objective of recommending products using sentiments and intentions learned from the online reviews is to utilize the extracted product features and their opinion orientations and creating the relationship between the sentiment and intention so that similarity between the query case and product case is determined solely through intentions with sentiments. In order to achieve this goal, a framework is presented in Figure 1. below.

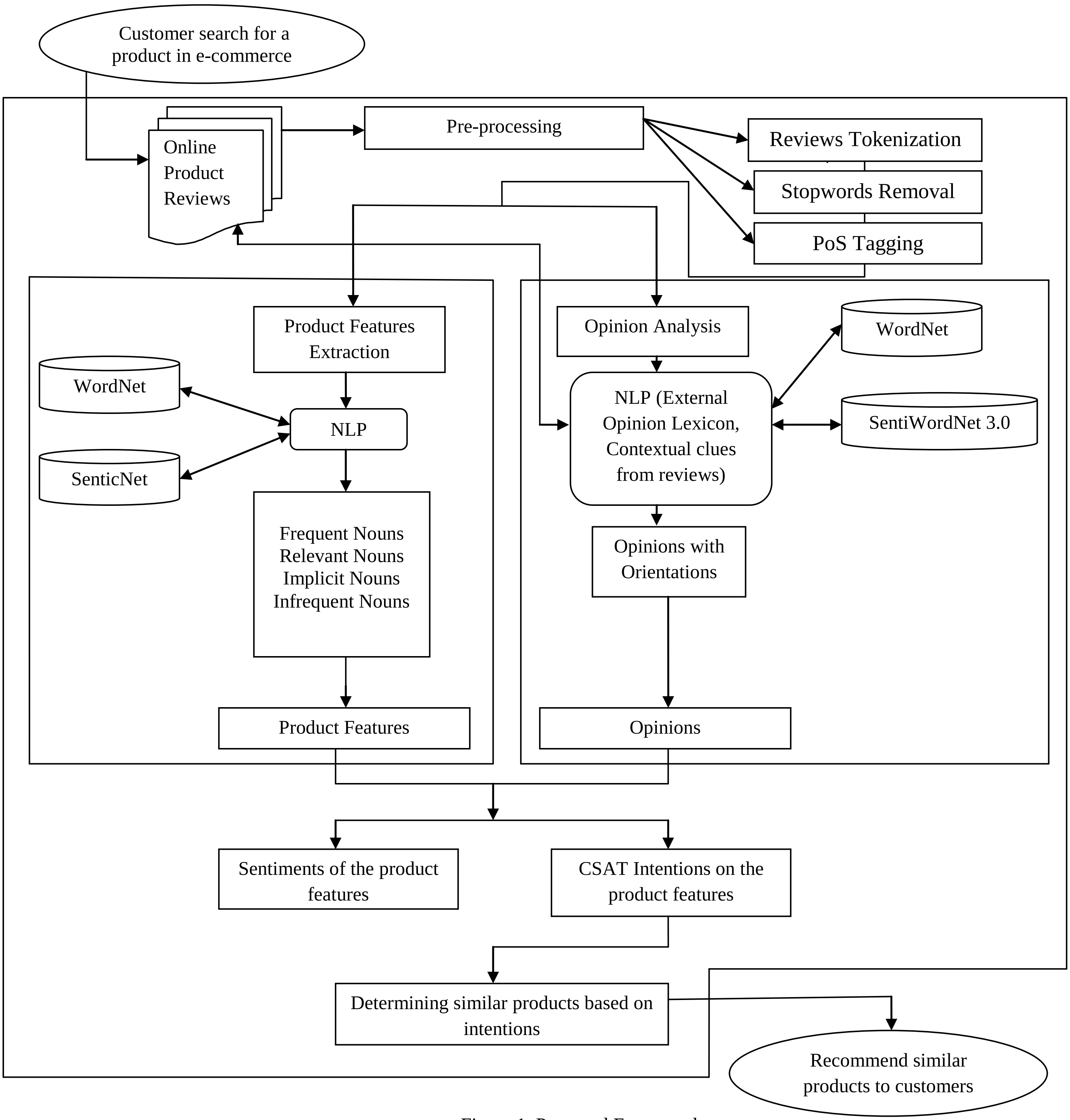

Figure 1. Proposed Framework

The framework is composed of four modules. The first being extraction of product features from the reviews based on the customer searched product and the similar class products. The second module is to extract opinions of the features which are extracted in the previous module and to determine their orientations. The third module is to calculate sentiments and intentions from the obtained opinion orientations. The product cases and the query case are generated based on the (feature, intention) pair. The last module determines the similarity between the query case and the product case based on the intentions and sentiments. The products are then recommended to the customer.

Briefly, a case for a product P is made up of a set of product features and their intention scores extracted from Reviews(P), the set of reviews written for product P. The intention of each feature is evaluated at the word level and is used at the case-level in relation with the sentiment score for that feature. At recommendation time suitable cases are retrieved and ranked based on their similarity and intention with respect to a given query case Q.

3.1 Extraction of product features from online reviews

A comprehensive product feature extraction approach from online reviews is specified in this section. This approach is based in natural language processing in which the language patterns are identified in each kind of feature extraction. This comprehensive model begins with extracting the frequent features, then finding the relevant features, and next the implicit features and finally extracting the infrequent features. The model is general and is applicable to any domain reviews collection.

Initially, the incoming product reviews are pre-processed. The steps in pre-processing are namely review tokenization, stopwords removal and Part-of-Speech (PoS) tagging. The process of review tokenization divides the sentence into individual tokens. Then, the stopwords list is applied on the tokens to remove those words which carry no meaning in the analysis. The stop words are compiled from the reviews itself. This compilation is carried out by sorting the terms in the decreasing order of collection frequency and thereby hand-filtering those terms for