Information Sciences 460–461 (2018) 23–41

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

Aspect-based opinion ranking framework for product reviews using a Spearman's rank correlation coefficient method

Ashok Kumar J*, Abirami S

Department of Information Science and Technology, Anna University, Chennai 600025, India

ARTICLE INFO

Article history:
Received 6 March 2017
Revised 30 April 2018
Accepted 2 May 2018
Available online 10 May 2018

Keywords:
Opinion mining
Sentiment analysis
Opinion ranking
Social media
Visualization
Spearman's rank correlation

ABSTRACT

Opinion mining (also called sentiment analysis) is a type of natural language processing for computing people's opinions and emotions. It detects opinions from structured, semi-structured, and unstructured social media contents at different levels, such as the document, word, sentence, and aspect levels. In all these levels except aspect, opinion mining identifies the overall subjectivity or sentiment polarities. An aspect level is described as a part or an attribute of an entity. It exactly describes people's likes and dislikes in social media contents. In this paper, we propose a new framework for ranking products based on aspects. First, the system identifies the aspects of products. Second, the aspects and their opinion words are identified and visualized from the products' reviews using a Harel–Koren fast multiscale layout. Third, the network visualization is constructed and modeled, and a Spearman's rank correlation coefficient based opinion ranking method is applied to rank the products based on positive and negative ranks. Fourth, the supervised learning methods (Naïve Bayes, Maximum Entropy, and Support Vector Machine) are employed for the aspect-based sentiment classification task. Finally, the performance of the system is measured by the experimental results.

1. Introduction

Opinion mining is one of the most important research areas in the fields of natural language processing, text mining, and information retrieval. It extracts subjective and objective information from structured, semi-structured, and unstructured texts. The subjective information expresses individuals', groups', and the public's opinions, such as love, joy, sadness, etc., about a product, service, government, and organization. The objective information expresses only the factual information. It does not express any personal feelings [14]. Opinion mining performs the task of classifying document polarities as positive or negative for ranking documents on specific topics of interest. In online social networking sites (OSNs), the vast amount of posts, comments, reviews, and tweets discuss products or brands [16,24]. As online social media contents increase rapidly, it becomes more difficult for the user or the product launcher to understand the valuable information. For instance, the user may be interested in buying an iPhone or laptop. The user may seek information from friends and relatives about the type of phone or laptop, configuration details, software's usability, durability, picture quality, product's company, etc. However, this information alone is not enough to make a decision to buy a product. Therefore, the user seeks more information from OSNs. As it is given more information, the user needs to filter the reviews that are relevant to the product. Analyzing these

* Corresponding author.
E-mail address: jashokkumar83@auist.net (A.K. J).

https://doi.org/10.1016/j.ins.2018.05.003
0020-0255/

reviews becomes a challenging task. Opinion mining is categorized into various levels, such as the document level, word level, aspect level, sentence level, concept level, phrase level, link-based level, clause level, and sense level.

The document level determines the sentiment polarity in a whole document. A document may contain one or more sentences. For example, the document, *"I bought a Samsung Galaxy S7 two months ago. The picture quality is amazing"* was determined as positive. The sentence level determines the sentimental polarity in each sentence. The first sentence in the above document, "*I bought a Samsung Galaxy S7 two months ago*" was determined as neutral and the second sentence "*The picture quality is amazing*" was determined as positive. The word level determines the sentiment of a single word [3]. For example, the sentence "*The dominant role of the European climate protection policy has benefits for our economy*" identifies the sentiment of the target word *policy* using the dependency parse tree. The aspect level is described as a part or an attribute of an entity. It exactly describes people's likes and dislikes in social media contents. For example, the document "*I love Samsung Galaxy S7. The picture quality is amazing*" determines the positive sentiment of the *Galaxy S7* and the *picture quality*. Concept-level sentiment analysis analyzes text beyond the word level with the use of semantic associations or semantic networks [4]. For example, the sentence "*The picture quality is amazing*" determines the concept *painting* with a positive sentiment.

The phrase level analyzes a group of two or more words [37]. It provides more contextual polarity to the sentences. A phrase may contain a noun without a verb or a verb without a noun, but it does not have a predicate. For example, take the sentence "*There is no reason at all to believe that the polluters are suddenly going to become reasonable*". The words *reason* and *reasonable* are tagged as positive in the sentimental lexicon, but the word *reason* is negated. Therefore, the phrase "*no reason at all to believe*" changes the contextual polarity of *reasonable*. The clause level analyzes a group of words that can act as a sentence, but they need not be a complete sentence. A clause can contain a subject and a verb with a predicate that tells what the subject is doing. Narayanan et al. [27] determined the sentimental polarity in a conditional sentence. For example, the clause level deals with the sentence, *if your Lenovo phone is not good, buy this great iPhone7*. The opinion holder is positive about the iPhone7 but does not express any opinion on the Lenovo phone. If there is no "if" in the first sentence, then it clearly indicates a negative opinion. The link-based level analyzes the relationship among entities. These links represent the information explicitly and implicitly. For example, take the phrases *the list of user friends in a social network* and *if many of my friends like a given product, chances are that I might like it well* [17]. A sense-level analyzes the characteristic of an entity at the basic practical knowledge level instead of the traditional word level. For example, in the sentence, "*a person may be observed to assert that God exists, but not to believe that God exists*, belief is the private sense [5]".

In all these levels, opinion mining identifies the overall subjectivity or sentimental polarities (positive, negative, and neutral) except at the aspect level. Thus, it plays an important role in identifying the potential customers for the product launcher in terms of the decision-making process. For instance, the review "*Owned it for only a few days and really enjoy driving it. Panel controls are complicated but not impossible. Will see about reliability later*" expresses the overall positive opinion at the document level and the sentence level. At the aspect level, the positive opinion is expressed about the reliability of the car. To rank the product based aspect and its opinion, a ranking mechanism needs to be used. A ranking is defined as the relationships between a set of items. It supports the individual's specific needs and influences the preferences. The aspect-based opinion ranking system addresses the needs of the individual or the product launcher. Therefore, we propose a new framework for a product ranking system based on the aspects and opinions using the rank correlation method. The remainder of this paper is organized as follows. Section 2 discusses the related works in opinion mining and aspect-based opinion rankings. Section 3 explains the proposed aspect-based opinion mining, the ranking system for products and the sentiment classification algorithms. Section 4 presents the experimental results and discussions. Section 5 presents the conclusions and future works.

2. Related works

Many ranking methods are used to mine aspects and their opinions. Eirinaki et al. [24] presented an opinion search engine with the high adjective count algorithm and the Max opinion score algorithm to analyze the sentiment at the document level and to identify the semantic orientation of components in product reviews. The search engine was built with the part-of-speech tagger module to process the text, reviews, posts, and comments for calculating the distance between nouns and adjectives. Bashir et al. [31] developed a framework for Opinion-Based Entity Ranking (OpER) to rank entities by matching the user's query keywords. In this framework, the author's features are combined by applying deep learning to the rank approach based on genetic programming (GP). The results have shown that GPRank outperforms the RankSVM and RankBoost learning methods. Lo et al. [32] presented a ranking mechanism to identify the top-k followers on Twitter with minimal annotation effort based on an index. The index was developed using a combination of supervised and semi-supervised learning methods. In this approach, the authors combined Fuzzy Matching, the Twitter LDA (Latent Dirichlet Allocation) and Support Vector Machine Ensembles (the bootstrapping SVM and the bagging SVM ensemble using TF and TF-IDF) to build seed words and training data sets. Krestel and Dokoohaki [28] presented a framework to rank product reviews with user-assigned ratings by finding latent topics (LDA), building language models (LM), and combining models and star ratings (LM + LDA). The best result was achieved with a combination of the LDA and the LM using a language model representation. In this method, the LDA models each product review as a mixture of various topics, and it learns the topics based on the co-occurrence of words in reviews (documents) during the training phase. The assignment of various topics to each review allows for the identification of similar reviews. This method adopts common sense-based topic modeling to provide topic assignments for each product review when there is no training corpus. The finding latent topics method is not dependent on a specific topic modeling algorithm. It can be replaced by common sense-based topic modeling, latent semantic indexing or random indexing using the bag-of-words model (BOW). LM computes a maximum likelihood for a document with words (the bag-of-words).

The combination of the LDA and the LM methods was performed on the user-assigned star rating, which indicates the sentiment expressed in a review instead of performing opinion mining on a word, sentence, or document. The authors had known which aspects or topics a review is about and a user's overall sentiment towards the reviewed product. In the LDA, each topic is assigned a star rating ranging from one to five and the probability that a particular topic occurs with a particular star rating in the whole corpus is determined. Then, the LM was combined with the star rating information (LDA) and a similar likelihood was computed based on the words instead of the topics. Similarly, each review is modeled as either a topic or a word for each rating class (the topic-rating model) with some overlap between the classes. For ranking reviews, the authors proposed three ranking strategies, namely, the summary-focused ranking, the sentiment-focused ranking, and the topic-focused ranking. The objective function of these strategies minimizes the Kullback–Leibler divergence of top-K rank entries and a target distribution. The LM + LDA covered both the relevant abstract topics and the relevant terms. The summary-focused ranking summarizes the sentiment in all reviews. The sentiment-focused ranking summarizes one particular class of rating. The topic-focused ranking strategy ranks the reviews based on latent topics (aspects) that encompass all the sentiments (positive, negative, and neutral). This ranking is not based on language models (LM).

