

Article

Multi-Attribute Online Decision-Making Driven by Opinion Mining

Azra Shamim ^{1,2,*}, Muhammad Ahsan Qureshi ¹, Farhana Jabeen ³, Misbah Liaqat ^{1,*}, Muhammad Bilal ^{4,5,*} , Yalew Zelalem Jembre ⁶  and Muhammad Attique ^{7,*} 

¹ Department of Information Technology, College of Computing and Information Technology at Khulais, University of Jeddah, Jeddah 23218, Saudi Arabia; maqureshi@uj.edu.sa

² Faculty of Computer Science and Information Technology, University of Malaya, Kuala Lumpur 50603, Malaysia

³ Department of Computer Science, COMSATS University Islamabad, Islamabad 44000, Pakistan; farhanakhan@comsats.edu.pk

⁴ School of Computer Science and Engineering, Taylor's University, Subang Jaya 47500, Malaysia

⁵ Centre for Data Science and Analytics (C4DSA), Taylor's University, Subang Jaya 47500, Malaysia

⁶ Department of Electronics, Keimyung University, Daegu 42601, Korea; zizutg@kmu.ac.kr

⁷ Department of Software, Sejong University, Seoul 05006, Korea

* Correspondence: asmajeed@uj.edu.sa (A.S.); mhaseeb@uj.edu.sa (M.L.); muhammadbilal@sd.taylors.edu.my (M.B.); attique@sejong.ac.kr (M.A.)

Abstract: With the evolution of data mining systems, the acquisition of timely insights from unstructured text is an organizational demand which is gradually increasing. The existing opinion mining systems have a variety of properties, such as the ranking of products' features and feature level visualizations; however, organizations require decision-making based upon customer feedback. Therefore, an opinion mining system is proposed in this work that ranks reviews and features based on novel ranking schemes with innovative opinion-strength-based feature-level visualization, which are tightly coupled to empower users to spot imperative product features and their ranking from enormous reviews. Enhancements are made at different phases of the opinion mining pipeline, such as innovative ways to evaluate review quality, rank product features and visualize opinion-strength-based feature-level summary. The target user groups of the proposed system are business analysts and customers who want to explore customer comments to gauge business strategies and purchase decisions. Finally, the proposed system is evaluated on a real dataset, and a usability study is conducted for the proposed visualization. The results demonstrate that the incorporation of review and feature ranking can improve the decision-making process.

Keywords: opinion mining; opinion visualization; sentiment analysis; feature ranking; review quality evaluation



Citation: Shamim, A.; Qureshi, M.A.; Jabeen, F.; Liaqat, M.; Bilal, M.; Jembre, Y.Z.; Attique, M. Multi-Attribute Online Decision-Making Driven by Opinion Mining. *Mathematics* **2021**, *9*, 833. <https://doi.org/10.3390/math9080833>

Academic Editor:
Vicente Coll-Serrano

Received: 11 March 2021
Accepted: 8 April 2021
Published: 11 April 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Improvements in information and communication technologies break down geographical boundaries, allowing for faster connection and communication worldwide [1]. Further, the proliferation of social networks due to the ubiquity of Web 2.0 revolutionizes the way people present their opinions by providing social interactions [2,3]. As a result, consumers from all over the world are sharing their emotions, opinions, evaluations and judgments to a wide ranging audience by connecting themselves to online platforms such as blogs, newsgroups, discussion boards, and social networking sites [4–6]. Consequently, the Web consists of huge volumes of publicly available opinion data about different objects, for instance, individuals, government, products, events, organizations, services, education, news [7,8]. The volume of opinion data about different entities (individuals, products, events, organizations, services) is growing rapidly on these platforms due to the accessibility, scalability, and enhanced user participation of Web 2.0. The fast-growing opinion

data are unstructured, freely available, and decision-oriented to fulfil the need of diverse stakeholders such as corporates and consumers [9–11]. Though there are various platforms of electronic word of mouth (e-WOM), the most highly focused are online reviews platforms. Enterprises are now focusing on customer online reviews to support their decision-making process, such as risk management, sale prediction, market intelligence, new product design, trend prediction, advertisement placement, threats from competitors and benchmarking [12–16]. From the customer point of view, e-WOM considerably impacts customers' product choice and adoption, purchase intentions, and use of products [17,18]. Moreover, social networks have increased the sophistication of customers, and so customers compare competing products before buying a product [19,20]. As a result, positive e-WOM improves the trust, satisfaction and loyalty of customers [2]. In contrast, negative e-WOM decrease customers' patronage and loyalty [3].

The volume of e-WOM has been growing at a remarkable pace as the universality of the Web offers easy users' participation through different online platforms [21]. These platforms provide valuable insights into different features of a target product, for instance, a digital camera. Consumers consult these platforms to compare features of different cameras (picture quality, battery life, zoom, flash) from competing brands before making a purchase [22]. The features of a product play a crucial role in the purchase decision. However, online reviews do not highlight the critical features of a product to facilitate consumers in their decision-making process. Moreover, users need to gravitate through a host of reviews to learn about features that need to be considered before making a purchase due to the varying quality, enormous availability, distributed nature, voluminosity, heterogeneity, and multi-dimensionality of online reviews. As a result, it is a time-consuming and tedious task to analyze and summarize online reviews to get information about competing features between products of different brands [23,24]. Further, the nature of the online reviews poses many challenges to the text mining community, such as filtering low-quality reviews from high-quality reviews (information overloading problem), ranking of products' prominent features and integrated visual views of consumers' opinions (consumers-to-consumers communication problem) [25–27]. The information overload problem has been resolved in literature by a variety of review ranking schemes. To overcome the consumer-consumers communication problem, opinion mining systems have been available in the literature aiming at providing automatic analysis and summarization of online reviews to pinpoint decision-oriented features of a target product [28].

Due to the characteristics of online reviews, it becomes difficult to identify high-quality reviews which cover a diverse set of opinions. The quality of a review is described by its embedded polarity or opinions [29] or how helpful a review is [30]. For example, the Amazon online reputation system in Ref. [31] asks users to vote for customer reviews they found "helpful" in the form of helpfulness votes and star rating. The helpfulness votes and 5-star rating represent the quality of the product that signifies customers' endorsement and impact other customers' shopping intentions [32–34]. Currently, high-quality reviews are identified when users explicitly filter reviews based on the star rating and/or their level of helpfulness [32,35]. Different review quality evaluation methods have been proposed in the literature that utilize combinations of metadata, textual, and social features [36–38]. In contrast to classical rating systems [39] in which users rate a product, community-driven review systems allow the customers to rate prominent features of a product. The customers rating may vary based on the different features of the product. Consequently, it emphasizes the need to develop mechanisms to identify prominent features of a product in Ref. [40] about which the customers are concerned and rank products based on identified features by excavating a large number of customer reviews. Commonly, feature frequency, semantic polarities and ratings are used for feature ranking to enhance consumers to consumer communication [41–46]. However, these methods overlook important factors that can improve feature ranking by focusing on (i) opinion strength, such as how positive or negative an opinion word is, (ii) quality of reviews, and (iii) consideration of user preferences. Further, the visualization of the opinion summary has the same importance as the assessment

of review quality and feature ranking. A feature-based opinion summary with ample visualization may be more valuable than a summary showing only an average rating for features of a target product [41]. However, existing review quality evaluation methods are not integrated with feature ranking, opinion visualizations, and user preferences and they overlook a few parameters, such as visitor count and title information. Therefore, there is a need for an integrated system that ranks, analyzes, summarizes and visualizes these online reviews to fulfil the requirements of consumers and enterprises.

In the light of above discussion, the motivation of the work is to (i) remove low-quality reviews from feature ranking as suggested by Ref. [47], (ii) enhancing the feature ranking by incorporating missing parameters, and (iii) improving existing opinion visualization to provide opinion-strength-based summary. Therefore, this study aims to propose and develop a reputation system to provides users a multi-level analysis and summarization of consumer opinions. Specifically, the objectives of this paper are to propose (i) a review ranking method incorporating vital parameters and user preferences, (ii) a feature ranking method based on indispensable parameters, and (iii) opinion strength-based visualization.

The main contributions include:

- (a) Scheme for the selection of high-quality reviews by incorporating users' preferences.
- (b) Feature ranking scheme based on multiple parameters for a deeper understanding of consumers' opinions.
- (c) Opinion-strength-based visualization based on high-quality reviews to provide high-quality information for decision-making. The proposed visualization provides a multi-level detail of consumers' opinions (ranging from -3 to $+3$) on critical products features at a glance, which allows entrepreneurs and consumers to highlight decisive product features having a key impact on the sale, product choice and adoption.
- (d) Reputation system is evaluated on a real dataset
- (e) Usability study for the evaluation of the proposed visualization

The rest of the paper is organized as follows. Section 2 presents existing work on review quality evaluation, feature ranking and opinion visualizations. Section 3 presents the proposed system. Section 4 presents the results and discussion, and finally, Section 5 concludes the paper.

2. Related Work

Existing studies on review quality evaluation, feature ranking, and opinion visualizations are presented in this section.

2.1. Review Quality Evaluation and Review Ranking

It is difficult for customers and enterprises to identify high-quality reviews projecting the true quality of a target product due to the massive volume of reviews. Existing studies of review quality evaluation have focused on a number of features such as helpfulness votes, rating, review length and term frequency [20,48,49]. In Ref. [30], five feature classes: Lexical (uni-gram and bi-gram), Structural (i.e. length, number of sentences), Syntactic (nouns, adjectives, verbs), Meta-data (rating) and Semantic (features and opinion words) of a review were explored to predict the quality of a review. Review length, user ratings, and term frequency were found to be significant in review quality prediction. Similarly, the experimental results performed by Ref. [29] highlighted the shallow syntactic features such as verbs, nouns, and interjections as the strongest predictor of review quality prediction [29]. The authors in Ref. [50] utilized three more feature sets reviewer history features, reviewer profile features, and readability features to identify the helpfulness of a review. Their results demonstrated a correlation between these feature sets and the perceived helpfulness of reviews. Ref. [38] pinpointed the helpfulness of a review from three different perspectives: the writing style of the review, the reviewer's expertise, and the timeliness of the review. The experimental results of Ref. [26] found two main features review length and the number of product features to be significant while ranking reviews [26]. Ref. [51] proposed a review ranking scheme for book reviews based on the number of features. A score is

assigned to each review on the basis of the number of features that appeared in the review. Then reviews were ranked according to the assigned score. The review ranking scheme outclassed the helpfulness votes-based ranking scheme. Ref. [42] extended the previous scheme by incorporating the number of opinion words with the number of features for review ranking. The extended scheme outperformed the term frequency-based scheme.

In Ref. [20], more weightage is given to the title of the review as compared to the body while computing review ranking. According to the authors, the title of the review conveys the overall mood of the reviewer and an effective summary of the review. However, this work only considers opinion words from the title and ignores product features. The pioneer work that included integration review quality evaluation, feature ranking, and user preferences were presented by Shamim et al. [32], wherein a review quality evaluation scheme based on user preferences and four other parameters including helpfulness ratio, review rating, number of features and opinion words were proposed. In terms of user preferences, the users are allowed to perform the following: (i) adjust the weight of each parameter and (ii) select reviews for features ranking.

However, Ref. [52] ignored the title of the review in ranking reviews and features, and Ref. [26] ignored product features in the review title for ranking reviews. To address these limitations, the method proposed in the current study to rank the reviews calculates separate weights for title and body of the based on both the number of features and opinion words expressed in the title and body of a review along with metadata feature (review rating and helpfulness ratio). We enhanced the previous review ranking scheme in Ref. [52] by including the title score. Further, in contrast to Ref. [26] in which authors considered rating for review raking, our work does consider multiple parameters: (i) feature frequency, (ii) number of opinion words, (iii) accumulated strength of associated opinion words, and (iv) title information of a review. The reason to include the title information of a review in ranking is that the title highly summarizes a review and presents the overall opinion of the reviewer [42]. Considering the significance of the title, the title score is included in the review ranking and is associated with the weight coefficient α .

