IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 23 May 2023, accepted 16 June 2023, date of publication 20 June 2023, date of current version 3 July 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3287830

RESEARCH ARTICLE

Dependency Structure-Based Rules Using Root Node Technique for Explicit Aspect Extraction From Online Reviews

PALLAVI MISHRA[1] AND SANDEEP KUMAR PANDA[2], (Member, IEEE)
[1]Department of Computer Science and Engineering, Faculty of Science and Technology (IcfaiTech), ICFAI Foundation for Higher Education, Hyderabad 501203, India
[2]Department of Artificial Intelligence and Data Science, Faculty of Science and Technology (IcfaiTech), ICFAI Foundation for Higher Education, Hyderabad 501203, India

Corresponding author: Sandeep Kumar Panda (sandeeppanda@ifheindia.org)

This work was supported by The ICFAI Foundation For Higher Education, Hyderabad, Telangana.

ABSTRACT Aspect-based sentiment analysis (ABSA) includes aspect extraction and sentiment analysis on those extracted aspects. This paper is mainly focused on the explicit aspect extraction from product based online customer reviews. Many research studies on explicit aspect extractions have adopted dependency rule-based techniques but limited their focus to only nouns and noun phrases as potential aspects. Moreover, extraction of multiple-aspects and multi-word aspects from online reviews are also not addressed by previous research studies. In this paper, we have proposed a Dependency structure-based rules using ROOT Node (DS-RN) technique using spaCy dependency parser to extract nouns, noun phrases, verbs, verb phrases in addition to single word, multiple aspects and multi word aspect extractions from customer review datasets. In our proposed methodology, 30 new dependency-based rules are formulated and implemented on 5 different product-based datasets. This study is based on the pattern analysis of dependency structures of review sentences to develop dependency-based rules for explicit aspect extraction. The proposed approach also incorporated lexicon-based pruning techniques to remove irrelevant aspects and retain correct aspects. The performance results on 5 different product-based customer review datasets demonstrate that our proposed DS-RN approach outperforms all other state-of-the-art baseline works with averaged value of precision as 87%, recall with 97% and 91% as F1-score.

INDEX TERMS Aspects, dependency structure-based rule, explicit aspect extraction, opinion lexicon, ROOT node, sentiment analysis.

I. INTRODUCTION

The emergence of e-commerce websites has rehabilitated the living style of common people that allows them to buy and sell physical goods, services, and digital products online rather than from a physical store. Unlike traditional shopping, online shopping has grown in popularity around the world, primarily owing to its ease of use. Online stores such as Amazon, Flipkart, and many others, have provided a digital portal for their customers to post and share their opinions in the form of reviews about different products and services which they have utilized [1]. These online reviews play a significant role and act as a valuable resource for both the potential customers and e-commerce companies. Many customers rely on online reviews to gain prior information about their target product aspects (for e.g., quality, design, color, battery-life, look, customer service, etc), that could facilitate their purchasing decisions [2], [3]. On the other hand, the e-commerce companies use these customer reviews to analyze the customers' sentiments towards their product and services and identify if any existence of product flaws, as early detection of flaws can boost a company's reputation, which in turn boosts sales as well as customer satisfaction [4]. People, however, find it difficult to read and summarize the large number of reviews that are relevant to their needs in order to make an informed decision. Therefore, an automated method for extracting and

The associate editor coordinating the review of this manuscript and approving it for publication was Arianna Dulizia.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

summarizing customer reviews is of paramount importance for the social purpose. Sentiment analysis (SA) plays a significant role in analyzing these reviews and producing a comprehensive summary. As a result, SA has gained popularity among the research community for its ability to process the massive amount of data available on social media. SA is a large suitcase that encapsulates several natural language processing (NLP) tasks, according to [5] such as sarcasm detection, aspect extraction, sentiment classification and subjectivity detection.

Broadly, Sentiment analysis is performed at three levels of granularity: document level, sentence level, and aspect level [6]. In document level sentiment analysis, the goal is to determine the overall sentiment (i.e., positive or negative) about the whole document. For example, let's consider one of the review comments ''I really like my overall experience with my MotoG Turbo Edition, but it runs out of battery power in no time. The design and quality of the display is however, amazing''. We have referred to this product-based review comment in each of the three levels of sentiment analysis as mentioned below to understand the advantages of fine-grained sentiment analysis (aspect level) over coarse-grained (document and sentence level). Using document level sentiment analysis, the overall sentiment is positive. However, it becomes difficult to extract irrelevant sentences from the document and also, not possible to separately extract the different sentiments about different aspects of an entity or object. In case of sentence level sentiment analysis, the task at this level goes to the individual sentences and determines the type of sentiment polarity of each sentences i.e,whether each sentences expressed a positive, negative or neural opinion. But, the complexity arises when users express different views in the same sentence. Considering the referred example, the sentiment polarity of the first sentence is turned as neutral while the sentiment polarity of the second sentence is positive. However, customers, on the other hand, are no longer satisfied with the overall sentiment polarity regarding products or services, rather a fine-grained representation of sentiment(s) towards various aspects or features of the product is highly valued. Mostly in case of product reviews, customers often express their opinions on various aspects (i.e., multiple aspects) of that product. For instance, the multiple aspect terms in case of the referred review comment are MotoG Turbo Edition, battery, design, and quality. The corresponding sentiments towards the aspect terms are positive, negative, positive and positive, respectively. Such a fine-grained approach is called Aspect-based sentiment analysis (ABSA) [7].

Aspect-based sentiment analysis (ABSA) studies the relationship between aspects (i.e, features) of a product and its associated sentiment(s) or opinion words. Among different tasks of ABSA, aspect extraction, is the first and most important task which involves identification and extraction of aspects related to a specific product (i.e., entity), and the second task is to perform sentiment analysis or opinion mining on those extracted aspects. Broadly, aspect is categorized either explicit or implicit [7]. For example, in the sentence ''the battery life is amazing'', the word ''amazing'' is associated with the explicit aspect term i.e., ''battery life'', on the other hand, implicit aspects are those that are not expressed directly in the review comment. Consider the following sentence: ''The phone is small''. This sentence expresses an opinion about the ''size'' (aspect) of the ''phone''. Both explicit and implicit aspect extraction methods can be conducted using three modes of machine learning i.e., either supervised [8], [9], [10], [11], [12], [13], [14] or unsupervised [22], [23], [24], [25], [26]. The significance of efficient extraction of aspects and its associated sentiment words become evident in many practical applications in following ways:

- It represents an important phase to determine feature popularity.
- Ranking of products in the market.
- Useful in Aspect retrieval system
- Sentiment Analysis and Classification using aspect-based extraction.
- useful to construct user and item profiles in decision-making processes, such as recommended systems [27].

A. MOTIVATION

Basically, most of the research studies based on aspect extraction have used either dependency rule-based technique [18], [26], [27], [28], [29] or sequential pattern-based approach [24], [30], [31], [32] for aspect extraction. In general, the dependency rule-based technique perform relatively well as compared to sequential rule approach in terms of handling informal review text. However, the sequential pattern-based rules have poor generalization capability due to its static pattern-based nature [33]. Moreover, the existing approaches tend to become liable by considering only noun or noun phases as aspects, whereas verbs can also be considered as potential aspect terms. In some of the cases, a customer might express his/her opinions about multiple aspects of product/services. Many previous studies have not focused on the extraction of multiple aspects, even if they exist. Apart from this, several studies on aspect extraction are often based on extraction of single-word aspect(s) (i.e., aspect word consists of only single word, for e.g., design, material, color etc.) and multi-word aspects with length not more than two (for e.g.: battery life, screen display, customer service, technical support, etc.). But, practically, the word size of aspect may vary depending upon the domain of interest.

B. MAIN CONTRIBUTION

We proposed a Dependency Structure based rules using Root Node (DS-RN) technique to address the extractions of multiple aspects and multi-word aspects from online reviews. In this paper, 30 new rules based on dependency structure are formulated to extract both single word aspect(s) and multi-word aspect(s) with length more than two, along with multiple aspect words from a sentence, based on its occurrences in a sentence. To implement these rules, we have used spaCy Dependency Parser which shows the dependency relationship between head words and their dependents. All these newly formulated dependency-based rules in reference to "ROOT Node" are useful for extraction of nouns, noun phrases and verbs as aspects. In addition, we have incorporated lexicon based pruning technique to extract the correct aspects by eliminating irrelevant aspects to reduce the false positive counts.

The dependency structure is like a directed acyclic graph with words as nodes and relationship between among two words as edges. Each word in the sentence either modifies another word or is modified by a word. For example, in the sentence "beautiful watch", here the adjective word "beautiful" is modifying the noun word "watch". Here, the adjective word is called as dependent, or child node and noun word (here "watch") is known as ROOT Node / governor /parent /head. In other words, the "ROOT Node" of the parsing tree is the only node that is modified but doesn't modify anything else and become the head of the entire structural graph of the whole sentence. In dependency-based approaches, the head-dependent relationship is made explicit by directly linking heads (i.e., incoming arcs) towards the words that are immediately dependent on them.

We have used annotated datasets containing customer reviews of five electronic products for explicit aspect extraction [6], [18], [27], [34]. Many researchers also have worked on this dataset for their studies on aspect extraction tasks.

C. ARTICLE ORGANIZATION

The rest of the article is organized as follows: in Section II, we present a literature survey of the explicit aspect extraction method, especially in the aspect-based opinion mining. In Section III, we describe the workflow of the proposed model i.e., Dependency structure-based rule using the ROOT node (DS-RN) approach to extract single-word aspect(s), multiple aspect(s), multi-word aspect(s), considering both noun and verbs as possible aspects. In addition, the significance of the utilization of opinion lexicon [7], self-created action-verb lexicon and non-aspect noun lexicon for aspect extraction are well illustrated in the form of figures. This is followed by presentation of materials (graphical visualization of the proposed rules i.e., Dependency based rules using ROOT Node (DS-RN)) and an observational study as shown in Table 1. Section IV describes the statistical information about the customer reviews dataset used in our study and discusses experimental results and compares our work with existing approaches. Finally, Section V concludes the paper and proposes some directions for future work.