Huang et al. [39] investigated dementia's disease severity prediction using a ranking and presented learning techniques for the disease severity prediction task. Marrese-Taylor et al. [9] presented an extension of Bing Liu's aspect-based opinion mining approach in tourism product reviews for sentiment classification at the aspect-level. The authors considered the new features (word rules, negation rules, and too rules) that were not covered by Bing Liu's aspect-based opinion mining approach. Xu et al. [12] designed a framework for ranking contents and time-sensitive documents in the scientific literature. The framework dynamically adjusts its random walk parameters. The results revealed that the scientific document ranking is more effective than the PageRank. Zhang et al. [7] proposed a Clustering-Based Fan-in Analysis (CBFA) to automatically recommend aspects. The lexical-based clustering approach was used for finding clusters and ranked those clusters based on the cluster fan-in analysis. Yau and Yin [35] presented a QoS-based service ranking and selection approach for determining users' satisfaction with service with a different aspect. This system selects the best service and improves the flexibility based on the QoS requirements and specifications. Zha et al. [42] proposed a framework for product aspect ranking to automatically identify important products from online reviews. In this framework, the product aspects are identified using a dependency parser and it determined the sentiments of the identified aspects. The authors achieved improved performance of product aspect ranking in real-world applications. Zhu et al [40] proposed a novel relational learning-to-rank approach to produce a summary of the important aspects in multi-documents. This method was implemented by applying the stochastic gradient descent and the greedy selection procedure for the learning process and summarization. Chutmongkolporn et al. [21] proposed a graph-based opinion entity ranking framework to rank the products with opinions in online customer reviews. The ranking algorithm was employed based on four paradigms, namely the entity-specific aspect ranking, the aspect-specific entity ranking, the entity ranking, and the reviewer ranking, to compute the ranking scores.

Zhang et al. [22] used the HITS (Hyperlink-Induced Topic Search) ranking algorithm to extract the important features and ranked those features by considering feature relevance and its frequency. The authors showed that the method shows promising results in real-life datasets. HITS is a link analysis algorithm that rates a web page based on providing information on a topic and providing links to other pages on a topic. It can be characterized in a two ways, namely, the authority and the hub. The authority (acts as feature sets) provides good information about a topic (or subject). The hub (acts as feature indicators) provides links to a good authority on a topic. A mutual enforcement relationship exists between the authority and the hub. Garcıa-Moya et al. [23] presented a method to identify product aspects from customer reviews. The authors identified the candidate features using the dependency relations between aspects and opinion words. Then, they ranked the aspects based on their relevance. The proposed method does not achieve the best results using the existing performance measures. Zhang et al. [20] proposed a novel approach to rank products. The authors identified subjective and comparative sentences from customer reviews to build a weighted, directed product graph. They used the PageRank algorithm to rank the products. Brody and Elhadad [30] presented an unsupervised method to extract features and determine the sentiments from online reviews. The Local LDA was used to infer the topics and aspects. The method was implemented by extracting the adjectives, building the graph, constructing a seed set, propagating the polarity, and setting aspect-specific gold standards. Yu et al. [15] identified important product aspects from a large number of online reviews based on a dependency parser, and determined opinions on these aspects using a sentiment classifier. The ranking algorithm was developed to rank those important aspects. The authors adopted the Normalized Discounted Cumulative Gain (NDCG) to measure the performance of the system. Popescu and Etzioni [2] introduced an unsupervised system OPINE to extract the product features and opinions from online reviews. The system extracts explicit features by applying various dependency rules. Rafi et al. [26] proposed a framework to analyze the importance of having relevant data and identifying important product aspects for rankings in sentiment analysis.

All of the research papers discussed in this section have focused their rankings based on opinion scores, summary based rankings, sentiment-based rankings, and rule-based rankings by identifying aspects and their opinions. None of the above authors discussed ranking the products based on aspects and their opinions by considering the positive and negative ranks using a Spearman's rank correlation coefficient method.

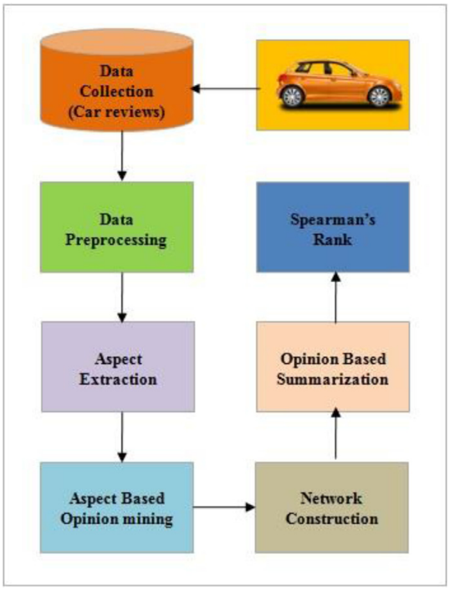


Fig. 1. Aspect-based opinion ranking system.

Table 1
Car dataset.

Products	No. of reviews
Acura MDX	38
Acura RDX	32
Acura TL	107
Acura TSX	157
Audi A4	93
Audi A5	25
Audi Q5	47
BMW 1 Series	19
BMW 3 Series	72
BMW 5 Series	21
Total	611

3. Aspect based opinion ranking system

In this section, an aspect-based opinion ranking system for products is proposed, as shown in Fig. 1. The system consists of different key components, namely, the data collection process, preprocessing, aspect extraction, aspect-based opinion mining, aspect-based opinion ranking and sentiment classification, and evaluation. The proposed system is explained as follows.

3.1. Data collection

Data collection is the method of gathering information on specific topics. In this system, the existing OpinRank review dataset has been used to rank the products (cars) based on opinion mining [19]. This dataset contains 42,230 reviews for approximately 140–150 cars from the model years 2007, 2008, and 2009. Only 611 reviews for 10 cars are collected from the model year 2009 for this opinion ranking system (Table 1). The cars include the Acura MDX, Acura RDX, Acura TL, Acura TSX, Audi A4, Audi A5, Audi Q5, BMW 1 Series, BMW 3 Series, and BMW 5 Series.

3.2. Data preprocessing

Real-world data are incomplete, noisy, and inconsistent, which includes lacking attribute values, containing errors and discrepancies, etc. The quality of data is more important for the decision-making process. To improve the quality of data, the raw data need to be preprocessed to improve the efficiency of the system. In this component, the raw data was in the format of "<DOC> <DATE>06/15/2009</DATE> <AUTHOR>The author</AUTHOR> <TEXT>The review goes here. </TEXT> <FAVORITE>What are my favorites about this car</FAVORITE> </DOC>". First, the tags, punctuation, and stop words are

Algorithm 1

Input: A set of preprocessed reviews
Output: A set of aspects
1 For each preprocessed review R_i
2 Extract aspects (Nouns and Pronouns)
3 Return aspects
4 End

Table 2
Aspect-based opinion polarity score.

Products	Positives	Negatives	Neutrals	None	Total no. of aspects
Acura MDX	165	93	18	262	538
Acura RDX	145	60	16	224	445
Acura TL	436	77	19	597	1129
Acura TSX	684	188	30	1016	1918
Audi A4	505	108	13	713	1339
Audi A5	117	24	11	163	315
Audi Q5	202	52	9	360	623
BMW 1 Series	85	7	1	97	190
BMW 3 Series	232	99	37	470	838
BMW 5 Series	105	19	0	123	247

removed. Second, stemming and lemmatization are used to reduce the words into root form and return the dictionary form of words, respectively. These processing techniques are applied to obtain the quality of the data.

3.3. Aspect extraction

Information extraction is the automatic extraction of structured information from the semi-structured or unstructured data. This is one of the most important challenges among the research communities. The aspect and entity extraction is the part or class of information extraction (Sunita [36]). It extracts structured information such as events, entities or relationships. In this paper, the named entities (people, organization, places, etc.), concepts, and relations are extracted from the textual reviews (Algorithm 1). For example, the review "*The 2009 MDX is a good mix of luxury and performance. Interior styling is visually pleasant, but Acura should not have skimped on the quality of material such as the plastic interior handles and plastic fake wood. The strengths of the MDX are definitely the handling quality compared to other SUVs*" is preprocessed and the aspects extracted, namely *Acura, plastic, forest, performance,* and *interior*. Similarly, the aspects and entities are extracted for 10 cars from 611 reviews.

3.4. Aspect-based opinion mining

The opinions expressed at the document level and the sentence level are insufficient for applications such as products, services, etc. These levels express the overall opinion of the application but are not exactly in the opinion target of the application. For example, the review (sentences numbered in parentheses) "*Nice car(1). Build quality is excellent(2). Dealer service has been excellent(3). Bluetooth cell phone pairing is easy and nice (4). The seats are hard for me to set up comfortably, though, and the headrests are absolutely awful(5). Very uncomfortable due b/c they tilt too far forward(6). Dash has too many buttons(7). Acura should consolidate the controls to make them easier to use, rather than trying to make you depend on voice input(8). GPS requires far too many steps(9). No iPod integration is disappointing(10). Needs better, more durable clear coat on the paint(11). Rear camera display hangs switching modes (dark to light) backing out of a driveway so often is useless when needed most on a sunny day(12)*" expresses a positive opinion at the document level. At the sentence level, the sentences *(1, 2, 3, 4, 8, 10), and (11)* express positive opinions and *(5), (6), (7), (9), and (12)* express negative opinions. Both the document and sentence level express an overall opinion but not an opinion on the targets. To overcome this issue, an aspect-based sentiment analysis is used to determine the opinions of the target entities. The entities and their aspects are extracted (Table 2) with opinion polarities (positive, negative, neutral) using MeaningCloud [25] . For the abovementioned example (Algorithm 2), the aspects car*(1)*, dealer*(2)*, service*(2)*, mobile phone*(4)*, Bluetooth*(4), Acura(8)*, and coat*(11)* are detected as positive opinions, dash *(7)*, button*(7), GPS(9)*, and camera*(12)* are detected as negative opinions, and sunny*(12)* is detected as no opinion.