2.2. Feature Ranking

The pioneer work on feature-based opinion mining was done by Hu and Liu, [53] to mine and summarize customers' reviews by developing a system called feature-based summarization (FBS). This work aimed to identify product features, and opinion orientation for each feature and present a summary of the identified features and corresponding opinions in a textual form. In this work, the authors' utilized the classification based on associations (CBA) system using an Apriori algorithm to extract frequent explicit features. The adjective synonyms and antonyms of Wordnet are utilized in FBS to identify the opinion orientation of opinion words. An average accuracy of 84 % was achieved by FBS for opinion orientation.

In the literature, a variety of feature ranking schemes are available that rank features on the basis of feature frequency, opinion words, star rating, and semantic polarities. The most popular feature ranking approach ranks features based on the frequency of a feature [44,45]. For instance, Ref. [44] utilized feature frequency for feature ranking. Likewise, the feature frequency-based PageRank algorithm was revised for product ranking [45], and the findings of this approach showed promising results. In Ref. [42], the authors utilized the number of associated opinion words with each feature to rank features. The experimental results of this ranking outperformed ranking based on feature frequency. The previous ranking approach was enhanced by Ref. [43] by integrating guidelines defined by the review website with the number of associated opinion words. The outcomes of this integration exhibited significant improvement in the accuracy of the existing system [53]. The ranking approach of Ref. [42] was also extended by Ref. [54] such that the authors incorporated review rating with opinion words for feature ranking. The results of the extended approach outclassed the frequency-based method. Correspondingly, the amalgamation of review rating and opinion polarity resulted in higher precision than the

frequency based method [41]. Semantic polarity with feature frequency was also deployed in ranking features, and this method achieved 92% precision [55]. Ref. [52] provided two types of feature ranking, positive and negative, based on the semantic orientation and intensity (strength).

In the context of feature ranking, Refs. [44,45] utilized only feature frequency, Ref. [43] targeted feature frequency and opinion words, Ref. [52] ignored opinion and feature frequency, Ref. [55] overlooked opinion strength and opinion frequency. Therefore, the current work proposed new methods for feature ranking based on imperative parameters (i) title count, (ii) review count, (iii) accumulated opinion strength, (iv) feature frequency, and (v) opinion orientation count. These parameters are described in detail in Section 3. The current work contributes to the literature of feature ranking by providing four types of ranking, ranking by weight, feature by positive credence, ranking by negative credence, and overall ranking based on novel ranking methods.

2.3. Opinion Visualizations

The existing literature highlights a variety of visualizations that have been utilized to show consumer opinions, including bar charts, radials, pie charts, graphs and maps. Radial visualization was deployed in the opinion wheel and rose plot to present hotel customer feedback and sentiment contents from a large number of documents, respectively [24,56,57]. Graphs are used for opinion visualization and include coordinated graphs [57], line graphs and pie charts [58], positioning maps [59], comparative relation maps [16] and bar charts [52]. The contradictory comments on the 'Da Vinci Code' (bestseller and controversial novel) was visualized using a coordinated graph [57]. The positioning map [59], comparative relation map [16], and bar chart [60] provide competitive intelligence by comparing competitive product based on key features.

A scalable method 'visual summary report' for comparing several products and features at a glance was proposed in Ref. [61]. The glowing bars [62] and bars with different shapes [31] present a visual analysis of the really simple syndication RSS news feed. The treemap in Ref. [39] presents a summary of car reviews in which prominent keywords are rendered as boxes. The size of the boxes indicates the number of sentences in which a specific keyword has appeared. The color of the boxes specifies the average opinion of a keyword ranging from red to green to encode the opinion tendency (red for negative opinion and green for positive opinion). The treemap provides multi-dimensional information, such as the most common keywords, the average semantic associated with keywords, and the most positive and negative keywords.

The treemap [39] and bar chart [52] are unable to present opinion strength (that ranges from -3 to $+3$) on each feature of a target product. Therefore, the treemap is enhanced in this work to present opinion strength (that ranges from -3 to $+3$) on each feature of a target product. We selected the treemap visualization based on the finding of a usability study with 146 participants performed in our previous work to identify a suitable opinion visualization [63]. This work contributed to opinion visualization literature by proposing opinion-strength-based visualization to provide a multi-dimensional view of consumers' opinions by displaying the comparison of positive and negative opinions at various levels of ($+3$ to -3) of opinion strength and significance of a feature.

3. Proposed System

3.1. Theoretical Framework

Let document D with product reviews contain n reviews $R = [r_1, r_2, r_3, \dots, r_n]$. Every review (r_k) is comprised of a set of feature-opinion pair, which consists of a feature f_j and an opinion word OPW_o . Each feature f_j may pair with more than one opinion words in a single review or over the set of n reviews. In our proposed system, each review r_k is represented by a tuple (termed as review tuple) of two elements $[MD_{r_k}, B_{r_k}]$. The review tuple is as follows:

$$Review = r_k = [MD_{r_k}, B_{r_k}]$$

where $MD_{r_k} = [MD_{r_k}HR, MD_{r_k}Rating, MD_{r_k}Title]$, $MD_{r_k}Title = [MD_{r_k}TitleF, MD_{r_k}TitleOPW]$, $B_{r_k} = [s_1, s_2, s_3, \dots, s_p] \forall s_i \in B_{r_k}, s_i = [B_{r_k}s_if_j, B_{r_k}s_if_jSP, B_{r_k}s_if_jOS, B_{r_k}s_iContent]$

The MD represents the metadata of a review, and B represents a set of sentences in the body of reviews. Table 1 depicts the description of the abbreviation used. Each sentence s_i in B_{r_k} is represented by a proposed tuple. The proposed tuple is an extension of tuples presented in Refs. [40,58]. As shown in Figure 1, each sentence s_i contains single feature f_j . The opinion related to the feature $f_j^{s_i}$ in a sentence can be positive (OSP^{POS}) or negative (OSP^{NEG}). Opinion polarity is estimated in the range of -3 to -1 for negative and $+1$ to $+3$ for positive (three for strongest and one for weakest). The opinion with positive semantic polarity can have opinion strength strong positive (OS^{POS_S}), mildly positive (OS^{POS_M}) or weak positive (OS^{POS_W}) [52,64,65]. Similarly, the opinion with negative semantic polarity can have opinion strength strong negative (OS^{NEG_S}), mildly negative (OS^{NEG_M}) or weak negative (OS^{NEG_W}) [52,64,65]. A feature tuple is also proposed in work, as shown below.

Table 1. Notations used in Review Tuple.

Notation	Description
D	Document with n product reviews
r_k	Review k
MD_{r_k}	Metadata of review k
B_{r_k}	Body of review k
$MD_{r_k}HR$	Helpfulness Ratio of review k (HR is MD element)
$MD_{r_k}Title$	Title of review k
$MD_{r_k}Rating$	Rating of review k (Rating is MD element)
$MD_{r_k}TitleFcount$	represents the number of features in the review title
$MD_{r_k}TitleOPWcount$	depicts the number of opinion words in the review title
$B_{r_k}FCount$	number of the features in the body of the review
$B_{r_k}s_if_j$	represents a product feature (f_j) in sentence s_i
$B_{r_k}s_if_jSP$	reflects the semantic polarity (SP) of the feature f_j in sentence s_i in B_{r_k}
$B_{r_k}s_if_jOS$	reflects the opinion strength (OS) of the feature f_j in sentence s_i
f_jfreq	Frequency of the feature
$B_{r_k}s_iContent$	reflects the content of the sentence s_i
OSP^{POS}	reflects that the semantic polarity of the opinion is positive
OSP^{NEG}	reflects that the opinion semantic polarity is negative
OS^{POS_S}	reflects that the opinion strength is strong positive
OS^{POS_M}	reflects that the opinion strength is mild positive
OS^{POS_W}	reflects that the opinion strength is weak positive
OS^{NEG_S}	reflects that the opinion strength is strong negative
OS^{NEG_M}	reflects that the opinion strength is mild negative
OS^{NEG_W}	reflects that the opinion strength is weak negative
$W_{OS^{POS_S}}$	reflects the weight of OS^{POS_S} (i.e., +3)
$W_{OS^{POS_M}}$	reflects the weight of OS^{POS_M} (i.e., +2)
$W_{OS^{POS_W}}$	reflects the weight of OS^{POS_W} (i.e., +1)
$W_{OS^{NEG_S}}$	reflects the weight of OS^{NEG_S} (i.e., -3)
$W_{OS^{NEG_M}}$	reflects the weight of OS^{NEG_M} (i.e., -2)
$W_{OS^{NEG_W}}$	reflects the weight of OS^{NEG_W} (i.e., -1)

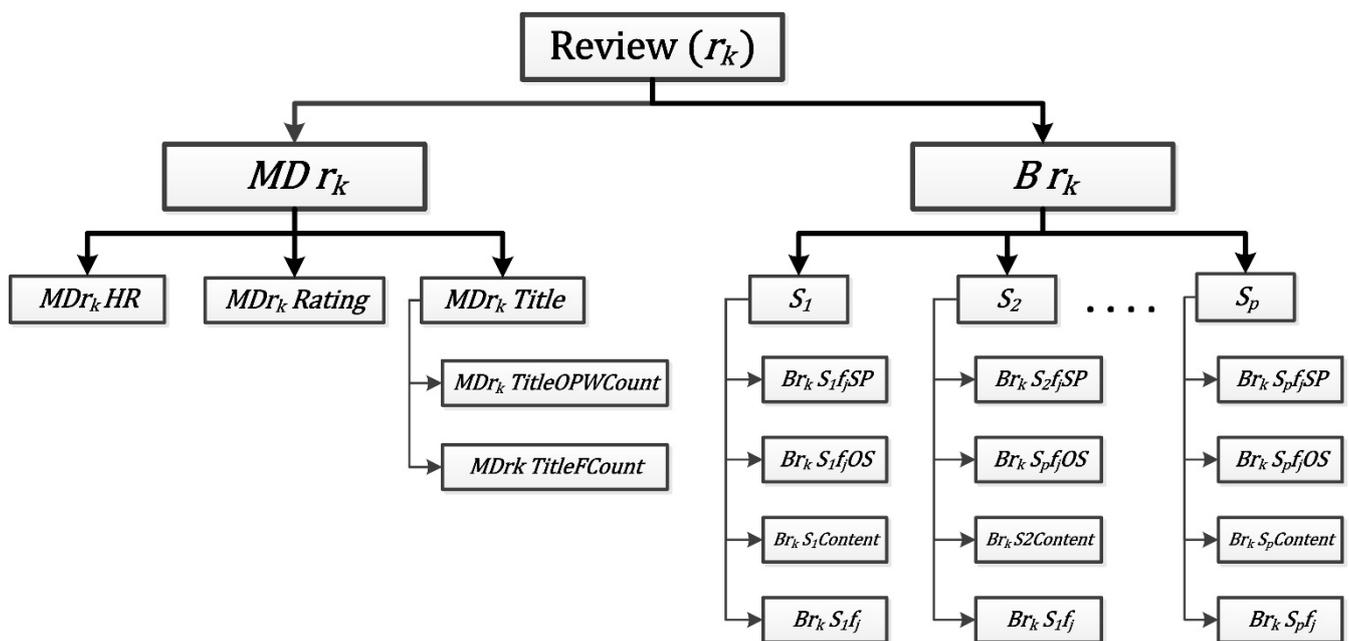


Figure 1. Hierarchy of the review tuple.

The mathematical model of the feature tuple that is part of the review tuple is shown below:

$$B_{r_k} s_{i f_j} = [s_{i f_j} \text{freq}, s_i \text{OSP}^{\text{POS}}, s_i \text{OSP}^{\text{NEG}}] \text{ where}$$

$$s_i \text{OSP}^{\text{POS}} = [OS^{\text{POS}_S}, OS^{\text{POS}_M}, OS^{\text{POS}_W}]$$

$$s_i \text{OSP}^{\text{NEG}} = [OS^{\text{NEG}_S}, OS^{\text{NEG}_M}, OS^{\text{NEG}_W}]$$

Consider the following review shown in Figure 2.

