RESEARCH ARTICLE

PRUS: Product Recommender System Based on User Specifications and Customers Reviews

NAVEED HUSSAIN[ID][1], HAMID TURAB MIRZA[ID][2], FAIZA IQBAL[3], AYESHA ALTAF[ID][3],
AHTSHAM SHOUKAT[ID][2], MÓNICA GRACIA VILLAR[4,5,6], EMMANUEL SORIANO FLORES[4,7,8],
MARCO ANTONIO ROJO GUTIÉRREZ[4,7,9], AND IMRAN ASHRAF[ID][10]

[1]Faculty of Information Technology and Computer Science, University of Central Punjab, Lahore 54000, Pakistan
[2]Department of Computer Science, COMSATS University Islamabad, Lahore Campus, Lahore 54000, Pakistan
[3]Department of Computer Science, University of Engineering and Technology Lahore, Lahore 54890, Pakistan
[4]Faculty of Social Science and Humanities, Universidad Europea del Atlántico, 39011 Santander, Spain
[5]Department of Extension, Universidade Internacional do Cuanza, Cuito, Bié, Angola
[6]Fundación Universitaria Internacional de Colombia, Bogotá 111311, Colombia
[7]Universidad Internacional Iberoamericana, Campeche 24560, Mexico
[8]Universidad de La Romana, La Romana 22000, Dominican Republic
[9]Universidad Internacional Iberoamericana, Arecibo 00613, Puerto Rico
[10]Department of Information and Communication Engineering, Yeungnam University, Gyeongsan 38541, Republic of Korea

Corresponding authors: Ayesha Altaf (ayesha.altaf@uet.edu.pk) and Imran Ashraf (ashrafimran@live.com)

ABSTRACT The rising popularity of online shopping has led to a steady stream of new product evaluations. Consumers benefit from these evaluations as they make purchasing decisions. Many research projects rank products using these reviews, however, most of these methodologies have ignored negative polarity while evaluating products for client needs. The main contribution of this research is the inclusion of negative polarity in the analysis of product rankings alongside positive polarity. To account for reviews that contain many sentiments and different elements, the suggested method first breaks them down into sentences. This process aids in determining the polarity of products at the phrase level by extracting elements from product evaluations. The next step is to link the polarity to the review's sentence-level features. Products are prioritized following user needs by assigning relative importance to each of the polarities. The Amazon review dataset has been used in the experimental assessments so that the efficacy of the suggested approach can be estimated. Experimental evaluation of PRUS utilizes rank score (RS) and normalized discounted cumulative gain ($nDCG$) score. Results indicate that PRUS gives independence to the user to select recommended list based on specific features with respect to positive or negative aspects of the products.

INDEX TERMS Customers reviews, product ranking, sentiment analysis, user specification, feature extraction.

I. INTRODUCTION

The e-commerce industry has witnessed rapid growth with the increasing trend of online shopping. Online sales have become dominant and customers tend to write reviews about the purchased products to express their opinions [1]. Several modes are used for the expression of views about products and brands which include social media posts, comments on portals of these products, and review-posting

The associate editor coordinating the review of this manuscript and approving it for publication was Muhammad Asif[ID].

sites over the internet [2]. Customers generally point out the advantages and disadvantages of the purchased products in the reviews. Many people rely on these opinions concerning product quality for finalizing their decision about a specific product [3].

Due to the growth in popularity of online shopping, a large number of opinions or reviews are being generated on a regular basis. These reviews are not only providing helpful insights for the companies which are planning to launch a new product but can also provide significant information for the user who intends to buy a product [4]. People often

consider buying a product with different specifications as per their needs and desire. Several similar products are offered by different brands and companies which makes it difficult for users to find the best products according to their specifications [5].

Customer reviews can be very useful for decision-making authorities at all tiers that might range from a decision regarding buying a specific product to stock price prediction in stock exchange markets. However, it is hard to extract relevant information from the huge number of reviews. This is because these reviews may contain different opinions and sentiments regarding different features of the same product. There exist several studies which have focused on summarizing user reviews to make it easy to have an idea of the user opinions. These methods include the extractive approach which selects the most important sentence from the input document, the abstractive method generates the summary, hybrid method which combines both the extractive and abstractive methods [6]. However, most of these works do not consider extracting user opinions concerning their sentiments [7].

Customer reviews consist of their opinions about a specific product which can represent varying sentiments about different features of that product [8]. For example, consider a review ''Phone has a great battery life and high-resolution camera but the screen resolution is a little low, though it was acceptable''. This review has overall positive sentiment but when you go deeper it is found that the review describes the screen resolution in a negative sense. The sentiment of a review plays an important role in decision-making; however, aspect-level sentiment is more important than the overall sentiment of the review [9]. For instance, the above review expresses user opinion about the camera, battery, and screen resolution but gives negative feedback for the screen resolution feature of the phone.

Aspect-based sentiment analysis deals with the sentiment about specific features represented in the review and generates the summary of the positive and negative polarity of the aspects [10]. To perform aspect-based sentiment analysis, the first task is to find the aspect whose sentiment is supposed to be extracted and then analyze its impact on overall opinion [11]. Several existing research studies have utilized fuzzy logic and unsupervised approach to identify aspect-level sentiment analysis and related the sentiments with aspects to find out the opinion of the user related to that particular product [12], [13], however, to identify user-specified aspects, there is a need to consider user query while extracting aspects from a large number of reviews and sort them according to positive and negative sentiments of aspects. The first widely-recognized work in this area was conducted by Ganesan et al. [14]. The authors ranked the products according to the specifications of users which are given in the form of a query. Several existing studies have ranked products according to user specifications based on reviews, however, most of these works have only highlighted positive sentiments of the aspects [15], [16]. The purpose of the current research is to propose an effective system to deal with selecting and ranking products with respect to user specifications by using the aspects considering positive and negative sentiments of these aspects. The main contributions of the study are enlisted below

- Developed the Product Recommendation based on User Specification (PRUS) framework to extract recommended product list from the review dataset based on user-specified features,
- Proposed a procedure to assign sentiments to the aspects or features by extracting the individual sentiment of an aspect from the overall sentiment of the review,
- Developed a procedure to assign weights to positive and negative sentiments counts of the aspects to generate recommended product list based on positive or the negative polarity of its respective aspects,
- Performed extensive experimental evaluation using the Amazon review dataset to estimate the effectiveness of the proposed PRUS approach.

