Journal of Global Information Management
Volume 30 • Issue 8

A Text Analytics Framework for Performance Assessment and Weakness Detection From Online Reviews

Amit Singh, Indian Institute of Technology Jodhpur, India*
Mamata Jenamani, Indian Institute of Technology, Kharagpur, India
Jitesh Thakkar, National Rail and Transportation Institute, Vadodara, India
Yogesh K. Dwivedi, Swansea University, UK

ABSTRACT

The present research proposes a framework that integrates aspect-level sentiment analysis with multi-criteria decision making (TOPSIS) and control charts to uncover hidden quality patterns. While sentiment analysis quantifies consumer opinions corresponding to various product features, TOPSIS uses the sentiment scores to rank manufacturers based on their relative performance. Finally, U and P control charts assist in discovering the weak aspects and corresponding attributes. To extract aspect-level sentiments from reviews, the authors developed the ontology of passenger cars and designed a heuristic that connects the opinion-bearing texts to the exact automobile attribute. The proposed framework was applied to a review dataset collected from a well-known car portal in India. Considering five manufacturers from the mid-size car segment, the authors identified the weakest and discovered the aspects and attributes responsible for its perceived weakness.

KEYWORDS

Artificial Intelligence, Automobile Industry, Consumer Review, Control Charts, Defect Discovery, Ranking, Reviews, Sentiment Analysis, Text Mining, TOPSIS, User-Generated Content, Weakness Detection

INTRODUCTION

Products such as automobiles may have both safety and performance defects. Government regulations and exposure to severe brand-value and financial losses compel manufacturers to be pro-active in detecting and eradicating safety defects. Traditionally, safety defects are identified through process improvement tools and service center feedbacks. Such approaches not only suffer from high cost, and incomprehensiveness; their applicability is limited in the case of performance defects (Law et al., 2017; Liu et al., 2018). In this regard, massive product review data generated from the web has turned out to be an important source to comprehend user experiences, reactions, and perceptions. While prospective consumers use them to analyse the peers' experience with the product, the organizations mine it to identify user requirements and expectations (Singh et al., 2020). However, it is beyond

DOI: 10.4018/JGIM.304069
*Corresponding Author

human cognition to scan the available reviews manually, summarize them, and use them for sensible decision making.

In this regard, artificial intelligence in general (Dwivedi et al., 2019; Grover et al., 2019; Dwivedi et al., 2020 Stieglitz et al., 2020) and sentiment analysis (SA) in particular has emerged as a tool to mine information from text. Its usefulness is well tested and validated in domains such as product promotions and marketing (Ting et al., 2014), demand and sales forecasting (Archak et al., 2011; Chong et al., 2017; Geva et al., 2013; Hou et al., 2017; Zhang et al., 2020), supply-chain performance evaluation (Swain & Cao, 2019), and product quality assessment (Abrahams et al., 2015; Law et al., 2017). Specifically, it assists the businesses in decision making in automotive industry (Abrahams et al., 2012, 2013, 2015; Gruss et al., 2018) electronic products (Abrahams et al., 2015), dishwasher appliances (Law et al., 2017), body wash products (Zhang et al., 2012), entertainment industry (Chintagunta et al., 2010; Yang & Chao, 2015) travel industry (Chang & Chen, 2019; Choi & Lee, 2017; Sann, & Lai, 2020), and the toy industry (Winkler et al., 2016; Saumya et al., 2019). However, there has been almost no effort to connect these results with traditional quality-control tools with which the manufacturing community is acquainted. Moreover, most of such studies focus on document or sentence level. More recently, aspect-level sentiment analysis (ASLSA) has emerged as a tool to identify product defects, more precisely targeting specific attributes and context (Schouten & Frasincar, 2016). In this research, the authors have contributed to this growing field by proposing an integrated automobile-defect detection framework that connects ASLSA with the Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) and traditional quality-control tools. The framework answers following research questions:

RQ1: What are the important product features, which customers frequently discuss in online reviews?
RQ2: How feature level consumer sentiments be used to quantify manufacturers' perceived performance rating?
RQ2: Are review embedded consumer sentiments useful in discovering products' perceived weakness?

The proposed framework consists of three phases. In *Phase-I*, the authors extract attribute-level consumer sentiments indices for specific car aspects. The authors define aspects with various systems and subsystems of a car, and attributes to the more specific parts, features, or service of the system under consideration. For example, regarding *EXTERIOR* as an aspect, the authors can consider *bumper* as an underlying attribute. In *Phase-II*, the TOPSIS, a multi-criteria decision-making tool, compare extracted sentiments to compute manufacturers' relative performance index. Traditionally, TOPSIS requires inputs from experts. In this research, expert inputs are replaced by the sentiments scores mined from reviews. In this phase, the authors compute a performance score for each manufacturer based on overall consumer perception, from which a manufacturer could find its perceived performance in the market. In *Phase-III*, the authors use control charts, the *U*-chart at the aspect level, and the *P*-chart at the attribute level, to discover the reasons for performance degradation. This discovery gives the manufacturer an opportunity to identify the reasons for consumer dissatisfaction and take action accordingly. The authors apply the framework to a review dataset from a car portal in India, compare the manufacturers within a car segment, and identify the worst-performing manufacturer. In addition, the authors delve into the data to find the reasons for performance degradation.

The contributions of this research are as follows. *First*, the authors have proposed a passenger car aspect ontology consisting of 16 aspects at the system level and 15 at the subsystem level. *Second*, a heuristic for attribute-level sentiment index generation has been proposed, which differs from contemporary approaches such as (Hu & Liu, 2004; Moghaddam & Ester, 2010), in the way opinion-bearing words are connected to the exact target attribute. Specifically, the proposed heuristic splits the sentences with more than one attribute into a number of sub-sentences, each containing one attribute and corresponding sentiment phrases. The proposed heuristic has also been vetted using human annotators. *Third*, to the authors' best knowledge, the this is the first attempt to integrate text analytics with the traditional tools of quality control, providing a new avenue for a company to visualize consumer perception of its product. *Fourth*, the authors extend TOPSIS, a well-accepted method for prioritizing alternatives based on expert views, to apply to consumer views for ranking automobile manufacturers. While the traditional approach can synthesize the opinions from a handful of expert comments, which may include their biases, the views synthesized from a large number of online reviews in this novel approach are likely to be more precise, following the law of large numbers. *Sixth*, the authors collect the data from CarWale[1] and processed it for automobile attribute extraction and attribute-level sentiment index generation. the authors have also tested the proposed framework using this data.

The rest of the paper is organized as follows. Section two reports the related research works on SA and its applications, especially on product defect identification. the authors report the proposed framework in Section three, followed by its application in Section four. Section five reports the evaluation measures of the algorithm/heuristic used. The authors discuss the results and the implications of the present research in Section six. Section seven concludes the research followed by the limitations and the potential future research directions in Section eight.

BACKGROUND

Sentiment analysis (SA) captures peoples' opinions, sentiments, evaluations, appraisals, attitudes, and emotions regarding products and their attributes (Lin et al., 2016). It can be performed at a) document level, b) sentence level, and c) attribute level. In the first case, the complete document is classified, whereas in the second case, individual sentences are analysed (Lyu et al., 2020; Araújo et al., 2020; Singh et al., 2020). In the third case, the specific attribute is targeted and the corresponding context is analysed to capture consumer sentiments. Technique wise SA can be broadly categorised into two categories, dictionary-based and machine learning-based approach. In the dictionary-based approach, sentiment dictionaries such as SentiWordNet (Baccianella et al., 2010), SentiStrength (Thelwall et al., 2012), SenticNet-3 (Cambria et al., 2014), etc., are used to quantify the text. Unlike the dictionary-based approach, the machine learning-based method is data-dependent. This approach learns many of the parameters from the available text only. It is categorised into two categories, supervised and unsupervised approaches. In the first approach, the classifier is trained through the annotated data, based on which it predicts the polarity of the new text. The second approach is frequently used, as it does not require manual annotation. Researchers use dependency parsing (Fernández-Gavilanes et al., 2016), semantic orientation of the phrases (Turney, 2002), and probabilistic modelling (Rustamov et al., 2013), to name a few as the unsupervised-learning approach for text classification.

In some cases supervised learning approaches are preferred over dictionary-based ones to extract user sentiments, the latter are more appropriate in the new area (Bhatia et al., 2015). Machine learning approaches require data labeling and training and testing data from the same domain, which is expensive and time-consuming (Gamon et al., 2005). As an alternative, lexicon-based approaches have been found effective in cross-domain applications (Kouloumpis et al., 2011; Mudinas et al., 2012). Moreover, such approaches are useful for aspect-level analysis because they provide more structured, readable results with aspect-oriented explanation and justification (Mudinas et al., 2012). These approaches also use rules for *sentiment shifters* and *and/but clauses* (Lin et al., 2016).

As per Mankad et al., (2016), contributions to the literature in SA lie in two categories, a) *methodological literature* and b) *managerial literature*, which the authors connected. The first category focuses on either evolving new algorithms or amending the existing ones. It fundamentally add value to the way sentiments are extracted by extending the numbers of features, enhancing the computational efficiency, and improving accuracy (Baccianella et al., 2009; Ray, & Chakrabarti, 2020). The key contributions to this category in the automobile domain have focused on designing algorithms to extract product attributes (Schouten & Frasincar, 2016). The second category, managerial literature, focuses on applications. In managerial literature well-established methods have been applied in different domains to draw specific insights such as weakness identification (Law et al., 2017; Liu et al., 2018; Wang & Wang, 2014), predicting recalls (Zhang et al., 2015), predicting sales (Chong et al., 2017; Hou et al., 2017), assisting the businesses in mining consumer requirements (Qi et al., 2016), analysing social media influence (Chang, 2019). As research in the *methodological literature* has progressed, more applications have evolved.

Text analytics-based product ranking, automatic product-weakness detection and product-recall prediction have attracted numerous researchers in recent years. A few classic examples include product and service evaluation (Guo et al., 2018; Peng et al., 2014) product weakness finder (Wang & Wang, 2014; Zhang et al., 2012; Singh, et al., 2020), automated product-defect discovery (Abrahams et al., 2012, 2015; Law et al., 2017; Liu et al., 2018), product recall prediction (Bhat & Culotta, 2017; Zhang et al., 2015) and product-improvement strategy development (Qi et al., 2016). Wang & Wang (2014) have proposed a text analytics-based method for weakness detection in digital camera domain. They found their method to be outperforming over baseline methods. Abrahams et al. (2015) have proposed a method that uses the smoke words to detect the defects of automobile and consumer electronics products. Law et al. (2017) have introduced unigram, bigram, and trigram smoke words and walidated their applicability in detecting the defects in home appliance products. A few attempts at product defect discovery in the automobile industry were made (Abrahams et al., 2012, 2015; Singh et al., 2020). For example, Abrahams et al. (2012) proposed a learning-based text-mining method that analyzes consumer sentiments inherent in social media to detect vehicle defects automatically. They have identified that negative sentiments are not correlated with the defects. Abrahams et al. (2015) proposed an integrated text-analytics method to automatically discover product defects using textual information. To detect product defects, they used the frequent keywords mined from vehicle complaints. Singh et al. (2020), used pareto analysis and analysed the consumer sentiments to discover product weakness from the reviews.