3.5. Network visualization and opinion-based summary

Network visualization aims to understand the structure and its characteristics between two nodes, entities, or actors. This network maps and measures the relationship and information flow between information or knowledge processing entities (or groups). It is a powerful tool to convey complex information. The network visualization method leads to analyzing and

Algorithm 2

Input 1: A set of preprocessed reviews
Output: Aspect-based opinion polarity, a set of aspects and opinion words
For each preprocessed review Ri
  Extract opinion word (Adjectives)
  Return aspects
End
For each aspect and opinions
  Return opinion polarity (P, P+, N, N+, None)
End

identifying the connection power among all nodes in a particular environment. The data (nodes and ties) can be explored in different layouts using color, size, and nodal properties. The network visualization method that is used in many practical applications includes behavioral analysis, data aggregation and mining, location-based interaction analysis, sentiment analysis, recommender systems, and network modeling. To understand the system, the textual corpus is converted into networks as aspects and its opinion polarities for products are as shown in Fig. 2 using the Harel-Koren Fast Multiscale layout [33]. It is a two force-directed algorithm designed to make all the edges have the same length and to minimize the edge crossings in the layout. The position of the selected vertices and nodes can be adjusted manually for better visualization of the graph. The aspects and their corresponding polarities are given as input data to generate this layout. The positive, negative, neutral, and none are highlighted in red, maroon, green, and pink colors. The aspects and their frequencies are represented as primary labels and secondary labels. These aspects and opinions are summarized based on centrality, which measures the most important aspects and opinions within the network. The degree centrality measures the number of ties between aspects and opinions. A cross-clique centrality (a node with high connectivity) is used to determine the positive and negative opinions as a subgraph. These positive and negative counts can be used to rank products in order to understand the products' importance and their relationships. Users can also delete any aspect that does not give importance to the product.

3.6. Aspect-based opinion ranking

The aspects and their opinions are ranked and measured based on the relationship between a set of products or items. Aspect-based opinions are more important to the end user for buying a product or selecting a service in the market. Let $P=\{p_1, p_2, \ldots, p_n\}$ be the set of products to be ranked. Let $R_i=\{r_{i1}, r_{i2}, \ldots, r_{in}\}$ be the set of collected reviews for each product. Let $A=\{a_1, a_2, \ldots, a_n\}$ be the set of aspects and $O_i=\{o_{i1}, o_{i2}, \ldots, o_{in}\}$ be the set of opinions for each aspect. Let $PS=\{pos_{p1}, pos_{p2}, \ldots, pos_{pn}\}$ be the set of positive opinion summaries of each product based on the aspects. Let $NS=\{neg_{p1}, neg_{p2}, \ldots, neg_{pn}\}$ be the set of negative opinion summaries of each product based on the aspects. Based on this opinion summary, the products are ranked using a Spearman's rank correlation coefficient algorithm.

Spearman's rank correlation coefficient (also called Spearman's rho) is a nonparametric measure. It is used when the data are not normally distributed between two variables. Spearman's rho is defined as the Pearson correlation coefficient between two ranked random variables. It ranks both positive and negative variables from the smallest to the largest [38]. The difference between the positive and negative ranks for each data point is recorded. For a data sample of size n, the raw data scores A_i and B_i are converted to ranks rgA_i and rgB_i, and r_s is calculated from the Pearson correlation coefficient (ρ) as in Eq. (1). If all ranks are distinct, then r_s is calculated as in Eq. (2). r_s takes a value from +1 to −1. If $r_s=+1$, then it indicates a perfect positive correlation. If $r_s=-1$, then it indicates a perfect negative correlation.

$$r_s = \rho_{rg_A, rg_B} = \frac{\mathrm{cov}(rgA,\ rgB)}{\sigma_{rg_A}\sigma_{rg_B}} \tag{1}$$

where ρ is the Pearson correlation coefficient between two ranked variables, $cov(rg_A, rg_B)$ is the covariance of the rank variables, and σ_{rgA} and σ_{rgB} are the standard deviations of the ranked variables.

$$r_s = 1 - \frac{6\sum d_i^2}{n(n^2-1)} \tag{2}$$

where $d_i = rg(A_i) - rg(B_i)$ is the significant difference between the two ranked variables, and n is the total number of observations.

3.7. Aspect-based sentiment classification

Sentiment classification is a task to classify a given text based on the opinion polarity, such as positive or negative. It can be performed at the document level, sentence level, or aspect level. Sentiment classification is mainly divided into three approaches, namely, the machine learning-based approach, the lexicon-based approach, and the hybrid approach. The machine learning-based approach is further divided into three types, namely, supervised learning (the system learns from labeled data), semi-supervised learning (the system learns from labeled and unlabeled data), and unsupervised learning

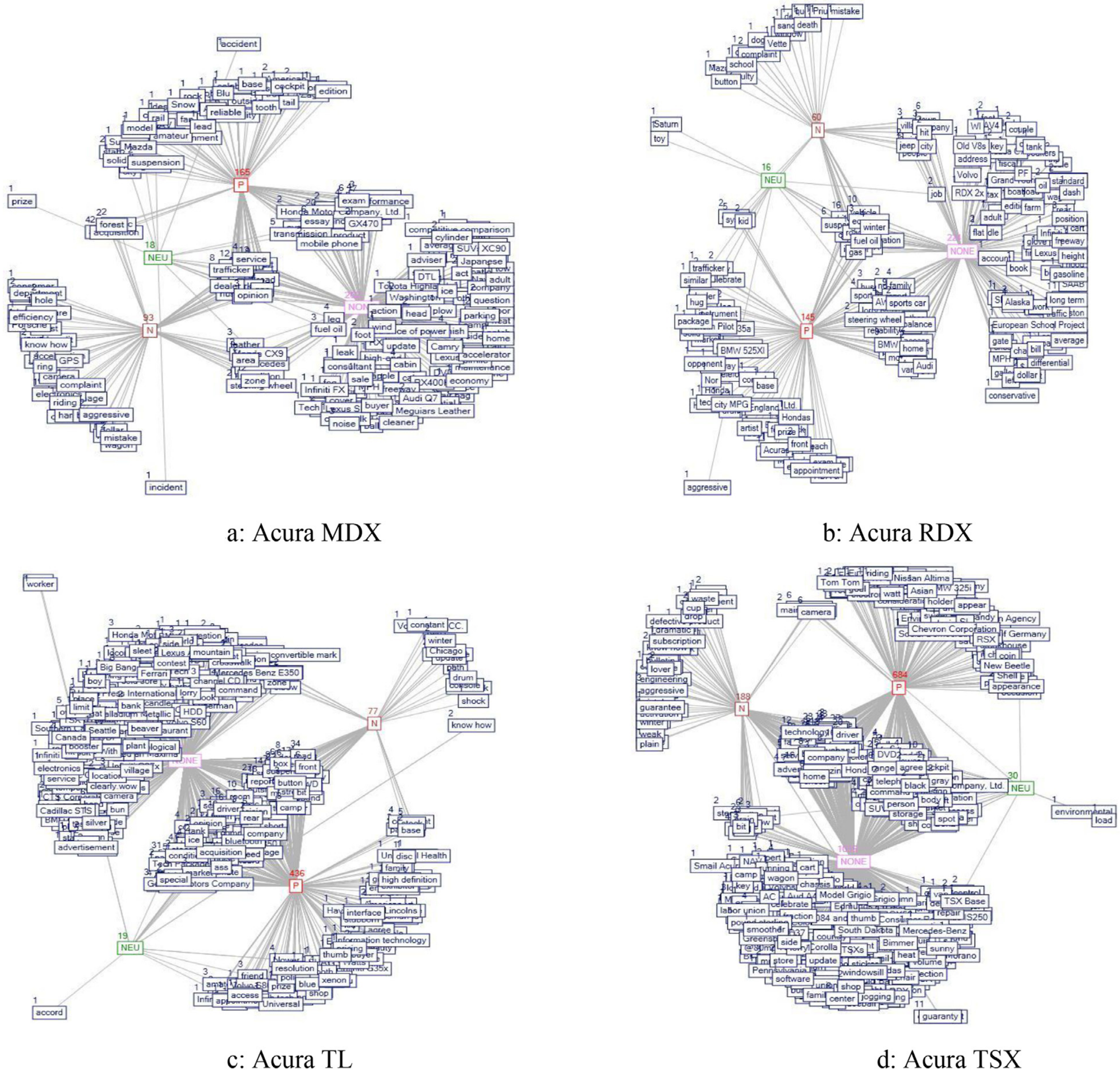


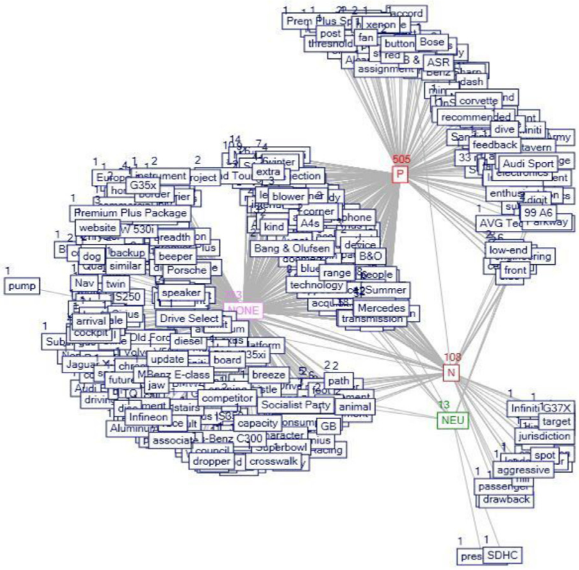
Fig. 2. (a–j). Visualization of aspects and opinions using Harel-Korel Fast Multiscale Layout.

methods (the system learns from unlabeled data). In this paper, the sentiment classification task is performed at the aspect level using three supervised learning methods, namely, Naïve Bayes (NB), Maximum Entropy (MaxEnt), and the Support Vector Machine (SVM).