II. RELATED WORK

Several research studies have been focused on aspect extraction especially on explicit aspect type using various modes of learning methods such as Unsupervised,Supervised and Semi-supervised. The studies based on customer review datasets [7] (i.e., product based reviews) are taken into consideration for the study. Hu and Liu [7] initiated the research work on aspect based opinion mining,considering nouns as only aspect terms. The author used frequency based approach i.e., Apriori Algorithm, extracting all the aspects which are above the threshold count value. Later on, the adjectives which are related to those extracted aspects are considered as opinion words. The performance of the proposed model on customer review datasets achieved results with 80% recall and 72% precision. The work by Popescu and Etzioni [22] further improved the previous work [7] by using both frequency and Pointwise mutual information (PMI) methods. They utilized web similarity measure in order to improve precision value by discarding incorrect extracted aspects. The experiment was conducted on customer review datasets and the best result obtainied was 77% recall and 94% precision.

Boot-strapping methods for aspect extraction was used by Li et al. [35] for extracting aspects and their related opinion words or sentiment words. In this paper, dependency rules are formulated using grammatical relationship between aspect(s) and opinion words. The newly extracted aspects were again used for generating relation rules so called bootstrapping approach. The results obtained 89% F-measure on Customer review dataset.

A study by Samha and Li [36] adopted dependency relation based rules for aspect extraction using customer review dataset. The author formulated five new dependency rules and combined 11 dependency rules from previous works [18], [28]. The author has applied Data pre-processing techniques such as removing punctuation symbols, after that performing lemmatization, and finally tagging each of the words in a sentence using POS tagger and finding the dependency relationship in-between these words using Stanford Parser and achieved 71.8% as F-measure.

The following are some examples of unsupervised works that used syntactic patterns for aspect extraction. The author Moghaddam and Ester [23] proposed a combined approach using both frequency and POS patterns for both extraction of aspects and pruning of incorrect extracted aspects. The model performed on customer review dataset as 80% precision and 87% recall. Htay and Lynn [24] proposed eight pattern based rules by perceiving all nouns and noun phrases as aspects and obtained F-measure as 79%. Maharani et al. [30] also utilized pattern based rules by defining new rules and also used previously defined rules from other studies and achieved best result as 67.2% F-measure on combining both the rules for aspect extraction.

The most recent works on explicit aspect extraction by Tubishat et al. [37] used the Whale optimization algorithm (WOA) in conjunction with a new local search algorithm to extract explicit aspects. The author used certain dependency rules and pattern rules from previous studies and also formulated new rules with total 126 rules for explicit aspect extraction. Also, this study used new prunning algorithm(PA) and achieved best result on customer review dataset with 92%

F-measure after using both Improved Whale Optimization and pruning algorithm.

Some of the research works have also focused on semi-supervised based learning mode: Wu et al. [16] work based on dependency relation only at phrase level such as noun phrases and verb phrases for extraction of aspects. The author considered the opinion words as neighbor to the extracted aspect words based on the existing opinion lexicons (i.e., dictionary of opinion words). Additionally, this study also utilized a tree kernel with Support vector machine (SVM) model in order to derive the relations between aspect and opinion words. The study used customer review dataset and performed result 57% using F-measure.

The author [18] and [38] proposed a model called Double Propagation(DP), based on dependency relation rules method. This study considered only nouns and noun phrases as potential aspect words whereas adjectives as opinion words and defined eight dependency based rules. However, seed opinion words are pre-requisite to initiate the propagation method. The seed opinion words are useful in extraction of new aspects using these formulated dependency based rules. In addition, the author used prunning techniques to improve the precision measure. The experiment of this study focused on customer review dataset and obtained result 86% as F-measure. By further improving this approach, another study by Liu et al. [39], integrated aspect association with semantic similarity approach and the best result achieved on customer review dataset with 87% F-measure.

Another approach were proposed by Kang and Zhou [27], which improvised DP algorithm by adding new dependency based rules. In addition, the author included comparative rules, indirect dependency rules, direct dependency rules for further improvement in result. Pruning methods were also used to filter incorrect aspects. Finally, the proposed model performed on customer review dataset where F-measure was 87%.

Several research works also adopted syntactic pattern based approach using semi-supervised mode of learning method. A study by Samha et al. [20]proposed eleven frequent patterns for aspect extraction. The author has used the opinion lexicon developed by Hu and Liu [7], where dictionary of opinion words are available. Based on this opinion lexicon, the author used to match any opinion words in the review sentence to one of the 11 formulated syntactic pattern rule, then finally extract the corresponding aspect. The proposed work is experimented on customer review dataset and achieved result with 77% F-measure. Also, Rana and Cheah [32] work based on syntactic pattern based method, where the author only considered noun and noun phrases as possible candidate aspects. In addition, this study also used prunning techniques such as frequency based and Normalized Google Distance (NGD). They reported F-measure results of aspect extraction on customer review dataset [7] as 89%. Pattern based rules were used by [40] where the author created ten pattern based rules by studying the correlation between aspect terms and their associated opinion words.However, the work is based on consideration of noun or noun phrases as possible aspects only. The performance resulted in F-measure as 89%.

One of the most recent work by Feng et al. [41] used topic modeling method such as Latent Dirichlet Allocation method along with synonym identification. The author used TF-IDF as a prunning technique to discard incorrect aspects. Based on reported results, the best achieved results were obtained on customer review dataset with 87% as F-measure.

In contrast to the approaches mentioned above, supervised approaches have been used in a variety of domains. In addition, the majority of these techniques make use of English language reviews. Work by Poria et al. [42] used seven layer deep neural network model (DNN) with word embedding, in addition with 5 linguistic rules for aspect extraction. The author experimented with customer review dataset and achieved result with 88% F-measure. Another study by Da'u et al. [43], the author used Convolutional Neural Network Model having multiple channels, combining both POS embedding and Word Embedding channel. The result achieved on customer review dataset was 89% as F-measure.

Some of the researchers also used Conditional Random Field (CRF) method as supervised mode of aspect extraction. Chen et al. [11] used CRF model for extraction of aspects. Data preprocessing techniques such as removing of symbols, spelling corrections and checker, stemming of noun words are incorporated in this study. Using Labelled datasets, CRF model got trained and applied on testing data for the performance measurement. The result obtained was 82% with F- measure.

Another work based on dependency based rule approach, Liu et al. [29] extended the work of DP by adding new dependency rules and used greedy algorithm for selection of best rules from the set of rules for aspect extraction. This experiment was carried out on customer review dataset and obtained result 87% F-measure. One of the recent based studies by Nawaz et al. [44] used both Normalized Google Distance and ConceptNet method for extraction of aspect words from the customer review dataset by discarding the incorrect aspects based on word similarity. The author perceived only noun and noun phrases as potential aspects. The study achieved its best result with 86% F-measure. Mubashar Hussauin et al. [33], one of the recent studies about explicit aspect extraction proposed multi-layered technique using both syntactic dependency parser based rules and sequential pattern based rules. The author used subset of sequential rules used in the study [40]. These rules were crafted to extracted both noun and verbs as potential aspects. The experiments were conducted on customer review datasets containing both five electronic products [7] and another dataset consists of three other products [29] with the best achievements of F1-score as 0.90% and 0.88%, respectively.

Dragoni et al. [46] created dependency graph and developed three dependency rules based on the dependency relationship for aspect extraction using SemEval 2015 and SemEval 2016 datasets about restaurants and laptop reviews. The results achieved on the laptop dataset and restaurant dataset were 60% F-measure and 51% F-measure, respectively. In addition, the outcomes obtained in SemEval 2016 dataset were 67% for the restaurant dataset and 57% for the laptop dataset. Another study [47] implemented seven-layer deep convolutional neural network (CNN) in addition to four dependency rules to enhance the performance of aspect extraction. The experiment was conducted on a Nikon camera dataset with 88.6% precision and 90% recall.

III. METHODOLOGY

In our study, we have proposed total 31 dependency-based rules based on the various dependency structural patterns of review sentences for explicit aspect extraction from online reviews. These rules are graphically represented in Table 8 (Appendix) with encircled aspect words(s). Among these 31 rules, Rule 1 to Rule 29 follow Type-01 (i.e., dependency structural based rules in reference to ROOT node), whereas, Rule 30 and Rule 31 follow Type-02 (i.e., dependency structural based rules non-reference to "ROOT Node") approach. In this work, we have developed such rules to address extraction of single word aspects(i.e., only one word aspect), multi-word aspects(i.e., more than one word aspect), multiple aspects(i.e., more than one aspect in single review sentence), considering nouns, noun phrases, verbs and verb phrases as possible aspect words.

Figure 1 shows the workflow of the proposed methodology i.e., Dependency Structure based rules using Root Node (DS-RN) technique. In Phase 1 of the proposed work-flow model, data-preprocessing technique is carried out to clean the review sentences by removing unnecessary symbols such as quotation marks, salutations, apostrophes, hyphens, exclamation marks, which have no significant role to play with aspect extraction. Further steps of data-preprocessing include conversion of certain contracted forms of words (such as 'll, 've, n't, 've, 'm, etc) to its corresponding formal style of writing words (i.e., will, would have, not, have, have, am), which proved useful to reduce the complexity of the dependency relation graph [48] of a review sentence. After that, sentence tokenization is performed followed by Part-Of-Speech (POS) tagging and extraction of dependency relations among the words in sentences using spaCy dependency parser. The built-in dependency visualizer of spaCy library shows the POS tags and syntactic dependencies of any given input sentence, which plays a key role in identifying the "Root Node" of the sentence.

After data-preprocessing step, our developed rules are implemented on the pre-processed review sentences in phase 2, which consists of Rules in reference to ROOT Node and Rules non-reference to ROOT Node to extract possible aspect words. Let's illustrate the Type-01 i.e., Rules in reference to ROOT Node. Consider one of the reviews given by the customer, i.e., "the manual is relatively clear". The main objective is to extract the target aspect word using dependency structure based rule as shown in Figure 2, where the auxiliary word "is" forms the Root word of this sentence. There exists nominal subject (nsubj) dependency relationship between the root word and left child (i.e., manual) of the root word. And, the right child of Root Node (i.e., clear) is in adjective complement relationship (acomp) with Root Node. According to Rule 1, if Part-of-Speech (POS) of left child of Root Node is NOUN and POS of right child of Root Node is adjective $ADJ \in O_w$, where (O_w) represents the Opinion Word Lexicon, which contains list of positive and negative sentiment words, then, extract the left child of Root Node and termed it as "aspect word".