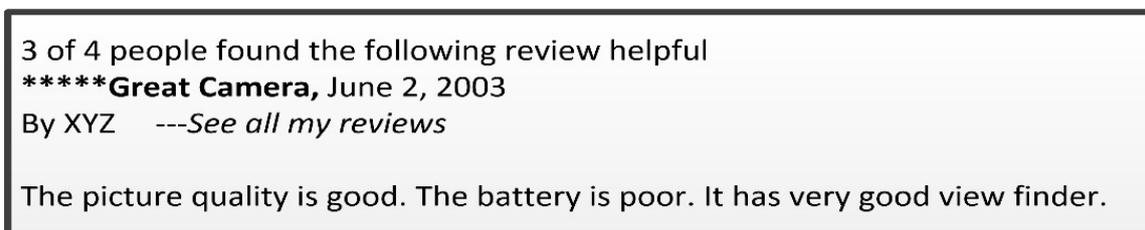


Figure 2. Example review.

The helpfulness ratio ($MD_{r_k} HR$) of the above-mentioned review is 75 ($3/4 \times 100$) with a 5-star rating. The title of the review indicates a positive opinion with ‘great’ opinion word in the title having an opinion strength strong positive (OS^{POS_S}) associated with the weight of +3 ($W_{OS^{\text{POS}_S}}$). The review presents opinions on battery, picture quality, and viewfinder features of a camera. The ‘battery’, ‘picture quality’, and ‘viewfinder’ features are described by the opinion word good, poor, and very good, respectively. The opinion word good is a positive word with the weak positive (OS^{POS_W}) opinion strength, which is associated with the weight +1 ($W_{OS^{\text{POS}_W}}$). However, the opinion word poor is a negative word with the OS^{NEG_W} strength (associated with the weight –1). The semantic orientation of the opinion word very good is positive with the opinion strength of OS^{POS_S} (where $W_{OS^{\text{POS}_S}} = 3$). The tuple of the review presented in Figure 2 is demonstrated in Figure 3.

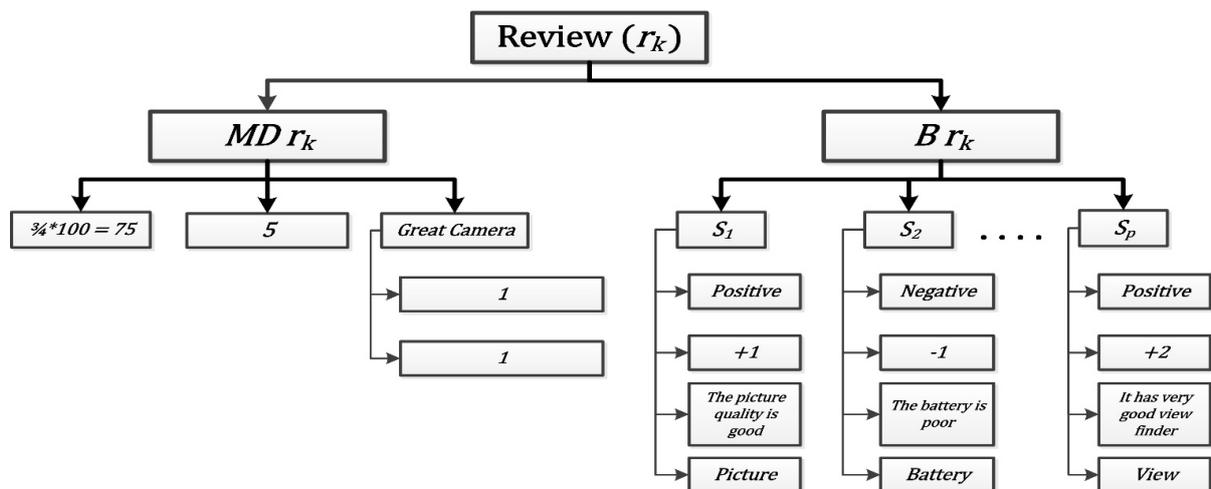


Figure 3. Review tuple based on the example review in Figure 2.

Consider the following reviews shown in Figure 4.



Figure 4. Example reviews.

Figure 4 shows three reviews containing opinions on three different features (picture quality, battery, viewfinder) of a digital camera. The resultant feature tuple of the battery feature is presented in Figure 5. The battery feature was mentioned in all of the reviews three times; therefore, its weight is three. Three opinion words (good, poor, disappointing) are associated with the battery feature. These opinion words are Weakly Positive, Weakly Negative and Mildly Negative with corresponding values of +1, -1, and -2, respectively. One positive opinion word having an opinion strength of +1 is associated with the battery feature, and hence the OSP^{POS} value of the battery is +1, while two negative words with -1 and -2 opinion strengths are connected with the feature battery, resulting in the value of OSP^{NEG} of the feature equals to -3 (-2 + -1).

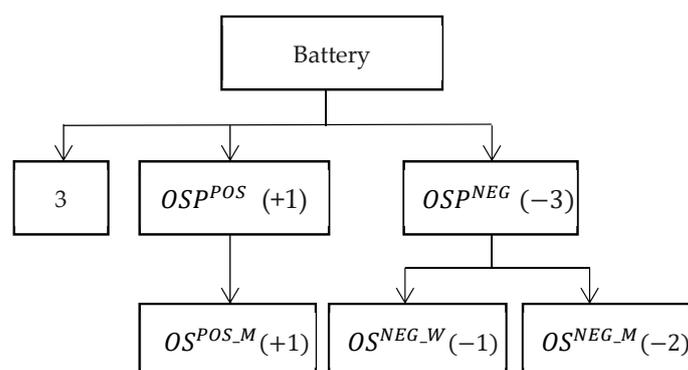


Figure 5. Feature tuple.

3.2. Architecture of the System

The proposed system consists of five components: pre-processor, feature and opinion extractor, review ranker, feature ranker, and opinion visualizer (see Figure 6). This architecture is based on a previous study [52].

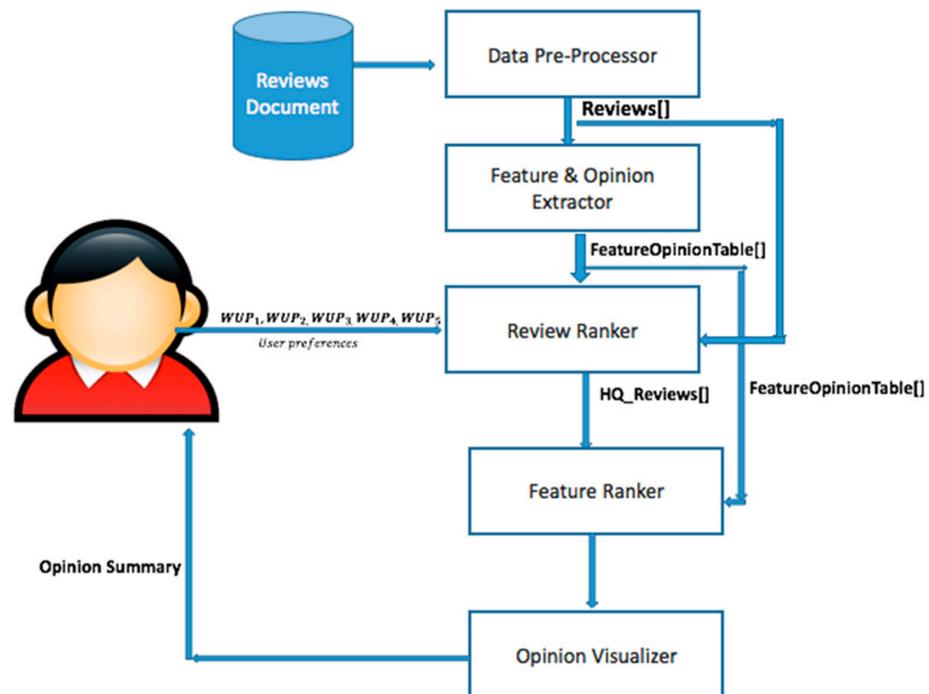


Figure 6. Architecture of proposed system.

3.2.1. Pre-Processor

A pre-processor prepares a document containing reviews for review and feature ranking. A variety of processes, including conversion of review text to lower case, removal of non-alphabetic characters, tokenization, stop word filter, spell checker, word stemming, and part of speech (POS) tagging, are performed by the component. Firstly, the text of the document is transformed into lower case. Secondly, stop words are eliminated from the document by applying a stop word filter using a defined list of stop words. Thirdly, word-stemming is performed to convert derivationally and inflectional forms of words into a base form. After that, noise is removed from the document by spell checking. Then POS tagging is employed to assign a POS category to each word in the document. In the end, the tokenization returns a list of words.

3.2.2. Feature and Opinion Extractor

Feature and opinion extractor extracts candidate features along with opinion words to generate a list of potential features. An existing study revealed that 60%–70% of the product features are represented by frequent explicit nouns [55]. The current study considered frequent nouns or noun phrases as candidate features based on the findings of existing studies [42,52,53,58]. Let us suppose there are q nouns (i.e., features) $(n_1, n_2, n_3, \dots, n_q)$ extracted from all the review tuples and stored in the list. A window based approach [41] is then utilized to extract opinion words associated with a particular feature in which opinion words discussed within K words of a particular feature are selected as associated opinion words. In contrast to existing studies [42–44,52] which extract nouns based only on feature frequency, we utilized three parameters to extract prominent features from a review document.

The noun j weight (n_j weight) is calculated based on the assumption that frequently discussed nouns in a large number of high-quality reviews associated with several opinion

words discussed in considerable review titles are significant product features. To identify potential features and associated opinion words from the list of nouns, an algorithm called feature and opinion extractor is proposed and presented in Figure 7. The inputs to the algorithm are the list of nouns (NounsList[]) extracted from reviews tuples (stored in the list Reviews[]) and review document. NounsList[] is an adjacency list [10] that can store the associated opinion words with a noun. Associated opinion words for feature n_j is searched in all reviews, and opinion words that are on a distance of K words from the selected noun are populated in the NounsList[]. The following five equations are used to calculate n_j weight. In Equation (1), the frequency of noun n_j (n_jfreq) is calculated based on its occurrence in the review document consisting of m reviews. In other words, n_jfreq in Equation (1) represents count of n_j occurrence in reviews.

$$n_jfreq = \sum_{k=1}^m countn_jOccurrence(r_k). \quad (1)$$

The number of opinion words associated with noun n_j in the whole review document is calculated in Equation (2). $n_jOPWCount$ (depicted in Equation (2)) is the number of opinion words associated with the noun n_j .

$$n_jOPWCount = \sum_{k=1}^m countDistinctOPWwithn_j(r_k). \quad (2)$$

The numbers of times noun n_j appeared in the titles of reviews is described by Equation (3) where $TitleCountwithn_j$ is the number of titles in which the noun n_j discussed. In Equation (3), the bracket value is 1 (using the inversion notation [10]), if the condition holds otherwise, it is 0 [10]. The condition $[n_jExistinReviewTitle(r_k) = True]$ in Equation (3), returns 1, if n_j exists in the review r_k title.

$$TitleCountwithn_j = \sum_{k=1}^m [n_jExistinReviewTitle(r_k) = True]. \quad (3)$$

Equation (4) computes the numbers of times noun n_j appeared in reviews.

$$ReviewCountwithn_j = \sum_{k=1}^m [n_jExistinReview(r_k) = True]. \quad (4)$$

Equation (5) shows the calculation of feature weight. The values of n_jfreq , $n_jOPWCount$, $TitleCountwithn_j$, and $ReviewCountwithn_j$ calculated in Equations (1)–(4) are utilized in the Equation (4) to calculate the weight of a feature. Therefore, our proposed method to calculate weight of a noun is based on four parameters: (i) frequency of noun n_j , (ii) the number of associated opinion words with noun n_j , and (iii) the number of times noun n_j appeared in reviews' title, and number of reviews in which n_j appeared ($ReviewCountwithn_j$). These parameters are calculated for each noun from (Lines 8–17) of the algorithm.

$$n_jweight = n_jfreq + n_jOPWCount + TitleCountwithn_j + ReviewCountwithn_j \quad (5)$$

The noun frequency n_jfreq , $n_jOPWCount$, and $TitleCountwithn_j$ are then summed up to compute the weight of a review called $n_jFinalScore$ in the algorithm (Line 20–21). This $n_jFinalScore$ is used to filter the nouns. Nouns having a $FinalScore_{n_j}$ above a threshold β are selected as potential features. After this, FeatureOpinionList[] containing the selected (most frequent) features with associated opinion words is built.

```

1 Procedure FeatureAndOpinionExtractor (NounsList[], Reviews [])
2 {
3   // Populating list with opinion words that are associated to each noun
4   For every review in the Reviews [] do
5     For every noun in the NounsList[] do
6       PopulateAssociatedOPWordsforEachNoun(Reviews[review], NounsList[noun])
7
8   For every review in the Reviews [] do
9     For every noun in the NounsList[] do
10      NounInfo[noun].njfreq ← NounInfo[noun].njfreq + FreqOfNounInReview(Reviews[review], NounsList[noun])
11      If (NounFoundInReviewTitle(Reviews[review], NounsList[noun])) then
12        NounInfo[noun].TitleCountwithnj ← NounInfo[noun].TitleCountwithnj + 1
13      If (FreqOfNounInReview(Reviews[review], NounsList[noun]) >= 1)
14        NounInfo[noun].ReviewCountwithnj ← NounInfo[noun].ReviewCountwithnj + 1
15
16      For every noun in the NounsList[] do
17        NounInfo[noun].njOPWCount ← NounWithAssociatedOPWordsCount(NounsList[noun]);
18
19      For every noun in the NounsList[] do
20        NounInfo[noun].njFinalScore ← NounInfo[noun].njfreq + NounInfo[noun].TitleCountwithnj +
21          NounInfo[noun].njOPWCount
22      // Extraction of the identified features with associated opinion words in the feature opinion table
23      For every noun in the NounsList[] do
24        If (NounInfo[noun].njFinalScore > β)
25          FeatureOpinionList[] ← ExtractNounWithAssociatedOpinionWords(NounsList[noun])
26 } |

```

Figure 7. Algorithm for identifying potential features with associated opinion words.