The rest of the paper is organized as follows. Section II presents the literature review. Section III describes the dataset and elaborates on the proposed PRUS framework. Section IV demonstrates and describes the experimental evaluation of the proposed method. Finally, Section V concludes the study.

II. RELATED WORK

This section reviews existing studies related to product ranking based on user reviews. The studies are categorized based on approaches that utilized user-specified features to rank products, rule-based models to extract aspects from reviews, opinion mining, aspect-based sentiment analysis, and semantic-based aspect extraction approaches.

Several existing studies have proposed approaches for product raking based on user-specified features. A novel method for ranking an entity based on user preferences was proposed by Ganesan et al. [14]. In this work, user-desired features of the entity are given in a query which is used to rank entities based on these specific features. The proposed approach first achieved entity ranking based on their aspects, then, user query-based features are used to rank entities. Another approach in this domain is proposed by Kumar et al. [17] which achieved product ranking based on aspects of opinions. The proposed approach obtained all the aspects of entities and calculated the polarity of these aspects in reviews. It utilized Spearman's rank correlation coefficient to rank entities. Another product ranking framework was derived by Bashir et al. [18] which ranked entities by matching opinions with the user's specified features. This study utilized several ranking features which are combined by applying a learning-to-rank approach to genetic programming.

Several studies utilized a rule-based model to extract aspects from entities. A two-fold rule-based model (TF-RBM) was developed by Rana et al. [19]. In this work, rules are created based on sequential patterns derived from the opinions of different customers. In a two-fold rule-based model, one-fold extracts aspects from domain-dependent opinions whereas the other fold extracts aspects which are

domain-independent. Similarly, Wu et al. [20] proposed a rule-based approach utilizing machine learning methods. In the first step, this approach used linguistic rules to get nominal phrases. Next, it discarded unnecessary aspects which were determined based on domain correlation. Finally, the remaining aspects were used to train the recurrent unit for the extraction of aspect terms and opinion targets. Luo et al. [21] proposed a framework that used reviews to extract K most prominent aspects. Experimental evaluations were conducted on different products or services collected from Amazon, TripAdvisor, and Yelp.

The fields of opinion mining and sentiment analysis are related as both deal with extracting opinions. In opinion mining, opinions are expressed, analyzed, and summarized whereas sentiment analysis differentiates the opinion into positive or negative categories. Several existing studies have proposed approaches in the domain of opinion mining. Sharmila et al. [22] utilized machine learning and a fuzzy approach for opinion mining to extract opinions available over social media sites. In another study [15], neural networks were used to obtain aspects from opinions. The study utilized a neural network with seven hidden layers to tag words from the given opinions about entities. The evaluation is performed on two word-embedding models, namely Google embedding and Amazon embedding [23]. The research conducted by Venugopalan et al. [24] proposed an approach for domain-dependent aspect extraction. It implemented an Opinion-monitoring service and used supervised techniques for aspects-based opinion mining.

In several existing studies, aspects-based sentiment analysis is performed by considering noun phrases as aspects. An approach proposed by Hu et al. [25] has considered nouns as potential aspects. The approach calculates the frequency of each aspect. In the same direction, the study by Popescu et al. [26], calculated the frequency of each aspect by using pointwise mutual information (PMI) and achieved better accuracy. Similarly, Li et al. [27] proposed a web-based similarity method along with the frequency pruning method which improved the aspects extraction accuracy. The study conducted by Zhichao et al. [28] identified the aspects of the reviews of the Chinese language. It combined NLP with statistical methods to extract these aspects. Similarly, the study proposed by Raju et al. [29] proposed a clustering-based approach for aspect identification. The proposed approach grouped similar noun phrases into clusters which are used for aspect identification. Similarly, Güneş et al. [30] proposed a method to extract structural aspects by using frequency-based noun clustering. In another research by Quan et al. [31], the association between aspects and opinions is explored by using PMI which is used along with document frequency. The proposed approach combined local and global context information in a document for aspect extraction and ranked them based on frequency and score. Similarly, approaches proposed by [32] and [33] used bootstrapping method to extract product aspects and opinions from reviews of customers.

To identify the most frequent aspects, a semantic-based approach is proposed by Wei et al. [34] which eliminates all those aspects which are irrelevant to the seed opinion words. Another study by Ma et al. [35] utilized Latent Dirichlet allocation (LDA) which is combined with the lexicon of synonyms to extract aspects from reviews of the Chinese language. Another model for aspect identification from the word position in the sentence was proposed by Liu et al. [36]. Similarly, [37] built a list of aspects from the manufacturers' information which is used to extract aspects from reviews. The approach used syntactic patterns to find associations among aspects, opinions, and products.

Deep learning-based model is proposed by Kaur et al. [38] which utilized hybrid feature extraction approach for analyzing semantics. The approach used NLP techniques for the preprocessing stage to eliminate the undesirable data from input text reviews. The semantic-enabled Markov model recommendation (SEMMRec) system is proposed by Nasir et al. [39]. The proposed model extract products' sequential and semantic knowledge according to their usage and textual features by finding similarity between products based on their characteristics using distributional hypothesis methods [39].

It has been observed from the literature review that most of the existing studies have utilized sentiment analysis for product ranking and only a few studies considered user-specified feature-based product ranking. It can be observed that none of these studies have combined the problem of product ranking based on user-specified features considering aspect-based sentiment analysis. Most existing studies have only considered positive aspects and mostly applied overall sentiment/ranking extracted from given reviews to all aspects of that specific product. To fulfill this gap, there is a need to develop an approach that can achieve product ranking based on user-specified features by considering positive as well as negative polarities of the aspects and calculating the ranking of each aspect individually by considering respective negative or positive polarity extracted from the product's review.