Similarly, let's consider another review sentence consisting of multiple aspect words (i.e., extraction of more than one aspects). For instance, "it has good color and price was great". Here, color and price are two target aspects which hold direct object (dobj) and conjunction (conj) dependency relationship with respect to "ROOT node", respectively. The task is to extract these target aspect words using dependency structure based rules. In this particular sentence, the word "has" has no incoming arrow dependency from other words and thus becomes the ROOT node of this dependency structural graph. There exists "nominal subject" (nsubj) dependency relationship between the Root word and left child of the root word. And here, the POS of the left child is "PRONOUN" (PRON). In addition, the right child of Root Node is in "direct object" relationship (dobj) with reference to "ROOT node". The word "good" which is one of the members of opinion lexicon is in "adjectival modifier" (amod) dependency relationship with the "color" word. The adjective word "good" shows the positive opinion about the aspect "color" and thus plays a vital role in identification of aspect word. Again, there exists another word which holds conjunction (conj) dependency relationship with the extracted aspect word (color). Based on the given dependency structural pattern in Figure 3, target aspects can be extracted using Rule 10 as shown in Table 2.

In our work, we have also addressed verbs as aspects using action verb lexicon approach as shown in Figure 3. Now, let's examine the case of extraction of multi-word aspect(s) from the review sentences. For instance, "the moveable lcd screen is great" as shown in Figure 5. In this particular sentence, "lcd screen" is the multi-word aspect, which is in "nsubj" relationship with the "ROOT node" (is) and its associated opinion word i.e., great is in "acomp" dependency relationship with respect to "ROOT node". The most important dependency relationship in multi-word aspect extraction includes "compound". In this case, the word(s) which hold "nsubj" as well as"compound" dependency relationship with respect to the "ROOT Node" should not be a member of Non-aspect Noun Lexicon, which includes Pronouns (she, he, it, they, etc.), time units (hour, minutes, seconds, days, months, years, seasons, time etc.), distance units, quantifiers(both countable and uncountable nouns) like- lots of,

TABLE 1. Summary of related works.

Work	Technique	Average Results on 5 product-based customer reviews datasets (in%)	Drawbacks
Hussain et al.[33]	Defined dependency based rules + sequential pattern based rules	Precision(P): 0.85 Recall(R):0.95 F1-score: 0.90	1. Considered only nouns as possible aspects, no rules were developed to extract noun phrases, verbs, adjectives as explicit aspects. 2. The developed rules lack ability to extract multiple aspects (more than one aspect words) from online customer reviews
Tubishat et al.[37]	Developed 126 dependency-based rules + sequential pattern-based rules	Precision(P): 0.92 Recall(R):0.93 F1-score: 0.92	1. Extraction of multi word aspects of length more than 2, is not explored in this paper. 2. Rules for verb as explicit aspect extraction remain unexplored.
Rana et al.[40]	Defined 10 sequential pattern-based rules	Precision(P): 0.86 Recall(R):0.91 F1-score: 0.89	1. These rules are limited to only extraction of nouns and noun phrases as possible aspects
Nawaz et al.[44]	Concept Net methods and Normalized Google Distance for word similarity index for aspect extraction	F1-score: 0.86	1. Extraction of multiple aspects need to be addressed
Feng et al.[41]	Implemented Latent Dirichlet Allocation method along with synonym identification and used TF-IDF as a pruning technique to discard incorrect aspects	Precision(P): 0.87 Recall(R):0.88 F1-score: 0.87	1. Similarity relationship is done among nouns only.
Rana et al. [32]	Pattern-based extraction rules and two pruning methods (NGD + frequency)	Precision(P): 0.87 Recall(R):0.92 F1-score: 0.89	1. Pattern based rules only focused on noun/noun phrases as possible aspects
Kang et al.[27]	Developed new rules by adding comparative rules, indirect-dependency rules, and part–whole relation rules + Pruning methods	Precision(P): 0.87 Recall(R):0.88 F1-score: 0.87	1. These rules lack the ability to handle long distance relations in a sentence
Samha [36]	Five new dependency relation rules with subjective opinion lexicon and used eleven dependency rules from previous studies[18, 28]	F1-score:0.71	1. This paper assumed only nouns and adjectives as possible aspects 2. Scope for further improvement in subjective opinion lexicon exists
Asghar[31]	Defined ten patterns rules	F1-score: 0.77	1. Multiple aspects extraction is not addressed in this study
Qiu et al.[18]	Adopted Double Propagation strategy using eight dependency rules for aspect and opinion extraction using seed opinion lexicon approach and also proposed pruning methods to remove incorrect aspect words.	Precision(P): 0.88 Recall(R):0.83 F1-score: 0.86	1. Proposed a limited set of dependency rules where single word nouns are only considered to be candidate term for aspects. 2. Also, lacks an effective method for pruning opinion words.

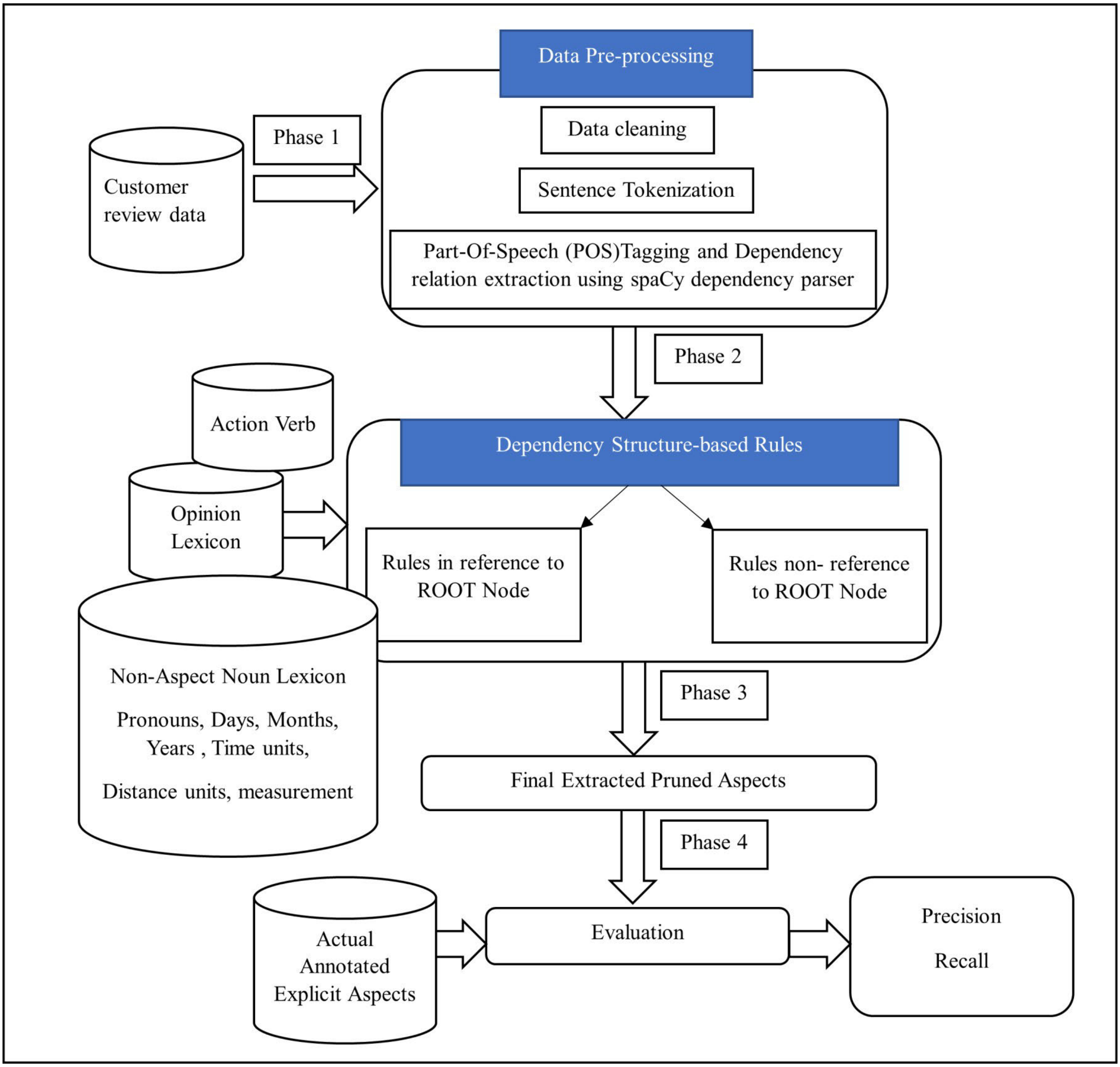


FIGURE 1. The workflow of the proposed Dependency Structure based Rules using ROOT Node (DS-RN) approach.