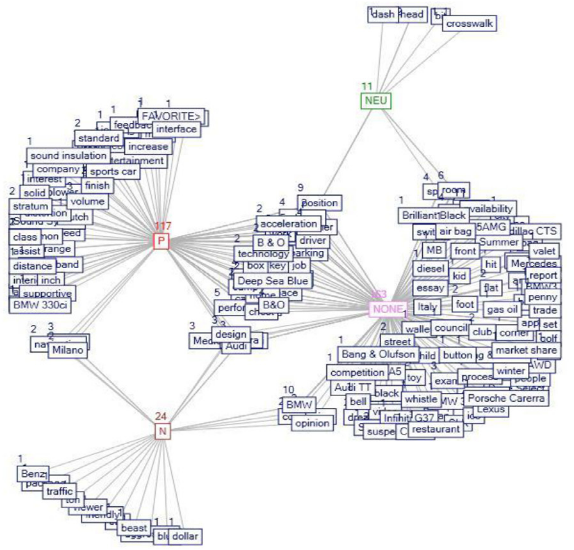
3.8. Naïve Bayes

Naïve Bayes is a simple probabilistic model based on Bayes theorem. It is also known as simple Bayes and independence Bayes in the field of statistics and computer science. Naïve Bayes classifier assigns class labels to instances or observations, which are represented as vectors in feature space. It assumes the feature values are independent. The Naïve Bayes classifier uses the bag-of-words model to classify the aspects into a positive or negative category. If the word is present in the positive category, then it assigns the positive score value and vice versa. The Naïve Bayes probability model is constructed as follows. In particular, the most likely class assignment is determined mathematically.

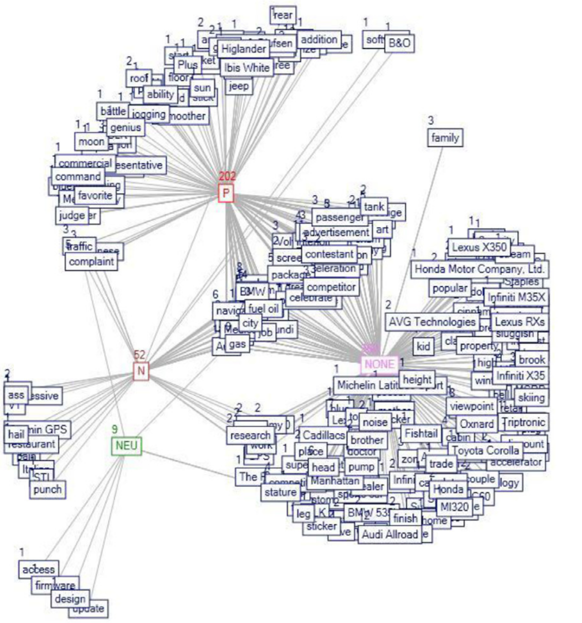
$$p(C_a)\prod_{i=1}^{n} p(x_i|C_a) \; > p(C_b)\prod_{i=1}^{n} p(x_i|C_b) \;\Rightarrow\; p(C_a|x_1, ..., x_n) \; > \; p(C_b|x_1, ..., x_n)$$



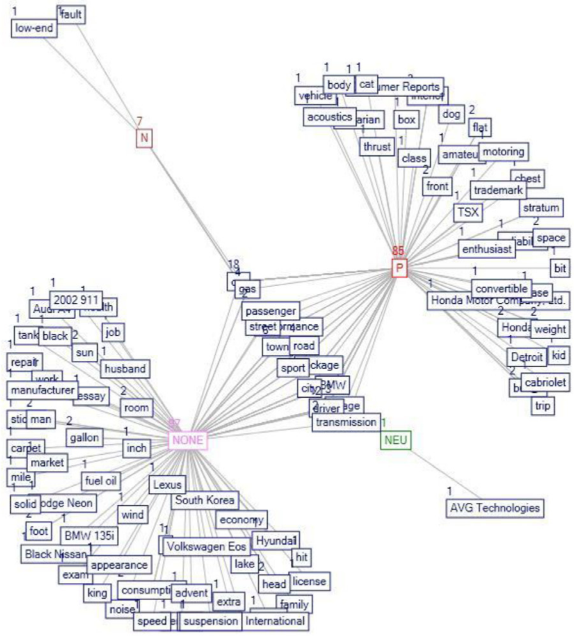
e: Audi A4



f: Audi A5



g: Audi Q5

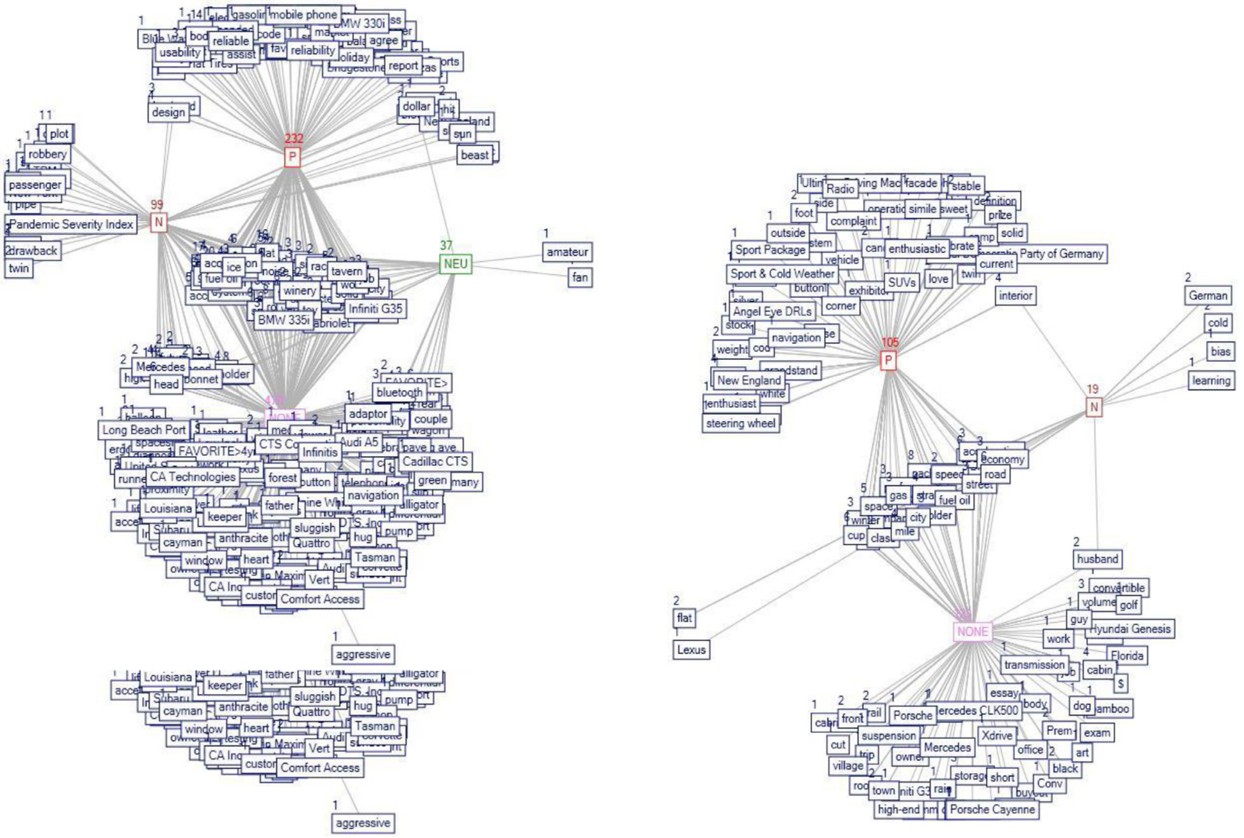


h: BMW 1 Series

Fig. 2. Continued

where $p(C_k)$ denotes the probability of class k. $p(x_i|C_k)$ denotes the conditional probability or likelihood of the feature set in the given class. Therefore, the most likely class assignment for a feature point $\vec{x} = x_1, \ldots, x_n$ is calculated by

$$p(C_k) \prod_{i=1}^{n} p(x_i|C_k) \qquad for \;\; k = 1, \ldots, K$$



i: BMW 3 Series

j: BMW 5 Series

Fig. 2. Continued

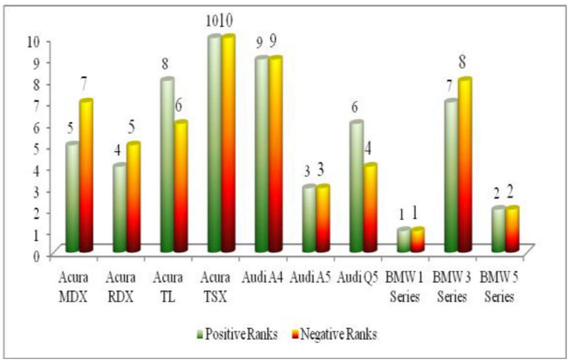


Fig. 3. Positive and negative ranks.

In addition, the feature point $\vec{x}$ is assigned to the class (C_k) based on the largest value. The mathematical expression is shown in Eq. (2).

$$\hat{C} = \underset{k \in \{1,\ldots K\}}{\arg\max}\, p(C_k) \prod_{i=1}^{n} p(x_i | C_k)$$

where $\hat{C}$ is the measured class for $\vec{x}$.

Table 3
The average of aspect-based opinion polarity score.

Products	Positives (%)	Negatives (%)	Neutrals (%)	None (%)
Acura MDX	30.67	17.29	3.35	48.70
Acura RDX	32.58	13.48	3.60	50.34
Acura TL	38.62	6.82	1.68	52.88
Acura TSX	35.66	9.80	1.56	52.97
Audi A4	37.71	8.07	0.97	53.25
Audi A5	37.14	7.62	3.49	51.75
Audi Q5	32.42	8.35	1.44	57.78
BMW 1 Series	44.74	3.68	0.53	51.05
BMW 3 Series	27.68	11.81	4.42	56.09
BMW 5 Series	42.51	7.69	0.00	49.80

3.9. Maximum Entropy (MaxEnt)

Maximum Entropy is a probabilistic classifier that uses training data to estimate the conditional distribution of the class in the given document [13,18,29]. This classifier is largely used in text classification problems where the features are words. The labeled features are converted into feature values. The maximum entropy classifier sets the constraint in the training data to express its characteristics. Maximum Entropy is parameterized in the form of exponential functions as

$$P_{ME}(y|x) = \frac{1}{Z(x)} \exp\left(\sum_i \lambda_{i,y} f_{i,y}(x,y)\right)$$

such that

$$f_{i,y}(x,y) = \begin{cases} 0 & if\ y \neq y' \\ \frac{N(x,i)}{N(x)} & otherwise \end{cases}$$

where P_{ME} $(x|y)$ denotes the probability of the class 'y' in the given document 'x', $f_{i,y}$ (x,y) is the feature or class function between the feature 'f_i' and the class 'y', $\lambda_{i,y}$ represents the estimation parameter, Z(x) denotes the normalizing factor, $N(x,i)$ denotes the number of occurrences of the feature 'i' in document 'x', and $N(x)$ represents the number of words in document 'd'.