Table 1 critically summarizes a few key studies that served as the basis of this research; the authors position their work in their context. The tables provide the Text analytics applications on product-defect discovery and recall prediction. As evident from the table, the authors position their research in a domain that combines algorithmic research and its application. Most of these attempts were at the product level and have not tried to integrate the traditional tools of quality control. By contrast, this work focused on the manufacturer level and integrated tools of statistical quality control with ASLSA.

PROPOSED RESEARCH FRAMEWORK AND METHODOLOGIES

The proposed framework integrates sentiment analysis (SA) with a multi-criteria decision-making technique (TOPSIS) and two kinds of control charts (*U*-chart and *P*-chart) that aims discover their weaknesses. In particular, the framework a) compares the automobile manufacturers operating in a segment and computes their perceived ranks, b) discovers the perceived weak aspects responsible for performance degradation, and c) identifies the attributes responsible for the perceived weakness. As presented in Figure 1, the framework comprised three phases. In *Phase I*, automobile reviews are processed to extract aspect level sentiment indices (ASLSIs). In Phase II, extracted indices are analyzed to carry out intra-segment comparison of manufacturers. The methodology employed here is TOPSIS, with modification to incorporate sentiment indices (SIs) instead of expert inputs as the decision matrix. In Phase III, control charts are developed to discover aspects and corresponding attributes responsible for consumer-perceived product weaknesses. The proposed method is detailed in Figure 1.

Phase I: Sentiment Index Generation

Car Aspects Ontology Preparation

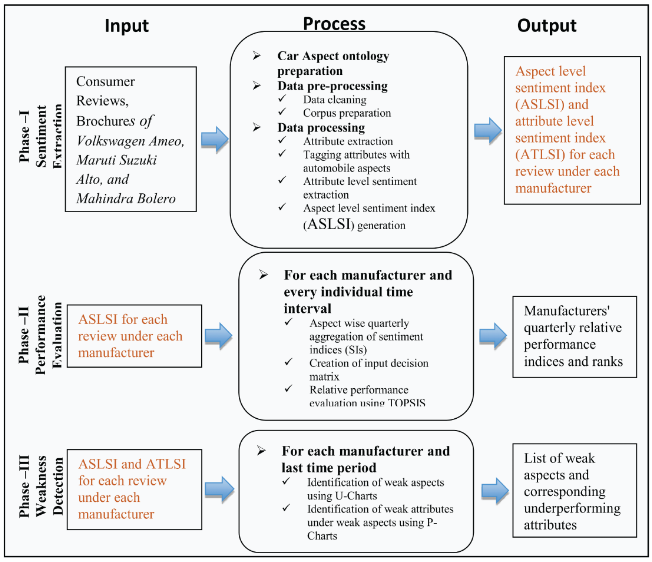
To develop a passenger car aspect ontology, the authors refer to brochures for the Volkswagen Ameo, the Maruti Suzuki Alto, and the Mahindra Bolero and prepare a list of car systems and subsystems,

Table 1. Literature on defect/weakness detection

Author	Contribution Type		Approach		Remarks	Tools Used
	Decision Making	Algorithmic	Learning-Based	Lexicon Based	Data Sources, Area, and Findings	
(Abrahams et al., 2012)	✓	×	✓	✓	**Area:** Automobile industry **Data source:** Complaint documents, Online discussion threads **Contributions:** They tested the usefulness of auto enthusiast discussion forums in discovering defects. Prepared a list of automotive smoke words for defect discovery. They found that negative sentiment is not positively correlated with defects.	OpinionFinder, Harvard General Inquirer, Automotive smoke words, Text Analytics, Logistic regression
(Wang & Wang, 2014)	✓	×	✓	✓	**Area:** Digital Camera **Data source:** Online reviews **Contributions:** They proposed a sentiment mining supported weakness finder, which was found outperforming over other methods.	Authority score, and Sentiment analysis
(Abrahams et al., 2015)	✓	×	✓	✓	**Area:** Automobile industry and consumer electronics **Data source:** NHTSA complaint documents, automotive and consumer electronics discussion forums **Contributions:** They identified that selection of distinctive terms, product features, and semantic factors were found to be the predictors of the defect. Smoke words were found very useful in defect detection.	Sentiment analysis, Harvard General Inquirer H4, Laswell LVD lexicon, PCA, Multivariate logistic regression, Learning based classifiers
(Zhang et al., 2015)	✓	×	✓	✓	**Area:** Automobile industry **Data source:** Online vehicle discussion threads, NHTSA complaint documents **Contributions:** They developed a vehicle recall prediction model. They also claimed that K-Nearest Neighbour (KNN) classifier performs better as compared to Naïve-Bayes and decision tree in predicting vehicle recalls.	Content analysis, KNN, Naive-Bayes and Decision tree, Smoke words, Chi-squared feature selection
(Singh et al., 2020)	✓	×	✓	✓	**Area:** Automobile industry **Data source:** CarWale, SIAM **Contributions:** An integrated framework that integrates sentiment analysis with quality function deployment (QFD) to evaluate the manufacturers and pareto analysis to discover the weakness.	Sentiment analysis QFD, TOPSIS, Pareto-chart, Fishbone diagram dictionaries. Harvard general inquirer dictionary
Present research	✓	✓	×	✓	**Area:** Automobile industry **Data source:** Consumer reviews from CarWale.com and brochure of the Volkswagen Ameo Maruti Suzuki Alto and Mahindra Bolero **Contributions:** The authors proposed a quality analytics framework that integrates multi-criteria decision making and tools of quality control with aspect level sentiment analysis to evaluate the manufacturers' and discover perceived weaknesses.	TOPSIS, U-chart, P-chart, Text analytics, ATLSA, SentiWordNet dictionary, RAKE

defined as *"aspects"* in the present research. The three brochures coming from different segments are used to generalize the ontology. The list is further refined and validated in consultation with the *deputy general manager of production* of a car manufacturing company situated in eastern India. The final list consists of 16 aspects at the system level and 15 at the subsystem level. To highlight which

Figure 1. Proposed research proposal flow chart for performance assessment and weakness detection

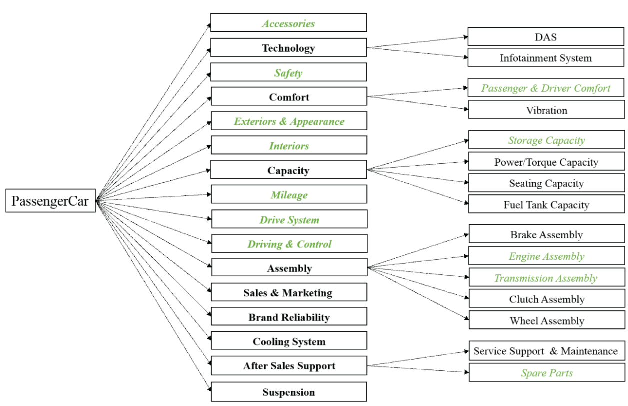


aspect is selected from which brochure, the authors have tabulated the aspects with a tick mark if they are present in the particular car brochure (Table A1 in appendix). From the table, one may note that 16 out of 26 aspects (61.5%) are common in all three models. Based on the list, the authors prepare the passenger car aspect ontology as shown in Figure 2.

Data Pre-Processing

Typical pre-processing tasks, such as removing mark-ups, HTML tags, numbers, and stop words, are performed to retain informative text only. Duplicate reviews, sentences, and words are identified and eliminated. In the reviews, spelling errors (e.g., "Verrrrrrrrryyyyyyyyyy Gooooooooooooooooood") are common; these are fixed with handwritten codes.

Figure 2. Passenger car aspect ontology



Attribute Identification and Aspect Tagging

The authors define an attribute as a word or a phrase representing an automobile component, feature, or service it provides. Identifying the target attributes is one of the core tasks in ASLSA. Among the various existing approaches, such as supervised and unsupervised machine learning, syntax-based, and frequency-based the latter option has been accepted as a straightforward and powerful tool for this task (Schouten & Frasincar, 2016). Despite their proven usefulness, however, frequency-based approaches possess certain shortcomings. First, they ignore less-frequent attribute phrases; second, they extract many phrases which do not represent actual product attributes (Hu & Liu, 2004; Schouten & Frasincar, 2016). To partially overcome these drawbacks, the authors use a frequency-based semi-automatic approach used in Singh et al. (2020). It is a three-step, semiautomatic approach comprising a) extraction of frequently used nouns and noun phrases, b) manual scanning of extracted attribute phrases, and c) tagging with appropriate car aspects. *Aspect* here has been defined as a group of similar automobile attributes connecting to an automobile system or subsystem (Figure 2). While the first step is automatic, the last two are manual. Manual intervention is essential to select the right noun phrases to represent given attributes and connect them to the passenger car ontology presented in Figure 2.

Attribute-Level Sentiment Extraction

Here, the authors have extracted consumer sentiments from the reviews and mapped them with the appropriate automobile attributes. The authors have proposed a rule-based heuristic for pairing sentiment phrases with their target attributes that breaks the sentence into sub-sentences in such a way that each sub-sentence comprises only one *AttributePhrase* and the corresponding *SentimentPhrase(s)*. The authors have defined an *AttributePhrase* as a (set of) word(s) representing an automobile attribute identified by the procedure mentioned in the last section. Similarly, a *SentimentPhrase* is defined as a set of consecutive words available in the sentiment dictionary or in the list of sentiment shifters reported by Yu et al., (2016). The heuristic shown in Figure 3 requires four inputs: *Pre-processed review data set* (D), *List of target automobile attribute phrases* (A), *Sentiment dictionary* (S), and *List of sentiment shifters* (F). The heuristic has five steps as detailed below. The authors have used two reviews as running examples to demonstrate the proposed heuristic.

Review 1: The car has good mileage but bad comfort, style and driving control. It has very good pickup and best class suspension.

Review 2: The look and interiors of the car are excellent whereas the boot space and air conditioning are troubling. Its driving is not so smooth.

Step 1: Tokenizing reviews to sentences.

A review typically comprises multiple sentences. Lines 5–7 in the heuristic (Figure 3) split the reviews into sentences and stored in *SentenceList*. The result of this step appears in Figure 4 (step 1).

Step 2: Creating pseudo words for attribute phrases and sentiment shifters.

Certain attributes and sentiment shifters contain more than one word (refer to Table A4 in appendix). However, the words in the sentence are the target of the sentiment extraction algorithm. Therefore, the authors have searched for such group of words and combined them to create single pseudo-words (Line 8 in Figure 3). In Review 1, "*driving control*", and in Review 2, "*boot space*", "*air conditioning*", and "*not so*", are examples of such pseudo words, as demonstrated in Figure 4 (output of step 2).