their semantic content relative to the domain of the product reviews. Finally, Part of Speech (PoS) tagging is carried out on the list of filtered tokens to unambiguously associate the word category with each of the token. The Stanford log-linear Part of Speech tagger is used [18] for tagging the tokens. The PoS tagger suffers with major problems [19] namely the unknown words which were not seen in the training phase of the PoS tagger, context level problems in assigning tags and the confusion state of the PoS tagger.

The step-by-step feature extraction approach is followed to reach the goal of extracting maximum number of product features. Various steps in feature extraction are namely frequent features extraction, relevant features extraction, implicit features extraction and infrequent features extraction. Nouns are extracted as product features as the research of Liu [20] confirmed that 60-70% of the features are explicit nouns. After the implementation of every step, the obtained features are added to the list of features so as to assist the count. In all the steps, WordNet is utilized [21] to finalize the extracted noun as a dictionary word.

3.1.1 Frequent Nouns Extraction

In general, a review sentence is the combination of a noun phrase and an adjective phrase. This sub module calculates the frequency count of each noun from the nouns and noun phrases which were earlier tagged by the PoS tagger. A noun is regarded as frequent if its occurrence in the reviews is within the three percent from the set of nouns that are found. The obtained frequent nouns are stored in a file and are used for further analysis.

3.1.2 Relevant Nouns Identification

The nouns which are written less in number in online reviews are relevant nouns and infrequent nouns respectively. The relevant nouns specify the associated information on the actual features of the product. A closer analysis of the reviews corpus revealed three important and interesting clues for identifying the relevant nouns. These are specifically, the nouns that are modified by multiple adjectives, the part-whole relation patterns among the product features, and the adjectives modifying the frequent nouns.

The collection of the adjectives that are available adjacent to the nouns and frequent nouns is carried out. Once these adjectives are collected, the corresponding nouns are extracted as relevant nouns. The PoS patterns that are learned for extracting the relevant nouns are given below.

word_JJ word_NN, word_JJ word_NN word_NNS

Also, the sub-features of the actual features are also extracted as relevant nouns. The obtained relevant nouns are added to the set of frequent nouns which are extracted and stored in the previous step for further analysis.