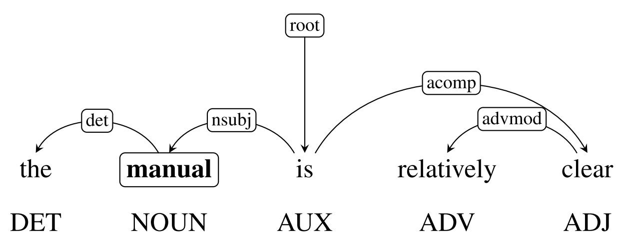
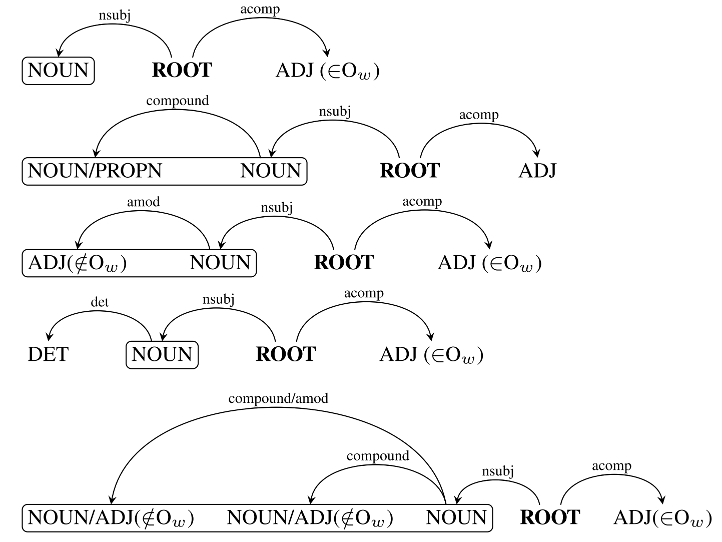
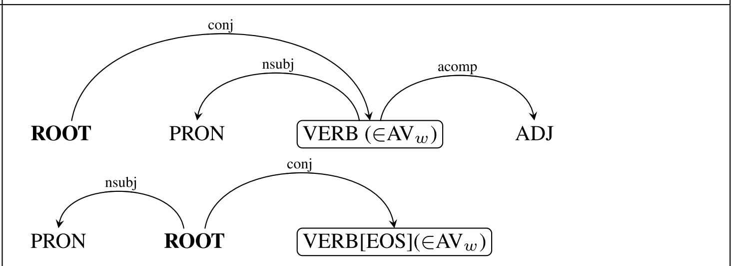
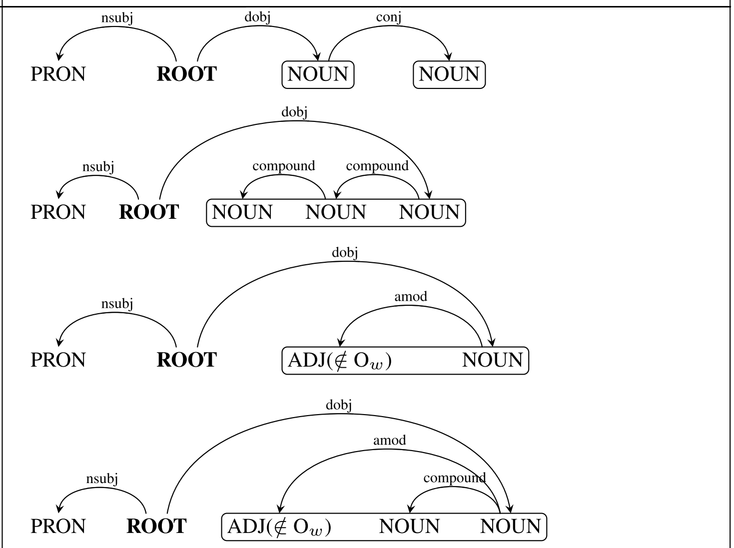
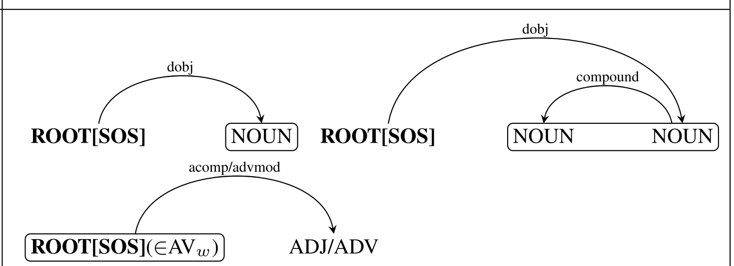


FIGURE 2. Single-word aspect extraction using Rule 1.

many, some, any, little, cardinal numbers (one, two, three etc.), ordinal numbers (first, second, third, fourth etc). Other non-aspect noun words such as way, issues, complaints, inconvenience, troubles, problems, tons, facts, difficulty, possibility, smell, hearings, couple, combination, arrangement and many more are being added in non-aspect noun lexicon. In general, the noun word which is either in direct or indirect relationship to the ''ROOT Node'' if represents to any of the Non-aspect Noun Lexicon, then it cannot be considered as an''aspect''.

In this paper, we have incorporated opinion lexicon and self-created Action verb word lexicon (such as look, play, read, set, install, freezes, operate, works, use etc.), Non-aspect Noun Lexicon as different pruning techniques to eliminate

TABLE 2. Rules in reference to ROOT node (Type-01).

RULES	ROOT NODE	DEPENDENCY STRUCTURAL PATTERN WITH REFERENCE TO ROOT NODE
Rule 1	VERB,AUX	
Rule 5	VERB	
Rule 10	VERB,AUX	
Rule 11	VERB	

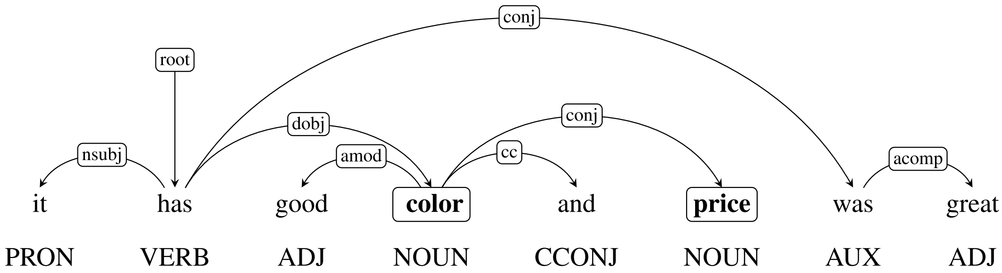


FIGURE 3. Multiple aspect extraction with application of opinion lexicon using Rule 10.

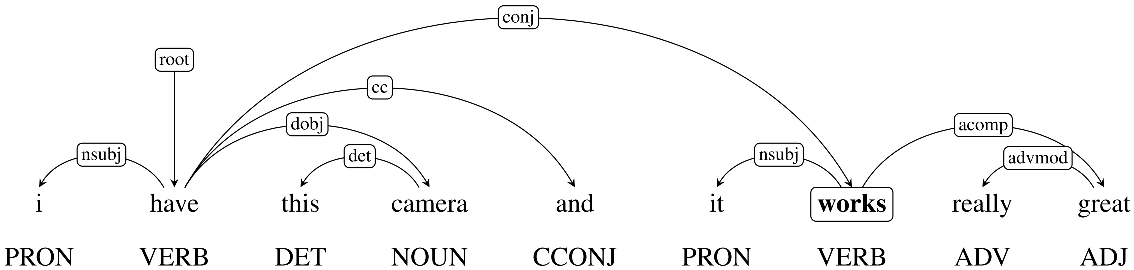


FIGURE 4. Extraction of verb as aspect with application of action verb (AV_W) verb as lexicon using Rule 5.

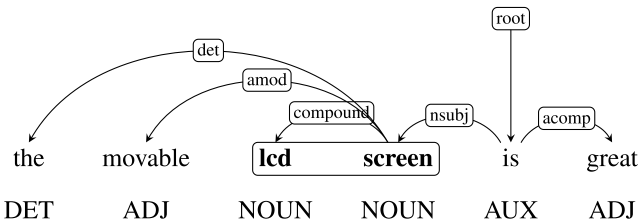


FIGURE 5. Multi-word aspect extraction using Rule 01.

incorrect extracted aspects (i.e., reducing false positive value) which results in better performance of the proposed model. We can use different pruning techniques such as normalized google distance (NGD) as similarity metrics, cosine similarity to increase the precision of the proposed model by reducing false positive counts.

An illustration regarding Non-aspect Noun lexicon is shown in Figure 6, where the word ''everything'' tagged as Pronoun (PRON), though is in ''nsubj'' dependency relationship in reference to ''ROOT Node'', still cannot be called as an aspect. In general, every noun with ''nsubj'' dependency relationship with the ''ROOT word'' is not always an aspect.

Previous studies have developed limited set of dependency based rules, which have not addressed extraction of multi-word aspects of length more than 2. However, in real life application, the length of the aspect may vary depending upon the domain and its associated features. In this paper, we have addressed this issue by formulating dependency rules as shown in Figure 7. Consider the following sentence as: ''the light auto-correction is awesome'', where light auto-correction is the multi-word aspect of length 3. Here, the word ''light'' though tagged as ''adjective'', but it is not truly representing member of Opinion lexicon ($\notin O_w$), thus, can be considered as part of the multi-word aspect.

As shown in Table 3, Rule 30 and Rule 31 are basically Type-02 category of the Proposed DS-RN model. Type-02 based rules (i.e., Rules Non-Reference to ROOT Node) are special heuristic based rules, that involve no role of ''ROOT Node'' in the dependency structural graph for aspect extraction. Upon, studying various patterns of review sentences from 5 different online product based datasets, some of the most likelihood of occurrences of patterns have enforced to create these special rules. Figure 8 illustrates Type-02 of Rule 30, where, the end word of sentence (EOS) word i.e., ''control'' is tagged as NOUN which is not a member of Non-aspect Noun Lexicon. And the preceding words (i.e., remote, and universal), though tagged as ''adjective'', however, not representing members of opinion lexicon (O_w), therefore, ''universal remote control'' syntactically forms a ''Multi-word Aspect'' of length more than 2. In otherwords, the End Word of the Sentence (EOS) having POS (NOUN), which is not a member of Non-Aspect Noun based lexicon and holding dependencies (dobj, pobj, conj) are most likely to get considered as an aspect.

Figure 9 illustrates the extraction of verb and verb phrase as aspects, which follow the dependency structural pattern as shown in Rule 31 of Table 3. In this case, the application of self-created action-verb word lexicon helps to extract the word ''install'' (verb) and ''set-up''(verb-phrase). Figure 10 and Figure 11 illustrate Type-02 based dependency structural pattern rules with no involvement of ''Root node'' concept for aspect extraction.

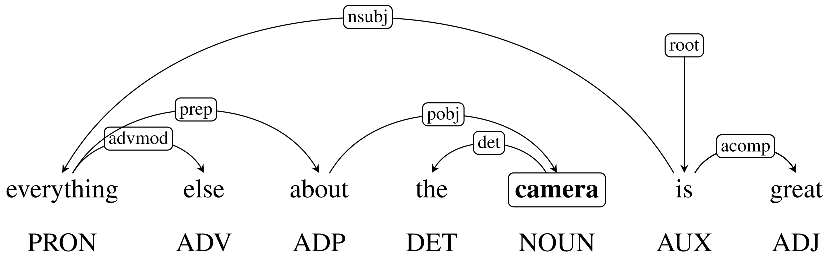


FIGURE 6. Illustrated the application of Non-Aspect Noun Lexicon for aspect extraction.

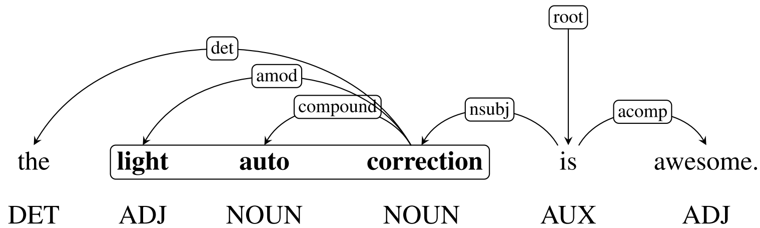


FIGURE 7. Multi-word aspect extraction of length more than two using Rule 01.