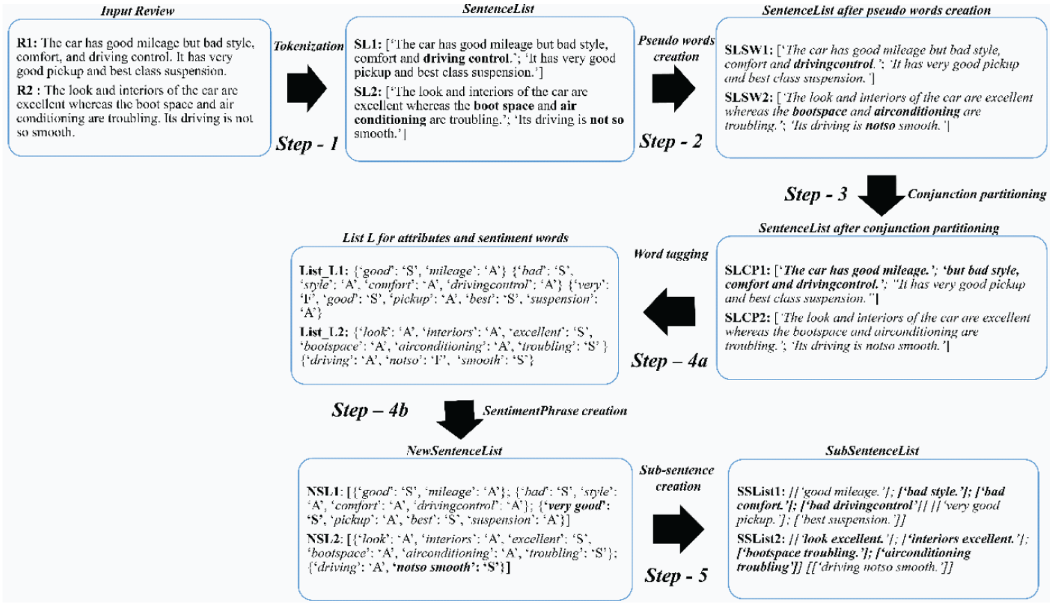
Step 3: Conjunction partitioning.

Figure 3. Heuristic for attribute-sentiment pair generation

```
1 SubSentenceGeneration:
2 Input: Pre-processed review data set: D, list of target automobile attribute phrases: A, Opinionated Words (Sentiment
  Dictionary): S, List of Sentiment Shifter:F
3 Output: Sub sentences with only one attribute and corresponding Opinionated Words and Sentiment Shifters
4 begin
// Tokenising reviews to sentences
5        for each review in D:
6                tokenize the reviews into sentences
7                store the sentences in SentenceList
// Creating pseudowords for attribute phrases and sentiment shifters
8        MakeWord (A,F)  //Search for the phrases in A and F and treat as words
// Conjunction partitioning at "but" and "however"
9        for each sentence in SentenceList
10               if but word is in the sentence,
11                       break the sentence at but
12                       create two sentences and put back in SentenceList
//  Word tagging and SentimentPhrase creation
13       for each sentence in SentenceList:
14               tokenise sentence into words
15               for each word in the sentence:
16                       If word in A
17                               Tag word as AttributePhrase and add in list L
18                       If word in S
19                               Tag word as OpinionatedWord and add in list L
20                       If word in F
21                               Tag word as SentimentShifter and add in list L
22               for all words in L
23                       combine consecutive opinionated words and sentiment shifters in between two
                         AttributePhrases, represent as a SentimentPhrase
24                       restore the sequence of AttributePhrases and SentimentPhrases in list L1

25               Store L1 as a sentence in NewSentenceList
// Generating subsentences using AttributePhrase and SentimentPhrase combination
26       for each sentence in NewSentenceList

// AttributePhrase-First Rule
27               if the sentence starts with SentimentPhrase
28                       If two or more consecutive AttributePhrases
29                               insert the first SentimentPhrase before the consecutive AttributePhrases in
                                 between each pair of AttributePhrases
30                               break the sentence after each AttributePhrases
31 Create a sub-sentence for each AttributePhrases  and SentimentPhrases pair store in SubSentenceList
32                       If no two consecutive AttributePhrases
33                               break the sentence after each AttributePhrases
34  Create a sub-sentence for each AttributePhrases  and SentimentPhrases pair store in SubSentenceList
// AttributePhrase-Second Rule

35               if the sentence starts with AttributePhrase
36                       If two or more consecutive AttributePhrases
37                               insert the first SentimentPhrases after the consecutive AttributePhrases in
                                 between each pair of AttributePhrases.
38                               break the sentence before each AttributePhrases
39  Create a sub-sentence for each AttributePhrases  and SentimentPhrases pair store in SubSentenceList
40                       If no two consecutive AttributePhrases
41                               break the sentence before each AttributePhrases
42 Create a sub-sentence for each AttributePhrases  and SentimentPhrases pair store in SubSentenceList
43        end

44 return SubSentenceLists
```

It is accepted in the literature that the polarity of the context represented by a text content gets typically reversed at conjunctions such as "but" and "however" (Shelke et al., 2017). To accommodate this effect, the authors broke a given sentence at each conjunction (i.e., but, yet, however, etc.). The authors have restored the broken sentences as individual entities in the *SentenceList* (Line 9–12, Figure 3). For the running examples, the revised sentence list appears in Figure 4 (output of step 3). It may be noted from the figure that two input sentences in review 1 became three because of the partitioning at "but" in the first sentence, whereas this step did not affect review 2.

Figure 4. Process diagram (the example of sub-sentence generation heuristic)



Step 4: Word tagging and *SentimentPhrase* creation.

In ASLSA, the authors identify the sentiment phrases and map them with the corresponding attributes (Schouten & Frasincar, 2016). Since sentiment shifters alter the polarity of texts (Polanyi & Zaenen, 2006; Xia et al., 2016), the authors need to locate them as well. To do so, the proposed heuristic first tags the words within these sentences with three categories: attribute phrases (A), sentiment words (S), and sentiment shifters (F) (Lines 13–21, Figure 3). The results of this step for the running examples appear in Figure 4 (output of step 4[a]). Next, all consecutive opinion bearing words and sentiment shifters are grouped and represented as what the authors termed a *SentimentPhrase* (Lines 22–25, Figure 3). The remaining words in a sentence are ignored. This step for both of the running examples appears in Figure 4 (step 4[b]), where the tags A, S, and F have been attached appropriately.

Step 5: Generating sub-sentences using *AttributePhrase* and *SentimentPhrase* combination.

In some sentences, more than one automobile attribute may be the target of a single sentiment phrase. For example, in Review 1, the string "bad comfort, style and driving control" indicate that the sentiment word *"bad"* is associated with three attributes: *"comfort"*, *"style"*, and *"driving control"*. The logical approach here is to pair the sentiment word with each of the attributes. Two widely discussed approaches in this regard are: a) parse syntactic dependencies (Popescu & Etzioni, 2007) and b) grammatical relations (Heemskerk et al., 2011; Zhuang et al., 2006). Either approach would require defining the set of relations between the phrases. Defining a specific set of relations would create the problem of high precision but low recall, while a more general set of rules would result in low precision and high recall (Schouten & Frasincar, 2016). The proposed heuristic has partially addressed such concerns. The authors have first searched for all sentiment and attribute phrases in a sentence and then utilized the sentence structure to map them appropriately using a rule-based approach. The rules have been created based on the observation of a few sample reviews.

The authors have noted that the *SentimentPhrases* are not randomly scattered in the sentence; rather, most of them follow one of the two structures appearing either before or after the target attribute. The authors designed two specific rules to address these structures separately.

Rule 1: *SentimentPhrase* precedes *AttributePhrase*: If the sentence in *NewSentenceList* starts with a *SentimentPhrase* and more than one *AttributePhrase* co-exist, insert the last *SentimentPhrase* just before them, in between each pair of *AttributePhrases*; break the sentence after each *AttributePhrase*. Otherwise, if the sentence starts with a *SentimentPhrase* and no two *AttributePhrases* co-exist, simply break the sentence after each *AttributePhrase*. Lines 27–34 in Figure 3 describe the rule.

Rule 2: *AttributePhrase* precedes *SentimentPhrase*: If the sentence in *NewSentenceList* starts with an *AttributePhrase* and two or more *AttributePhrases* co-exist, insert the first *SentimentPhrase* just after them, in between each pair of *AttributePhrases*, and break the sentence before each *AttributePhrase*. Otherwise, if the sentence starts with an *AttributePhrase* and no two attributes co-exist, simply break the sentence before each *AttributePhrase*. Lines 35–42 in Figure 3 describes the rule. The results of the applications of these rules for the running examples appear in Figure 4 (output of step 5). (Review 1: First Rule and Review 2: Second Rule)

Each *SubSentence* generated in the previous section contains only one attribute and corresponding *SentimentPhrase*. Following a dictionary-based approach, the authors propose to compute the sentiment index (SI) of each *SubSentence* and assign it to the attribute therein to determine the attribute-level sentiment indices (ATLSIs) for each review in the dataset.

Aspect-Level Index Generation

The automobile aspects contain more than one attribute (refer to Table A4 in appendix). Therefore, the authors need to add up the attribute level sentiment index (ATLSI) values corresponding to each aspect to compute the ASLSIs. Here, the authors have followed the weighted summation method, and used term-frequency (TF) of attributes as their weight (Abualigah et al., 2017). The intuition behind using TF rating is- "*the more the discussion about the attribute in the text the higher will be its corresponding weight*".

Phase II: Manufacturers' Performance Evaluation

Traditionally, in TOPSIS (Gong, 2017; Yoon & Hwang, 1981), experts evaluate various *alternatives* based on predefined *evaluation criteria* to create the input *decision matrix*. The matrix is used to prioritize the alternatives. the authors have proposed to modify TOPSIS by regarding reviewers as experts, manufacturers as *alternatives*, and automobile aspects as *evaluation criteria*. For a detailed description of the steps involved in modified TOPSIS, refer to Appendix. The proposal for creating the decision matrix to be used as the input for TOPSIS is as follows:

- Divide the available time-span into a number of equal time intervals.
- For the target manufacturers, segregate the aspect-level sentiment indices (ASLSIs) according to time interval.
- Select the inspection criteria (aspects).
- For each manufacturer, for each inspection criterion and for each time interval, calculate the *consumer-perceived performance rating* as:

$$x_{ijt} = \frac{\sum_{k=1}^{l} P_{ijtk}}{\sum_{k=1}^{l} P_{ijtk} - \sum_{k=1}^{l} N_{ijtk}} \tag{1}$$

where x_{ijt} represents the consumer-perceived performance rating of i^{th} manufacturer with respect to j^{th} aspect for t^{th} time interval; l denotes the number of reviewers in the same time-interval; P_{ijtk} and N_{ijtk} represent the positive and negative sentiment index (SI) with respect to i^{th} manufacturer, j^{th} aspect, t^{th} time interval, and k^{th} reviewer respectively. Hence, a $m * n$ decision matrix is obtained as:

$$X_t = \begin{pmatrix} x_{11t} & \cdots & x_{1nt} \\ \vdots & \ddots & \vdots \\ x_{m1t} & \cdots & x_{mnt} \end{pmatrix}$$

where n and m represent the number of evaluation criterion (automobile aspects) and the number of manufacturers in the segment, respectively, for t^{th} time interval.