3.1.3 Implicit Nouns Identification

In some of the reviews, the product features are not written in an explicit manner. The features in such reviews are called as implicit features. The nouns pertaining to these features are called as implicit nouns. The identification of these nouns is a complex task. In order to carry out this task, the feature indicators which are present in the implicit featured reviews are identified, and with the help of SenticNet knowledgebase [22], the nouns specific to the identified feature indicators are extracted. The identification of implicit feature indicators is performed using the Conditional Random Field (CRF) [23] sequence labelling model based CRF++ framework. Similar kind of work on identifying the implicit feature indicators was carried out by Cambria et al. [24] in their work. The obtained implicit nouns are also added to the previous list of frequent nouns and relevant nouns and are used for further analysis.

3.1.4 Infrequent Nouns Extraction

As specified earlier, the infrequent nouns are also present less in number in the online reviews. These nouns are found to be interesting for certain section of customers who want to purchase the product. A noun is regarded as infrequent if its occurrence in the reviews is less than three percent from the set of nouns that are found. The obtained infrequent nouns are stored finally in the previously updated file. The updated file with all the kinds of nouns is considered as the product features.

3.2 Extraction of opinion words and determination of opinion orientations

The process is composed of following steps:

1. A standard opinion lexicon [25] in which two sets of adjectives is present is considered as input for bootstrapping. These sets are representative of the two categories namely Positive and Negative. Two seed terms 'good' and 'bad' representative of the two categories are taken into consideration.

2. The sizes of the Positive and Negative adjective sets are increased by adding the synonyms of the available adjectives using WordNet.

3. The increased sizes of Positive and Negative adjective sets is used to compare with the obtained adjectives from the dataset. Once the dataset adjectives are matched with the opinion lexicon adjectives the dataset adjectives are considered as opinion words. This completes the identification of opinion words from the dataset.

4. The opinion word and the seed terms are assigned with the sentiment scores available under adjective category from Sentiwordnet by finding the contextual clues surrounding the opinion word to disambiguate its sense. The contextual clues are finalized based on the typed dependency grammatical relations.

5. The distance between the opinion word and the seed term and the distance between the seed terms is calculated as given below.

$$\text{distance}(w_i,w_j) = \text{sentiwordnetscore}(w_i) - \text{sentiwordnetscore}(w_j)$$

where w_i is either the opinion word or the seed term and w_j is the seed term.

6. The semantic orientation (SO) of the opinion word is determined as given below.

SO(opinion word) = distance(opinion word, bad) - distance(opinion word, good)

distance(good, bad)

7. The opinion word is deemed to be positive if the orientation measurement is greater than zero, and negative otherwise.

Step 2 is based on the premise that the lexical relations used in this expansion task define a relation of orientation. It is possible that two synonyms may have same orientation and two antonyms have opposite orientation. In step 4, the basic assumption is that the terms with a similar orientation tend to have similar glossaries. The similarity or difference between the opinion word and the seed term is based on identifying the appropriate senses in the context in which the opinion word is written in the document. The senses of the seed terms change based on the context of the opinion word under analysis. The replacement of the number of synonyms in the synonymy graph with the sentiwordnet scores in step 5 enables to determine the orientation of any opinion word with the help of SO measure specified in step 6.

3.3 Evaluation of Feature Intentions and Feature Sentiments

The intentions on the product features as written by the reviewers are the opinions held by them on those features. CSAT metric is used to determine the satisfactory value of the customer on the product brand. CSAT uses the statistical intentions to do this. The feature intentions are captured using CSAT. The CSAT formula is;

$$\text{CSAT} = p / p+n$$

where p = positive opinion count on the product feature and,
n = negative opinion count on the product feature

From the CSAT formula it is well understood that customer satisfaction mainly depends on the positive opinion count (praise intention) on the product feature.
The sentiment of the product feature as formed by the reviewers is calculated as;

$$\text{Sentiment(Feature, Product)} = p - n / p + n + q$$

Where p = positive opinion count on the product feature,
n = negative opinion count on the product feature and
q = neutral opinion count on the product feature

The Sentiment(Feature, Product) returns a value between -1(negative sentiment) and +1(positive sentiment) on the product feature.