3.2.3. Review Ranker

The job of the review ranker is to calculate the rank of reviews stored in the review document. To calculate the rank of a review, first, the weight of each review is calculated based on five parameters. A user can define the contribution of these parameters by assigning them a weight. After assigning the weight to the reviews, the reviews are classified into five classes: (a) excellent, (b) good, (c) average, (d) fair, and (e) poor according to their weights.

To compute the class of each review stored in the review document, a ReviewRanking algorithm is proposed and presented in Figure 8. The core of the algorithm is to assign weights to reviews. The parameters used to compute the weight of a review tuple are: (i) title score (*TitleScore*), (ii) number of features in the review body ($B_{r_k}Fcount$), (iii) number of opinion words in the review body ($B_{r_k}OPWcount$), (iv) helpfulness ratio ($MD_{r_k}HR$), and (v) users' rating $MD_{r_k}rating$. The Title score depicts the sum of the number of feature and opinion words in the review title. Firstly, for each review tuple, the algorithm computes the number of features ($MD_{r_k}TitleFcount$) and opinion words ($MD_{r_k}TitleOPWcount$) appearing in the review title and then these computations are used to calculate title score (*TitleScore*) (Line 4–12). In other words, the Title score represents the sum of $MD_{r_k}TitleFcount$, and $MD_{r_k}TitleOPWcount$. Moreover, for each review tuple the number of opinion words ($B_{r_k}OPWcount$) and features ($B_{r_k}Fcount$) appearing in the body are calculated (Line 14–21). Weight of each review ($r_kWeight$) is computed based on

the values of the parameters ($B_{r_k}Fcount$, $MD_{r_k}rating$, $MD_{r_k}HR$, $B_{r_k}OPWcount$, $TitleScore$). Users' preferences are incorporated in review ranking by defining the weight of each parameter (W_{UP_1} , W_{UP_2} , W_{UP_3} , W_{UP_4} , and W_{UP_5}) by users. User preferences weights (W_{UP_1} , W_{UP_2} , W_{UP_3} , W_{UP_4} , and W_{UP_5}) and the weights assigned to $TitleScore$ (α) are presented in (Line 23-25). Title weight coefficient (α) can be adjusted depending on the size and nature of experimental data. We set the value of α to 10 based on the conclusion of Ref. [42]. Maximum weight ($MaxWeightReview$) among the m reviews is computed in Line 27. After calculating the weights of reviews, class of each review is calculated (based on review own weight and maximum weight ($MaxWeightReview$) among all reviews weights). Based on $r_kWeight$ and $MaxWeightReview$; r_k can be classified as one of the following review classes: (i) Excellent, (ii) Good, (iii) Average, (iv) Fair, and (v) Poor. We utilized these five review classes of Ref. [52] to depict the quality of each review in the review document and to distinguish high-quality reviews (HQ_reviews) from low-quality reviews to improve feature ranking. The presented scheme requires the user to decide which of the classes to be selected. The reviews that belong to selected classes (termed as high-quality reviews) are considered for feature ranking and opinion summary. For example, if the user selects classes Excellent, and Good; then all the reviews with r_kClass equals Excellent, and Good are declared as high-quality reviews.

```

1  Procedure ReviewRanking (FeatureOpinionTable[], Reviews[], W_UP1, W_UP2, W_UP3, W_UP4, W_UP5,  $\alpha$ )
2  {
3      opinionwords[] ← UniqueOpinionWords(FeatureOpinionTable [])
4      For every review in the Reviews [] do
5          For every feature in the FeatureOpinionList [] do
6              If (FeatureExistinTitle(review, feature))
7                  ReviewInfo[review].MDrkTitleFcount ← ReviewInfo[review].MDrkTitleFcount + 1
8          For every opinionword in the opinionwords [] do
9              If (OPWordExistinTitle(review, opinionword))
10                 ReviewInfo[review].MDrkTitleOPWcount ← ReviewInfo[review].MDrkTitleOPWcount + 1
11                 ReviewInfo[review].TitleScore ← ReviewInfo[review].MDrkTitleFcount +
12                 ReviewInfo[review].MDrkTitleOPWcount
13
14     For every review in the Reviews [] do
15         For every sentence in the SentencesInBody(review)
16             For every feature in the FeatureOpinionTable [] do
17                 If (FeatureExistinSentence(review, sentence))
18                     ReviewInfo[review].BrkFcount ← ReviewInfo[review].BrkFcount + 1
19             For every opinionword in the opinionwords [] do
20                 If (opinionword exist in Sentence(review, sentence))
21                     ReviewInfo[review].BrkOPWcount ← ReviewInfo[review].BrkOPWcount+ 1
22
23                 ReviewInfo[review].rkWeight ← ((W_UP1 * ReviewInfo[review].BrkFcount) +
24                 (W_UP2 * ReviewInfo[review].BrkOPWcount) + (W_UP3 * ReviewInfo[review].MDrkHR)
25                 + (W_UP4 * ReviewInfo[review].MDrkrating) + (W_UP5 * ReviewInfo[review].TitleScore *  $\alpha$ ))
26
27     MaxWeightReview ← ReviewWithMaxWeight (ReviewInfo[].rkWeight)
28     For every review in the Reviews [] do
29         ReviewInfo[review].rkClass ← (ReviewInfo[review].rkWeight / MaxWeightReview) * 5
30 }

```

Figure 8. Algorithm for identifying high-quality reviews.

Consider the review shown in Figure 9 as an example. In this review, features are highlighted in red while opinion words are highlighted in green. The $TitleScore$ of the review is

two as the title of the review comprises of one feature (Picture) and one opinion word (Brilliant). There are four features in the body of the review (picture quality, viewfinder, zoom, battery), and as a result, the $B_{r_k}FCCount$ score of the review is four. Four opinion words (excellent, poor, good and fantastic) are expressed in the review, so the $B_{r_k}OPWCount$ of the review is four. Putting the values of the $r_kWeight$ computation equation results in 25.4. In the calculation of the $r_kWeight$ of the review r_k shown in Figure 9, assuming the value of 0.20 for all preferences (W_{UP_1} , W_{UP_2} , W_{UP_3} , W_{UP_4} , and W_{UP_5}).

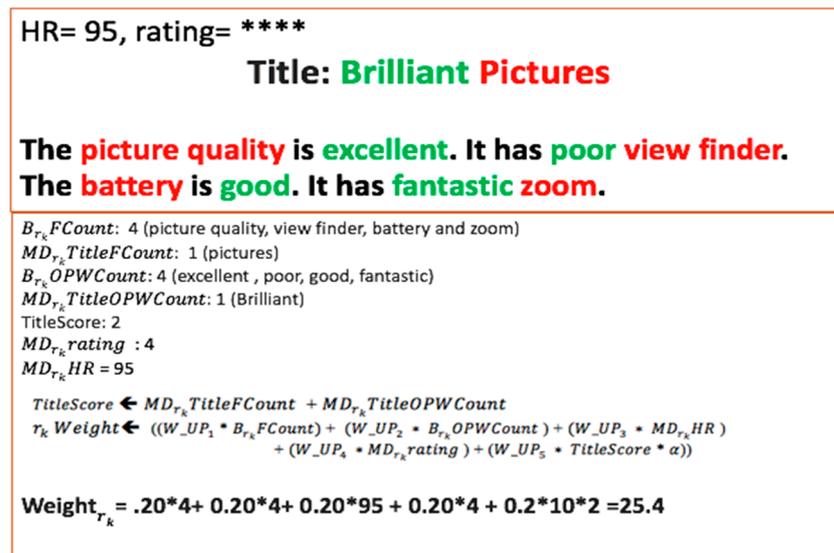


Figure 9. An example of review weight calculation.

3.2.4. Feature Ranker

After discarding low-quality reviews (with $Class_{r_k}$ below a certain threshold θ specified by user) among n reviews, we are left with m high-quality reviews (HQ-Reviews[]). The feature ranker ranks the extracted features (FeatureOpinionList[]) by utilizing high-quality reviews provided by the review ranker. In contrast to Ref. [52], we are enhancing the feature ranking by incorporating opinion and feature frequency along with opinion strength and orientation. The proposed feature ranker computes four rankings for every feature f_j based on the information presented in high-quality reviews: (i) feature f_j weight ($f_jweight$), (depicted in Equation (10)); (ii) positive credence of f_j ($f_jPOSCred$), (depicted in Equation (12)); (iii) negative credence ($f_jNEGCred$), (depicted in Equation (14)); and (iv) overall credence (f_jRank), (depicted in Equation (15)). An algorithm is proposed and presented in Figure 10 to calculate these ranking of a feature.

The $f_jweight$ is calculated based on the idea that frequently discussed features in a large number of high-quality reviews associated with many opinion words that appeared in substantial reviews' titles are decisive product features. Therefore, the value of $f_jweight$ is calculated using four parameters; (i) count of f_j occurrence in high-quality reviews (f_jfreq), (ii) the number of opinion words associated with the f_j ($f_jOPWCount$), (iii) number of reviews which discussed the feature in title or body ($ReviewCountwithf_j$), and (iv) the number of reviews' titles that contains the feature ($TitleCountwithf_j$). These parameters are computed using Equations (6)–(9), respectively as shown (Lines 3–10) of the algorithm. Moreover, $TitleCountwithf_j$ and $RevCountwithf_j$ are exploited in calculating the weight of a feature f_j on the ground that if a feature f_j is discussed in many reviews and titles, then it is significant. The calculation of $f_jweight$ is depicted in Equation (10).