III. MATERIALS AND METHODS
This section describes the dataset used for this study along with a discussion of the proposed approach and its functionality.

A. DATASET
The dataset used in this study consists of 413840 reviews and was primarily compiled by a data crawler from Amazon. The dataset contains reviews of mobile phones of various brands and companies. It consists of three attributes which include 'product title', 'brand', and 'review text'. Data preprocessing is considered the first step to preparing data by removing inconsistency and existing noise in the dataset. As a preprocessing of the dataset used in this study, those reviews were removed in which either the product title or brand name was missing. Moreover, the reviews in which the
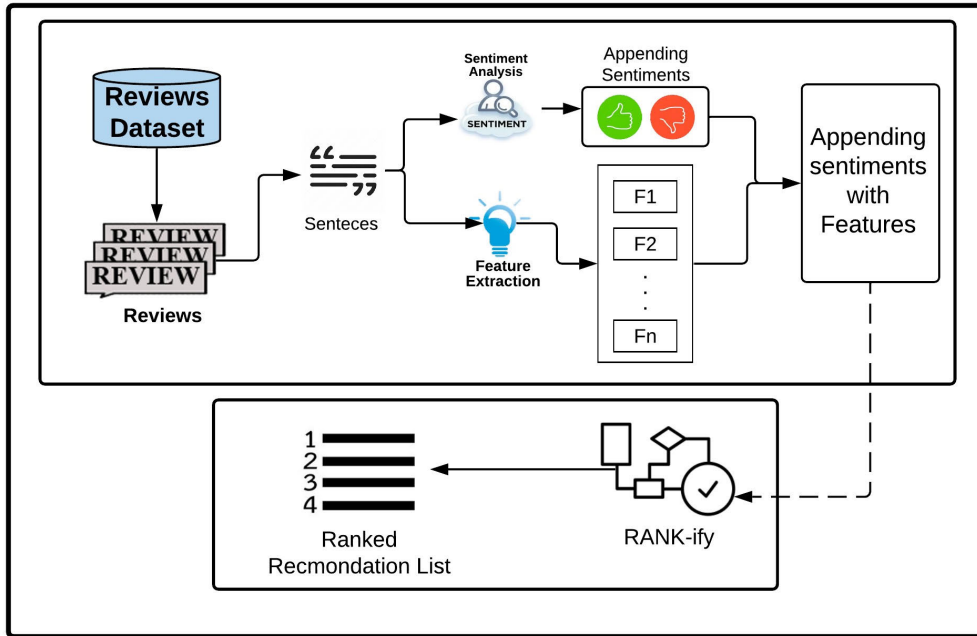
FIGURE 1. Framework of the Product Ranking based on the User Specification (PRUS) method.

TABLE 1. Detailed distribution of amazon dataset used in proposed method.

Category	Total Reviews	Total Product
Cell Phones	158,405	2,150
Clothing and Shoes	59,392	998
Sports and Outdoor	53,608	745
Toys and Games	45,406	407
Total	316,811	4300

review text was less than five words were eliminated because such reviews might not contain any feature or sentiment or both. Before performing experiments, the dataset was cleaned to make maximum use of the review text by removing variations. For cleaning purposes, stop-words and punctuation marks were removed. Integers were converted to words and tokens were stemmed and lemmatized using WordNet Lemmatizer [40]. The purpose of this cleaning was to facilitate feature extraction by reducing the possible word variations and writing style differences.

After preprocessing, the dataset is reduced to 316811 reviews which are used for experimental evaluation of the proposed approach. The analysis of the dataset shows that it contains reviews of 371 mobile phone brands. It can also be observed that about 40% of the reviews are covering three brands only which include Samsung, BLU, and Apple. A total of about 4300 products were reviewed by the users. With further analyses, it was observed that the top ten most reviewed products have received 800 to 1000 reviews each. Table 1 presents a detailed distribution of the dataset in terms of categories, reviews, and products.

B. PROPOSED APPROACH PRODUCT RANKING BASED ON USER SPECIFICATION

This section elaborates on the proposed approach PRUS. The aim is to rank products by considering the specifications provided by the users. PRUS approach works in three phases. In the first phase, reviews are divided into sentences to obtain respective features and sentiments from each sentence. This step helps in extracting aspects from the reviews of products which can be used to calculate the sentiments of products on the sentence level. In the second phase, the extracted features are appended with obtained sentiments extracted from the specific sentence of the review. Finally, in the last phase, the product's features obtained from the second phase are ranked according to user specifications based on positive and negative sentiments. Specifications are provided by the user in the form of a search query from which the desired features are extracted.

1) FRAMEWORK OF PRUS METHOD

In this section, the complete process for the proposed PRUS has been presented in Figure 1 The proposed approach is initialized by obtaining reviews from the review dataset. These reviews are then broken down into sentences. Each sentence is processed to extract its existing features and sentiments. The proposed framework utilized information gain which is an evaluation method based on entropy. Information gain is defined as the amount of information that a certain feature is able to provide for the whole classification. Once obtained, the respective features of each sentence are appended with its sentiment. These features are then

matched with the features obtained from user specifications. The ranking of reviews is performed using the RANK-ify algorithm which assigns weight according to positive and negative sentiments. Finally, the ranked recommendation list presents the features weighted by their respective sentiments as per user specifications.

2) PROCESS OF PRUS METHOD

This sub-section elaborates the proposed PRUS approach by considering an example review as demonstrated in Table 2. The process starts by breaking the example review of a specific product into sentences. This procedure is implemented by using Python Library NLTK. In the second phase, the sentiment analysis and feature extraction of broken sentences of the example review is performed. The second phase is implemented by using the Python library TextBlob [41]. Similarly, the respective features and sentiments for all the product reviews in the dataset are obtained. The obtained features of each review are appended with the corresponding sentiment (Table 2). In this way, based on the sentiment of the sentence, each feature of the review is labeled as positive or negative. This means that if a sentence is categorized with positive sentiment, then its extracted features will be labeled as positively reviewed features. Similarly, if a sentence is categorized with a negative or neutral sentiment, then its extracted features will be labeled as negatively or neutrally reviewed features, respectively. As a result, total counts are calculated based on the number of times a feature is reviewed as positive, negative, or neutral in all reviews of a specific product in the dataset. These scores are obtained by calculating the ratio of the difference between positive counts and negative counts with the total count of a feature in a product. Finally, the recommended ranked list is calculated using RANK-ify algorithm.