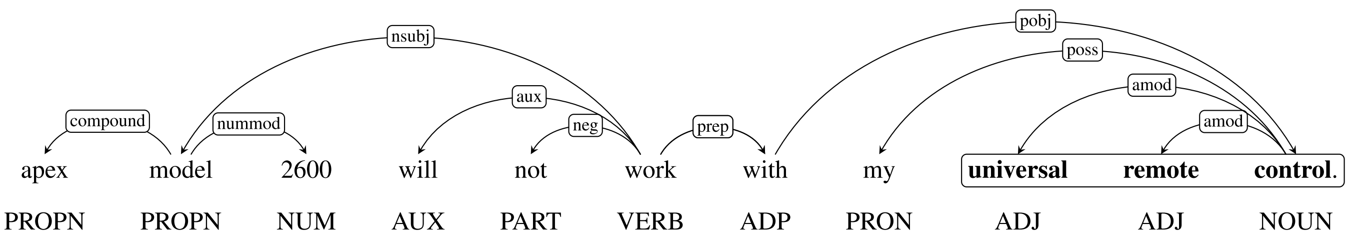


FIGURE 8. Illustrated Type-02 Rule.

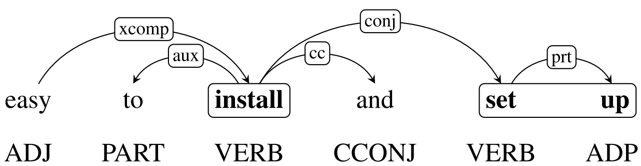


FIGURE 9. Extraction of verb and verb phrase as aspect using Rule 31.

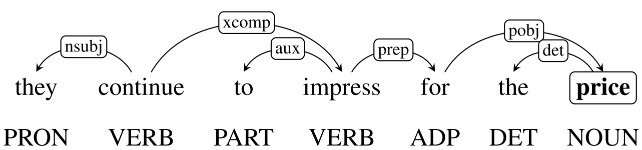


FIGURE 10. Type-02 Non-reference ROOT Node Example with single NOUN in sentence.

In order to understand the notion (i.e., algorithmic analysis) of Type- 01 based rules in Table 2, let us have a brief explanation about the dependency structural pattern of Rule 1, so that rest can be interpreted in a similar way based on the relationship structural patterns among the words with the ROOT node. Now, a sentence to follow the dependency structural pattern of Rule 1, the foremost criteria is that the Part-of-speech (POS) of the ROOT node should be either a VERB or AUX (auxiliary). Again, the left child and the right child of the ROOT Node must hold "nsubj" and "acomp" dependency relationship, respectively, in reference to the "ROOT Node" of the dependency graph, which forms the base structure of Rule 1. The most challenging task is to derive and generalize various forms of the dependency structural graphs of the base form, so that irrespective of writing style of reviews (formal or informal), our proposed DS-RN Model can capture the target aspect(s) by matching the dependency structural patterns as shown in Table 1.

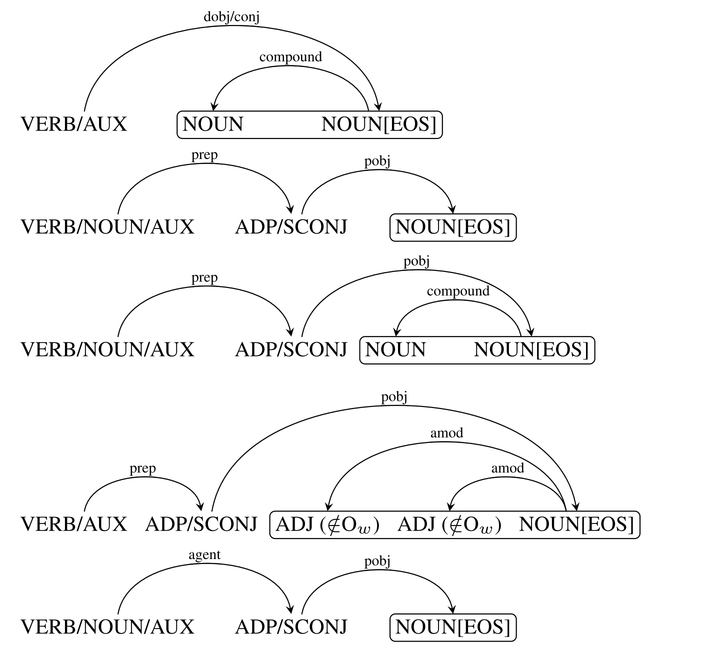
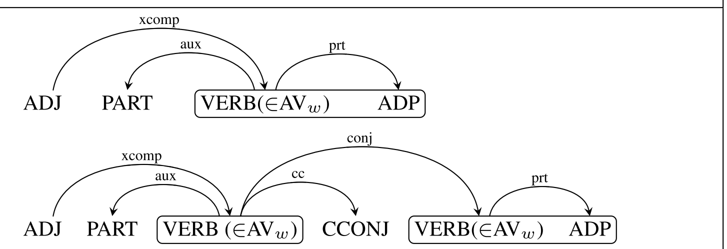
IV. EXPERIMENTAL RESULTS AND ANALYSIS

A. DATASETS

In this study, we have collected publicly available customer review datasets from amazon website which contains five review datasets of four domains: Apex DVD Player(Data-A),Creative MP3 Player(Data-B), Canon Digital Camera (Data-C), Nikon Digital Camera (Data-D) and Nokia Cell Phone(Data-E). It is a benchmark dataset used by many research works based on aspect extraction.

In addition, these electronic product based datasets were manually annotated by Hu and Liu by labeling each sentence

TABLE 3. Rules with non-reference to ROOT node (Type-02).

RULES	ROOT NODE	DEPENDENCY STRUCTURAL PATTERN WITH NON-REFERENCE TO ROOT NODE
Rule 30	—	
Rule 31	—	

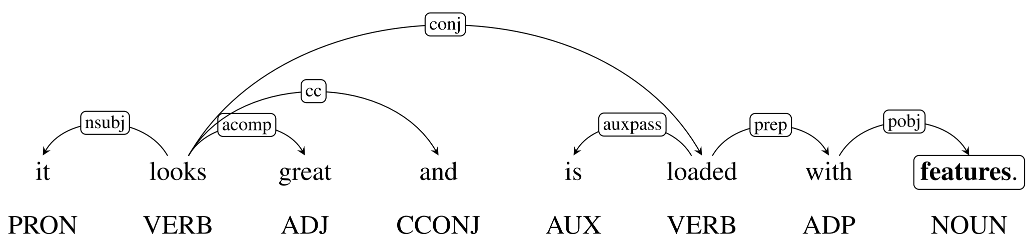


FIGURE 11. Type-02 rule: Non-reference to ROOT Node.

with the target aspect(s). Table 2 summarises the attributes of product-based electronic datasets in terms of total number of explicit words and total number of sentences in particular dataset.

B. EVALUATION METRICS

The metrics that were used in this study for evaluating the performance of the proposed DS-RN technique on 5 different electronic products based datasets include Recall, Precision, and F1- Score measures. Basically, there are two ways to compute the results regarding explicit aspect extraction: (a) based on multiple occurrences of each aspect word, and (b) based on unique occurrences of each aspect word.

In a dataset, an important aspect often occurs multiple times based on the context of the opinion of a user. For e.g., the aspect "battery" occurred 15 times in a set of cell phones reviews. For (a), if any occurrence of "battery" is extracted, then all occurrences of "battery" are considered extracted, i.e., 15. If none of its occurrences is extracted, it is considered as 15 losses. In (b), if any occurrence of "battery"

TABLE 4. Detailed statistic of the dataset.

Data	Product	# of explicit aspects	# of sentences
Data-A	Apex DVD Player	296	740
Data-B	Creative MP3 Player	674	1716
Data-C	Canon Digital Camera	237	597
Data-D	Nikon Digital Camera	174	356
Data-E	Nokia Cell Phone	302	546

is extracted, it is considered as one extraction. If the word is extracted at least once, it is considered as one loss. In this paper, we have considered only (a), as it is crucial to get those important aspects extracted.

Consider a set X as set of distinct (i.e., unique) aspect words extracted by proposed model and set Y as set of labeled aspect words manually done by human annotators. Let's analyze the Confusion matrix terms i.e., True Positive (TP), False Positive (FP), False Negative (FN) with context to aspect extraction task. TP $= \mid X \cap Y \mid$, FP $= \mid X \backslash Y \mid$, FN $= \mid Y \backslash X \mid$.

For type(a) based computation result i.e., based on multiple occurrences of each aspect word, the evaluation metrics are mathematically defined as:

- Recall= $\frac{TP}{TP+FN}$
- Precision= $\frac{TP}{TP+FP}$
- F1-Score= $\frac{2*Precision*Recall}{Precision+Recall}$

C. PERFORMANCE OF PROPOSED (DS-RN) MODEL ON PRODUCT BASED DATASETS

The performance of the proposed DS-RN technique in terms of the aforementioned evaluation metrics on product based datasets as discussed in section IV-A is shown in Table 3. The average results of recall is much higher as compared to precision. It is due to the fact that, the DS-RN model has formulated generalized dependency based rules with respect to both "ROOT node" and "Non-ROOT node" concept, which ultimately increases the probability of extracting more number of potential aspects. Nevertheless, the average F1-score of the dataset is also quite promising.

D. PERFORMANCE COMPARISON OF DS-RN PROPOSED MODEL WITH BASELINE MODEL TECHNIQUES

In this section, the performance of the proposed DS-RN methodology is compared with other approaches such as dependency based rule techniques, pattern based rules and hybrid techniques(i.e., combination of two or more techniques). All these existing approaches used the same set of datasets i.e., electronic(product) based customer review datasets for their work. Though many works have been done for explicit aspect extraction using dependency based rules, sequential pattern based rules, hybrid based techniques.

TABLE 5. Performance comparison with existing works based on Dependency rule based approach.