$$f_jfreq = \sum_{k=1}^m countf_jOccurence(r_k), \tag{6}$$

$$f_jOPWCount = \sum_{k=1}^m countDistinctOPWwithf_j(r_k), \tag{7}$$

$$ReviewCountwithf_j = \sum_{k=1}^m [f_j ExistinReview(r_k) = True], \tag{8}$$

$$TitleCountwithf_j = \sum_{k=1}^m [f_j ExistinReviewTitle(r_k) = True], \tag{9}$$

$$f_j weight = f_j freq + f_j OPWCount + ReviewCountwithf_j + TitleCountwithf_j. \tag{10}$$

```

1  Procedure FeatureRanking (FeatureOpinionList [], HQ-Reviews [])
2  {
3      For every feature in the FeatureOpinionList [] do
4          For every review in the HQ_Reviews [] do
5              If (feature exist in Title(review))
6                  FeatureInfo[feature]. TitleCountwithf_j ← FeatureInfo[feature]. TitleCountwithf_j +1
7              If (feature exist in Body(review))
8                  FeatureInfo[feature]. BodyCountwithf_j ← FeatureInfo[feature]. BodyCountwithf_j +1
9              If (featureExistInReview(feature))
10                 FeatureInfo[feature]. ReviewCountwithf_j ← FeatureInfo[feature]. ReviewCountwithf_j +1
11
12             For every sentence in the SentencesInBody(review)
13                 If (FeatureExistInSentence (feature, review, sentence))
14                     opinionWord ← OWAssociatedWithFeatureInSentence(feature, sentence)
15                     If (StrongPositiveOW(opinionword))
16                         FeatureInfo[feature]. OSPOS,S ← FeatureInfo[feature]. OSPOS,S + 1
17                     If (MildPositiveOW(opinionword))
18                         FeatureInfo[feature]. OSPOS,M ← FeatureInfo[feature]. OSPOS,M + 1
19                     If (WeakPositiveOW(opinionword))
20                         FeatureInfo[feature]. OSPOS,W ← FeatureInfo[feature]. OSPOS,W + 1
21                     If (StrongNegativeOW(opinionword))
22                         FeatureInfo[feature]. OSNEG,S ← FeatureInfo[feature]. OSNEG,S + 1
23                     If (MildNegativeOW(opinionword))
24                         FeatureInfo[feature]. OSNEG,M ← FeatureInfo[feature]. OSNEG,M + 1
25                     If (WeakNegativeOW(opinionword))
26                         FeatureInfo[feature]. OSNEG,W ← FeatureInfo[feature]. OSNEG,W + 1
27
28                 FeatureInfo[feature]. Countf_jOSPPOS ← FeatureInfo[feature]. OSPOS,S + FeatureInfo[feature]. OSPOS,M +
29                     FeatureInfo[feature]. OSPOS,W
30
31                 FeatureInfo[feature]. Countf_jOSPNEG ← FeatureInfo[feature]. OSNEG,S + FeatureInfo[feature]. OSNEG,M +
32                     FeatureInfo[feature]. OSNEG,W
33
34                 FeatureInfo[feature]. f_j freq ← FeatureInfo[feature]. TitleCountwithf_j + FeatureInfo[feature]. BodyCountwithf_j
35
36                 FeatureInfo[feature]. f_j OPWCount ← CountOPWordsAssociatedWithFeature(feature)
37
38                 FeatureInfo[feature]. f_j weight ← FeatureInfo[feature]. f_j freq + FeatureInfo[feature]. TitleCountwithf_j +
39                     FeatureInfo[feature]. ReviewCountwithf_j + FeatureInfo[feature]. f_j OPWCount
40
41                 FeatureInfo[feature]. f_j POSCred ← FeatureInfo[feature]. Countf_jOSPPOS + W_OSPOS,S * FeatureInfo[feature]. OSPOS,S +
42                     W_OSPOS,M * FeatureInfo[feature]. OSPOS,M + W_OSPOS,W * FeatureInfo[feature]. OSPOS,W
43
44                 FeatureInfo[feature]. f_j NEGCred ← FeatureInfo[feature]. Countf_jOSPNEG + W_OSNEG,S * FeatureInfo[feature]. OSNEG,S +
45                     W_OSNEG,M * FeatureInfo[feature]. OSNEG,M + W_OSNEG,W * FeatureInfo[feature]. OSNEG,W
46
47                 FeatureInfo[feature]. f_j Rank ← FeatureInfo[feature]. f_j POSCred - FeatureInfo[feature]. f_j NEGCred
48     }

```

Figure 10. An algorithm for feature ranking.

Count_{f_j}OSP^{POS} in Equation (11) represents the count of positive opinions for the feature f_j in all sentences of m reviews. In Equation (11), the condition $[s_k f_j OSP = OSP^{POS}]$ returns to 1 if $s_k f_j OSP$ is positive, otherwise it would return zero. In Equation (12), the condition $[s_k f_j OS = OS^{POS_M}]$ returns to 1 if the feature opinion strength in sent s_k is OS^{POS_M} , otherwise it would return zero. Moreover, $\sum_{r=1}^m \sum_{k=1}^{sentences(r)} OS^{POS_M} [s_k f_j OS = OS^{POS_M}]$ sums the opinion strength OS^{POS_M} the number of times it appears in the sentences of all m reviews. Positive credence ($f_j POSCred$) of a feature f_j in Equation (12), denotes the number of positive opinion words used to describe a feature f_j and the accumulated strength of

associated positive opinion words. The larger the value of $f_jPOSCred$ denotes that the feature f_j was discussed more positively and many times.

$$Count f_jOSP^{POS} = \sum_{r=1}^m \sum_{k=1}^{sentences(r)} [s_k f_jOSP = OSP^{POS}], \tag{11}$$

$$f_jPOSCred = Count f_jOSP^{POS} + \sum_{r=1}^m \sum_{k=1}^{sentences(r)} W_{OS^{POS_S}} * [s_k f_jOS = OS^{POS_S}] + \sum_{r=1}^m \sum_{k=1}^{sentences(r)} W_{OS^{POS_M}} * [s_k f_jOS = OS^{POS_M}] + \sum_{r=1}^m \sum_{k=1}^{sentences(r)} W_{OS^{POS_W}} * [s_k f_jOS = OS^{POS_W}]. \tag{12}$$

Equation (13) depicts the total number of occurrences of negative opinions of a feature f_j in the body of m high-quality reviews. $f_jNEGCred$ in Equation (14) reflects the number of negative opinion words used to describe a feature and the total strength of these negative opinion words. The idea behind $f_jNEGCred$ is that the rank of a feature f_j should be higher than other features if f_j is associated with more negative words. The high value of $f_jNEGCred$ indicates that the features are discussed negatively by a large number of users.

$$Count f_jOSP^{NEG} = \sum_{r=1}^m \sum_{k=1}^{sentences(r)} [s_k f_jOSP = OSP^{NEG}], \tag{13}$$

$$f_jNEGCred = Count f_jOSP^{NEG} + \sum_{r=1}^m \sum_{k=1}^{sentences(r)} W_{OS^{NEG_S}} * [s_k f_jOS = OS^{NEG_S}] + \sum_{r=1}^m \sum_{k=1}^{sentences(r)} W_{OS^{NEG_M}} * [s_k f_jOS = OS^{NEG_M}] + \sum_{r=1}^m \sum_{k=1}^{sentences(r)} W_{OS^{NEG_W}} * [s_k f_jOS = OS^{NEG_W}]. \tag{14}$$

$Count f_jOSP^{NEG}$ is subtracted from $f_jPOSCred$ of a feature f_j to obtain an overall rank (f_jRank) as shown in Equation (15).

$$f_jRank = f_jPOSCred - f_jNEGCred. \tag{15}$$

The count of $OS^{POS_S}, OS^{POS_M}, OS^{POS_W}, OS^{NEG_S}, OS^{NEG_M}, OS^{NEG_W}$ in each review is computed (Lines 12–25). $f_jWeight$ is computed in Line 38 and f_jRank is computed in Line 47.

3.2.5. Opinion Visualizer

An extensive literature review followed by a usability study with 146 participants was performed by the authors in their previous work to identify a suitable opinion visualization [63]. In Ref. [63], a questionnaire survey was performed to get feedback from users about existing opinion visualizations. Users’ preferred visualization (tree map [39]) is adapted for the current study based on the findings of a previous study. The proposed visualization provides a multi-dimensional view of consumer opinions. The proposed visualization is discussed in the result section.

4. Evaluation of Proposed System

4.1. Dataset

Python 2.7 using a natural language toolkit (NLTK) was used to implement the proposed system. For the evaluation of the proposed system, experiments were performed on a real dataset (from amazon.com) utilized by Refs. [52,53,64,66,67]. The dataset contains user reviews of five digital devices, as shown in Table 2.

Table 2. Detailed information of dataset.

	Product Type	Product Name	Number of Reviews	Number of Sentences	Length in Words	Length in Characters
1.	Digital Camera 1	Canon G3	45	597	11,280	48,714
2.	Digital Camera 2	Nikon Coolpix 4300	34	346	6749	29,763
3.	Cellular Phone	Nokia 6610	44	546	9681	42,795
4.	MP3 Player	Creative Labs Nomad Jukebox Zen Xtra 40 GB	95	1716	12,719	54,872
5.	DVD Player	Apex AD2600 Progressive-scan DVD player	100	740	32,553	138,301
	Total		318	489	72,982	314,445

The evaluation of the proposed system was performed by computing the accuracy of review quality evaluation, $f_jPOSCred$, $f_jNEGCred$, and f_jRank .

Manually calculated class (actual class) is compared with a system generated class (extracted class) to calculate the accuracy of reviews, as shown in the formula below:

$$Accuracy = \frac{Extracted\ Value}{Actual\ Value} * 100 \quad (16)$$

The accuracy reveals how accurate the proposed review ranking scheme is in calculating the review quality class. Correspondingly, the actual values of $f_jPOSCred$, $f_jNEGCred$ and f_jRank are compared with extracted values to find the accuracy of the proposed feature ranking scheme. An example of accuracy calculation of $f_jPOSCred$, $f_jNEGCred$, and f_jRank is given in Table 3.

Table 3. Calculation of accuracy.

Metrics	Picture Quality
Actual $f_jPOSCred$	15
Extracted $f_jPOSCred$	12
Accuracy of $f_jPOSCred$	$12/15 * 100 = 80\%$
Actual $f_jNEGCred$	10
Extracted $f_jNEGCred$	9
Accuracy of $f_jNEGCred$	$9/10 * 100 = 90\%$
Actual f_jRank	5
Extracted f_jRank	3
Accuracy of f_jRank	$3/5 * 100 = 60\%$

4.2. Results and Discussion

4.2.1. Review Quality Classification

The classification of the review quality of 'digital camera 1' shows mixed quality reviews (Table 4). The majority of the reviews are classified as good, presenting sufficient opinions on digital camera 1. It is interesting to note that only a few reviews are labelled as excellent. Furthermore, 64% of the reviews belong to 'good' and 'average' reviews, delivering ample opinions on different features of digital camera 1. Only 14 reviews out of 45 (31%) were found to be 'fair' and 'poor'. The review quality classes of 'DVD Player' are illustrated in Table 4. Forty-four percent of the reviews were collectively classified as 'Excellent' and 'Good'. However, many reviews belong to the poor class, showing the low quality of these reviews. Notably in Table 4, 58% of reviews are categorized as the top three classes (Excellent, Good, Average) of review quality.

Table 4. Review quality evaluation of Digital Camera 1 and DVD Player.

Review Classes	Digital Camera 1	DVD Player
Excellent	3	6
Good	17	37
Average	11	14
Fair	5	6
Poor	9	36

The average accuracy of the review classification of all five products is presented in Figure 11. The system accomplished greater than 80% accuracy for all products. Further, the system achieved an average accuracy of 85% for all products.

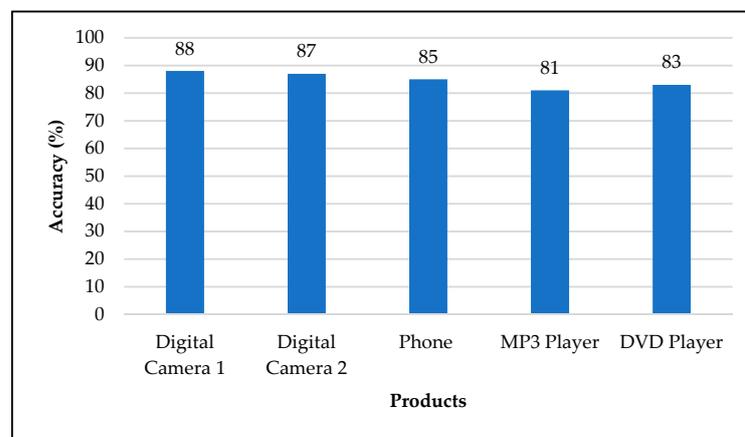


Figure 11. Accuracy of Reviews Quality Classification.