This decision matrix is used as the input for TOPSIS to compute manufacturers' relative performance indices. Based on these index values, manufacturers are ranked in the segment.

Phase III: Weakness Detection and Root Cause Analysis

Degradation in performance, if any, for a specific manufacturer calls for further analysis of the data to identify the root cause. The authors propose to statistically examine the products' perceived underperformance over the time, the literature suggest the use of statistical quality control tools *(i.e., P*-chart, *U*-chart*)* for the same (Chukhrova, & Johannssen, 2019). Accordingly, the authors propose to use two categories of control chart, *U*-chart and *P*-chart (Laney, 2002), at the aspect and attribute level, respectively. Since an aspect has more than one attribute and can have more than one weakness/defect per unit (in case customers are dissatisfied with respect to more than one attributes corresponding to an aspect), using *U*-chart is appropriate. Whereas an attribute represents only one feature, proportion of nonconformity makes sense here, therefore the authors propose to use *P*-chart at attribute level. While developing control charts, the authors have defined the nonconformity in accordance with the problem as discussed below.

Weakness Detection at Aspect Level Using U-Chart

Considering each time-interval as an inspection sample, the authors have recorded number of nonconformities and the sample size for each aspect. In this context, the authors have defined the number of nonconformities as the *number of attributes bearing negative SI* per review. Based on the recorded data, the authors have developed separate *U*-charts for all aspects. If the data point with respect to the interval of analysis interest crosses the *Upper Control Limit (UCL)*, it indicates consumer-perceived weakness. Such incidences with respect to each of the aspects require analysis at the attribute level as discussed below.

Root Cause Analysis at Attribute Level Using P-Chart

Once weak aspects are identified, the authors need to process the SIs to record the time-interval-wise nonconformities and sample size for each individual underlying attribute. The authors define a nonconformity as an *attribute bearing negative sentiment*. Separate *P*-charts are need to be developed for individual attributes. If the data point with respect to the interval of interest deviate from UCL, the attribute is considered to be weak. This analysis provides a list of underperforming attributes, which would be of interest to manufacturers.

APPLICATION OF THE FRAMEWORK

The authors have applied the proposed framework to a dataset containing 36,558 automobile reviews received from Carwale.com, a well-known car portal in India. The dataset comprises consumer reviews for 53 different car manufacturers for 2006 to 2016.

Phase I: Sentiment Index Generation

In the data set the authors found 315 duplicate reviews, which are deleted. Remaining reviews are pre-processed to eliminate HTML tags, mark-ups, stop words, and duplicate reviews. In addition, the authors have eliminated duplicate sentences and words from the text.

Attribute Identification and Aspect Tagging

Referring to Singh et al. (2020), the pre-processed dataset is processed with the rapid automatic keywords extraction (RAKE) (Rose 2012) algorithm. The authors have extracted 10,000 frequent noun phrases, which are manually checked by three engineering graduates, to create a short list of car attributes. Though the manual intervention is both expensive and time consuming but is necessary to a) eliminate phrases that are not intended to identify car attributes, and b) simplify the task of researchers and practitioners associated with aspect level sentiment analysis in automobile industry as the prepared list of frequent car attributes can easily be reused. Such practices have been adopted widely in ASLSA (Schouten & Frasincar, 2016). During manual scanning, the authors note that many attributes are represented by their synonyms or by different, misspelled words. For instance, the word *maintenance* is present in 21 forms, such as *maintainace, maintained, maitanance*, etc. The authors have fixed the issue by devising rules (in form of python codes) to replacing them with the correct, most standard word. In this phase 260 car attribute phrases are shortlisted and tagged with the respective aspects, following the car ontology the authors have prepared (Figure 2). A particular attribute could belong to one or more listed aspects. To tag the attributes-aspect pair, the authors have circulated the list of all shortlisted attribute phrases, along with the list of target aspects, to a group of three undergraduate students. The authors have collected their responses and opted for a voting-based approach to prepare the list of car aspects and corresponding underlying attributes. The list is further amended by two automobile experts. Their amendments are tabulated in Table A3 in appendix. The able A3 comprising three columns: first, the aspects; second, the common attributes in both lists; third, the uncommon attributes in both lists. Table A3 is then shared with the third expert (*deputy general manager of production of an automobile manufacturer*) with a request to reallocate/include/exclude the attributes as needed. From the list amended by the third expert, the authors note that many similar attributes are represented with different words; the authors have replaced all such words with the standard *AttributePhrase.* The final list is compiled as Table A4 (appendix). From Table A4, it is apparent that few attributes are present with more than one aspect. Such ambiguous terms are made bold and italic in the list. The final list has 252 attributes tagged with 26 aspects.

Aspect-Level Index Generation

Among the various car segments in the dataset, the authors have selected the mid-size segment for performance evaluation. The authors have carried out the intra-segment comparison of manufacturers based on the suggestion of the deputy general manager of *a car manufacturer*. To demonstrate the framework, the authors have compared the manufacturers from mid-sized segment. For segmentation, the authors have referred to the Society of Indian Automobile Manufacturers (SIAM) report. For the list of mid-sized car manufacturers in India and corresponding reviews (irrespective of their models) in the dataset readers are referred to Table III in Singh et al. (2020). From the dataset, it is apparent that very small number of reviews pertained to certain manufacturers, which are supposed to have scope of a huge margin of error. Therefore, the authors have computed maximum possible margin of errors for each manufacturer and tabulated the results in Table A2 (appendix). The computation is based on the average number of reviews per quarter and total vehicles sold in the last quarter by the particular manufacturer (for details readers are referred to Singh et al., 2020). To demonstrate the proposed framework, the authors have selected the five manufacturers with the smallest margin of error from the list and renamed them $\left(M_1, M_2, \ldots, M_5\right)$. The authors have purposefully given pseudonyms to the manufacturers to avoid any kind of market confusion. For further analysis, the authors have separated 21,852 reviews corresponding to the five abovementioned manufacturers from the dataset.

The authors have tokenized the review into sentences using the *sent_tokenize* module of *NLTK package in Python* (Bird et al., 2009) and applied the *SubSentenceGeneration heuristic* (Figure 3) to split them into sub-sentences containing *"only one attribute and corresponding opinionated phrase(s)"*. These sub-sentences are further tokenized into words with the *word_tokenise* (Bird et al., 2009) module. Next, the authors have lemmatized each word using *WordNetLemmatizer* and compared the word and its lemma with the same in *SentiWordNet*, a general-purpose lexicon. The authors have created an updated list of sentences by keeping either the original word or its lemma, whichever is present in *SentiWordNet.* These updated sentences are processed to compute sentence-level *SI* by aggregating the sentiment scores of individual words. Next, the sentences are checked if sentiment shifters are present. In case of shifters are present, the sentiment index of sentence is modified referring to Yu et al. (2016), and Singh et al. (2020). The sentiment score is assigned to the attribute present in sub-sentence. The sentiment quantification is done for all the sentences in a review and all the reviews in the dataset.

Finally, referring to the Table A4 (appendix), the authors have aggregated the SIs of all attributes within a particular aspect to generate ASLSI for all 26 aspects. The authors have used a weighted summation method to compute the aspect level index. As discussed in previous section, the TF rating of the attribute determined from the entire corpus is used as the weight. The authors have generated such aspect-level indices for the individual reviews of all five manufacturers involved in present analysis. With these indices, the authors have generated a table comprising 21,852 rows and 26 columns, where the rows indicate the reviews and the columns refer to the aspects. The table turns out to be sparse because many aspects are not discussed frequently in this set of reviews. Based on the frequency of occurrence of the aspects in the review dataset, the authors have selected 12 most frequent aspects for further analysis (for frequency table refer to Table A5 appendix). The selected aspects are, ACCESSORIES (C_1), PASSENGER AND DRIVER COMFORT (C_2), SAFETY (C_3), EXTERIORS AND APPEARANCE (C_4), INTERIORS (C_5), MILEAGE (C_6), DRIVE SYSTEM (C_7), DRIVING & CONTROL (C_8), STORAGE CAPACITY (C_9), ENGINE ASSEMBLY (C_{10}), TRANSMISSION ASSEMBLY (C_{11}), and SPARE PARTS (C_{12}).

Phase II: Manufacturers' Performance Evaluation

The abovementioned aspects are used as the evaluation criteria in modified TOPSIS (Deng et al., 2000) to compare manufacturers' performance. The dataset includes reviews from June 2006 to May 2016 (41 quarters). The dataset is incomplete for the second quarters of the years 2006 and 2016; therefore, the authors have excluded those quarters from the analysis. From the remaining data, for demonstration purposes, the authors have selected the last five consecutive quarters for analysis in this phase. As discussed in last section, the TOPSIS input decision matrices for the selected quarters are prepared using Equation 1. The resulting matrix appears in Table A6 (appendix). It is further normalized with Equation 2 (appendix) to obtain positive and negative ideal solutions using Equations 4 and 5 (appendix) for all five quarters (refer to Table 2). The intermediate results can be found in Table A7 in appendix. The criterion weights are also obtained from the decision matrix (Table A6 in appendix) by computing row-wise standard deviation (SD). Table 3 shows the criteria-wise weights for each quarter.

Table 2. Ideal solutions

Quarters	Ideal Solution	Criterion											
		C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12
Q1	Positive	0.2381	0.2069	0.2240	0.2034	0.2031	0.2355	0.2389	0.2654	0.2069	0.2358	0.2434	0.2539
	Negative	0.1186	0.1841	0.1530	0.1964	0.1924	0.1609	0.0956	0.1056	0.1781	0.1067	0.1664	0.1274
Q2	Positive	0.2206	0.2134	0.2778	0.2076	0.2211	0.2111	0.2197	0.2536	0.2325	0.2191	0.2185	0.2110
	Negative	0.1689	0.1569	0.0937	0.1928	0.1753	0.1820	0.1698	0.1155	0.1347	0.1715	0.1644	0.1792
Q3	Positive	0.2268	0.2040	0.2260	0.2036	0.2046	0.2130	0.2356	0.2177	0.2026	0.2350	0.2192	0.2193
	Negative	0.1591	0.1920	0.1707	0.1967	0.1940	0.1674	0.1598	0.1561	0.1977	0.1594	0.1580	0.1717
Q4	Positive	0.2404	0.2159	0.2152	0.2092	0.2175	0.2261	0.2199	0.2417	0.2364	0.2202	0.2157	0.2206
	Negative	0.1097	0.1758	0.1848	0.1932	0.1834	0.1687	0.1484	0.1298	0.1444	0.1485	0.1813	0.1787
Q5	Positive	0.2481	0.2060	0.2615	0.2082	0.2118	0.2750	0.2214	0.2314	0.2117	0.2217	0.2717	0.2361
	Negative	0.1396	0.1858	0.0922	0.1885	0.1832	0.0630	0.1477	0.1597	0.1795	0.1480	0.1254	0.1189