Data	Metric	DP[18]	RubE[27]	RSLS[29]	DS-RN
Data-A	P	**0.92**	0.90	0.86	0.86
	R	0.86	0.85	0.90	**0.96**
	F1 Score	0.89	0.87	0.88	**0.91**
Data-B	P	0.81	**0.87**	0.82	**0.87**
	R	0.84	0.90	0.91	**0.96**
	F1 Score	0.82	0.88	0.86	**0.91**
Data-C	P	**0.87**	**0.87**	0.85	0.86
	R	0.81	0.86	0.91	0.96
	F1 Score	0.84	0.86	0.88	0.91
Data-D	P	**0.90**	**0.90**	0.89	0.85
	R	0.81	0.86	0.94	0.98
	F1 Score	0.85	0.88	0.91	0.91
Data-E	P	**0.90**	**0.90**	0.83	0.89
	R	0.86	0.91	0.90	**0.97**
	F1 Score	0.88	0.90	0.86	**0.93**
Average	P	0.88	0.87	0.85	0.87
	R	0.83	0.88	0.91	**0.97**
	F1 Score	0.86	0.87	0.88	**0.91**

TABLE 6. Performance comparison with other works based on Sequential Pattern rule-based approach.

Data	Metric	Htay[24]	SPR[40]	TF-RBM[32]	DS-RN
Data-A	P	0.78	0.86	**0.88**	0.86
	R	**0.97**	0.81	0.90	0.96
	F1 Score	0.87	0.84	0.89	**0.91**
Data-B	P	0.70	**0.89**	0.86	0.87
	R	0.76	0.91	0.93	**0.96**
	F1 Score	0.73	0.89	0.90	**0.91**
Data-C	P	0.74	0.81	0.80	**0.86**
	R	0.92	0.92	0.89	**0.96**
	F1 Score	0.82	0.87	0.84	**0.91**
Data-D	P	0.71	0.86	**0.87**	0.85
	R	0.81	0.96	0.93	**0.98**
	F1 Score	0.76	0.90	0.90	**0.91**
Data-E	P	0.74	0.89	**0.92**	0.89
	R	0.82	0.95	0.93	**0.97**
	F1 Score	0.78	0.92	0.92	**0.93**
Average	P	0.73	0.86	**0.87**	**0.87**
	R	0.86	0.91	0.92	**0.97**
	F1 Score	0.79	0.89	0.89	**0.91**

TABLE 7. Performance comparison with other works based on hybrid techniques.

Data	Metric	Feng[41]	IWOA+PA[37]	ML-RB[33]	DS-RN
Data-A	P	0.88	**0.92**	0.85	0.86
	R	0.90	0.91	0.93	**0.96**
	F1 Score	0.89	**0.91**	0.89	**0.91**
Data-B	P	0.88	**0.93**	0.86	0.87
	R	0.87	0.95	0.95	**0.96**
	F1 Score	0.87	0.94	0.90	**0.91**
Data-C	P	0.85	**0.91**	0.83	0.86
	R	0.84	0.93	0.95	**0.96**
	F1 Score	0.84	**0.92**	0.89	0.91
Data-D	P	0.89	**0.91**	0.84	0.85
	R	0.87	0.92	0.96	**0.98**
	F1 Score	0.88	**0.91**	0.90	**0.91**
Data-E	P	0.87	**0.93**	0.88	0.89
	R	0.90	0.93	**0.97**	0.97
	F1 Score	0.88	**0.93**	0.92	**0.93**
Average	P	0.87	**0.92**	0.85	0.87
	R	0.88	0.93	0.95	**0.97**
	F1 Score	0.87	**0.92**	0.90	0.91

However, in this paper, some of the recent done works based on the aforementioned techniques are already described in

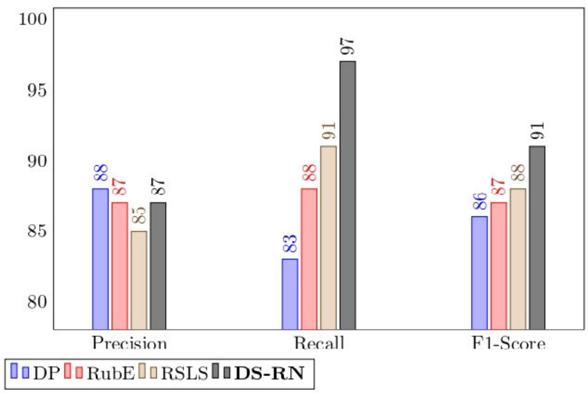


FIGURE 12. Average precision, recall and F1-score comparison chart based on dependency rules work on 5 product based dataset.

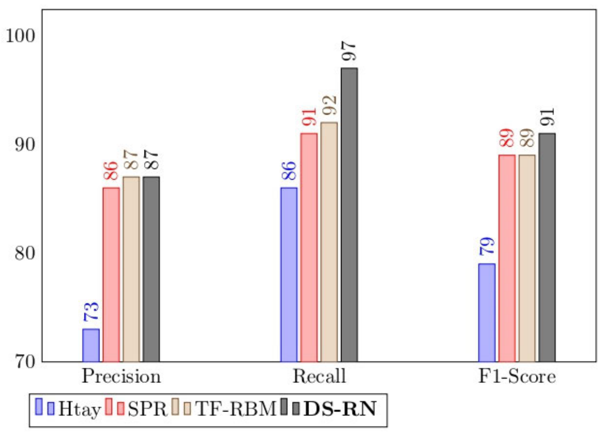


FIGURE 13. Average precision, recall and F1-score comparison chart based on sequential pattern based rules work on 5 product based dataset.

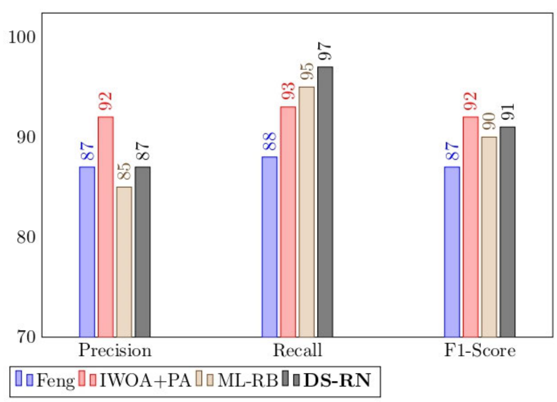


FIGURE 14. Average precision, recall and F1-score comparison chart based on hybrid techniques on 5 product based dataset.

Section II and are listed below for comparison of performance with our proposed DS-RN model:

- Works using Dependency based rules include DP [18], RubE [27], RSLS [29].
- Sequential Pattern rule based papers include Htay [24], SPR [40],TF-RBM [32].
- Studies using Hybrid based techniques include Feng [41], IWOA+PA [37], ML-RB [33].

E. DISCUSSIONS

The performance comparison of our proposed DS-RN Model is done separately with the works done on dependency rule based, sequential pattern rule based and hybrid techniques as shown in Table 3, Table 4 and Table 5, respectively. Overall, the DS-RN yield higher recall and F1-score compared to all other approaches. The comparative results are also shown graphically in Figure 11, Figure 12 and Figure 13.

V. CONCLUSION AND FUTURE DIRECTIONS

Previous works on aspect extraction mostly adopted dependency rule based techniques or syntactic pattern based approach. However, most of these studies mainly considered only noun/noun phrases as target aspects. In addition, extracting multiple aspects from single review sentences, multi-word aspect extraction are rarely concerned. In this study, we have proposed a Dependency based Rules suing ROOT Node (DS-RN) technique for the extraction of explicit aspects from customer reviews. This proposed model can identify and extract single aspect word, multi-word aspect(s), multiple aspect words form the sentence using 31 rules that have been formulated using the concept of ROOT Node as reference as well as Non-ROOT node. Three types of lexicon based pruning techniques i.e., Opinion-lexicon, Non-aspect based Noun lexicon and Action verb lexicon are incorporated manually into these rules in order to eliminate irrelevant aspects. Five benchmark datasets from different domain types were used in this experiment. The experimental results conducted on DS-RN proposed model have shown the significance improvement in the Recall, Precision, and F1-score of the proposed model compared to other approaches worked on the same field and on the same dataset.

In future, lexicons(i.e., dictionary) of opinions, action-verb words, non-aspect based nouns can be created automatically using reinforcement learning. Another possible future direction of this study is to formulate more robust dependency based rules in different domains of dataset other than product-based so as to develop more generalized rules for aspect extraction.

FUNDING AND/OR CONFLICTS OF INTERESTS/COMPETING INTERESTS

The authors declare that they have no competing interests or personal relationships that could have appeared to influence the work reported in this paper.

APPENDIX

See Table 8.

TABLE 8. List of Type-01 and Type-02 Rules.