3.10. Support vector machines (SVM)

The SVM is used for classification and regression problems in statistical learning. It is widely used in binary text classification and sentiment classification problems. The SVM examines the data and identifies the hyperplane that separates the data from one class to other classes by using a function as large as possible. Let (x_i, y_i) be a set of training data with a labeled class where $x_i \in R^n$ for each data point, and $y_i \in \{+1, -1\}$ for two data classes. The SVM can be solved using the following optimization problem [1,41]:

$$\min_{w,b,\xi} \frac{1}{2} W^T W + C \sum_{i=1}^{l} \xi_i$$
$$subject\ to\ y_i\left(W^T \phi(X_i) + b\right) \geq 1 - \xi_i,$$
$$\xi_i \geq 0.$$

where 'W' denotes the weight parameter, ξ denotes the slack variables that need to be very small (nearly equal to 0), and C represents the regularization factor. The SVM method becomes the minimization problem where $y_i(W^T \phi(X_i) + b)$ is greater than $1 - \xi$. The training vector X_i is used to map the higher dimensional feature space by 'φ'.

4. Experimental result and discussion

In this section, we present the experimental results to rank products and to measure the effectiveness of the proposed ranking system. Table 1 shows the details of the product dataset. This dataset contains 611 reviews for 10 cars from the model year 2009. These reviews were preprocessed and the textual reviews were extracted. The extracted reviews were processed through MeaningCloud to detect the aspects and their opinion polarities. This tool classified opinion polarities into 6 categories, namely highly positive (P+), positive (P), negative (N), highly negative (N+), neutral (NEU), and objective statements (NONE). The highly positive and positive classes were considered as one positive class, and the highly negative and negative classes were considered as one negative class. The aspect-based opinion polarity detection is shown in Table 2. The average of the aspect-based opinion polarity score is shown in Table 3. For instance, Acura MDX obtains 93 negative opinions score out of 538 aspects, as shown in Table 2. The average of this negative score is 17.29% without ranking the

Table 4
Ranked products for proposed method.

Products	Positive ranks	Negative ranks
Acura MDX	5	7
Acura RDX	4	5
Acura TL	8	6
Acura TSX	10	10
Audi A4	9	9
Audi A5	3	3
Audi Q5	6	4
BMW 1 Series	1	1
BMW 3 Series	7	8
BMW 5 Series	2	2

Table 5
Spearman's rank measure.

Correlation coefficients			
Pearson	0.868339	Alpha	0.05
Spearman	0.915152	Tails	2
corr	0.868339	t	4.951957
std err	0.175353	*p*-value	0.001118
lower	0.463975	upper	1.272702

Table 6
The LDA method for products.

Products	Positive ranks	Negative ranks
Acura MDX	278	73
Acura RDX	268	46
Acura TL	210	83
Acura TSX	234	68
Audi A4	196	52
Audi A5	400	61
Audi Q5	334	94
BMW 1 Series	177	92
BMW 3 Series	268	71
BMW 5 Series	289	71

products. A Spearman's rank correlation coefficient algorithm was employed to rank the products and measure the relationship between a set of products. This ranking method represents the relationship between a set of products or items based on the statistical dependence between the positive and negative opinions. For any two products, the first product is higher than the second product, the first product is lower than the second product or the first product is equal to the second product. The positive and negative polarities are only used to rank the products. The neutral and none polarities are ignored for the ranking purposes. By taking the statistics from the reviews, the products can be ranked individually as either positive or negative without considering the statistical dependence between two variables. In many cases, people talk about entities that have many aspects (also called attributes, facets, or features). These aspects have different opinions on each review. Therefore, the proposed ranking method is better than taking statistics from the reviews. It is also compared with the LDA (Latent Dirichlet Allocation) method. The obtained positive and negative ranking and relationships for products are based on the aspects shown in Table 4. The learned aspects and features are visualized in Fig. 4(a–t).

In this proposed method, the positive and negative opinions are ranked from smallest to largest. The highest positive opinions represent that the number of likes is greater for the product, and the highest negative opinions represent that the number of dislikes is greater for the product. The negative opinions are ranked at the top when there is a very small number of dislikes. BMW 1 Series is the topmost positively ranked and negatively ranked among the 10 products. The positive ranks indicate that the BMW 1 Series is the most liked car. The negative ranks indicate that BMW 1 Series is disliked only by a few people or users. The Acura TSX is the lowest ranked in both positive and negative ranks. Table 5 shows the correlation coefficient measures of the two-tailed test. The null hypothesis H_0 considers that there is no monotonic relationship between the positive and negative ranks. Both the Pearson's and Spearman's correlation coefficients show a positive correlation. The achieved p-value ($p = 0.001118$) is less than the alpha value (0.05). Based on this observation, the null hypothesis H_0 is rejected. This indicates that there is a strong positive association or relationship between the individual ranks obtained in the positive and negative polarities, as shown in Fig. 5.

The proposed method is compared with the Latent Dirichlet Allocation (LDA) Model [8,28]. The LDA is a probabilistic generative model that determines which words are generated from a specific topic and topics are generated from a specific document [34]. We considered 10 topics for implementing the LDA, namely the Acura MDX, Acura RDX, Acura TL, Acura TSX, Audi A4, Audi A5, Audi Q5, BMW 1 Series, BMW 3 Series, and BMW 5 Series. The probability distribution over words

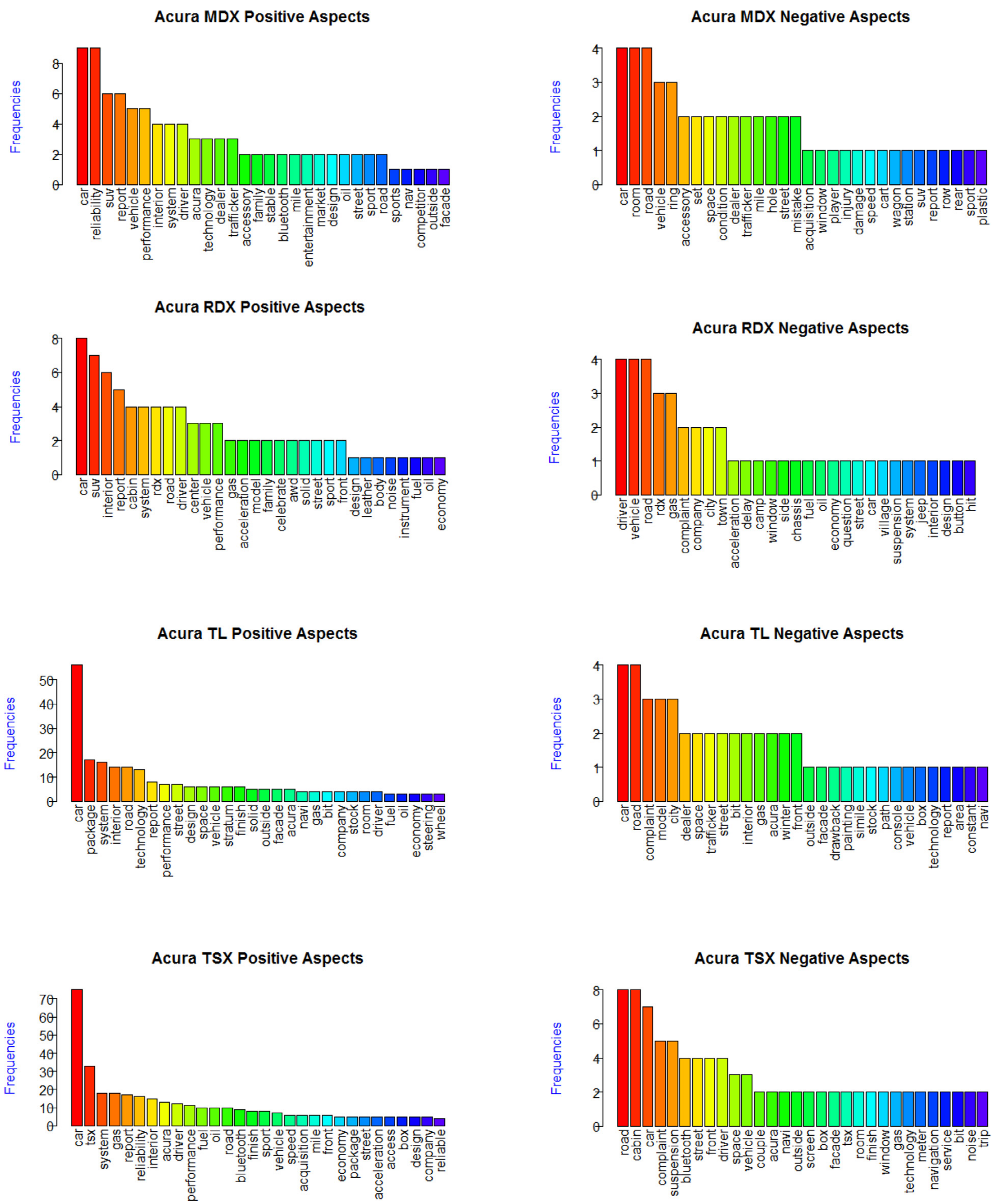


Fig. 4. (a–t). Visualization of learned aspects and features.