4.2.2. Feature Ranking

This section reports the $f_jPOSCred$, $f_jNEGCred$, and f_jRank of the data files along with the accuracy achieved by the System in Table 5. The $f_jPOSCred$, $f_jNEGCred$ and f_jRank of each feature f_j were computed using Equations (12), (14), and (15), respectively, given in Section 3.2.4. The accuracy of the feature ranking scheme was calculated using Equation (16) given in Section 4.1. Due to word limits, only the results of DVD Player are presented here.

Table 5. POSCred, NEGCred, and OverallCred of top ten features of DVD Player.

POSCred			NEGCred			Overall Cred		
Features	Weight	Accuracy	Features	Weight	Accuracy	Features	Weight	Accuracy
Player	144	87	Player	196	91	Feature	23	100
Play	31	90	Play	35	81	Price	17	61
Price	28	61	Picture	27	69	Work	7	71
Feature	23	100	Apex	22	100	Product	3	67
Apex	14	93	Quality	11	58	Unit	−3	100
Picture	13	77	Video	9	64	Service	−4	100
Work	8	88	Disc	8	67	Play	−7	58
Product	7	100	Button	7	100	Button	−7	100
Unit	4	100	Unit	7	100	Disc	−8	67
Service	0	100	Product	4	80	Apex	−9	89

The top ten features of ‘DVD Player’ are highlighted in Table 5 according to the positive credence ($f_jPOSCred$). The ‘Player’ received the highest $f_jPOSCred$, indicating its appreciation by many users. The next three features (Play, Price, Feature) show users’ endorsement with positive ranks of 31, 28, and 23, respectively. The features Apex, Picture, Work, ‘Product’, and ‘Unit’ are also acknowledged positively by some users. The accuracy of the top 10 features of DVD Player according to the $f_jPOSCred$ are shown in Table 5. The accuracy of features ‘Product’, ‘Unit’, ‘Service’, and ‘Feature’ was found to be 100%. Moreover, another four features, ‘Player’, ‘Play’, ‘Apex’, and ‘Work’, achieved accuracy of 86%, 90%, 92%, and 87.5%, respectively. Further, the accuracy of only 60% belongs to the feature ‘Feature’, resulting in an average accuracy of 90%.

The top ten features of DVD Player are shown in Table 5 according to the $f_jNEGCred$ (negative credence). DVD Player received $f_jNEGCred$ of 196, indicating its inadequacy. Users also disapproved the ‘Play’, ‘Picture’, ‘Apex’ and ‘Quality’ features of DVD Player, as indicated by their larger $f_jNEGCred$. The features, namely ‘Video’, ‘Unit’, ‘Disc’, ‘Button’, and ‘Product’, also were negatively discussed by some users. The accuracy of the top 10 features of DVD Player is shown in Table 5, according to the $f_jNEGCred$. Three features (apex, button, unit) achieved 100% accuracy. ‘Player’, ‘Play’, and ‘Product’ features showed more than 80% accuracy resulting in overall accuracy of 81%. However, the accuracy of one feature ‘Quality’ is only 58%.

The top ten features of DVD Player are highlighted in Table 5, according to the f_jRank . The top four features, namely ‘Feature’, ‘Price’, ‘Work’, ‘Product’, have positive f_jRank , describing users’ satisfaction about these features. Conversely, the negative f_jRank score of features Unit’, ‘Service’, ‘Play’, ‘Button’, ‘Disc’, and ‘Apex’ illustrate users’ dissatisfaction. The accuracy of DVD Player’s top ten features, according to the O_{rank} , is shown in Table 5, illustrating four features achieved 100% accuracy (feature, unit, service, button). However, the average accuracy of the system was found to be 81%.

4.3. Comparison of Proposed System with FBS System and Opinion Analyzer

We compared the results of the proposed system with two state-of-the-art systems, namely, the opinion analyzer (our previous work that is enhanced in the current study) [52] and the FBS system [53]. These systems are selected for the comparison, as the dataset utilized is same. In addition, the objectives of these systems are feature ranking based on consumers’ opinions. It is notable that the top ten features of Digital Camera 1 of the proposed system and opinion analyzer are different as the methods used to extract features in both systems are different. To compare these systems, firstly, we extracted the common features from top ten features of these systems according to positive and negative ranks. There are eight and nine common features in positive and negative rank, respectively. Secondly, we compared the accuracy of these common features for positive and negative ranks separately, as shown in Figures 12 and 13. The average accuracy of the proposed system (95%) is slightly better than opinion analyzer (92%) for positive rank (Table 6). However, the proposed system showed a little degradation on average accuracy for negative rank (Table 6). This is might due to the fact that we utilized more parameters for feature extraction as compared to opinion analyzer.

Table 6. Comparison of average accuracy of positive and negative ranks.

	Proposed System	Opinion Analyzer
Average Accuracy for Positive Rank	95	93
Average Accuracy for Negative Rank	94	96
Average Accuracy for Negative and Positive Rankings	95	95

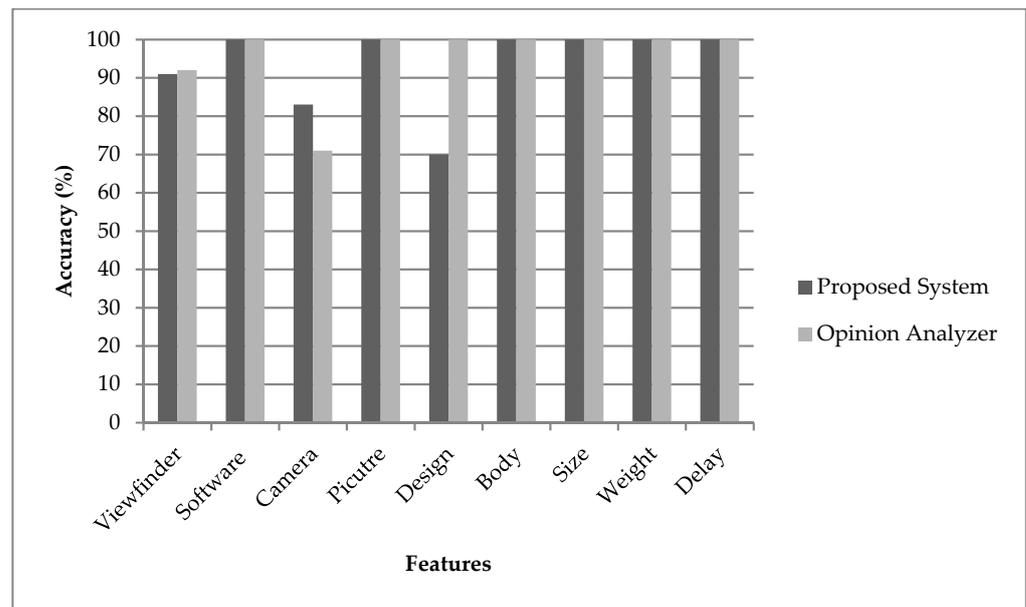


Figure 12. Comparison of proposed system with opinion analyzer on positive ranking.

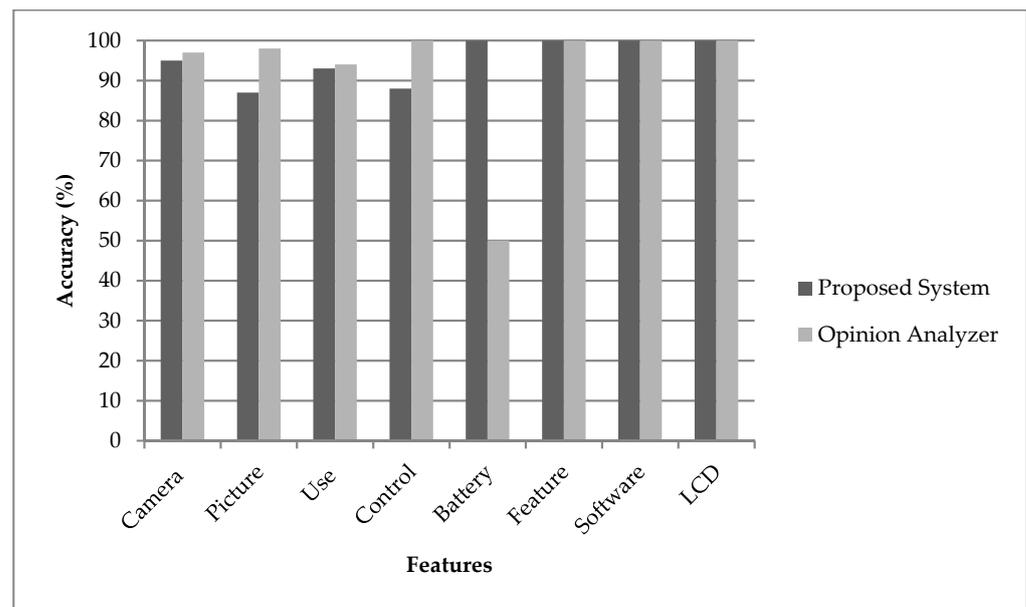


Figure 13. Comparison of proposed system with opinion analyzer on negative ranking.

Similarly, we compared the average accuracy of the top ten features of five products based on the positive and negative credence’s with the accuracy of the FBS system [53]. The proposed system outclassed FBS system on the accuracy of four products (cellular phone, Digital Camera 1, MP3 Player, DVD Player), as shown in Figure 14. In the case of Digital Camera 2, the proposed exhibited a little accuracy deprivation; however, the proposed system surpassed FBS based on average accuracy. Table 6 shows the average accuracy achieved by the proposed system and opinion analyzer for positive and negative ranks. It can be seen that the average accuracy for positive and negative ranks are same for both systems.

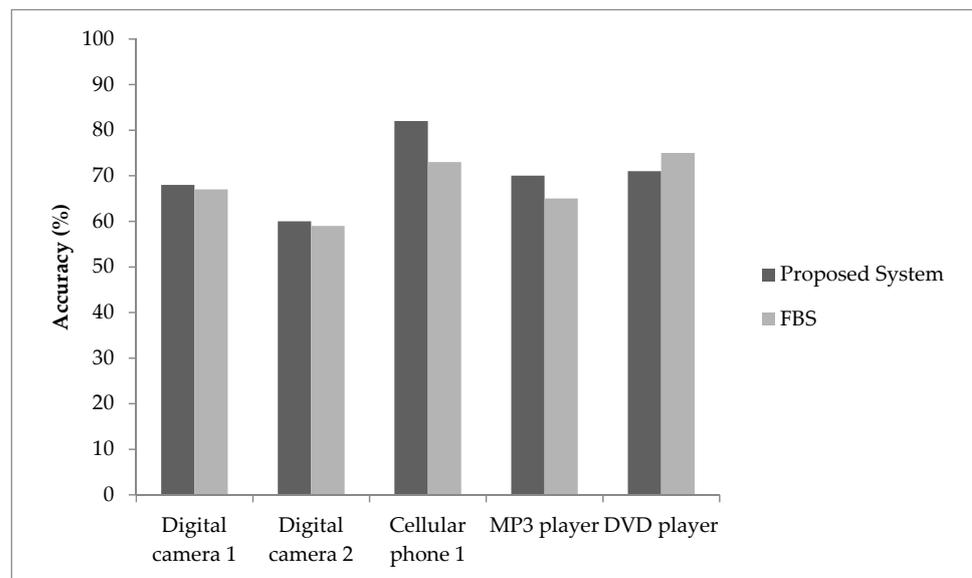


Figure 14. Comparison of average accuracy FBS and proposed system.

4.4. Opinion Visualizer

In this work, due to space constraints, we are presenting the opinion summary of Digital Camera 1 only. The proposed tree map visualization is shown in Figure 15. The tree map consists of ten rectangles. Each rectangle represents one feature. The weight of a feature is depicted by the size of the rectangle. Each rectangle is further divided into various sections according to opinion orientation and strength. Positive and negative opinions on a feature are expressed by the rectangle at 6 levels: three for positive (weakly positive, mildly positive, strongly positive) and three for negative (weakly negative, mildly negative, and strongly negative), using different shades of red and green colors. Figure 16 shows the color scheme used in the tree map. The proposed tree map presents the comparison of opinions at six levels of opinion strength as compared to the tree map of Reference [39].