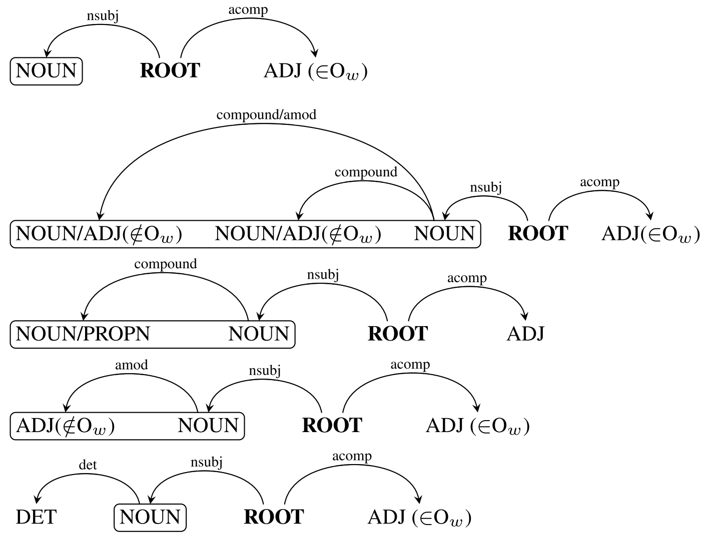
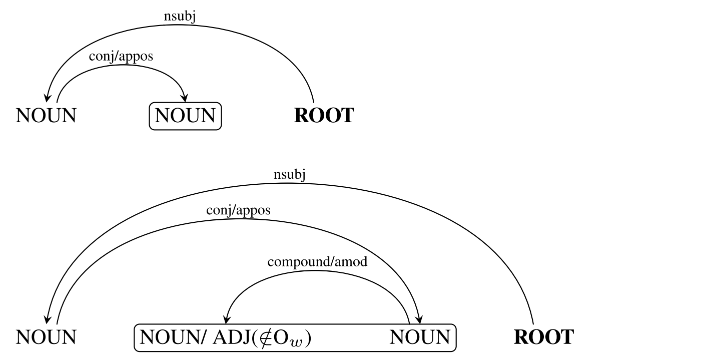
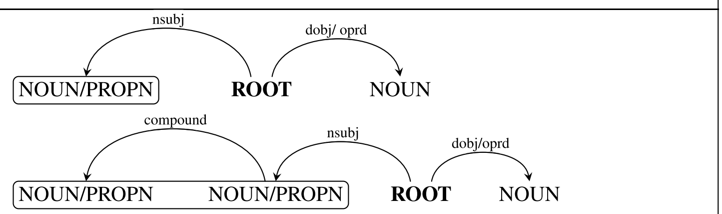
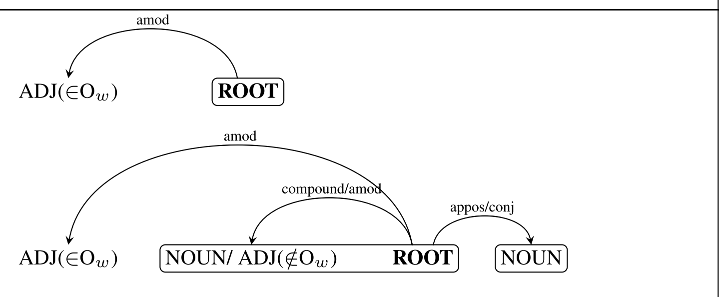
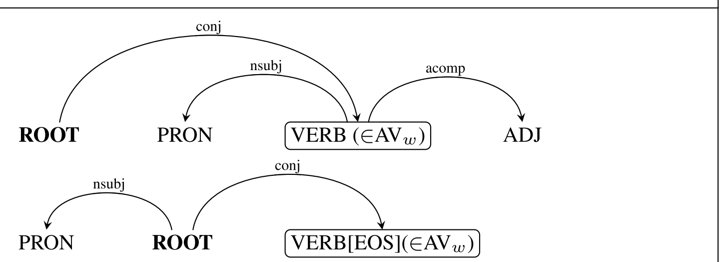
<table>
<tr><th>RULES</th><th>ROOT NODE</th><th>DEPENDENCY STRUCTURAL PATTERN WITH REFERENCE TO ROOT NODE</th></tr>
<tr><td>Rule 1</td><td>VERB,AUX</td><td>
</td></tr>
<tr><td>Rule 2</td><td>AUX</td><td>
</td></tr>
<tr><td>Rule 3</td><td>VERB,AUX</td><td>
</td></tr>
<tr><td>Rule 4</td><td>NOUN</td><td>
</td></tr>
<tr><td>Rule 5</td><td>VERB</td><td>
</td></tr>
</table>

TABLE 8. ***(Continued.)*** **List of Type-01 and Type-02 Rules.**

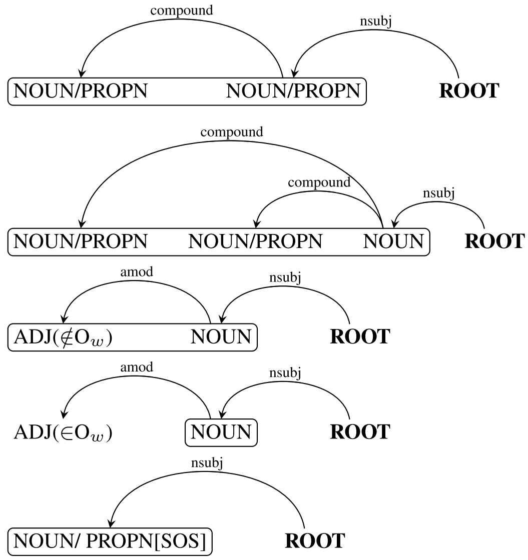
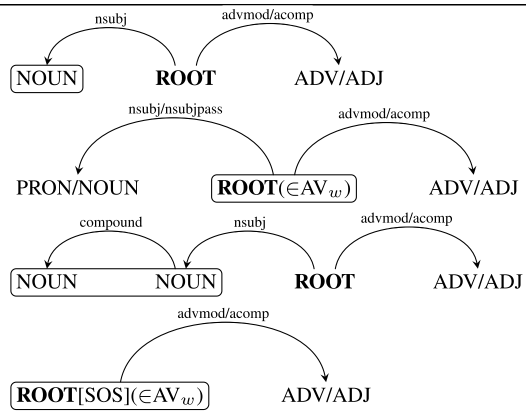
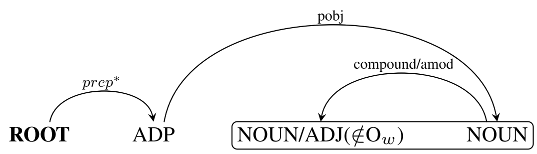
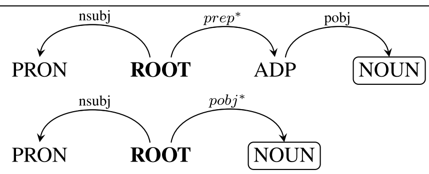
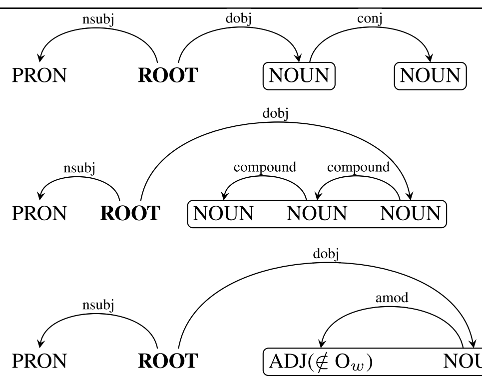
Rule 6	VERB, AUX	
Rule 7	VERB	
Rule 8	NOUN	
Rule 9	VERB	
Rule 10	VERB,AUX	

TABLE 8. *(Continued.)* List of Type-01 and Type-02 Rules.

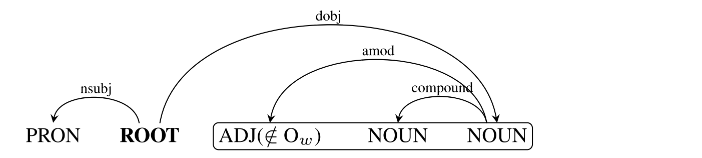
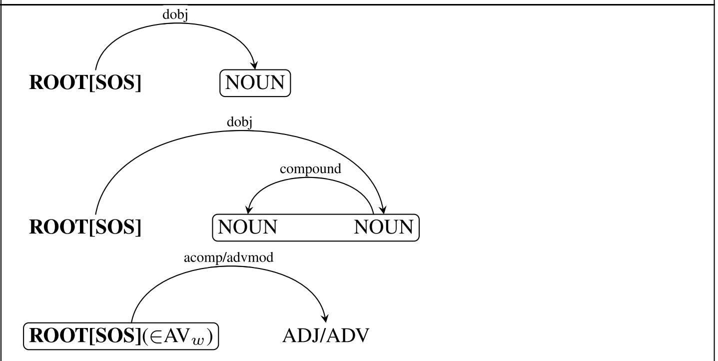
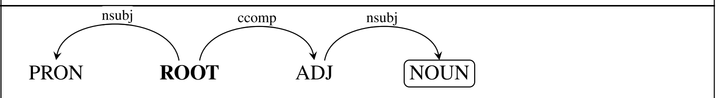
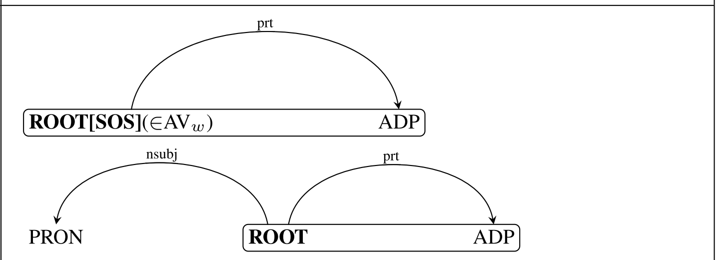
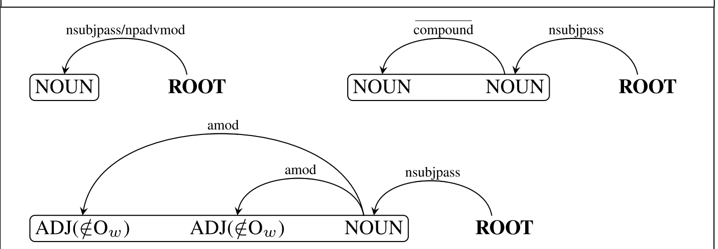
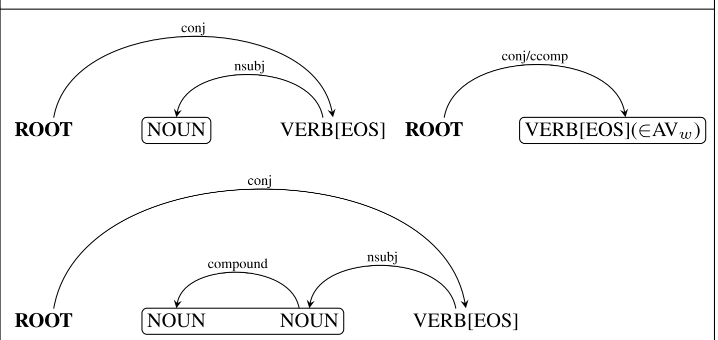
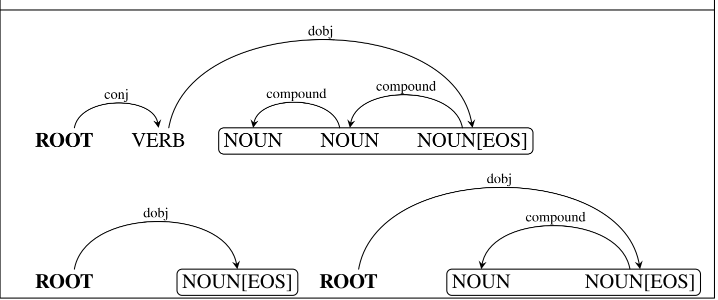
		
Rule 11	VERB	
Rule 12	VERB	
Rule 13	VERB	
Rule 14	VERB	
Rule 15	VERB	
Rule 16	VERB	