3) PROPOSED RANK-IFY ALGORITHM

This sub-section describes the working of the proposed RANK-ify algorithm. The goal of the RANK-ify algorithm (Algorithm 1) is to rank the products with respect to the feature score in accordance with the features requested by the user specification through a query. The algorithm takes three parameters as input and returns a ranked list of the products which cover the user-specified feature. In line 1, k represents the maximum number of products to be returned to the user in descending order, i.e., from highly relevant to the least relevant with user specification. Lp is the set of all the products which are reviewed in the dataset (line 2). This set also refers to all relevant features of respective products and their corresponding positive and negative occurrence counts. In line 3, Q represents the list of the features extracted from the user query. The algorithm starts by iterating the features of every single product (Line 4) and features specified in the user query (line 5). If the user-queried feature exists in the features of the product (Line 6), the Feature Score (FeaSco) of that feature is calculated using Equation 1 (Line 7).

$$FeaSco(f) = (w_1 c(f^+) - w_2 c(f^-))/(c(f^+) + c(f^-)) \quad (1)$$

Algorithm 1 RANK-Ify Algorithm

1: k = Number of Products in Dataset
2: L_p = List of Products in Dataset
3: Q = Features in User Query
4: **for** $p \in L_p$ **do**
5:     **for** $U_f \in Q$ **do**
6:         **if** $U_f \in p$ **then**
7:             $FS(U_f) = FeaSco(U_f)$
8:         **end if**
9:     **end for**
10:     $RS(p) = \sum_{U_f \in Q} FS(U_f)$
11: **end for**
12: Sort L_p w.r.t. RS
13: Return top-k from L_p

Feature score *FeaSco* for every feature is calculated using the occurrence count of the feature separately based on its positive or negative occurrence. Here $c(f^+)$ means how many times the feature f has appeared positively and $c(f^-)$ represents the count of the feature that appeared as negative. As the user may not be interested in positive and negative occurrences equally, therefore, weight attribute w is introduced to control the degree of positive and negative occurrence impact on the final ranked list. This weight is used with both positive and negative counts. It is pertinent to mention here that the sum of the positive and negative weight should not be exceeded from 1 to keep the dominance normalized, i.e., $w1 + w2 = 1$.

Once, the *FeaSco* for every user-queried feature is calculated for every single product individually, these feature scores are summed up to get the overall Rank-Score of all features of the products (Line 10). Resultantly, the Rank-SCORE is available for the products in the dataset, keeping in view the user's preferences provided in the form of the user query. Finally, the products are sorted based on this Rank-SCORE (Line 12) and the top-k products from that sorted list are returned to the user (Line 13).

IV. EXPERIMENTS AND RESULTS

This section presents the experimental evaluation of the proposed PRUS approach. The proposed approach is evaluated using the preprocessed dataset of Amazon containing reviews about mobile phones of various brands and companies. PRUS is evaluated by extracting features from a user query which are then used to obtain a sorted and recommended ranked list based on Discounted Cumulative Gain (DCG) measure. The experiments are conducted using varying values of positive and negative weights of review sentiments.

In existing studies, discounted cumulative gain (*DCG*) is used to evaluate the sentiment analysis approaches [42]. In most of these studies, the *DCG* value of a ranked product is calculated by using the rating given by the user to an aspect. However, most of the scenario's dataset does not contain the features and users' ratings of these features [43]. Mostly, datasets contain overall ratings along with the review text. Nevertheless, this overall rating may not be useful for all

TABLE 2. An example review text using the proposed PRUS approach.

Review Text		Review Sentence	Review Sentiment	Reveiw Sentence Feature
I do line the new design. And with 765g the battery power has vastly improved. In a world of huge phones, the 6.0-inch screen is a bit refreshing, compact, and lightweight. The display resolution is bad.	S1	I do like the new design	Positive	new design
	S2	And with 765g the battery power has vastly improved	Neutral	765g battery power
	S3	In a world of huge phones, the 6.0-inch screen is a bit refreshing, compact, and lightweight	Positive	huge phones, 6.0-inch, bit, compact, light weight
	S4	The display Resolution is bad	Neutral	Display Resolution

the features addressed in that review [44]. For example, the review "The screen quality is good but battery timing is not great" is rated as 4 out of 5. The review covers two features i.e., 'screen quality' and 'battery timing' and it is not justifiable if the 4 scores are assigned to both features as one of the features is appreciated while the other one is disapproved by the user. To overcome this issue, the proposed PRUS approach utilizes the sentiment score and assigns it to each feature instead of the overall rating of the review.

This study utilized the TextBlob library for sentiment analysis. The obtained sentiment scores, assigned to each feature, are normalized between the scales of 1 to 10. These scores assigned to each feature are averaged using Equation 2. This is referred to as the average aspect score (AAS) for that specific feature.

$$AAS(p, Q) = 1/Count\left(\sum_{i=1}^{count} S(i)\right) \quad (2)$$

Here, for a single product p, all the features that are required by the user query Q, are taken into consideration. Their scores are summed and averaged based on the total count. In this way, a score is allocated to each product with respect to the user-preferred features i.e., Q. Once AAS values are obtained for all the ranked products, the value of DCG is calculated for the products at the particular rank position p using Equation 3 as below

$$DCG_p = AAS_1(p, Q) + \sum_{a=1}^{p} \frac{AAS_a(p, Q)}{log_2 a} \quad (3)$$

In this way, the DCG score for all the products is obtained. The products are then sorted based on their AAS scores and ranking are generated using this score. This re-ranked product list is used for normalizing the DCG score which is calculated using the above equation. These sorted scores are named Ideal discounted cumulative gain (IDCG) which is obtained using Equation 4. Finally, nDCG is calculated using Equation 5.

$$IDCG_p = \sum_{i=1}^{p} \frac{AAS_i}{log_2(i+1)} \quad (4)$$

$$nDCG_P = \frac{DCG_P}{IDCP_P} \quad (5)$$

Based on this procedure, the previous example explained at the start of this section, the sentiment score of screen quality will be assigned as 8.65 whereas a sentiment score of 3.7 is assigned to the battery timing.