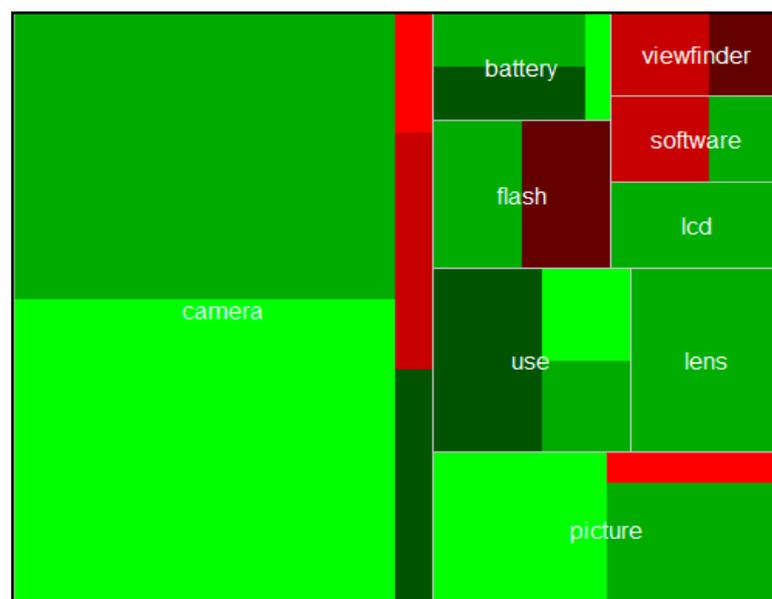


Figure 15. Proposed Tree Map Visualization of Digital Camera 1.

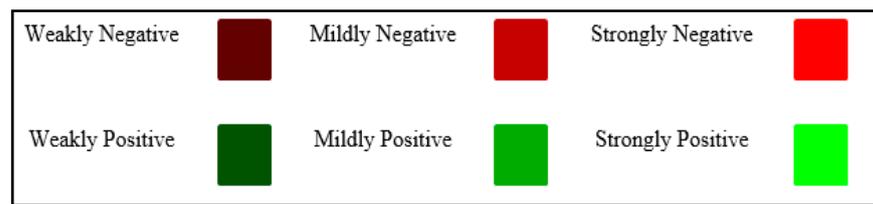


Figure 16. Color scale.

A large number of users discussed about the camera as shown by the size of camera rectangle in Figure 15. Two types of negative opinions (strongly negative and mildly negative) expressed on the camera. However, users appreciated the camera by stating strongly, mildly, and weakly positive opinions. The second feature is picture according to the size of rectangle (weight). It received only three types of opinions: strongly positive, mildly positive, and strongly negative. The features ‘Battery’ and ‘Use’ acknowledged by only positive opinions. On the other hand, ‘Viewfinder’ of the camera is discussed negatively with mildly negative or weakly negative comments. Only mildly positive opinions were expressed by users on the features ‘LCD’ and ‘Lens’. The features ‘Software’ and ‘Flash’ of Digital Camera 1 were considered both positively and negatively by the users. The overall opinion of users on Digital Camera 1 found to be positive.

Case Study

The proposed opinion-strength-based visualization was evaluated by conducting a usability study. The aim of the usability study is to identify the effectiveness and usefulness of the visualization. A total of ten participants (6 Male, 4 Female) was participated in the study. At first, the concepts of the proposed visualization were presented to the participants. After that the participants were asked to provide their feed back about user-friendliness, visual appeal, informativeness, understandability, and intuitiveness of the visualization. A 5-point Likert scale (Strongly Disagree to Strongly Agree) was utilized to get the feedback. Figure 17 demonstrates the result of the usability study.

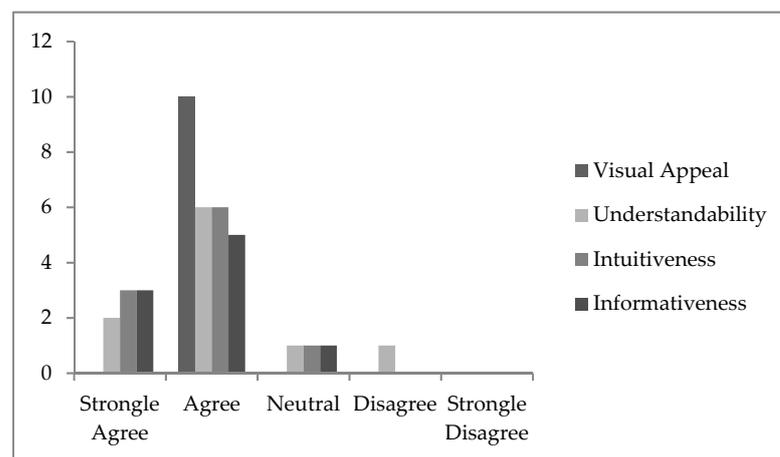


Figure 17. Result of usability study.

None of the participants strongly disagreed with the visual appeal, understandability, intuitiveness, and informativeness. Figure 17 shows that most of the participants reported strong agreement or agreement on the usability of the proposed visualization. The use of a color scale to increase the understanding of the visualization was suggested by two participants. The suggestion was incorporated. Another suggestion provided by many participants is to increase the width of borders and this suggestion is amalgamated. Last modification was done is the increase of font size based on the results of the usability study.

5. Conclusion, Limitation and Future work

In this paper, authors proposed novel ranking schemes for users' reviews and product features along with an opinion-strength-based visualization to present users high quality information from massive reviews. The focus is to improve existing ranking schemes of reviews and features by incorporating users' preferences with enhanced parameters set that are not considered in previous studies. In contrast to existing opinion mining system, the proposed system integrates review ranking and feature ranking by utilized only high-quality reviews based on users' preferences for feature ranking that result in enhanced product feature ranking. First, the information overload problem (selecting high quality reviews) was addressed by proposing a new review ranking scheme. Second, a new scheme for feature ranking based on an enhanced parameter set was proposed. Third, binary classification-based visualization was improved by the introduction of opinion-strength-based visualization that present users' opinions on critical product features at multiple levels according to opinion intensity. Four, the accuracy of the system is accessed using a real dataset of 332 reviews of five products from amazon.com. Finally, a usability study is performed to evaluate the quality of the proposed visualization. Our results show an average accuracy of 85% for review quality classification. Moreover, the results of Digital Camera 1 and DVD Player show five classes of reviews, presenting the insight about the quality of reviews. Player, play, and feature are found to be the top three features of the DVD Player according to positive credence, whereas player, play, and pictures are the top three features according to negative credence. The proposed system achieved promising results over existing systems. The study has some limitations as the system is evaluated on 332 reviews of one domain (electronic product). Future research should target more products having a large number reviews from different domains.

Author Contributions: Conceptualization, A.S., M.A.Q. and M.L.; Data curation, F.J. and M.B.; Formal analysis, M.L. and Y.Z.J.; Funding acquisition, M.A.; Investigation, Y.Z.J.; Resources, M.A.; Software, M.B.; Supervision, M.A.Q. and M.A.; Visualization, M.B. and Y.Z.J.; Writing—original draft, A.S., M.A.Q., F.J. and M.L.; Writing—review & editing, M.B. and M.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the National Research Foundation of Korea grant funded by the Korean Government (2020R1G1A1013221).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The online customer reviews used in this study are taken from the publicly available [Amazon.com](https://www.amazon.com) dataset.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Li, G.; Wang, H.; Hardjawana, W. New advancement in information technologies for industry 4.0. *Enterp. Inf. Syst.* **2020**, 402–405. [[CrossRef](#)]
2. Khan, S.S.; Khan, M.; Ran, Q.; Naseem, R. Challenges in Opinion Mining, Comprehensive Review. *Sci. Technol. J.* **2018**, 33, 123–135.
3. Reyes-Menendez, A.; Saura, J.R.; Thomas, S.B. Exploring key indicators of social identity in the# MeToo era: Using discourse analysis in UGC. *Int. J. Inf. Manag.* **2020**, 54, 102129.
4. Na, J.-C.; Thet, T.T.; Khoo, C.S.G. Comparing sentiment expression in movie reviews from four online genres. *Online Inf. Rev.* **2010**, 34, 317–338. [[CrossRef](#)]
5. Al-Natour, S.; Turetken, O. A comparative assessment of sentiment analysis and star ratings for consumer reviews. *Int. J. Inf. Manag.* **2020**, 54, 102132. [[CrossRef](#)]
6. Lo, Y.W.; Potdar, V. A review of opinion mining and sentiment classification framework in social networks. In Proceedings of the 3rd IEEE International Conference on Digital Ecosystems and Technologies, Istanbul, Turkey, 1–3 June 2009; pp. 396–401.
7. Hassani, H.; Beneki, C.; Unger, S.; Mazinani, M.T.; Yeganegi, M.R. Text mining in big data analytics. *Big Data Cogn. Comput.* **2020**, 4, 1. [[CrossRef](#)]