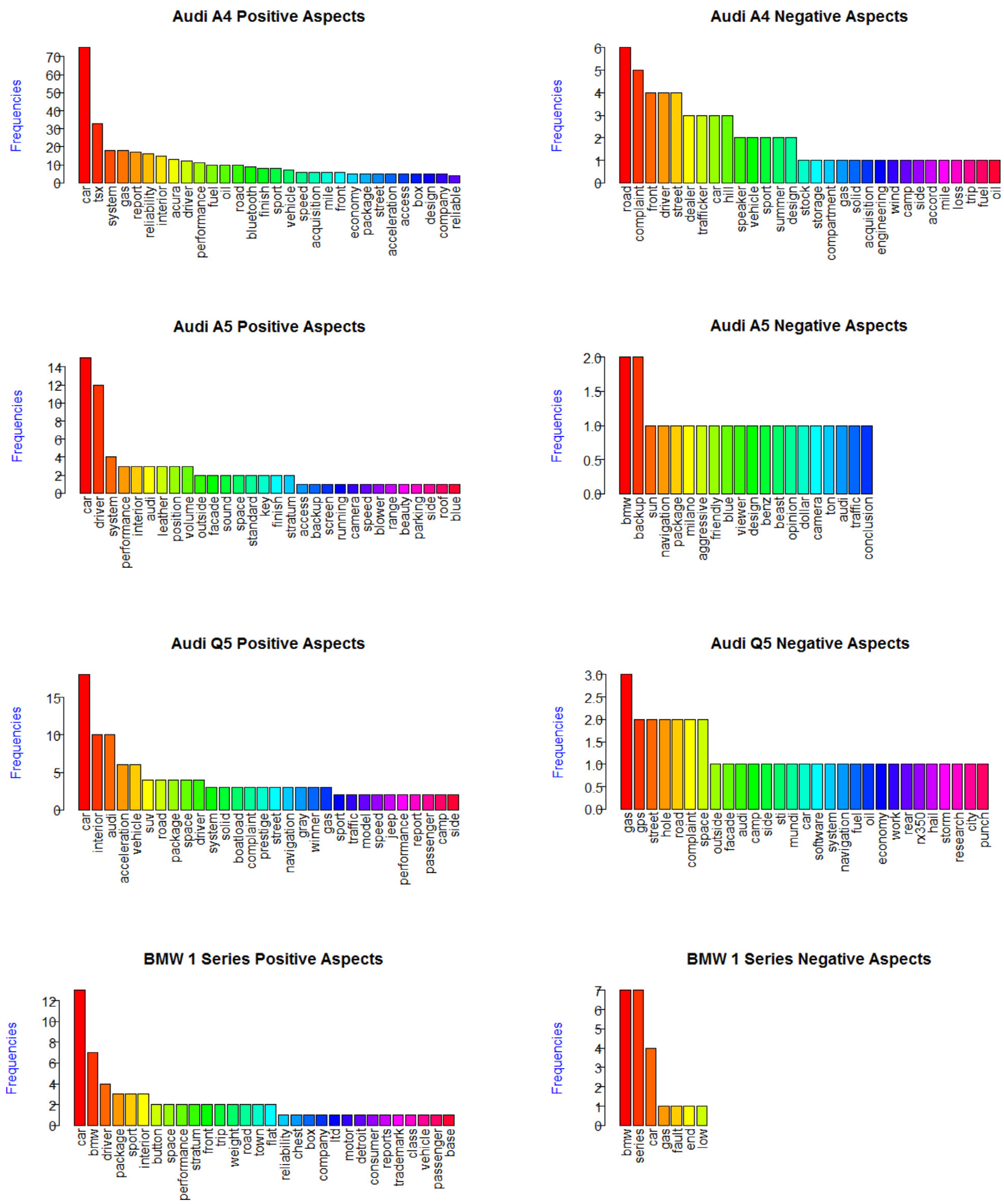


Fig. 4. Continued

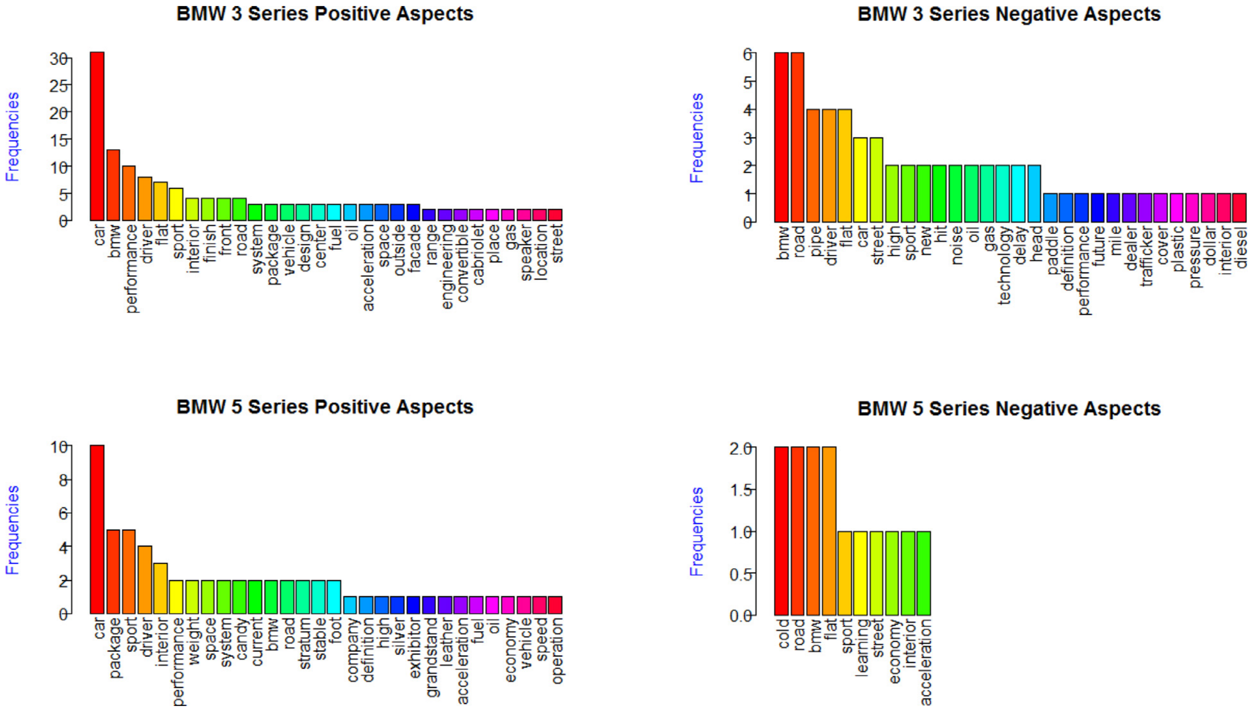


Fig. 4. Continued

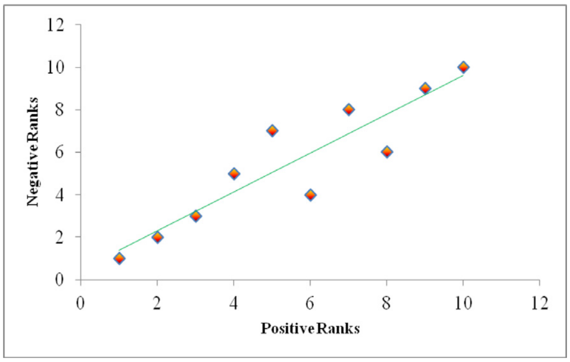


Fig. 5. Positive correlation.

per topic for positive and negative aspects is shown in Table 6. Spearman's rank method is used to rank the products for both positive and negative variables (smallest to largest), as shown in Table 7. The identical values are converted into fractional ranks by taking the average in order to compute Spearman's rank. The performance of the LDA is shown in Table 8. It shows that the p-value ($p = 0.883541$) is greater than the alpha value (0.05). Therefore, the null hypothesis H_0 is not rejected for the LDA. This indicates that there is inconclusive evidence about the significance of the relationship between the individual ranks obtained for both the positive and negative ranks in the LDA. Based on this observation, the proposed method performs better than the LDA for ranking products.

For the aspect-based sentiment classification, the NB, MaxEnt, and SVM methods are employed to measure the classification performance of the system. The standard evaluation metrics, namely, the confusion matrix, precision, recall, F-measure and accuracy are used to evaluate the performance of the sentiment classification algorithms [6,10,11]. The confusion matrix for two classes is defined in Table 9. The diagonal elements tpX and tpY are the correctly classified data and eXY and eYX are the incorrectly classified data for each class in the confusion matrix. The precision, recall, and F-measure are used to measure the correctness of classes X and Y. The accuracy is used to measure the overall sentiment classification correctness. The precision is the proportion of the number of correctly identified observations or instances (true positives) and the total number of identified observations (true positives + false positives). Mathematically, they are as follows:

$$Precision(X) = \frac{tpX}{tpX + eYX} \text{ and } Precision(Y) = \frac{tpY}{tpY + eXY}$$

Table 7
The ranked products for LDA.

Products	Positive ranks	Negative ranks
Acura MDX	4	7
Acura RDX	5	1
Acura TL	8	8
Acura TSX	7	4
Audi A4	9	2
Audi A5	1	3
Audi Q5	2	10
BMW 1 Series	10	9
BMW 3 Series	5	5
BMW 5 Series	3	5

Table 8
Spearman's rank measure for LDA.

Correlation coefficients			
Pearson	0.053398	Alpha	0.05
Spearman	0.009146	Tails	2
corr	0.053398	t	0.151248
std err	0.353048	*p*-value	0.883524
lower	−0.760734	upper	0.867530

Table 9
Confusion matrix.

		Predicted Class	
		X	Y
Known Class	X	tpX	eXY
	Y	eYX	tpY

where tpX and tpY are the numbers of true positives for the classes X and Y, respectively, and eYX and eXY are the false positives for the classes X and Y, respectively.

The recall is defined by the fraction of the number of correctly identified observations (true positives) and the total number of observations in that particular class (true positives + false negatives). Mathematically, they are as follows:

$$Recall(X) = \frac{tpX}{tpX + eXY} \text{ and } Recall(Y) = \frac{tpY}{tpY + eYX}$$

where tpX and tpY are the numbers of true positives for the classes X and Y, respectively, and eXY and eYX are the number of false positives for the classes X and Y, respectively.