To conduct a comprehensive experimental evaluation, the proposed PRUS approach is demonstrated using an example. Table 3 shows an example where a user query is provided as input. The user query consists of two features; a good camera and the high screen resolution of a mobile phone. In response to this user query, PRUS based system displayed the top-10 ranked products that are recommended to the user based on rank score (RS) and nDCG values. RS values are calculated using the RANK-ify algorithm and nDCG values are calculated using equation 5. While calculating these values, the PRUS approach specifically utilized positive and negative aspects assigned to each feature of that product. As in this example, those mobile phone products are highly ranked which consist of a good camera coupled with high screen resolution. The recommended list contains all the products having the same highest rank score (0.9); however, these are sorted based on varying nDCG values. The list represents recommended products in a sorted way; the first product is highly recommended and the last product is the least recommended one.

A. PERFORMANCE ANALYSIS OF THE PRUS METHOD

This section analyzes the performance of the PRUS method. Several experiments have been conducted to observe the change in the ranking of the retrieved products. In these experiments, the weights $w1$ and $w2$ were assigned alternate values such that their sum equals 1 i.e., $w1 + w2 = 1$. The results of these experiments in terms of rank score (RS) and normalized discounted cumulative gain (nDCG) score are listed in Table 3. Here, $P1$ to $P10$ represent the product list which is different for all values of RS and nDCG scores obtained by varying positive and negative weight in the range of 0.1 to 0.9. The detailed product list for the variations of weights is represented in Table 3.

It can be observed from the results that there was no prominent change in the overall sum score of nDCG as long as the value of the positive weight ($w1$) remains between 0.9 to 0.5. However, as soon as negative weight ($w2$) increases from 0.5, the nDCG scores show a sudden increase in terms of the sum. Now consider the value of RC, the results show that as the negative weight crossed the mid value i.e., 0.5, the rank scores for the products start to show a decline. This is because the FeaSco formula (Equation 1) of the proposed PRUS approach is derived from the concept that users intend to buy a product by considering to get its useful features with respect to its positive and negative

TABLE 3. Performance analysis of the PRUS method.

| Product		P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
Rank		1	2	3	4	5	6	7	8	9	10
PW=0.9	Rank Score	2.7	2.7	1.8	1.8	1.8	1.8	1.8	1.8	1.8	1.8
NW=0.1	nDCG Score	0.94	2.06	2.81	3.59	4.06	4.44	4.94	4.94	5.81	6.67
PW=0.8	Rank Score	2.24	2.24	1.61	1.61	1.61	1.61	1.61	1.61	1.61	1.61
NW=0.2	nDCG Score	1.00	2.06	2.92	3.65	4.15	4.55	4.90	5.37	5.91	6.36
PW=0.7	Rank Score	2.1	2.1	1.4	1.4	1.4	1.4	1.4	1.4	1.4	1.4
NW=0.7	nDCG Score	1.00	2.06	3.06	3.59	3.77	4.57	4.85	5.27	5.50	6.44
PW=0.6	Rank Score	1.8	1.8	1.2	1.2	1.2	1.2	1.2	1.2	1.2	1.2
NW=0.4	nDCG Score	0.94	2.06	2.79	3.60	4.02	4.44	4.58	5.15	5.97	6.72
PW=0.5	Rank Score	1.5	1.5	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
NW=0.5	nDCG Score	0.94	2.06	3.07	3.39	4.02	4.57	4.84	5.47	5.70	6.03
PW=0.4	Rank Score	1.2	1.2	1	1	1	0.85	0.84	0.82	0.80	0.80
NW=0.6	nDCG Score	1	2.06	2.73	3.29	3.47	4.57	5.41	6.13	7.36	8.07
PW=0.3	Rank Score	1	1	1	0.9	0.9	0.71	0.7	0.7	0.68	0.65
NW=0.7	nDCG Score	0.488	1.04	1.97	4.05	5.40	4.62	2.23	2.65	14.9	15.8
PW=0.2	Rank Score	1	1	1	0.8	0.8	0.6	0.6	0.6	0.57	0.52
NW=0.8	nDCG Score	0.48	1.00	2.09	1.42	1.99	6.67	8.35	6.00	14.7	17.2
PW=0.1	Rank Score	1	1	0.9	0.9	0.63	0.5	0.42	0.360	0.36	0.52
NW=0.9	nDCG Score	1.89	2.65	1.93	3.60	3.97	5.64	13.2	8.01	9.13	17.2
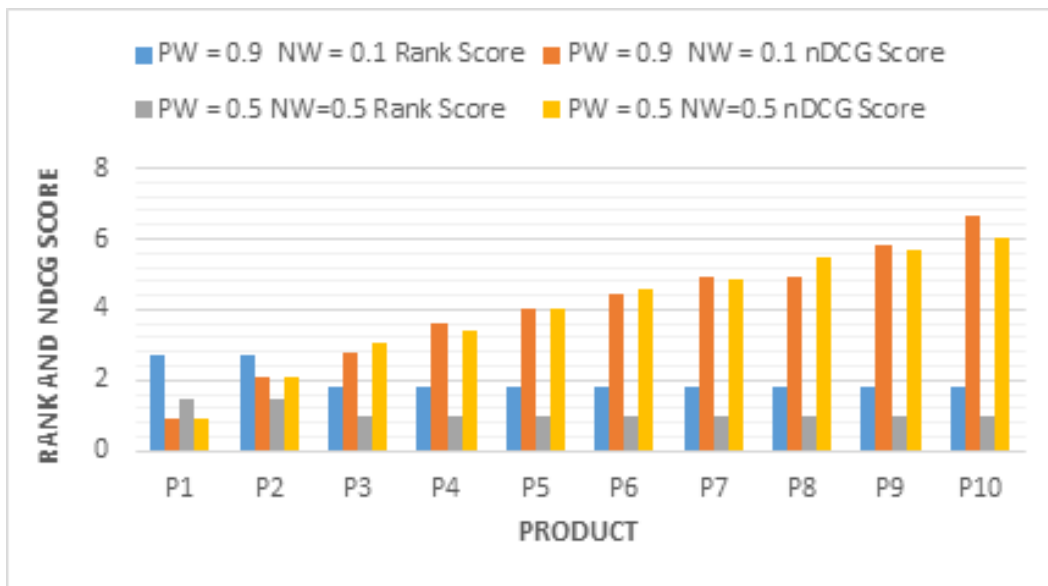
FIGURE 2. Performance analysis of the Product Ranking based on User Specification (PRUS) method.