3.4 Generation of Product cases and the Query case

From each review R_i the above approaches as described in sections 3.1, 3.2, 3.3 generate a set of valid features F_1, ..., F_{mi}, the opinion orientations with their counts, the associated CSAT values and sentiment scores (positive, negative, or neutral) on the features. The product cases and the query case are constructed in a straightforward fashion, as a set of product features paired with corresponding intention scores as;

$$Case(P_i) = \{(F_j, CSAT(F_j, P_i)) : F_j \in Features(P_i)\}$$
$$Case(Q) = \{(F_j, CSAT(F_j, Q)) : F_j \in Features(Q)\}$$

where i = 1,2,3,... and j = k-common features between product cases and the query case.

Where P is the similar product for the customer searched product (product cases) and Q is the customer searched product (query case).

The case features (Features(P)) for a product P are the union of the valid features extracted from its reviews. Each of these features is present in a number of P's reviews and with different sentiment scores across the similar products.

3.5 Understanding statistical intentions with the help of sentiments towards similar product recommendations

The product recommendations that depend on the sentiments of the static features suffers with the shift in the sentiment value as and when the size of the reviews database scales up drastically and the comparative product features are better than the preferred one. The second reason leads to a problem called "customer churn." This rate of decrease in the number of customers of the product affects the productivity of the good leading to downfall of the product brand in the market. The reviewer intention on the product feature is important in finalizing the recommended list of products. The k-common features identified after the customer searched for Iphone 6s plus are tabulated below in Table 1. The value of k is found is 13.

Table 1. List of k-common features

k-common features
Phone
Rom
Battery
Performance
Os
Brand
network connectivity
Camera
Price
build quality
Touch
Screen
battery life

The sentiment scores and the CSAT values for these 13 features are tabulated in Table 2 given below.

Table 2. sentiment scores and the CSAT values of the k-common features

	Iphone 6s plus(Q)		Oppo f1 plus(P1)		Samsung galaxy j7(P2)	
Product Feature	**Sentiment Score**	**CSAT**	**Sentiment Score**	**CSAT**	**Sentiment Score**	**CSAT**
Phone	0.88	1.067	0.83	1.1	0.66	1.25
Rom	0.33	2	1	1	0.71	1.2
Battery	1	1	0.2	3	1	1
Performance	0.6	1.3333	1	1	0.5	1.5
Os	1	1	0.14	4	1	1
Brand	1	1	0.42	1.666	0.6	1.333
network connectivity	-1	0	1	1	1	1
Camera	1	1	0.23	2.6	0.77	3
Price	0.57	1.375	-0.33	-1	0.07	7
build quality	1	1	1	1	1	1
Touch	1	1	1	1	-1	0
Screen	1	1	1	1	1	1
battery life	1	1	-1	0	1	1

From the above table it is observed that the sentiment score of +1 (upper threshold) has the corresponding CSAT value as 1 and the sentiment score of -1 (lower threshold) has the corresponding CSAT value as 0. The CSAT values for those features whose sentiment values which are in between -1 and +1 have values below 1 and above 1 respectively. The sentiment values of the features in between -1 and +1 specify that shift in the sentiment value occurs over the timeline when a competing product comes with better features.

The features whose CSAT values are floating point numbers are *floored*. The features whose CSAT values are zeroes are not used in the product similarity determination process as these features do not play important role in the product recommendation. The features namely network connectivity, battery life and touch are removed from the analysis. The k-common features are now reduced to k^1-common features. The k^1-common features are found to be 10. These are tabulated in Table 3 given below.