The F-measure is the harmonic mean of both the recall and precision. It is also called the F_1 measure, which gives equal weights to both the precision and recall.

$$F - measures = 2 * \frac{Precision * Recall}{Precision + Recall}$$

The weighted averages of the precision, recall, and F-measure are calculated to check the performance of the classification algorithms for both classes. Each of these is weighted based on the number of observations with a class label. Mathematically, they are as follows:

$$Weighted\ Precision = \frac{((Precision\ for\ X\ class * number\ of\ observations\ from\ X\ class) + (Precision\ for\ Y\ class * number\ of\ observations\ from\ Y\ class))}{Total\ observations\ in\ the\ dataset}$$

$$Weighted\ Recall = \frac{((Recall\ for\ X\ class * number\ of\ observations\ from\ X\ class) + (Recall\ for\ Y\ class * number\ of\ observations\ from\ Y\ class))}{Total\ observations\ in\ the\ dataset}$$

$$Weighted\ F - measure = \frac{((F - measure\ for\ X\ class * number\ of\ observations\ from\ X\ class) + (F - measure\ for\ Y\ class * number\ of\ observations\ from\ Y\ class))}{Total\ observations\ in\ the\ dataset}$$

The accuracy is the proportion between the number of all correctly classified observations (true positives and true negatives) and the total number of observations (true positive, false positive, true negative, and false negative).

$$Accuracy = \frac{\neq tpX + tpY}{\neq tpX + eXY + tpY + eYX}$$

Table 10
Confusion matrix for classifiers.

Datasets		NB		MaxEnt		SVM	
		NEG	POS	NEG	POS	NEG	POS
Acura MDX	NEG	2	91	21	72	23	70
	POS	0	165	0	165	0	165
Acura RDX	NEG	0	60	4	56	4	56
	POS	0	145	0	145	0	145
Acura TL	NEG	0	77	8	69	10	67
	POS	1	435	1	435	1	435
Acura TSX	NEG	5	183	41	147	41	147
	POS	2	682	6	678	6	678
Audi A4	NEG	9	99	23	85	23	85
	POS	7	498	7	498	7	498
Audi A5	NEG	0	24	0	24	0	24
	POS	0	117	0	117	0	117
Audi Q5	NEG	0	52	6	46	6	46
	POS	0	202	0	202	0	202
BMW 1 Series	NEG	0	7	0	7	1	6
	POS	0	85	0	85	0	85
BMW 3 Series	NEG	4	95	18	81	20	79
	POS	0	232	0	232	0	232
BMW 5 Series	NEG	4	15	4	15	4	15
	POS	0	105	0	105	0	105

Table 11
The detailed precision, recall, and F-measure by class.

Datasets	NB			MaxEnt			SVM		
	P	R	FM	P	R	FM	P	R	FM
Acura MDX	100.0	2.2	4.2	100.0	22.6	36.8	100.0	24.7	39.7
	64.5	100.0	78.4	69.6	100.0	82.1	70.2	100.0	82.5
Acura RDX	0	0	0	100.0	6.7	12.5	100.0	6.7	12.5
	70.7	100.0	82.9	72.1	100.0	83.8	72.1	100.0	83.8
Acura TL	0	0	0	88.9	10.4	18.6	90.9	13.0	22.7
	85.0	99.8	91.8	86.3	99.8	92.6	86.7	99.8	92.8
Acura TSX	71.4	2.7	5.1	87.2	21.8	34.9	87.2	21.8	34.9
	78.8	99.7	88.1	82.2	99.1	89.9	82.2	99.1	89.9
Audi A4	56.3	8.3	14.5	76.7	21.3	33.3	76.7	21.3	33.3
	83.4	98.6	90.4	85.4	98.6	91.5	85.4	98.6	91.5
Audi A5	0	0	0	0	0	0	0	0	0
	83.0	100.0	90.7	83.0	100.0	90.7	83.0	100.0	90.7
Audi Q5	0	0	0	100.0	11.5	20.7	100.0	11.5	20.7
	79.5	100.0	88.6	81.5	100.0	89.8	81.5	100.0	89.8
BMW 1 Series	0	0	0	0	0	0	100.0	14.3	25.0
	92.4	100.0	96.0	92.4	100.0	96.0	93.4	100.0	96.6
BMW 3 Series	100.0	4.0	7.8	100.0	18.2	30.8	100.0	20.2	33.6
	70.9	100.0	83.0	74.1	100.0	85.1	74.6	100.0	85.5
BMW 5 Series	100.0	21.1	34.8	100.0	21.1	34.8	100.0	21.1	34.8
	87.5	100.0	93.3	87.5	100.0	93.3	87.5	100.0	93.3

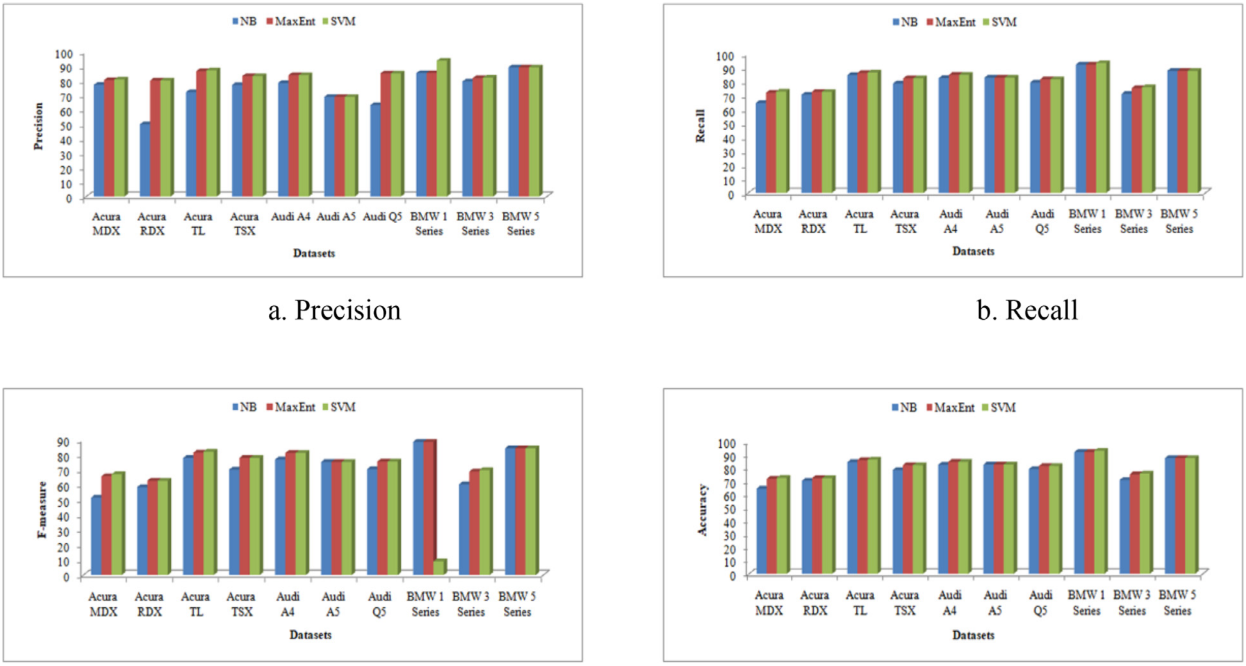
*P-Precision, R-Recall, FM - F-measure.

In this experiment, we employed the word embedding concept for aspects, which converts text into numbers. The word embedding can be classified into two types, namely, the frequency-based embedding (the count vector, TF-IDF vector, and Co-occurrence vector), and the prediction-based embedding (Word2vec – Continuous bag of words and skip-gram model). In particular, the frequency-based embedding (count vector) method is applied to the identified aspects shown in Table 2. Only the positive and negative categories are considered for the classification task. Then, the three machine learning algorithms NB, MaxEnt, and SVM are employed for sentiment classification task. The most relevant features were selected based on Information Gain (IG) and 10-fold cross-validation was conducted to estimate the sentimental classification performance. The confusion matrix for the three learning algorithms is shown in Table 10. The detailed precision, recall, and F-measure by class are shown in Table 11, and the weighted precision, weighted recall, weighted f-measure, and the obtained accuracy values are shown in Table 12 and Fig. 6(a–d). These results show that the SVM (82.21%) method outperforms the NB (79.59%) and MaxEnt (81.93%) with 83.54% precision, 82.22% recall, and 68.58% F-measure. In particular, the SVM method achieves better results for the Acura MDX (72.87%), Acura RDX (72.68%), Acura TL (86.74%), Acura TSX (82.45%), Audi A4 (84.99%), Audi A5(82.98%), Audi Q5 (81.89%), BMW 1 Series (93.48%), BMW 3 Series (76.13%), and BMW 5 Series (87.90%).

Table 12
The weighted precision, recall, F-measure, and accuracy for classifiers.

Datasets	NB				MaxEnt				SVM			
	P	R	FM	A	P	R	FM	A	P	R	FM	A
Acura MDX	77.3	64.7	51.6	64.73	80.6	72.1	65.8	72.09	81.0	72.9	67.1	72.87
Acura RDX	50.0	70.7	58.6	70.73	80.3	72.7	62.9	72.68	80.3	72.7	62.9	72.68
Acura TL	72.2	84.8	78.0	84.80	86.7	86.4	81.5	86.35	87.3	86.7	82.2	86.74
Acura TSX	77.2	78.8	70.2	78.78	83.3	82.5	78.0	82.45	83.3	82.5	78.0	82.45
Audi A4	78.6	82.7	77.0	82.71	83.9	85.0	81.3	84.99	83.9	85.0	81.3	84.99
Audi A5	68.9	83.0	75.3	82.98	68.9	83.0	75.3	82.98	68.9	83.0	75.3	82.98
Audi Q5	63.2	79.5	70.5	79.53	85.2	81.9	75.6	81.89	85.2	81.9	75.6	81.89
BMW 1 Series	85.4	92.4	88.7	92.39	85.4	92.4	88.7	92.39	93.9	93.5	9.11	93.48
BMW 3 Series	79.6	71.3	60.5	71.30	81.9	75.5	68.9	75.53	82.2	76.1	69.9	76.13
BMW 5 Series	89.4	87.9	84.4	87.90	89.4	87.9	84.4	87.90	89.4	87.9	84.4	87.90
Average	74.18	79.58	71.48	79.59	82.56	81.94	76.24	81.93	83.54	82.22	68.58	82.21

*P-Precision, R-Recall, FM - F-measure, A - Accuracy.



a. Precision b. Recall

c. F-measure d. Accuracy

Fig. 6. (a–d). Comparision of precision, recall, f-measure, and accuracy for NB, MaxEnt, and SVM.