aspects. As the user may not be interested in positive and negative occurrences equally, therefore, weight attribute (w) is introduced to control the degree of positive and negative occurrence impact on the final ranked list [45]. As the value of the negative weight ($w2$) is increased from 0.5, the list contains those products which are having features with more emphasis on negative aspects. These experiments show the impact of varying positive and negative aspects of products based on user-specified features considering the relevance of recommended products ($nDCG$) coupled with its RS. Although the higher positive weight gives a higher rank score to those products which have gained positive sentiments concerning their features however, a decreasing trend of RS can be observed if the negative weight value is increased which gives higher weightage to those products which have more counts of negative sentiments with respects to their features. The difference in the product list for the occurrences

of varying positive and negative weights can be observed in Table 3. This is the major contribution of the proposed PRUS approach that it gives independence to the user to select recommended list based on its specified features with respect to positive or negative aspects of the products.

Figure 2 shows the performance analysis of the proposed PRUS method. It is observed that when negative weight ($w2$) increases then nDCG scores increase, while if positive weight ($w1$) increases then Rank Score starts decreasing. This shows the impact of varying positive and negative aspects of products on user-specified features considering the relevance of recommended products (nDCG) coupled with its ranked score (RS).

V. CONCLUSION
This study proposed a framework, called PRUS, to generate recommended ranked list of products from customer reviews

based on features extracted from a user query. Existing literature in this domain has not considered negative aspects while recommending ranked products list. Moreover, existing approaches generally rank and display the products to all the users in a similar manner irrespective of the actual needs of the individual user. The proposed PRUS approach worked in three phases. In the first phase, features and sentiments from each sentence of a review have been extracted which are used to calculate the positive and negative sentiments of the products on the sentence level. In the second phase, the extracted features are appended with obtained sentiments, and in the last phase, features from the user query are extracted which are used to rank products considering the calculated Rank Score based on the positive and negative sentiments of their individual features. The positive and negative sentiments of each feature are attributed with a weighted parameter that can be used to control the impact of positive and negative sentiments on the recommended ranked list. The proposed PRUS approach is evaluated by conducting comprehensive experiments. This research opens several new opportunities in the product ranking domain. A future direction in this domain is to explore product ranking on multilingual product review datasets. Another interesting future direction is to provide more flexibility to the user to generate recommended ranked product lists based on negative and positive polarities of user-specified features.

REFERENCES

[1] P. K. Jain, R. Pamula, and G. Srivastava, "A systematic literature review on machine learning applications for consumer sentiment analysis using online reviews," *Comput. Sci. Rev.*, vol. 41, Aug. 2021, Art. no. 100413.

[2] S. Moro and P. Rita, "Data and text mining from online reviews: An automatic literature analysis," *WIREs Data Mining Knowl. Discovery*, vol. 12, no. 3, May 2022, Art. no. e1448.

[3] N. Hussain, H. T. Mirza, G. Rasool, I. Hussain, and M. Kaleem, "Spam review detection techniques: A systematic literature review," *Appl. Sci.*, vol. 9, no. 5, p. 987, Mar. 2019.

[4] L. He, X. Wang, H. Chen, and G. Xu, "Online spam review detection: A survey of literature," *Hum.-Centric Intell. Syst.*, vol. 2, nos. 1–2, pp. 14–30, Jun. 2022.

[5] S. K. Maurya, D. Singh, and A. K. Maurya, "Deceptive opinion spam detection approaches: A literature survey," *Int. J. Speech Technol.*, vol. 53, no. 2, pp. 2189–2234, Jan. 2023.

[6] E. Lee, F. Rustam, P. B. Washington, F. E. Barakaz, W. Aljedaani, and I. Ashraf, "Racism detection by analyzing differential opinions through sentiment analysis of tweets using stacked ensemble GCR-NN model," *IEEE Access*, vol. 10, pp. 9717–9728, 2022.

[7] W. S. El-Kassas, C. R. Salama, A. A. Rafea, and H. K. Mohamed, "Automatic text summarization: A comprehensive survey," *Expert Syst. Appl.*, vol. 165, Mar. 2021, Art. no. 113679.

[8] M. Umer, I. Ashraf, A. Mehmood, S. Kumari, S. Ullah, and G. S. Choi, "Sentiment analysis of tweets using a unified convolutional neural network-long short-term memory network model," *Comput. Intell.*, vol. 37, no. 1, pp. 409–434, Feb. 2021.

[9] P. K. Soni and R. Rambola, "A survey on implicit aspect detection for sentiment analysis: Terminology, issues, and scope," *IEEE Access*, vol. 10, pp. 63932–63957, 2022.

[10] N. Aslam, K. Xia, F. Rustam, A. Hameed, and I. Ashraf, "Using aspect-level sentiments for calling app recommendation with hybrid deep-learning models," *Appl. Sci.*, vol. 12, no. 17, p. 8522, Aug. 2022.