Table 3. List of k^1-common features

k^1-common features
Phone
Rom
Battery
Performance
Os

Brand
Camera
Price
build quality
Screen

4. RESULTS AND DISCUSSIONS

The product reviews obtained from Amazon were used for this experiment. Three product reviews were considered. Iphone 6s plus, Oppo f1 plus and Samsung galaxy j7 prime smart phones were the products for which the reviews were considered for analysis. The labels specified to these three datasets were Q, P1 and P2 respectively. Table 4 presents the details of the dataset.

Table 4. Dataset details

Document attributes	Values
Number of review documents	723
Min. sentences per review	9
Max. sentences per review	15

Table 5 below presents the first available online date of the product and the reviews obtained date for the analysis. The table also contains the count of the reviews on current date.

Table 5. Products and their reviews statistical details

Product	First available online to reviews obtained date	Reviews count as on 4^{th} May 2017
Apple iphone 6s plus	16-Oct-15 to 04-May-17	165
Oppo f1 plus	02-Feb-16 to 04-May-17	171
Samsung Galaxy j7 prime	12-Sept-16 to 04-May-17	387

In order to compare the sentiments of the k-common features and the intentions of the k^1- common features of the three products for providing recommendations, Cosine Similarity [26] is considered. The variations in the number of k-common features on the similar products using sentiments are tabulated in Table 6 below.

Table 6. Cosine Similarity

k	Cosine(Q,P1)	Cosine(Q,P2)
4	0.87	0.79
8	0.45	0.38
12	0.54	0.51
13	0.29	0.48

The product similarity with the sentiments of the features is displayed in Figure 2 below.

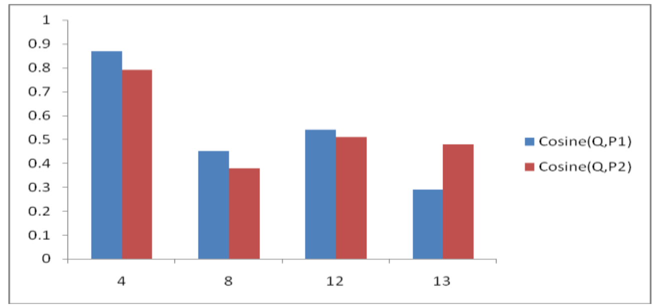

Figure 2. Sentiments based product similarities with the query case

From the results in above table it is observed that for different values of 'k' (4,8,12) the cosine similarity returned the similar products as recommendations in the same order (product P1 comes first in the list and then the product P2) by using the sentiments on k-common features. The product with higher cosine value between two similar products is shown as first product in the recommendations list. But for k value of 13, the order in the product recommendations has been changed. This is because the product P2 has higher cosine value and P1 has lower cosine value when compared with the searched product. The online comparison [27] between oppo f1 plus and Samsung galaxy j7 revealed that Samsung galaxy j7 has better ratings than oppo f1 plus. So, from the result, sentiment based product recommendations is not a better way to suggest the products to the customers.

Different values for 'k^1' provide the useful understanding about the products comparison for eventual recommendations. The variations in the number of k^1-common features on the similar products using intentions are tabulated in Table 7 below.

Table 7. Cosine Similarity with k^1 number of product fetures

k^1	Cosine(Q,P1)	Cosine(Q,P2)
3	0.73	0.94
6	0.74	0.95
10	0.69	0.64

The product similarity with the intentions in the form of CSAT is displayed in Figure 3 below.