8. Singh, R.K.; Sachan, M.K.; Patel, R.B. 360 degree view of cross-domain opinion classification: A survey. *Artif. Intell. Rev.* **2021**, *54*, 1385–1506. [[CrossRef](#)]
9. Hao, M.C.; Rohrdantz, C.; Janetzko, H.; Keim, D.A.; Dayal, U.; Haug, L.E.; Hsu, M.; Stoffel, F. Visual sentiment analysis of customer feedback streams using geo-temporal term associations. *Inf. Vis.* **2013**, *12*, 273–290. [[CrossRef](#)]
10. Rohrdantz, C.; Hao, M.C.; Dayal, U.; Haug, L.-E.; Keim, D.A. Feature-Based Visual Sentiment Analysis of Text Document Streams. *ACM Trans. Intell. Syst. Technol.* **2012**, *3*, 1–25. [[CrossRef](#)]
11. Bilal, M.; Gani, A.; Lali, M.I.U.; Marjani, M.; Malik, N. Social profiling: A review, taxonomy, and challenges. *Cyberpsychol. Behav. Soc. Netw.* **2019**, *22*, 433–450.
12. Chevalier, J.A.; Mayzlin, D. The Effect of Word of Mouth on Sales: Online Book Reviews. *J. Mark. Res.* **2006**, *43*, 345–354. [[CrossRef](#)]
13. Moghaddam, S.; Ester, M. ILDA: Interdependent LDA Model for Learning Latent Aspects and their Ratings from Online Product Reviews Categories and Subject Descriptors. In Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval-SIGIR '11, Beijing, China, 24–28 July 2011; pp. 665–674.
14. Lu, Y.; Tsaparas, P.; Ntoulas, A.; Polanyi, L. Exploiting social context for review quality prediction. In Proceedings of the 19th International Conference on World Wide Web-WWW '10, Raleigh, NC, USA, 26–30 April 2010.
15. Shamim, A.; Balakrishnan, V.; Tahir, M. Opinion Mining and Sentiment Analysis Systems: A Comparison of Design Considerations. In Proceedings of the 20th IBIMA Conference, Kuala Lumpur, Malaysia, 25–26 March 2013; pp. 1–7.
16. Xu, K.; Liao, S.S.; Li, J.; Song, Y. Mining comparative opinions from customer reviews for Competitive Intelligence. *Decis. Support Syst.* **2011**, *50*, 743–754. [[CrossRef](#)]
17. Jalilvand, M.R.; Samiei, N. The effect of electronic word of mouth on brand image and purchase intention: An empirical study in the automobile industry in Iran. *Mark. Intell. Plan.* **2012**, *30*, 460–476. [[CrossRef](#)]
18. Bilal, M.; Marjani, M.; Lali, M.I.; Malik, N.; Gani, A.; Hashem, I.A.T. Profiling Users' Behavior, and Identifying Important Features of Review "Helpfulness". *IEEE Access.* **2020**, *8*, 77227–77244. [[CrossRef](#)]
19. Dalal, M.K.; Zaveri, M.A. Opinion Mining from Online User Reviews Using Fuzzy Linguistic Hedges. *Appl. Comput. Intell. Soft Comput.* **2014**, *2014*, 735942. [[CrossRef](#)]
20. Bilal, M.; Marjani, M.; Hashem, I.A.T.; Gani, A.; Liaqat, M.; Ko, K. Profiling and predicting the cumulative helpfulness (Quality) of crowd-sourced reviews. *Information* **2019**, *10*, 295. [[CrossRef](#)]
21. Benlahbib, A. Aggregating customer review attributes for online reputation generation. *IEEE Access.* **2020**, *8*, 96550–96564. [[CrossRef](#)]
22. Xu, X. How do consumers in the sharing economy value sharing? Evidence from online reviews. *Decis. Support Syst.* **2020**, *128*, 113162. [[CrossRef](#)]
23. Wang, R.; Zhou, D.; Jiang, M.; Si, J.; Yang, Y. A survey on opinion mining: From stance to product aspect. *IEEE Access* **2019**, *7*, 41101–41124. [[CrossRef](#)]
24. Shamim, A.; Balakrishnan, V.; Tahir, M. Evaluation of Opinion Visualization Techniques. *Inf. Vis.* **2015**, *14*, 339–358. [[CrossRef](#)]
25. Ding, X.; Liu, B.; Yu, P.S.; Street, S.M. A holistic lexicon-based approach to opinion mining. In Proceedings of the International Conference on Web Search and Web Data Mining, Palo Alto, CA, USA, 11–12 February 2008; pp. 231–240.
26. Binali, H.; Potdar, V.; Wu, C. A State Of The Art Opinion Mining and Its Application Domains. In Proceedings of the IEEE International Conference on Industrial Technology (ICIT '09), Churchill, Australia, 10–13 February 2009.
27. Wu, Y.; Wei, F.; Liu, S.; Au, N. OpinionSeer: Interactive Visualization of Hotel Customer Feedback. *IEEE Trans. Vis. Comput. Graph.* **2010**, *16*, 1109–1118.
28. Moghaddam, S.; Jamali, M. ETF: Extended Tensor Factorization Model for Personalizing Prediction of Review Helpfulness Categories and Subject Descriptors. In Proceedings of the Fifth ACM International Conference on Web Search and Data Mining, Seattle, WA, USA, 8–12 February 2012; pp. 163–172.
29. Liu, J.; Cao, Y.; Lin, C.; Huang, Y.; Zhou, M. Low-Quality Product Review Detection in Opinion Summarization. In Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL), Prague, Czech Republic, 28–30 June 2007; pp. 334–342.
30. Ngo-Ye, T.L.; Sinha, A.P. Analyzing Online Review Helpfulness Using a Regression Relief-Enhanced Text Mining Method. *ACM Trans. Manag. Inf. Syst.* **2012**, *3*. [[CrossRef](#)]
31. Balaji, P.; Haritha, D.; Nagaraju, O. An Overview on Opinion Mining Techniques and Sentiment Analysis. *Int. J. Pure Appl. Math.* **2018**, *118*, 61–69.
32. Zhang, Z.; Varadarajan, B. Utility scoring of product reviews. In Proceedings of the 15th ACM International Conference on Information and Knowledge Management-CIKM '06, Arlington, VA, USA, 6–11 November 2006.
33. Kim, S.; Pantel, P.; Chklovski, T.; Pennacchiotti, M. Automatically Assessing Review Helpfulness. In Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing-EMNLP '06, Sydney, Australia, 22–23 July 2006; pp. 423–430.
34. Mudambi, S.M.; Schuff, D. What Makes a Helpful Online Review? A Study of Customer Reviews on Amazon.com. *MIS Q.* **2010**, *34*, 185–200. [[CrossRef](#)]
35. Korfiatis, N.; García-Bariocanal, E.; Sánchez-Alonso, S. Evaluating content quality and helpfulness of online product reviews: The interplay of review helpfulness vs. review content. *Electron. Commer. Res. Appl.* **2012**, *11*, 205–217. [[CrossRef](#)]

36. Walther, J.B.; Liang, Y.J.; Ganster, T.; Wohn, D.Y.; Emington, J. Online Reviews, Helpfulness Ratings, and Consumer Attitudes: An Extension of Congruity Theory to Multiple Sources in Web 2.0. *J. Comput. Commun.* **2012**, *18*, 97–112. [[CrossRef](#)]
37. Tsaparas, P.; Ntoulas, A.; Terzi, E. Selecting a comprehensive set of reviews. In Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining-KDD '11, San Diego, CA, USA, 21–24 August 2011; p. 168.
38. O'Mahony, M.P.; Smyth, B. Learning to recommend helpful hotel reviews. In Proceedings of the Third ACM Conference on Recommender Systems-RecSys '09, New York, NY, USA, 23–25 October 2009.
39. Kard, S.T.; Mackinlay, J.D.; Scheiderman, B. *Reading in Information Visualization, Using Vision to Think*; Morgan Kaufmann Publishers Inc.: Burlington, MA, USA, 1999.
40. Liu, Y.; Huang, X.; An, A.; Yu, X. Modeling and Predicting the Helpfulness of Online Reviews. In Proceedings of the 2008 Eighth IEEE International Conference on Data Mining, Pisa, Italy, 15–19 December 2008; pp. 443–452.
41. Yang, J.-Y.; Kim, H.-J.; Lee, S.-G. Feature-based Product Review Summarization Utilizing User Score. *J. Inf. Sci. Eng.* **2010**, *26*, 1973–1990.
42. Eirinaki, M.; Pisal, S.; Singh, J. Feature-based opinion mining and ranking. *J. Comput. Syst. Sci.* **2012**, *78*, 1175–1184. [[CrossRef](#)]
43. Moghaddam, S.; Ester, M. Opinion Digger: An Unsupervised Opinion Miner from Unstructured Product Reviews. In Proceedings of the 19th ACM International Conference on Information and Knowledge Management-CIKM '10, Toronto, ON, Canada, 26–30 October 2010; pp. 1825–1828.
44. Lei, Z.; Liu, B.; Lim, S.H.; Eamonn, O.-S. Extracting and Ranking Product Features in Opinion Documents. In Proceedings of the 23rd International Conference on Computational Linguistics: Posters, Beijing, China, 23–27 August 2010; pp. 1462–1470.
45. Kungpeng, Z.; Narayanan, R.; Choudhary, A. Voice of the Customers: Mining Online Customer Reviews for Product Feature-based Ranking. In Proceedings of the 3rd Conference on Online Social Networks, Boston, MA, USA, 22 June 2010.
46. Kauffmann, E.; Peral, J.; Gil, D.; Ferrández, A.; Sellers, R.; Mora, H. Managing marketing decision-making with sentiment analysis: An evaluation of the main product features using text data mining. *Sustainability* **2019**, *11*, 4235. [[CrossRef](#)]
47. Hao, M.; Rohrdantz, C.; Janetzko, H.; Dayal, U.; Keim, D.A.; Haug, L.; Hsu, M. Visual Sentiment Analysis on Twitter Data Streams. In Proceedings of the 2011 IEEE Conference on Visual Analytics Science and Technology (VAST), Providence, RI, USA, 23–28 October 2011; pp. 275–276.
48. Bilal, M.; Marjani, M.; Hashem, I.A.T.; Abdullahi, A.M.; Tayyab, M.; Gani, A. Predicting helpfulness of crowd-sourced reviews: A survey. In Proceedings of the 2019 13th International Conference on Mathematics, Actuarial Science, Computer Science and Statistics (MACS), Karachi, Pakistan, 14–15 December 2019; pp. 1–8. [[CrossRef](#)]
49. Bilal, M.; Marjani, M.; Hashem, I.A.T.; Malik, N.; Lali, M.I.U.; Gani, A. Profiling reviewers' social network strength and predicting the 'Helpfulness' of online customer reviews. *Electron. Commer. Res. Appl.* **2021**, *45*, 101026. [[CrossRef](#)]
50. Ghose, A.; Ipeirotis, P.G. Estimating the Helpfulness and Economic Impact of Product Reviews: Mining Text and Reviewer Characteristics. *IEEE Trans. Knowl. Data Eng.* **2011**, *23*, 1498–1512. [[CrossRef](#)]
51. Tsur, O.; Rappoport, A. R EV R ANK: A Fully Unsupervised Algorithm for Selecting the Most Helpful Book Reviews. In Proceedings of the Third International Conference on Weblogs and Social Media, ICWSM 2009, San Jose, CA, USA, 17–20 May 2009.
52. Shamim, A.; Balakrishnan, V.; Tahir, M.; Shiraz, M. Critical Product Features ' Identification Using an Opinion Analyzer. *Sci. World J.* **2014**, *2014*, 1–9. [[CrossRef](#)] [[PubMed](#)]
53. Hu, M.; Liu, B. Mining and summarizing customer reviews. In Proceedings of the 19th National Conference on Artificial Intelligence, San Jose, CA, USA, 25–29 July 2004; pp. 168–177.
54. Li, S.; Chen, Z.; Tang, L. Exploiting Consumer Reviews for Product Feature Ranking Categories and Subject Descriptors. In Proceedings of the 3rd Workshop on Social Web Search and Mining (SWSM'11), Beijing, China, 28 July 2011.
55. Ahmad, T.; Doja, M.N. Ranking System for Opinion Mining of Features from Review Documents. *IJCSI Int. J. Comput. Sci. Issues* **2012**, *9*, 440–447.
56. Gregory, M.L.; Chinchor, N.; Whitney, P.; Carter, R.; Hetzler, E.; Turner, A. User-directed Sentiment Analysis: Visualizing the Affective Content of Documents. In Proceedings of the Workshop on Sentiment and Subjectivity in Text, Sydney, Australia, 22 July 2006; pp. 23–30.
57. Chen, C.; Ibekwe-sanjuan, F.; Sanjuan, E.; Weaver, C. Visual Analysis of Conflicting Opinions. In Proceedings of the IEEE Symposium on Visual Analytics and Technology (2006), Baltimore, MD, USA, 31 October–2 November 2006; pp. 59–66.
58. Miao, Q.; Li, Q.; Dai, R. AMAZING: A sentiment mining and retrieval system. *Expert Syst. Appl.* **2009**, *36*, 7192–7198. [[CrossRef](#)]
59. Morinaga, S.; Yamanishi, K.; Tateishi, K.; Fukushima, T. Mining product reputations on the Web. In Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Edmonton, AB, Canada, 23–26 July 2002; pp. 341–349.
60. Liu, B.; Hu, M.; Cheng, J. Opinion observer: Analyzing and comparing opinions on the Web. In Proceedings of the 14th International Conference on World Wide Web, Chiba, Japan, 10–14 May 2005; pp. 342–351.
61. Oelke, D.; Hao, M.; Rohrdantz, C.; Keim, D.A.; Dayal, U.; Haug, L.; Janetzko, H. Visual Opinion Analysis of Customer Feedback Data. In Proceedings of the IEEE Symposium on Visual Analytics Science and Technology, Atlantic City, NJ, USA, 12–13 October 2009; pp. 187–194.
62. Gamon, M.; Basu, S.; Belenko, D.; Fisher, D.; Hurst, M.; König, A.C. BLEWS: Using Blogs to Provide Context for News Articles. In Proceedings of the International Conference on Weblogs and Social Media, Seattle, WA, USA, 30 March–2 April 2008.

63. Wanner, F.; Rohrdantz, C.; Mansmann, F.; Oelke, D.; Keim, D.A. Visual Sentiment Analysis of RSS News Feeds Featuring the US Presidential Election in 2008. In Proceedings of the Workshop on Visual Interfaces to the Social and the Semantic Web (VISSW2009), Sanibel Island, FL, USA, 8 February 2009.
64. Gamon, M.; Aue, A.; Corston-Oliver, S.; Ringger, E. Pulse: Mining customer opinions from free text. In Proceedings of the International Symposium on Intelligent Data Analysis, Madrid, Spain, 8–10 September 2005; pp. 121–132.
65. Liu, B. *Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data*; Springer: Berlin/Heidelberg, Germany, 2006.
66. Pan, S.J.; Ni, X.; Sun, J.-T.; Yang, Q.; Chen, Z. Cross-domain sentiment classification via spectral feature alignment. In Proceedings of the 19th International Conference on World Wide Web-WWW '10, Raleigh, NC, USA, 26–30 April 2010.
67. Qiu, G.; Liu, B.; Bu, J.; Chen, C. Expanding Domain Sentiment Lexicon through Double Propagation. In Proceedings of the 21st International Joint Conference on Artificial Intelligence, Pasadena, CA, USA, 11–17 July 2009.