[11] G. Brauwers and F. Frasincar, "A survey on aspect-based sentiment classification," *ACM Comput. Surv.*, vol. 55, no. 4, pp. 1–37, Apr. 2023.

[12] M. Sivakumar and S. R. Uyyala, "Aspect-based sentiment analysis of mobile phone reviews using LSTM and fuzzy logic," *Int. J. Data Sci. Anal.*, vol. 12, no. 4, pp. 355–367, Oct. 2021.

[13] J. Zhang, X. Lu, and D. Liu, "Deriving customer preferences for hotels based on aspect-level sentiment analysis of online reviews," *Electron. Commerce Res. Appl.*, vol. 49, Sep. 2021, Art. no. 101094.

[14] K. Ganesan and C. Zhai, "Opinion-based entity ranking," *Inf. Retr.*, vol. 15, no. 2, pp. 116–150, Apr. 2012.

[15] L. Akoglu, "Opinion fraud detection in online reviews by network effects," in *Proc. ICWSM*, vol. 7, no. 1, 2013, pp. 2–11.

[16] E. Marrese-Taylor, J. D. Velásquez, and F. Bravo-Marquez, "A novel deterministic approach for aspect-based opinion mining in tourism products reviews," *Expert Syst. Appl.*, vol. 41, no. 17, pp. 7764–7775, Dec. 2014.

[17] A. Kumar and S. Abirami, "Aspect-based opinion ranking framework for product reviews using a Spearman's rank correlation coefficient method," *Inf. Sci.*, vols. 460–461, pp. 23–41, Sep. 2018.

[18] S. Bashir, W. Afzal, and A. R. Baig, "Opinion-based entity ranking using learning to rank," *Appl. Soft Comput.*, vol. 38, pp. 151–163, Jan. 2016.

[19] T. A. Rana and Y.-N. Cheah, "A two-fold rule-based model for aspect extraction," *Expert Syst. Appl.*, vol. 89, pp. 273–285, Dec. 2017.

[20] C. Wu, F. Wu, S. Wu, Z. Yuan, and Y. Huang, "A hybrid unsupervised method for aspect term and opinion target extraction," *Knowl.-Based Syst.*, vol. 148, pp. 66–73, May 2018.

[21] Z. Luo, S. Huang, F. F. Xu, B. Y. Lin, H. Shi, and K. Zhu, "ExtRA: Extracting prominent review aspects from customer feedback," in *Proc. Conf. Empirical Methods Natural Lang. Process.*, 2018, pp. 3477–3486.

[22] S. Sharmila and S. Vijayarani, "Association rule mining using fuzzy logic and whale optimization algorithm," *Soft Comput.*, vol. 25, no. 2, pp. 1431–1446, Jan. 2021.

[23] S. Gurumoorthy, B. N. K. Rao, X.-Z. Gao, B. V. Krishna, A. K. Pandey, and A. S. Kumar, "Feature based opinion mining and sentiment analysis using fuzzy logic," in *Cognitive Science and Artificial Intelligence* (SpringerBriefs in Forensic and Medical Bioinformatics). Singapore, 2018, pp. 79–89.

[24] M. Venugopalan and D. Gupta, "An enhanced guided LDA model augmented with BERT based semantic strength for aspect term extraction in sentiment analysis," *Knowl.-Based Syst.*, vol. 246, Jun. 2022, Art. no. 108668.

[25] M. Hu and B. Liu, "Mining and summarizing customer reviews," in *Proc. 10th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2004, pp. 168–177.

[26] A.-M. Popescu and O. Etzioni, "Extracting product features and opinions from reviews," in *Proc. Conf. Hum. Lang. Technol. Empirical Methods Natural Lang. Process. (HLT)*, 2005, pp. 9–28.

[27] S. Li, L. Zhou, and Y. Li, "Improving aspect extraction by augmenting a frequency-based method with web-based similarity measures," *Inf. Process. Manage.*, vol. 51, no. 1, pp. 58–67, Jan. 2015.

[28] Z. Li, M. Zhang, S. Ma, B. Zhou, and Y. Sun, "Automatic extraction for product feature words from comments on the web," in *Proc. Asia Inf. Retr. Symp.* Sapporo, Japan: Springer, Oct. 2009, pp. 112–123.

[29] S. Raju, P. Pingali, and V. Varma, "An unsupervised approach to product attribute extraction," in *Proc. Eur. Conf. Inf. Retr.* Toulouse, France: Springer, Apr. 2009, pp. 796–800.

[30] Ö. F. Güneş, T. Furche, and G. Orsi, "Structured aspect extraction," in *Proc. COLING 2016, 26th Int. Conf. Comput. Linguistics, Tech. Papers*, 2016, pp. 2321–2332.

[31] C. Quan and F. Ren, "Unsupervised product feature extraction for feature-oriented opinion determination," *Inf. Sci.*, vol. 272, pp. 16–28, Jul. 2014.

[32] W. Maharani, D. H. Widyantoro, and M. L. Khodra, "Clue propagation based on non-adjective opinion words for handling disconnected propagation in product reviews," *IEEE Access*, vol. 10, pp. 56785–56795, 2022.

[33] X. Luo, "Efficient english text classification using selected machine learning techniques," *Alexandria Eng. J.*, vol. 60, no. 3, pp. 3401–3409, Jun. 2021.

[34] C.-P. Wei, Y.-M. Chen, C.-S. Yang, and C. C. Yang, "Understanding what concerns consumers: A semantic approach to product feature extraction from consumer reviews," *Inf. Syst. e-Bus. Manage.*, vol. 8, no. 2, pp. 149–167, Mar. 2010.

[35] B. Ma, D. Zhang, Z. Yan, and T. Kim, "An lda and synonym lexicon based approach to product feature extraction from online consumer product reviews," *J. Electron. Commerce Res.*, vol. 14, no. 4, p. 304, 2013.

[36] K. Liu, L. Xu, and J. Zhao, "Co-extracting opinion targets and opinion words from online reviews based on the word alignment model," *IEEE Trans. Knowl. Data Eng.*, vol. 27, no. 3, pp. 636–650, Mar. 2015.