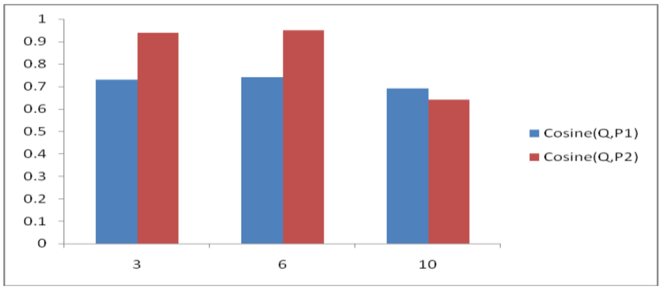

Figure 3. Intentions based product similarities with the query case

From the results in above table it is observed that for two different values of 'k[1]' (3,6) the cosine similarity returned product P2 as first in the list and then the product P1 in the list. The product with higher cosine value between two similar products is shown as first product in the recommendations list. But for k value of 10, the order in the product recommendations has been changed. This is because the product P1 has higher cosine value and P2 has lower cosine value when compared with the searched product. The results are in tune with the online comparison [27] rating values.

In order to evaluate the utility of the recommendations produced by the recommender system, Precision, Recall and F-1 score metrics are used. The formulae for precision, recall and F-1 score are given below.

$$\text{Precision} = \frac{|\text{good products recommended}|}{|\text{all recommendations}|}$$

$$\text{Recall} = \frac{|\text{good products recommended}|}{|\text{all good recommendations}|}$$

$$\text{F-1 score} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}$$

The parameters provided in Table 8 below compares the information retrieval metrics on the product recommendations between the works done by Wang and Wang [28] with the results obtained from the current work. They used user opinions that are written in online reviews as preferences to recommend the products through sentiment analysis. The sentiments calculated were based on the offline experimental setup. Whenever the size of the reviews database increases, the product recommendations are provided in the accurate manner. They used Collaborative Filtering based recommender system which is prone to cold start problem. The 'k' value is the number of users with similar preferences. In the current work, the feature level intentions were calculated in terms of customer satisfaction. These intentions purely depend on the positive opinion orientations on the product features. This kind of experimental setup is not affected with the dynamic reviews database environment. The recommendation system implemented was case based recommender model which eliminates the cold start problem. The 'k' value is the number of common product features considered for calculating the similar product recommendations.

Table 8. Recommender systems comparative information

RS type	'k' type	'k' value	Precision (%)	Recall (%)	F1-Score (%)
Opinion-enhanced CF based model through SA [28]	No. of users with similar product preferences	20	10	6	75
Intentions based RS model (Our work)	No. of common features among the similar products	10	50	100	67

The recall value from the Table 6 specify that the recommender system was able to provide better recommendations when compared with the recommendations produced by Wang and Wang [28] in their work. This shows that intentions based product recommendations are better than sentiment based product recommendations.

5. CONCLUSIONS AND FUTURE WORK

The improved product recommendations using the mined customer intentions on the product features were carried out successfully. The objectives are to improve the product recommendations using CSAT metric and to support the customer with better purchase decisions. The experimental results indicate that the proposed model is effective.

In future, the product cases retrieval is further improved by working on different feature weighting approaches.

REFERENCES

[1] Dong, Ruihai, et al. "Opinionated product recommendation." International Conference on Case-Based Reasoning. Springer Berlin Heidelberg, 2013.

[2] http://www.infosysblogs.com/data-analytics/2013/12/_my_personal_experience_of.html. Accessed on 10-04-2017.

[3] Liu, Yang, et al. "Riding the tide of sentiment change: sentiment analysis with evolving online reviews." World Wide Web 16.4 (2013): 477-496.