Where: accessories (C_1), passenger and driver comfort (C_2), safety (C_3), exteriors and appearance (C_4), interiors (C_5), mileage (C_6), drive system (C_7), driving & control (C_8), storage capacity (C_9), engine assembly (C_{10}), transmission assembly (C_{11}), and spare parts (C_{12}),

Table 3. Criterion weights

Quarters	Criterion											
	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12
Q1	0.1197	0.0225	0.0785	0.0075	0.0109	0.0676	0.1458	0.1881	0.0303	0.1307	0.0740	0.1245
Q2	0.0691	0.0741	0.2058	0.0214	0.0510	0.0404	0.0656	0.1677	0.1226	0.0630	0.0788	0.0405
Q3	0.1183	0.0229	0.1224	0.0129	0.0209	0.0882	0.1311	0.1195	0.0114	0.1305	0.1255	0.0965
Q4	0.1690	0.0468	0.0441	0.0199	0.0459	0.0749	0.0940	0.1666	0.1358	0.0937	0.0502	0.0591
Q5	0.0944	0.0184	0.1592	0.0238	0.0262	0.1829	0.0698	0.0659	0.0313	0.0700	0.1459	0.1121

Where: accessories (C_1), passenger and driver comfort (C_2), safety (C_3), exteriors and appearance (C_4), interiors (C_5), mileage (C_6), drive system (C_7), driving & control (C_8), storage capacity (C_9), engine assembly (C_{10}), transmission assembly (C_{11}), and spare parts (C_{12}),

Finally, the relative performance indices for each quarter are computed using Equations 6–8 (appendix). The relative performance indices, along with the corresponding ranks of the manufacturers, appear in Table 4. In analysing the results presented in Table 4, it is observed that M3 is the worst-performing manufacturer in the last quarter and its performance is continuously deteriorating over the previous three quarters. Next, the authors have considered this particular case to explain the weakness detection phase of the proposed approach.

Table 4. Manufacturers' relative performance index and ranks

Manufacturer	Quarter									
	Q1		Q2		Q3		Q4		Q5	
	Index	Rank	Index	Rank	Index	Rank	Index	Rank	Index	Rank
M1	0.5088	4	0.5536	4	0.7464	1	0.8661	1	0.6862	2
M2	0.6785	2	0.4517	5	0.2591	5	0.6405	3	0.6161	4
M3	0.5251	3	0.6344	3	0.7430	2	0.5299	4	0.2363	5
M4	0.5007	5	0.6565	1	0.6067	3	0.3901	5	0.6578	3
M5	0.7917	1	0.6351	2	0.5868	4	0.6491	2	0.8123	1

Phase III: Weakness Detection

To analyze the root cause of performance degradation of the manufacturer M3, the authors have performed two tasks. First, the authors have identified the weak aspects using the *U*-chart. Second, the authors have explored their attribute-level details and discovered the weak attributes using the *P*-chart. The control charts are prepared using Minitab 17. Traditionally, the control charts have used actual defects as the input. As discussed in last section and shown below, the authors have used SIs as inputs. Since the weakness are encountered within the last few quarters, the perceived performance degradation might have begun at an earlier time and the manufacturer have failed to adopt remedial measures. Therefore, in the analysis presented below, the authors have used the data from all past quarters beginning in 2006. Altogether, the authors have 39 data points.

Weakness Detection at Aspect Level Using U-Chart

Treating each quarter as an individual inspection sample, number of reviews as the sample size, and *number of attributes* bearing negative SIs corresponding to a particular aspect as the nonconformities, the authors have prepared separate *U*-charts for each of the 12 aspects as presented in Figure 5 (a-l). The authors have analyzed the charts in Figure 5 for the last three quarters because the manufacturer's performance was continuously degrading for those quarters. It is noted that the second last-quarter data point for the aspect *DRIVING & CONTROL* has crossed the UCL, indicating that manufacturer M3 to be underperforming with respect to this particular aspect. As discussed in last section, each aspect is represented by a set of attributes. Therefore, attributes pertaining to the aspect *DRIVING & CONTROL* must be analyzed to detect the underperforming attributes.

Root Cause Analysis at Attribute Level Using P-Chart

The authors have used *P*-chart for attribute-level weakness detection. This chart plots the proportion of non-conformity over time. The authors have treated a negative perception about an automobile attribute as equivalent to non-conformity. Therefore, the proportion of negative reviews to total reviews in a specific quarter is considered while constructing the *P*-chart. In the last section, the authors have observed that *DRIVING & CONTROL* is an especially weak aspect. This aspect contains many attributes; however, only a few attributes contain negative sentiments. The attributes bearing negative SI are: *grip*, *control*, *ground clearance*, *parking*, *ride handling*, *steering*, *suspension*, and *touch screen.* The authors have prepared separate *P*-charts for these attributes, as presented in Figure 6. Interpreting these charts, it is noted that the second last-quarter data point crossed the UCL in Figure 6 (f). This indicates underperformance with respect to the attribute *{Steering}* under *DRIVING & CONTROL.*

Figure 5. U-charts for Accessories (a), Passenger and Driver Comfort (b), Safety (c), Exteriors and Appearance (d), Interiors (e) Mileage (F), Drive System (g), Driving & Control (h), Storage Capacity (i), Engine Assembly (j) Transmission Assembly (k), and Spare Parts (l)

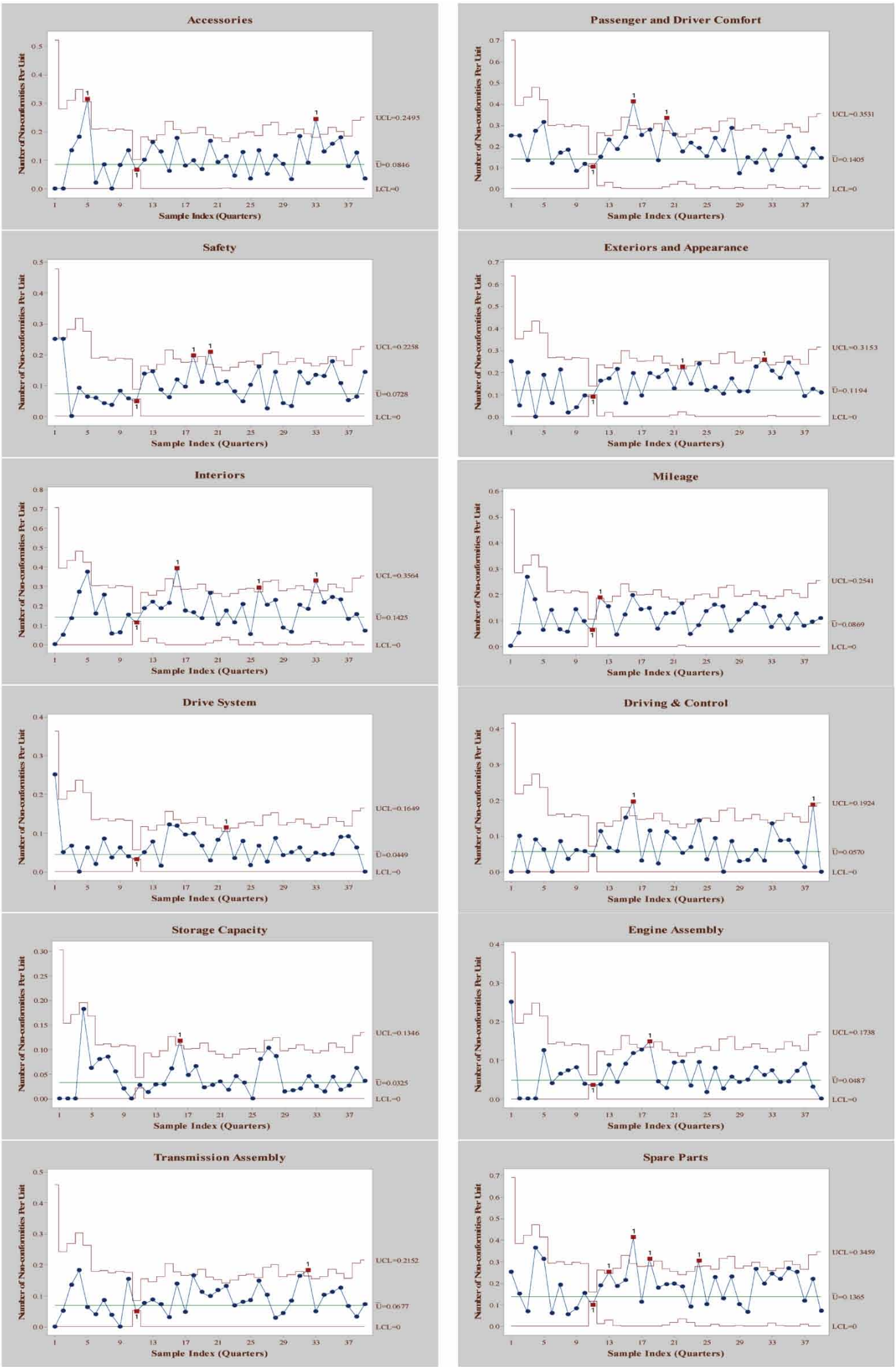
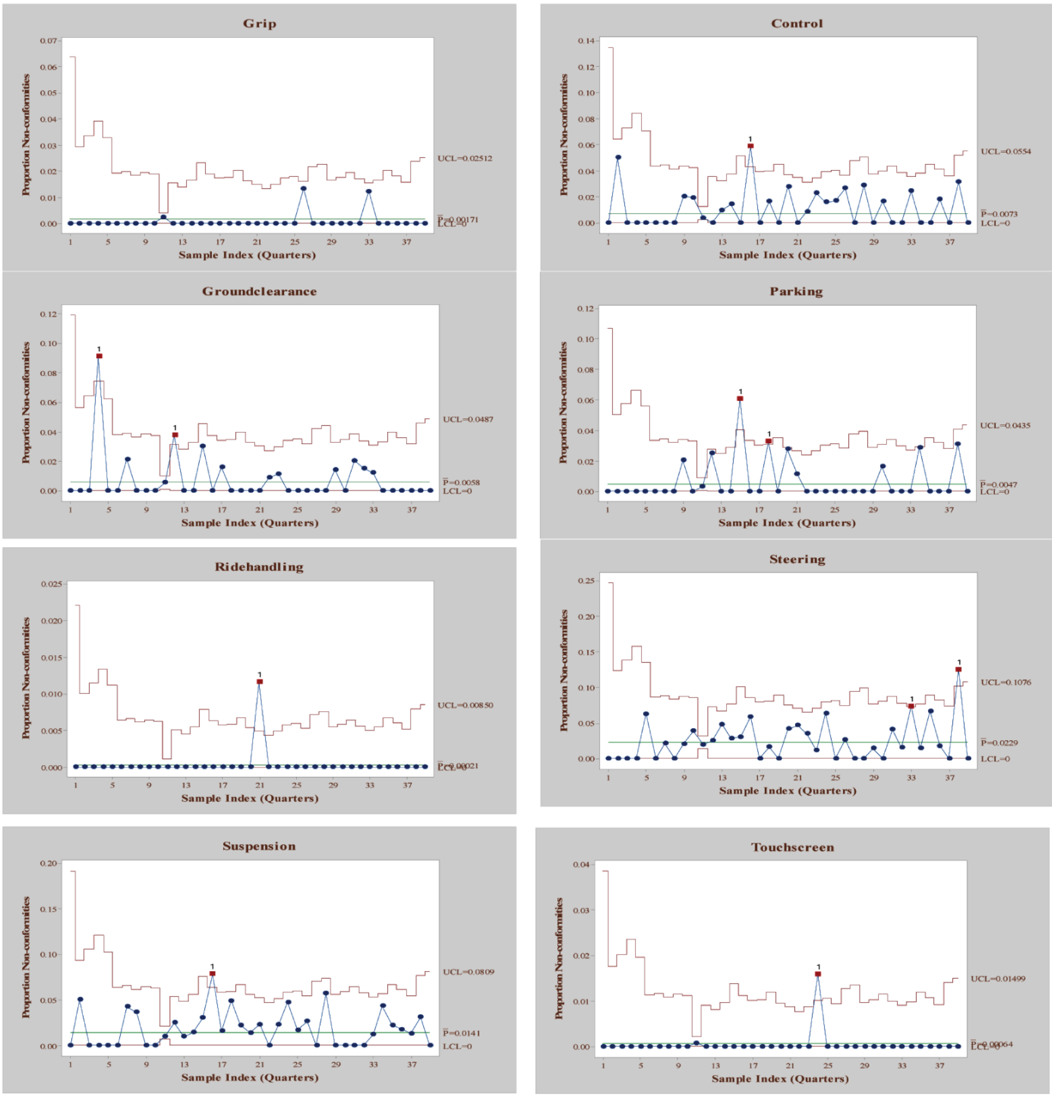