5. Conclusion

In this study, we proposed a new framework for ranking the products based on their aspects and opinion polarity. The core of the proposed system has ranked the products based on aspects using the rank correlation coefficient method. The system was employed based on four main components. First, the aspects were identified for the products. Second, the corresponding opinions were detected for each aspect. Third, the network was constructed for visualization and modeling, the Spearman's correlation coefficient based opinion ranking method was applied to rank the products, and three supervised learning methods (NB, MaxEnt, and SVM) were employed for classification tasks. Finally, the performance of the proposed system was measured by the experimental results. The results showed that there is a strong positive relationship between the individual ranks obtained in the positive and negative aspects of the products. The positive ranks of the product increase and the negative ranks of the product increase. Therefore, the proposed method achieves better results than the LDA. Moreover, the SVM method achieves better performance than the NB and MaxEnt methods. In the future, this work will be compared with lexicon-based approaches, namely SentiWordNet, Sentistrength, Sentiment140, Sentic LDA, and deep learning methods.

Acknowledgments

This work was supported by the UGC (University Grants Commission) National Fellowship (F1-17.1/201617/RGNF-2015-17-SC-TAM-4711). We are very grateful to the Editors and anonymous reviewers for their valuable guidance, comments and insights to improve the earlier version of this paper.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ins.2018.05.003.

References

[1] A. Tripathy, A. Agrawal, S.K. Rath, Classification of sentiment reviews using n-gram machine learning approach, Expert Syst. Appl. 57 (2016) 117–126.
[2] A.-M. Popescu, O. Etzioni, Extracting product features and opinions from reviews, in: Proceedings of Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing (HLT/EMNLP), Vancouver, 2005, pp. 339–346.
[3] A.B. Sayeed, J. Boyd-Graber, Grammatical structures for word-level sentiment detection, in: Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 2012, pp. 667–676.
[4] B. Agarwal, S. Poria, N. Mittal, A. Gelbukh, A. Hussain, Concept-level sentiment analysis with dependency-based semantic parsing: a novel approach, Cognit. Comput. 7 (4) (2015) 487–499.
[5] C. Banea, R. Mihalcea, J. Wiebe, Sense-level subjectivity in a multilingual setting, Comput. Speech Lang. 28 (2014) 7–19.
[6] C. Zhang, D. Zeng, J. Li, F.-Y. Wang, W. Zuo, Sentiment analysis of chinese documents: from sentence to document level, J. Am. Soc. Inf. Sci. Technol. 60 (12) (2009) 2474–2487.
[7] D. Zhang, Y. Guo, X. Chen, Automated aspect recommendation through clustering-based fan-in analysis. ASE 2008, in: 23rd IEEE/ACM International Conference on Automated Software Engineering, 2008, pp. 278–287.
[8] D.M. Blei, A.Y. Ng, M.I. Jordan, Latent dirichlet allocation, J. Mach. Learn. Res. 3 (2003) 993–1022.
[9] E. Marrese-Taylor, J.D. Velásquez, F. Bravo-Marquez, A novel deterministic approach for aspect-based opinion mining in tourism products reviews, Expert Syst. Appl. 41 (2014) 7764–7775.
[10] E. Frank, M.A. Hall, I.H. Witten, The WEKA workbench, Online Appendix for "Data Mining: Practical Machine Learning Tools and Techniques", Fourth Ed., 2016, Morgan Kaufmann, 2016.
[11] F.H. Khan, S. Bashir, U. Qamar, TOM: Twitter opinion mining framework using hybrid classification scheme, Decis. Support Syst. 57 (2014) 245–257.
[12] H. Xu, E. Martin, A. Mahidadia, Contents and time sensitive document ranking of scientific literature, J. Inf. 8 (2014) 546–561.
[13] H. Wang, L. Wang, L. Yi, Maximum entropy framework used in text classification, in: IEEE International Conference on Intelligent Computing and Intelligent Systems (ICIS), 2010, pp. 828–833.
[14] I. Chaturvedi, E. Ragusa, P. Gastaldo, R. Zunino, E. Cambria, Bayesian network based extreme learning machine for subjectivity detection, J. Franklin Inst. B 355 (2018) 1780–1797.
[15] J. Yu, Z.-J. Zha, M. Wang, T.-S. Chua, Aspect ranking: identifying important product aspects from online consumer reviews, in: Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics, Portland, Oregon, 2011, pp. 1496–1505. June 19-24.
[16] J.M. Chenlo, A. Hogenboom, D.E. Losada, Rhetorical structure theory for polarity estimation: an experimental study, Data Knowl. Eng. 94 (2014) 135–147.
[17] J.C.B. Rabelo, R.B.C. Prudêncio, F.A. Barros, Using link structure to infer opinions in social networks, in: 2012 IEEE International Conference on Systems, Man, and Cybernetics, 2012, pp. 681–685.
[18] K. Nigam, J. Lafferty, A. McCallum, Using maximum entropy for text classification, in: IJCAI-99 Workshop on Machine Learning for Information Filtering, 1999, pp. 61–67.
[19] K. Ganesan, C. Zhai, Opinion-based entity ranking, Inf. Retriev. 15 (2) (2012) 116–150.
[20] K. Zhang, R. Narayanan, & A. Choudhary (2009). CUCIS Technical Report Mining Online Customer Reviews for Ranking Products.
[21] K. Chutmongkolporn, B. Manaskasemsak, A. Rungsawang, Graph-based opinion entity ranking in customer reviews, in: IEEE 15th International Symposium on Communications and Information Technologies (ISCIT), 2015, pp. 161–164.
[22] L. Zhang, B. Liu, S.H. Lim, E. O'Brien-Strain. Extracting and ranking product features in opinion documents. Coling 2010: Poster (2010) Volume, pages 1462–1470.
[23] L. Garcıa-Moya, R. Berlanga-Llavori, M.J. Aramburu-Cabo, Extraction and ranking of product aspects based on word dependency relations, CERI (2012).
[24] M. Eirinaki, S. Pisal, J. Singh, Feature-based opinion mining and ranking, J. Comput. System Sci. 78 (2012) 1175–1184.
[25] MeaningCloud™ (http://www.meaningcloud.com) have been used for Text Analytics purposes in the development/testing/validation of this research/prototype/software.
[26] M. Rafi, U. Noman, M.R. Farooq, A.R. Farooq, Study of automatic extraction, classification, and ranking of product aspects based on sentiment analysis of reviews, (IJACSA) Int. J. Adv. Comput. Sci. Appl. 6 (10) (2015).
[27] R. Narayanan, B. Liu, A. Choudhary, Sentiment analysis of conditional sentences, in: Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing, 1, 2009, pp. 180–189. Association for Computational Linguistics.
[28] R. Krestel, N. Dokoohaki, Diversifying customer review rankings, Neural Netw. 66 (2015) 36–45.
[29] R. Xia, C. Zong, S. Li, Ensemble of feature sets and classification algorithms for sentiment classification, Inf. Sci. 181 (2011) 1138–1152.
[30] S. Brody, N. Elhadad, An unsupervised aspect-sentiment model for online reviews, in: Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the ACL, Los Angeles, California, 2010, pp. 804–812.
[31] S. Bashir, W. Afzal, A.R. Baig, Opinion-based entity ranking using learning to rank, Appl. Soft Comput. 38 (2016) 151–163.
[32] S.L. Lo, R. Chiong, D. Cornforth, Ranking of high-value social audiences on Twitter, Decis. Support Syst. 85 (2016) 34–48.
[33] M. Smith, A. Ceni, N. Milic-Frayling, B. Shneiderman, E. Mendes Rodrigues, J. Leskovec, C. Dunne, NodeXL: a free and open network overview, discovery and exploration add-in for Excel 2007/2010/2013/2016, 2010. http://nodexl.codeplex.com/ from the Social Media Research Foundation http://www.smrfoundation.org.
[34] S. Poria, I. Chaturvedi, E. Cambria, F. Bisio, Sentic LDA: improving on LDA with semantic similarity for aspect-based sentiment analysis, in: 2016 International Joint Conference on Neural Networks (IJNN), 2016, pp. 4465–4473. IEEE 2016.
[35] S.S. Yau, Y. Yin, QoS-based service ranking and selection for service -based systems, in: IEEE International Conference on Services Computing, 2011, pp. 56–63.
[36] S. Sarawagi, Information extraction, Found. Trends Databases 1 (3) (2007) 261–377.
[37] T. Wilson, J. Wiebe, P. Hoffmann, Recognizing contextual polarity in phrase-level sentiment analysis, in: HLT '05 Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing, 2005, pp. 347–354.
[38] T.D. Gauthier, Detecting trends on using Spearman's rank correlation co-efficient, Enviorn. Forens. 2 (2001) 359–362.
[39] W. Huang, S. Zeng, M. Wan, G. Chen, Medical media analytics via ranking and big learning: A multi-modality image-based disease severity prediction study, Neurocomputing 204 (2016) 125–134.

[40] Y. Zhu, Y. Lan, J. Guo, P. Du, X. Cheng, Novel relational learning-to-rank approach for topic-focused multi-document summarization, in: IEEE 13th International Conference on Data Mining, 2013, pp. 927–936.
[41] Y. Liua, J.-W. Bi, Z.-P. Fana, Multi-class sentiment classification: The experimental comparisons of feature selection and machine learning algorithms, Expert Syst. Appl. 80 (2017) 323–339.
[42] Z.-J. Zha, J. Yu, J. Tang, M. Wang, T.-S. Chua, Product aspect ranking and its applications, IEEE Trans. Knowl. Data Eng. 26 (5) (2014).