[4] Carlos, Cohan Sujay, and Madhulika Yalamanchi. "Intention Analysis for Sales, Marketing and Customer Service." COLING (Demos). 2012.
[5] Turney, Peter D., and Michael L. Littman. "Measuring praise and criticism: Inference of semantic orientation from association." ACM Transactions on Information Systems (TOIS) 21.4 (2003): 315-346.
[6] http://www.aiaioo.com/whitepapers/intention_analysis_use_cases.pdf. Accessed on 10-04-2017.
[7] Resnick, Paul, et al. "GroupLens: an open architecture for collaborative filtering of netnews." Proceedings of the 1994 ACM conference on Computer supported cooperative work. ACM, 1994.
[8] Stavrianou, Anna, and Caroline Brun. "Expert Recommendations Based on Opinion Mining of User-Generated Product Reviews." Computational Intelligence 31.1 (2015): 165-183.
[9] Lang, Ken. "Newsweeder: Learning to filter netnews." Proceedings of the 12th international conference on machine learning. 1995.
[10] Zhou, Jia, Tiejian Luo, and Fuxing Cheng. "Modeling learners and contents in academic-oriented recommendation framework." Dependable, Autonomic and Secure Computing (DASC), 2011 IEEE Ninth International Conference on. IEEE, 2011.
[11] Kolodner, Janet. Case-based reasoning. Morgan Kaufmann, 2014.
[12] Burke, Robin D., Kristian J. Hammond, and B. C. Yound. "The FindMe approach to assisted browsing." IEEE Expert 12.4 (1997): 32-40.
[13] Holland, Stefan, Martin Ester, and Werner Kießling. "Preference mining: A novel approach on mining user preferences for personalized applications." European Conference on Principles of Data Mining and Knowledge Discovery. Springer Berlin Heidelberg, 2003.
[14] Chen, Guanliang, and Li Chen. "Recommendation based on contextual opinions." International Conference on User Modeling, Adaptation, and Personalization. Springer International Publishing, 2014.
[15] Gurini, Davide Feltoni, et al. "Analysis of sentiment communities in online networks." Proceedings of the International Workshop on Social Personalization & Search, co-located at SIGIR. 2015.
[16] Li, Xiu, Huimin Wang, and Xinwei Yan. "Accurate recommendation based on opinion mining." Genetic and Evolutionary Computing. Springer International Publishing, 2015. 399-408.
[17] Dong, Ruihai, et al. "Combining similarity and sentiment in opinion mining for product recommendation." Journal of Intelligent Information Systems 46.2 (2016): 285-312.
[18] Toutanova, Kristina, and Christopher D. Manning. "Enriching the knowledge sources used in a maximum entropy part-of-speech tagger." Proceedings of the 2000 Joint SIGDAT conference on Empirical methods in natural language processing and very large corpora: held in conjunction with the 38th Annual Meeting of the Association for Computational Linguistics-Volume 13. Association for Computational Linguistics, 2000.
[19] Manning, Christopher. "Part-of-speech tagging from 97% to 100%: is it time for some linguistics?." Computational Linguistics and Intelligent Text Processing (2011): 171-189.
[20] Liu, Bing. Web data mining: exploring hyperlinks, contents, and usage data. Springer Science & Business Media, 2007.
[21] Christiane Fellbaum (1998, ed.) WordNet: An Electronic Lexical Database. Cambridge, MA: MIT Press.
[22] Cambria, Erik, et al. "SenticNet: A Publicly Available Semantic Resource for Opinion Mining." AAAI fall symposium: commonsense knowledge. Vol. 10. No. 0. 2010.
[23] Lafferty, John, Andrew McCallum, and Fernando Pereira. "Conditional random fields: Probabilistic models for segmenting and labeling sequence data." Proceedings of the eighteenth international conference on machine learning, ICML. Vol. 1. 2001.
[24] Poria, Soujanya, et al. "A rule-based approach to aspect extraction from product reviews." Proceedings of the second workshop on natural language processing for social media (SocialNLP). 2014.
[25] Hu, Minqing, and Bing Liu. "Mining and summarizing customer reviews." Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2004.
[26] Huang, Anna. "Similarity measures for text document clustering." Proceedings of the sixth new zealand computer science research student conference (NZCSRSC2008), Christchurch, New Zealand. 2008.
[27] http://gadgets.ndtv.com/samsung-galaxy-j7-prime-3735-vs-oppo-f1-plus-3414-vs-apple-iphone-6s-plus-2955 . Accessed on 10-04-2017.
[28] Wang, Wei, and Hongwei Wang. "Opinion-enhanced collaborative filtering for recommender systems through sentiment analysis." New Review of Hypermedia and Multimedia 21.3-4 (2015): 278-300.