Figure 6. P-charts for Grip (a), Control (b), Ground Clearance (c), Parking (d), ride handling (e) Steering (f), Suspension (g), and Touchscreen



Evaluation Measures

The correctness of the analysis depends on the quantitative performance of attribute extraction ans sentiment quantification algorithm. Researchers measure them separately using precision, recall, accuracy, and F_1 measures (Lu et al., 2011; Zhao et al., 2011; Yu et al., 2011). In order to measure the performance of the algorithm adopted in the present research, the authors have randomly selected 500 reviews for manual annotation and made two copies of the reviews. One copy is distributed among five mechanical engineering undergraduate students (100 reviews each). The authors requested them to i) tag the automobile attributes present in the review and ii) detect the polarity of the text corresponding to each specific attribute. The polarity indicates whether a text possesses a positive or negative sentiment (Kar, & Dwivedi, 2020). The second copy of reviews is annotated by the research lead. The authors have received 200 mutually annotated reviews with a total of 1,416 attributes tagged.

In several cases, conflicting annotations between the research lead and the other annotators are noted, which the authors resolved by preferring the annotation of the research lead.

Finally, based on the annotators' responses, the authors separately measured the performance for attribute extraction and sentiment index generation. During the attribute annotation, annotators have detected 1,416 attributes, whereas the algorithm proposed in the present research could detect only 1,266 attributes, out of which only 1,206 are correctly detected. Based on these numerical values, the authors have computed the performance evaluation metrics and compiled the results in Table 6. The metrics used in the computation appear in the tables itself: where, N is the total number of reviews; C_I is the number of actual attributes in review I; E_I is the number of attributes extracted by the algorithm from review I; EC_I is the number of correctly extracted attributes from review I. From the table, the authors have noted that the values for *Precision*, *Recall*, *and F1-Measure* are 95.62%, 85.16%, and 89.93%, respectively. *Precision* value shows that out of total attributes extracted by the algorithm, 95.62% were correctly identified. *Recall* value shows that out of total attributes present in the reviews 85.16% were correctly identified by the algorithm. Whereas *F1-Measure* represents the geometric mean of *Precision*, *and Recall.* For sentiment index generation, based on the responses from annotators and the heuristic, the authors prepared a confusion matrix as presented in Table 5, wherein the number rue positive (TP), true negative (TN), false positive (FP), and false negative (FN) instances are found to be 671, 119, 113, and 115, respectively. Based on these values, the authors computed the performance matrices as presented in Table 6. The values for *Precision, Recall*, *and* F_1*-Measure*, were 85.59%, 85.37%, and 85.48%, respectively. This algorithm is also used by Singh, et al. (2020). *Precession* score here indicates that out of the total positive sentiments identified by the algorithm, 85.59% were correctly captured. Whereas the *Recall* score suggests that out of total exact positive sentiment scores in the reviews, the algorithm could only discover 85.37%. Here one may note total correctly identified attributes were 1206, whereas in confusion matrix total attributes are 1018, the difference is because 188 out of 1206 reviews bear neutral sentiment index.

DISCUSSION

In this research the authors have proposed and tested a text analytics framework that not only ranks the manufacturers among the competitors but also discovers their consumer perceived weaknesses. The authors have selected five manufacturers from the mid-sized segment in India, compared them based on the consumers' perceptions with respect to the features, ACCESSORIES, PASSENGER AND DRIVER COMFORT, SAFETY, EXTERIORS AND APPEARANCE, INTERIORS, MILEAGE, DRIVE SYSTEM, DRIVING & CONTROL, STORAGE CAPACITY, ENGINE ASSEMBLY, TRANSMISSION ASSEMBLY, AND SPARE PARTS, and found M3 to be the least performing. The authors have also discovered the reason for the underperformance and found that customers were not happy with the feature driving and control. Probably because of customers' dissatisfaction with the aforementioned aspect/feature, the manufacturer performed weak.

Table 5. The confusion matrix for sentiment index generation heuristic

N=1206	Predicted as negative	Predicted as positive	
Annotated as negative	TN = 119	FP = 113	232
Annotated as positive	FN = 115	TP = 671	786
	234	784	

Table 6. Algorithm evaluation measures

Evaluation measures	Attribute extraction algorithm		Sentiment index generation heuristic	
	Metric (Liu et al., 2005)	Values	Metric (Tripathy et al., 2016)	Values
Precession	$P = \sum_{I=1}^{N} EC_I \Big/ \sum_{I=1}^{N} E_I$	0.9526	$P = \frac{TP}{TP + FP}$	0.8559
Recall	$R = \sum_{I=1}^{N} EC_I \Big/ \sum_{I=1}^{N} C_I$	0.8516	$R = \frac{TP}{TP + FN}$	00.8537
F_1-Measure	$F_1 = \frac{2 \times P \times R}{P + R}$	0.8993	$F_1 = \frac{2 \times P \times R}{P + R}$	0.8548
Accuracy	NA	NA	$A = \frac{TP + TN}{TP + TN + FP + FN}$	0.7760

THEORETICAL IMPLICATIONS

The present research makes several contributions to the existing body of knowledge. *First,* it has proposed a customized aspect level sentiment quantification algorithm. *Second,* it has also proposed a sub-sentence generation heuristic that deals with the sentences with more than one attributes. *Third,* it has proposed a replacement of expert inputs in TOPSIS through the sentiments extracted from reviews. *Fourth,* it has seamlessly integrated the aspect level sentiment analysis with P- and U- charts to statistically discover the consumer perceived weaknesses of the manufacturers.

MANAGERIAL IMPLICATIONS

The proposed framework can be used as decision support by customers, manufacturers, and component suppliers. A customer can rank the manufacturers to identify the best option in a specific product line and compare the performance of various cars at an individual system or subsystem level. The proposed framework can also help manufacturers in- i) keeping track of consumers' current interests by extracting most frequently discussed features from the online discussions; ii) comparing their performance with their competitors using the method proposed in the Phase II of the framework; discovering the performance, weaknesses, and strengths of their competitors using the last phase of the framework. A manufacturer can use the proposed framework to monitor its consumer-perceived market performance over time to facilitate informed decision making for improvement. If degradation in ranking occurs, the manufacturer can identify the origin of the problem at the aspect level and the root cause at the attribute level. This retrospective analysis may help them with continuous improvement. If the manufacturers are informed by their customers through online reviews that they have mistaken somewhere which may lead to hazardous events in the long run, they may initiate product recalls. In a typical automobile manufacturing setup, many subsystems and components are sourced from various suppliers. The proposed framework can facilitate early warning for such component suppliers regarding the weaknesses encountered in the components they provide, which would help not only in building a better relationship with the manufacturer by enabling the suppliers to correct their weaknesses promptly, but also in discovering new business prospects by identifying the weaknesses of competitors.

CONCLUSION

This research represents an attempt to connect aspect-level consumer sentiment extracted from online reviews with statistical tools for quality control to summarize the reviews in form of control-charts which are easily comprehensible by the operations management community. The proposed three-phase framework integrates sentiment analysis to mine aspect-level consumer sentiments and analyze them for i) manufacturers' market performance evaluation, using TOPSIS, ii) weakness detection at the aspect level, using a *U*-chart, and iii) root cause analysis at the attribute level using a *P*-chart. To demonstrate the proposed framework, the authors have selected five manufacturers from the mid-size car segment, identified the weakest-performing one, and discovered its weaknesses at the aspect and attribute levels. The proposed semi-automatic approach for aspect identification was validated with human annotators. The major contributions of the present research include an ontology for passenger cars, a semi-automatic method for aspect identification and sentiment index generation, use of TOPSIS with inputs as a sentiment index, and corroborating control charts as visualization tools for aggregating the perceived market sentiment.

LIMITATIONS AND FUTURE WORK

The present research possesses certain limitations:

- *U*-charts and *P*-charts are usually developed on the assumption that the random sample is selected from a large population, which is violated in the case of online reviews. Moreover, in the present analysis, the sample size (the number of reviews per quarter) was not large enough, leaving a margin of error from 9% to 14%. Therefore, the results drawn from these charts cannot directly be used as representatives of consumer-perceived weakness; rather, they can only be treated as indicators requiring further investigation.
- Based on a random sample of 100 reviews, the authors observe that 33.46% of the text was expressed non-emotively which was found to be useful by research community. The present study did not account for it.
- The authors collected reviews from Carwale the present study may be subject to self-selection bias.
- The authors do not use the reviews that compare products since the proposed approach is not appropriate for multi-product aspect extraction.
- This research focuses towards consumer review-based quality analytics framework which extracts qualitative patterns from automobile reviews. Since, this research focuses on application of sentiment analysis, the authors did not compare the performance of the proposed sentiment extraction algorithm with the other sentiment dictionaries (i.e. SenticNet-3 (Cambria et al., 2014), Opinionfinder (Wilson et al., 2005) etc.). Hence, leaving a scope for the future researchers to compare such dictionaries.
- The authors have adopted a dictionary-based approach due to the scarcity of annotated data in the automobile industry. Future researchers may consider manual annotation and apply machine learning approaches to investigate the same issue.
- *The F1-Measure* for the algorithm used in the present research is only 89.93%, for attribute extraction and 85.48%, for sentiment index generation. Hence, leaving a scope for the future researchers to improve it.
- In present the scenario, managing with the Misinformation is an issue (Song et al., 2017; Aswani et al., 2018; Aswani et al., 2019). In case of online reviews, it is noted that sometimes organizations manage fake positive reviews for themselves and negative reviews for their competitors. Present research did not account for detecting and eliminating them. Researchers may extend it in future.