[37] C.-H. Lai, D.-R. Liu, and K.-S. Lien, "A hybrid of XGBoost and aspect-based review mining with attention neural network for user preference prediction," *Int. J. Mach. Learn. Cybern.*, vol. 12, no. 5, pp. 1203–1217, May 2021.

[38] G. Kaur and A. Sharma, "A deep learning-based model using hybrid feature extraction approach for consumer sentiment analysis," *J. Big Data*, vol. 10, no. 1, pp. 1–23, Jan. 2023.

[39] M. Nasir and C. I. Ezeife, "Semantic enhanced Markov model for sequential E-commerce product recommendation," *Int. J. Data Sci. Anal.*, vol. 15, no. 1, pp. 67–91, Jan. 2023.

[40] S. Bird, E. Klein, and E. Loper, *Natural Language Processing With Python: Analyzing Text With the Natural Language Toolkit*. Sebastopol, CA, USA: O'Reilly Media, 2009.

[41] S. Loria, *Textblob Documentation*, document Release 0.15, 2018, vol. 2, no. 8.

[42] D. Valcarce, A. Bellogín, J. Parapar, and P. Castells, "Assessing ranking metrics in top-N recommendation," *Inf. Retr. J.*, vol. 23, no. 4, pp. 411–448, Aug. 2020.

[43] P. Nitu, J. Coelho, and P. Madiraju, "Improvising personalized travel recommendation system with recency effects," *Big Data Mining Anal.*, vol. 4, no. 3, pp. 139–154, Sep. 2021.

[44] M. K. Singh and O. P. Rishi, "Event driven recommendation system for E-commerce using knowledge based collaborative filtering technique," *Scalable Comput., Pract. Exper.*, vol. 21, no. 3, pp. 369–378, Aug. 2020.

[45] Z.-P. Fan, G.-M. Li, and Y. Liu, "Processes and methods of information fusion for ranking products based on online reviews: An overview," *Inf. Fusion*, vol. 60, pp. 87–97, Aug. 2020.

AYESHA ALTAF received the M.S. and Ph.D. degrees in information security from NUST, Pakistan, in 2009 and 2021, respectively. She is currently an Academician and a Researcher in cyber security. She is currently an Assistant Professor with the Department of Computer Science, University of Engineering and Technology Lahore, Lahore, Pakistan. Her professional services include industry consultation, a workshops organizer/resource person (workshops/seminars), and a technical program committee member, teaching (U.G./P.G./Ph.D.) courses, research and development, and reviewing for various international journals/conferences. She has almost 14 years of experience teaching with the university level. She has published more than 12 scientific research publications in major international journals (ISI-Indexed), such as IEEE Internet of Things Journal, *JNCA*, *JSA*, and *CEE* with a cumulative IF of more than 33.

AHTSHAM SHOUKAT received the M.S. degree in computer science from COMSATS University Islamabad, in 2020. He is currently a software engineer in a reputed software house in Lahore, Pakistan. His research interests include artificial intelligence, machine learning ubiquitous computing, accessibility, and location-based services.

MÓNICA GRACIA VILLAR is currently with Universidad Europea del Atlántico, Santander, Spain. He is also affiliated with Universidade Internacional do Cuanza, Cuito, Bié, Angola, and Fundación Universitaria Internacional de Colombia, Bogotá, Colombia.

EMMANUEL SORIANO FLORES received the bachelor's degree (Hons.) in business administration, the master's degree in educational innovation, the master's degree in international business administration, the master's degree in China–Asia Pacific business, the master's degree in corporate business communication, the master's degree in financial management, and the Ph.D. degree in higher education. He has experience as a professor and a researcher at various universities in Mexico and Spain. He is currently a Coordinator of the Master in Business Administration with the European University of the Atlantic, Spain.

MARCO ANTONIO ROJO GUTIÉRREZ is currently with Universidad Europea del Atlántico, Santander, Spain. He is also affiliated with Universidad Internacional Iberoamericana Campeche, Mexico, and Universidad Internacional Iberoamericana Arecibo, Puerto Rico, USA.

NAVEED HUSSAIN received the Ph.D. degree in computer science from COMSATS University Islamabad (Lahore Campus), Pakistan. He is currently an Assistant Professor with the Faculty of Information Technology and Computer Science, University of Central Punjab, Pakistan. He has published several articles in reputed journals. His research interests include data mining, sentimental analysis, machine learning, and opinion mining.

HAMID TURAB MIRZA received the M.Sc. degree (Hons.) in information systems from The University of Sheffield, England, in 2005, and the Ph.D. degree in computer science from Zhejiang University, China, in 2012. He has more than 15 years of research, teaching, and information systems development experience in the aviation, telecommunication, and academic sectors of both Pakistan and the U.K. His research interests include data mining, machine learning, and human–computer interaction. He has published several papers at reputed international conferences and journals in these areas.

FAIZA IQBAL received the Ph.D. degree from NUST, Pakistan. She directs multiple IoT-based software development initiatives. She received Pakistan's Higher Education Commission Indigenous Scholarship for M.S. leading to the Ph.D. degree. She is currently with the Department of Computer Science, University of Engineering and Technology Lahore, Lahore, Pakistan. Her current research interests include the IoT-based smart applications, AI for cyber security, network optimization modeling, and data analytics for high-performance protocol design. Professionally, she serves as a member of the program committee, technical committee, and reviewer panel for numerous international journals and conferences.

IMRAN ASHRAF received the M.S. degree (Hons.) in computer science from the Blekinge Institute of Technology, Karlskrona, Sweden, in 2010, and the Ph.D. degree in information and communication engineering from Yeungnam University, Gyeongsan, South Korea, in 2018. He was a Postdoctoral Fellow with Yeungnam University, where he is currently an Assistant Professor with the Information and Communication Engineering Department. His research interests include positioning using next-generation networks, communication in 5G and beyond, location-based services in wireless communication, smart sensors (LIDAR) for smart cars, and data analytics.

● ● ●