Keeping such biases and assumptions in mind, the authors recommend supplementing the findings of the present research with other diagnostic tools available in the domain of quality assessment and market research.

ACKNOWLEDGMENT

The authors are thankful to the Ministry of Human Resource Development (MHRD), New Delhi, for sponsoring this research. Sanction letter No: F. NO. 5-5/2014-TS, VII. The authors are also grateful to CarWale.com and SIAM for providing the datasets. The authors are also thankful to the reviewers and editor of the manuscript for their suggestions to improve it.

FUNDING AGENCY

The publisher has waived the Open Access Processing fee for this article.

REFERENCES

Abrahams, A. S., Fan, W., Wang, G. A., Zhang, Z. J., & Jiao, J. (2015). An Integrated Text Analytic Framework for Product Defect Discovery. *Production and Operations Management*, *24*(6), 975–990. doi:10.1111/poms.12303

Abrahams, A. S., Jiao, J., Fan, W., Wang, G. A., & Zhang, Z. (2013). What's buzzing in the blizzard of buzz? Automotive component isolation in social media postings. *Decision Support Systems*, *55*(4), 871–882. doi:10.1016/j.dss.2012.12.023

Abrahams, A. S., Jiao, J., Wang, G. A., & Fan, W. (2012). Vehicle defect discovery from social media. *Decision Support Systems*, *54*(1), 87–97. doi:10.1016/j.dss.2012.04.005

Abualigah, L. M., Khader, A. T., Al-Betar, M. A., & Alomari, O. A. (2017). Text feature selection with a robust weight scheme and dynamic dimension reduction to text document clustering. *Expert Systems with Applications*, *84*, 24–36. doi:10.1016/j.eswa.2017.05.002

Araújo, M., Pereira, A., & Benevenuto, F. (2020). A comparative study of machine translation for multilingual sentence-level sentiment analysis. *Information Sciences*, *512*, 1078–1102. doi:10.1016/j.ins.2019.10.031

Archak, N., Ghose, A., & Ipeirotis, P. G. (2011). Deriving the Pricing Power of Product Features by Mining Consumer Reviews. *Management Science*, *57*(8), 1485–1509. doi:10.1287/mnsc.1110.1370

Aswani, R., Kar, A. K., & Ilavarasan, P. V. (2018). Detection of spammers in twitter marketing: A hybrid approach using social media analytics and bio inspired computing. *Information Systems Frontiers*, *20*(3), 515–530. doi:10.1007/s10796-017-9805-8

Aswani, R., Kar, A. K., & Ilavarasan, P. V. (2019). Experience: Managing Misinformation in Social Media—Insights for Policymakers from Twitter Analytics. *Journal of Data and Information Quality*, *12*(1), 1–18. doi:10.1145/3341107

Baccianella, S., Esuli, A., & Sebastiani, F. (2009). Multi-facet rating of product reviews. Lecture Notes in Computer Science, 5478, 485–496. doi:10.1007/978-3-642-00958-7_43

Bhat, S. K., & Culotta, A. (2017, March 1). Identifying leading indicators of product recalls from online reviews using positive unlabeled learning and domain adaptation. *ArXiv Preprint ArXiv:1703.00518*. https://arxiv.org/abs/1703.00518

Bhatia, P., Ji, Y., & Eisenstein, J. (2015). Better document-level sentiment analysis from RST discourse parsing. *Conference Proceedings - EMNLP 2015: Conference on Empirical Methods in Natural Language Processing*, 2212–2218. doi:10.18653/v1/D15-1263

Bird, S., Klein, E., & Loper, E. (2009). Natural Language Processing with Python. Journal of Endodontics. doi:10.1097/00004770-200204000-00018

Cambria, E., Olsher, D., & Rajagopal, D. (2014). SenticNet 3: A common and common-sense knowledge base for cognition-driven sentiment analysis. *Proceedings of the National Conference on Artificial Intelligence*, *2*, 1515–1521.

Chang, W. L. (2019). The Impact of Emotion: A Blended Model to Estimate Influence on Social Media. *Information Systems Frontiers*, *21*(5), 1137–1151. doi:10.1007/s10796-018-9824-0

Chang, W. L., & Chen, Y. P. (2019). Way too sentimental? a credible model for online reviews. *Information Systems Frontiers*, *21*(2), 453–468. doi:10.1007/s10796-017-9757-z

Chintagunta, P. K., Gopinath, S., & Venkataraman, S. (2010). The effects of online user reviews on movie box office performance: Accounting for sequential rollout and aggregation across local markets. *Marketing Science*, *29*(5), 944–957. doi:10.1287/mksc.1100.0572

Choi, Y., & Lee, H. (2017). Data properties and the performance of sentiment classification for electronic commerce applications. *Information Systems Frontiers*, *19*(5), 993–1012. doi:10.1007/s10796-017-9741-7

Chong, A. Y. L., Ch'ng, E., Liu, M. J., & Li, B. (2017). Predicting consumer product demands via Big Data: The roles of online promotional marketing and online reviews. *International Journal of Production Research*, *55*(17), 5142–5156. doi:10.1080/00207543.2015.1066519

Chukhrova, N., & Johannssen, A. (2019). Hypergeometric p-chart with dynamic probability control limits for monitoring processes with variable sample and population sizes. *Computers & Industrial Engineering*, *136*, 681–701. doi:10.1016/j.cie.2019.07.056

Deng, H., Yeh, C. H., & Willis, R. J. (2000). Inter-company comparison using modified TOPSIS with objective weights. *Computers & Operations Research*, *27*(10), 963–973. doi:10.1016/S0305-0548(99)00069-6

Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., & Galanos, V. et al. (2019). Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management*, *101994*. Advance online publication. doi:10.1016/j.ijinfomgt.2019.08.002

Dwivedi, Y. K., Ismagilova, E., Hughes, D. L., Carlson, J., Filieri, R., Jacobson, J., & Kumar, V. et al. (2020). Setting the future of digital and social media marketing research: Perspectives and research propositions. *International Journal of Information Management*, *102168*. Advance online publication. doi:10.1016/j.ijinfomgt.2020.102168

Fernández-Gavilanes, M., Álvarez-López, T., Juncal-Martínez, J., Costa-Montenegro, E., & Javier González-Castaño, F. (2016). Unsupervised method for sentiment analysis in online texts. *Expert Systems with Applications*, *58*, 57–75. doi:10.1016/j.eswa.2016.03.031

Gamon, M., Aue, A., Corston-Oliver, S., & Ringger, E. (2005). Pulse: Mining customer opinions from free text. Lecture Notes in Computer Science, 3646, 121–132. doi:10.1007/11552253_12

García-Díaz, J. A., Cánovas-García, M., & Valencia-García, R. (2020). Ontology-driven aspect-based sentiment analysis classification: An infodemiological case study regarding infectious diseases in Latin America. *Future Generation Computer Systems*, *112*, 641–657. doi:10.1016/j.future.2020.06.019 PMID:32572291

GevaT.Oestreicher-SingerG.EfronN.ShimshoniY. (2013). Using Forum and Search Data for Sales Prediction of High-Involvement Products. Ssrn. 10.2139/ssrn.2294609

Grover, P., Kar, A. K., Dwivedi, Y. K., & Janssen, M. (2019). Polarization and acculturation in US Election 2016 outcomes–Can twitter analytics predict changes in voting preferences. *Technological Forecasting and Social Change*, *145*, 438–460. doi:10.1016/j.techfore.2018.09.009

Gong, Y. (2017). Estimating participants for knowledge-intensive tasks in a network of crowdsourcing marketplaces. *Information Systems Frontiers*, *19*(2), 301–319. doi:10.1007/s10796-016-9674-6

Gruss, R., Abrahams, A. S., Fan, W., & Wang, G. A. (2018). By the numbers: The magic of numerical intelligence in text analytic systems. *Decision Support Systems*, *113*, 86–98. doi:10.1016/j.dss.2018.07.004

Guo, C., Du, Z., & Kou, X. (2018). Products Ranking Through Aspect-Based Sentiment Analysis of Online Heterogeneous Reviews. *Journal of Systems Science and Systems Engineering*, *27*(5), 542–558. doi:10.1007/s11518-018-5388-2

Heemskerk, C. J. M., Elzendoorn, B. S. Q., Magielsen, A. J., & Schropp, G. Y. R. (2011). Verifying elementary ITER maintenance actions with the MS2 benchmark product. *Fusion Engineering and Design*, *86*(9–11), 2064–2066. doi:10.1016/j.fusengdes.2011.04.056

Hou, F., Li, B., Chong, A. Y. L., Yannopoulou, N., & Liu, M. J. (2017). Understanding and predicting what influence online product sales? A neural network approach. *Production Planning and Control*, *28*(11–12), 964–975. doi:10.1080/09537287.2017.1336791

Hu, M., & Liu, B. (2004). Mining opinion features in customer reviews. *American Association for Artificial Intelligence*, *4*(4), 755–760. doi:10.1145/1014052.1014073

Kar, A. K., & Dwivedi, Y. K. (2020). Theory building with big data-driven research–Moving away from the "What" towards the "Why". *International Journal of Information Management*, *54*, 102205. doi:10.1016/j.ijinfomgt.2020.102205

Kouloumpis, E., Wilson, T., & Moore, J. (2011). Twitter Sentiment Analysis: The Good the Bad and the OMG! *Int'l AAAI Conference on Weblogs and Social Media (ICWSM).* http://www.aaai.org/ocs/index.php/ICWSM/ICWSM11/paper/download/2857/3251?height%253D90%25%2526iframe%253Dtrue%2526width%253D90%25

Laney, D. B. (2002). Improved control charts for attributes. *Quality Engineering*, *14*(4), 531537. doi:10.1081/QEN-120003555

Law, D., Gruss, R., & Abrahams, A. S. (2017). Automated defect discovery for dishwasher appliances from online consumer reviews. *Expert Systems with Applications*, *67*, 84–94. doi:10.1016/j.eswa.2016.08.069

Lin, Y., Wang, X., & Zhou, A. (2016). Opinion spam detection. *Opinion Analysis for Online Reviews*, *5*(1), 79–94. doi:10.1142/9789813100459_0007

Liu, B., Hu, M., & Cheng, J. (2005). Opinion observer. *Proceedings of the 14th International Conference on World Wide Web - WWW '05*, 342. doi:10.1145/1060745.1060797

Liu, Y., Jiang, C., & Zhao, H. (2018). Using contextual features and multi-view ensemble learning in product defect identification from online discussion forums. *Decision Support Systems*, *105*, 1–12. doi:10.1016/j.dss.2017.10.009

Lu, B., Ott, M., Cardie, C., & Tsou, B. K. (2011). Multi-aspect sentiment analysis with topic models. *Proceedings - IEEE International Conference on Data Mining, ICDM*, 81–88. doi:10.1109/ICDMW.2011.125

Lyu, C., Foster, J., & Graham, Y. (2020). *Improving Document-Level Sentiment Analysis with User and Product Context*. arXiv preprint arXiv:2011.09210. 10.18653/v1/2020.coling-main.590

Mankad, S., Han, H., Goh, J., & Gavirneni, S. (2016). Understanding online hotel reviews through automated text analysis. *Service Science*, *8*(2), 124–138. doi:10.1287/serv.2016.0126

Moghaddam, S., & Ester, M. (2010). Opinion digger: an unsupervised opinion miner from unstructured product reviews. *Proceedings of the 19th ACM International Conference on Information and Knowledge Management*, 1825–1828. https://dl.acm.org/citation.cfm?id=1871739

Mudinas, A., Zhang, D., & Levene, M. (2012a). Combining lexicon and learning based approaches for concept-level sentiment analysis. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. doi:10.1145/2346676.2346681

Peng, Y., Kou, G., & Li, J. (2014). A Fuzzy PROMETHEE Approach for Mining Customer Reviews in Chinese. *Arabian Journal for Science and Engineering*, *39*(6), 5245–5252. doi:10.1007/s13369-014-1033-7

Polanyi, L., & Zaenen, A. (2006). Contextual Valence Shifters. In Computing Attitude and Affect in Text: Theory and Applications (pp. 1–10). doi:10.1007/1-4020-4102-0_1

Popescu, A. M., & Etzioni, O. (2007). Extracting product features and opinions from reviews. In Natural Language Processing and Text Mining (pp. 9–28). doi:10.1007/978-1-84628-754-1_2

Qi, J., Zhang, Z., Jeon, S., & Zhou, Y. (2016). Mining customer requirements from online reviews: A product improvement perspective. *Information & Management*, *53*(8), 951–963. doi:10.1016/j.im.2016.06.002

Ray, P., & Chakrabarti, A. (2020). A mixed approach of deep learning method and rule-based method to improve aspect level sentiment analysis. *Applied Computing and Informatics*. 10.1016/j.aci.2019.02.002

Rose, S. J., Cowley, W. E., Crow, V. L., & Cramer, N. O. (2012). *Rapid automatic keyword extraction for information retrieval and analysis.* U.S. Patent 8,131,735. https://patents.google.com/patent/US8131735B2/en

Rustamov, S., Mustafayev, E., & Clements, M. A. (2013). Sentiment analysis using Neuro-Fuzzy and Hidden Markov models of text. *Conference Proceedings - IEEE SOUTHEASTCON*. doi:10.1109/SECON.2013.6567382

Sann, R., & Lai, P. C. (2020). Understanding homophily of service failure within the hotel guest cycle: Applying NLP-aspect-based sentiment analysis to the hospitality industry. *International Journal of Hospitality Management*, *91*, 102678. doi:10.1016/j.ijhm.2020.102678

Saumya, S., Singh, J. P., & Dwivedi, Y. K. (2019). Predicting the helpfulness score of online reviews using convolutional neural network. *Soft Computing*, 1–17. doi:10.1007/s00500-019-03851-5

Schouten, K., & Frasincar, F. (2016). Survey on Aspect-Level Sentiment Analysis. *IEEE Transactions on Knowledge and Data Engineering*, *28*(3), 813–830. doi:10.1109/TKDE.2015.2485209

Shelke, N. M., Thakre, V., & Deshpande, S. (2017). Identification of scope of valence shifters for sentiment analysis of product reviews. *Proceedings - 2016 6th International Symposium on Embedded Computing and System Design, ISED 2016*, 265–269. doi:10.1109/ISED.2016.7977094

Singh, A., Jenamani, M., & Thakkar, J. (2020). Do online consumer reviews help to evaluate the performance of automobile manufacturers? *Journal of Enterprise Information Management*, *33*(5), 1153–1198. doi:10.1108/JEIM-09-2019-0292

Song, L., Lau, R. Y. K., Kwok, R. C. W., Mirkovski, K., & Dou, W. (2017). Who are the spoilers in social media marketing? Incremental learning of latent semantics for social spam detection. *Electronic Commerce Research*, *17*(1), 51–81. doi:10.1007/s10660-016-9244-5

Stieglitz, S., Meske, C., Ross, B., & Mirbabaie, M. (2020). Going back in time to predict the future-the complex role of the data collection period in social media analytics. *Information Systems Frontiers*, *22*(2), 395–409. doi:10.1007/s10796-018-9867-2

Swain, A. K., & Cao, R. Q. (2019). Using sentiment analysis to improve supply chain intelligence. *Information Systems Frontiers*, *21*(2), 469–484. doi:10.1007/s10796-017-9762-2

Thelwall, M., Buckley, K., & Paltoglou, G. (2012). Sentiment strength detection for the social web. *Journal of the American Society for Information Science and Technology*, *63*(1), 163–173. doi:10.1002/asi.21662

Ting, T., Davis, J., & Pettit, F. (2014). *Online marketing research utilizing sentiment analysis and tunable demographics analysis*. US Patent 8,694,357.

Tripathy, A., Agrawal, A., & Rath, S. K. (2016). Classification of sentiment reviews using n-gram machine learning approach. *Expert Systems with Applications*, *57*, 117–126. doi:10.1016/j.eswa.2016.03.028

Turney, P. D. (2002). Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews. *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics*, 417–424. arXiv:cs/0212032

Wang, H., & Wang, W. (2014). Product weakness finder: An opinion-aware system through sentiment analysis. *Industrial Management & Data Systems*, *114*(8), 1301–1320. doi:10.1108/IMDS-05-2014-0159

Wilson, T., Hoffmann, P., Somasundaran, S., Kessler, J., Wiebe, J., Choi, Y., Cardie, C., Riloff, E., & Patwardhan, S. (2005). Opinionfinder: A system for subjectivity analysis. *HLT/EMNLP 2005 - Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing, Proceedings of the Conference*, 34–35.

Winkler, M., Abrahams, A. S., Gruss, R., & Ehsani, J. P. (2016). Toy safety surveillance from online reviews. *Decision Support Systems*, *90*, 23–32. doi:10.1016/j.dss.2016.06.016 PMID:27942092

Xia, R., Xu, F., Yu, J., Qi, Y., & Cambria, E. (2016). Polarity shift detection, elimination and ensemble: A three-stage model for document-level sentiment analysis. *Information Processing & Management*, *52*(1), 36–45. doi:10.1016/j.ipm.2015.04.003

Yang, H. L., & Chao, A. F. Y. (2015). Sentiment analysis for Chinese reviews of movies in multi-genre based on morpheme-based features and collocations. *Information Systems Frontiers*, *17*(6), 1335–1352. doi:10.1007/s10796-014-9498-1

Yoon, K., & Hwang, C. L. (1981). *TOPSIS (technique for order preference by similarity to ideal solution): A multiple attribute decision making, w: Multiple attribute decision making--methods and applications, a state-of-the-at survey. A State-of-the-At Survey*. Academic Press.

Yu, H., Shang, J., Hsu, M., Castellanos, M., & Han, J. (2016). Data-driven contextual valence shifter quantification for multi-theme sentiment analysis. *International Conference on Information and Knowledge Management, Proceedings*, 939–948. doi:10.1145/2983323.2983793

Yu, J., Zha, Z.-J., Wang, M., & Chua, T.-S. (2011). Aspect Ranking: Identifying Important Product Aspects from Online Consumer Reviews. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies* (vol. 1). Association for Computational Linguistics. doi:10.1109/CC.2013.6488828

Zhang, C., Tian, Y. X., Fan, Z. P., Liu, Y., & Fan, L. W. (2020). Product sales forecasting using macroeconomic indicators and online reviews: A method combining prospect theory and sentiment analysis. *Soft Computing*, *24*(9), 6213–6226. doi:10.1007/s00500-018-03742-1

Zhang, W., Xu, H., & Wan, W. (2012). Weakness Finder: Find product weakness from Chinese reviews by using aspects based sentiment analysis. *Expert Systems with Applications*, *39*(11), 10283–10291. doi:10.1016/j.eswa.2012.02.166

Zhang, X., Niu, S., Zhang, D., Wang, G. A., & Fan, W. (2015). Predicting vehicle recalls with user-generated contents: A text mining approach. Lecture Notes in Computer Science, 9074, 41–50. doi:10.1007/978-3-319-18455-5_3

Zhuang, L., Jing, F., & Zhu, X. Y. (2006). Movie review mining and summarization. *International Conference on Information and Knowledge Management, Proceedings*, 43–50. doi:10.1145/1183614.1183625

ENDNOTE

[1] https://www.carwale.com/

Amit Singh is an Assistant Professor in School of Management and Entrepreneurship, Indian Institute of Technology Jodhpur, India. He has completed his M. Tech in Industrial & Management Engineering. His Ph.D. work is concerned with Sentiment Analysis Applications in Automobile Industry. His research interests include Text Mining, Sentiment Analysis, Multi-criteria Decision Making, Empirical Research, and Cloud-based Technology Solutions for MSMEs.

Mamata Jenamani is a Professor in the Department of Industrial & Systems Engineering, Indian Institute of Technology Kharagpur, India. She is a Ph.D. from IIT Kharagpur. She is a recipient of the Emerald/EFMD Outstanding Doctoral Research Award in Enterprise Application of Internet Technology. Her broad areas of interest are information system and e-business. Her current research interest includes web data analytics applications such as clickstream analysis, sentiment analysis of consumer reviews and social network-based recommender systems; e-procurement and reverse auction; and supply chain traceability.

Jitesh J. Thakkar is a Professor at National Rail and Transportation Institute, Vadodara, India. He has published research in the areas of Lean & Sustainable Manufacturing, Supply Chain Management, Quality Management, Small and Medium Enterprises, and Performance Measurement. The publications have appeared in the leading journals – International Journal of Production Economics, Journal of Cleaner Production, Production Planning and Control, Computers & Industrial Engineering, International Journal of Advanced Manufacturing Technology, Journal of Manufacturing Technology Management, and International Journal of Productivity and Performance Measurement.

Yogesh K. Dwivedi is a Professor of Digital Marketing and Innovation, Founding Director of the Emerging Markets Research Centre (EMaRC) and Co-Director of Research at the School of Management, Swansea University, Wales, UK. Professor Dwivedi is also currently leading the International Journal of Information Management as its Editor-in-Chief. His research interests are at the interface of Information Systems (IS) and Marketing, focusing on issues related to consumer adoption and diffusion of emerging digital innovations, digital government, and digital and social media marketing particularly in the context of emerging markets. Professor Dwivedi has published more than 300 articles in a range of leading academic journals and conferences that are widely cited (more than 18 thousand times as per Google Scholar). Professor Dwivedi is an Associate Editor of the Journal of Business Research, European Journal of Marketing, Government Information Quarterly and International Journal of Electronic Government Research, and Senior Editor of the Journal of Electronic Commerce Research. More information about Professor Dwivedi can be found at: https://www.swansea.ac.uk/staff/som/academic-staff/y.k.dwivedi/.