TABLE 8. ***(Continued.)*** **List of Type-01 and Type-02 Rules.**

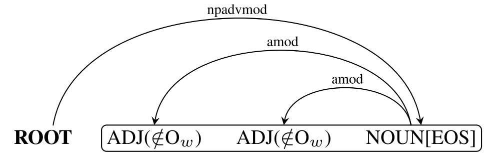
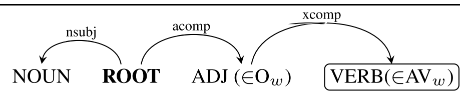
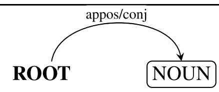
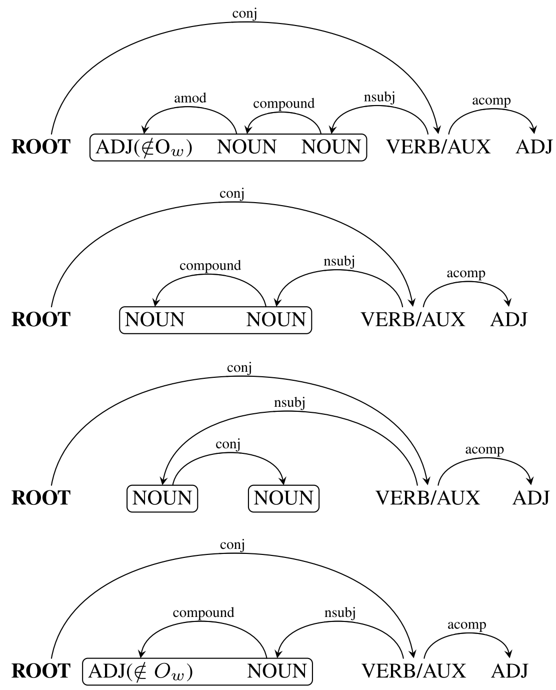
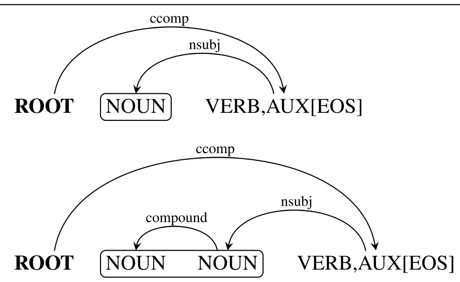
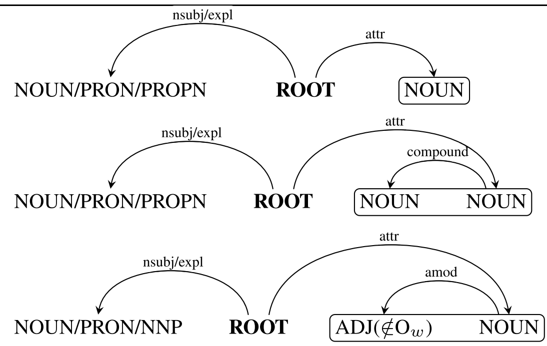
		
Rule 17	AUX	
Rule 18	NOUN	
Rule 19	VERB,AUX	
Rule 20	AUX	
Rule 21	AUX,VERB	

TABLE 8. ***(Continued.)*** **List of Type-01 and Type-02 Rules.**

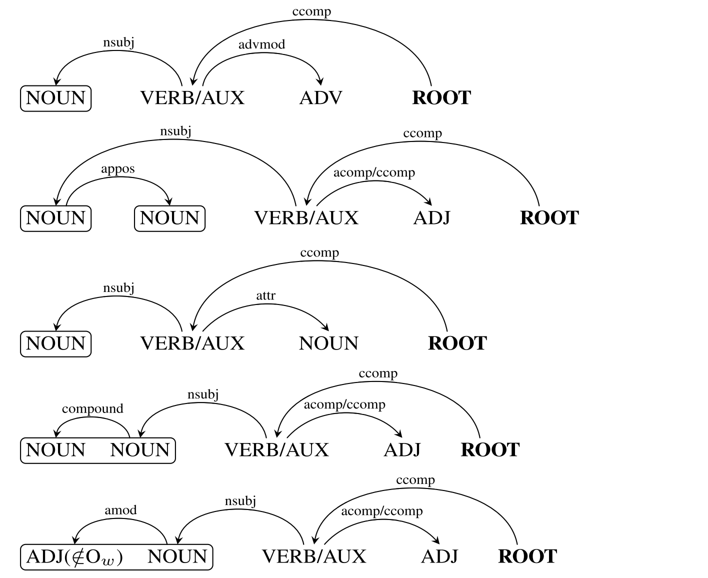
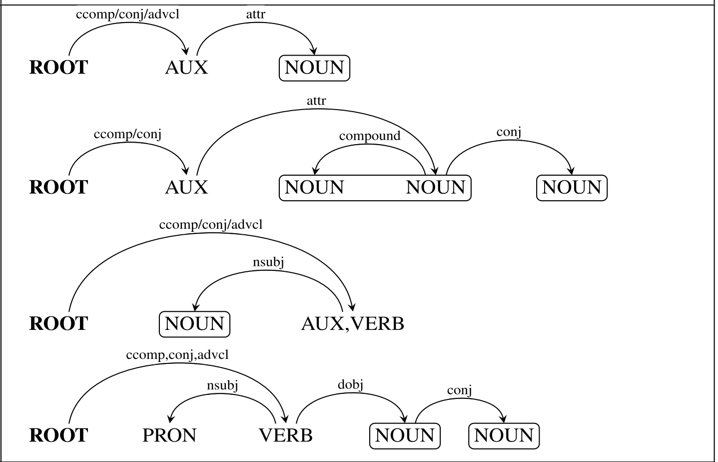
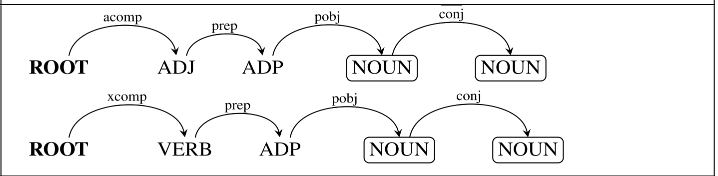
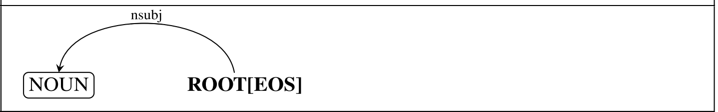
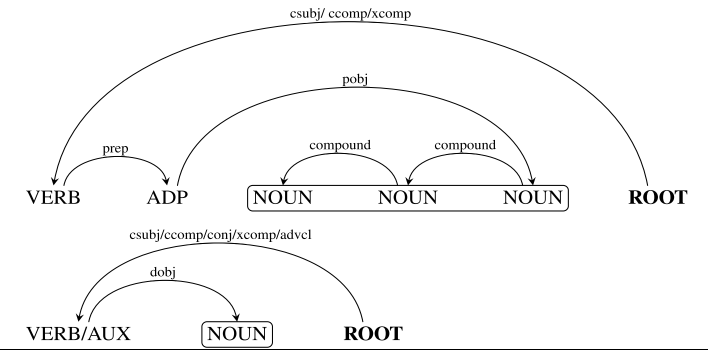
Rule 22	VERB, AUX	
Rule 23	VERB,AUX	
Rule 24	AUX	
Rule 25	VERB	
Rule 26	AUX,VERB	

TABLE 8. ***(Continued.)*** **List of Type-01 and Type-02 Rules.**

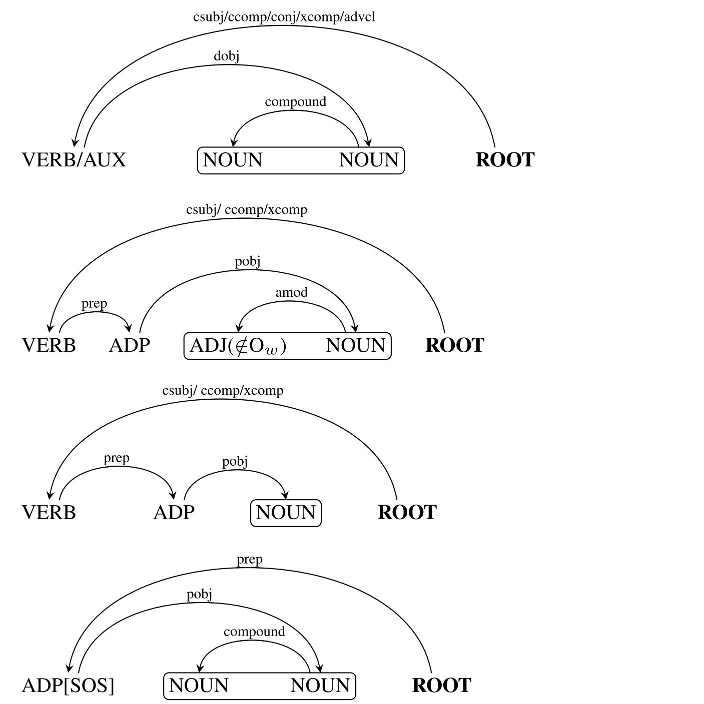
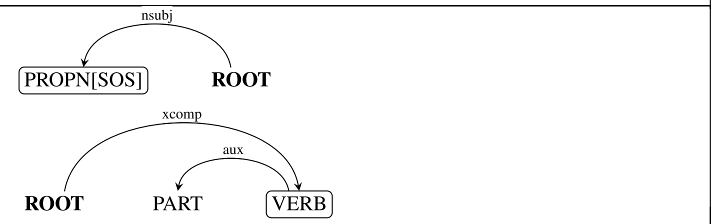
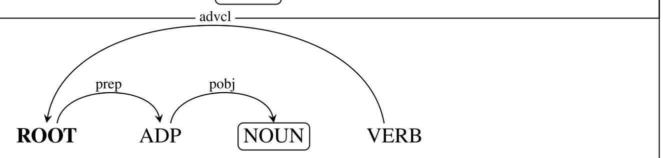
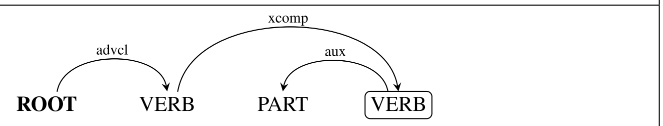
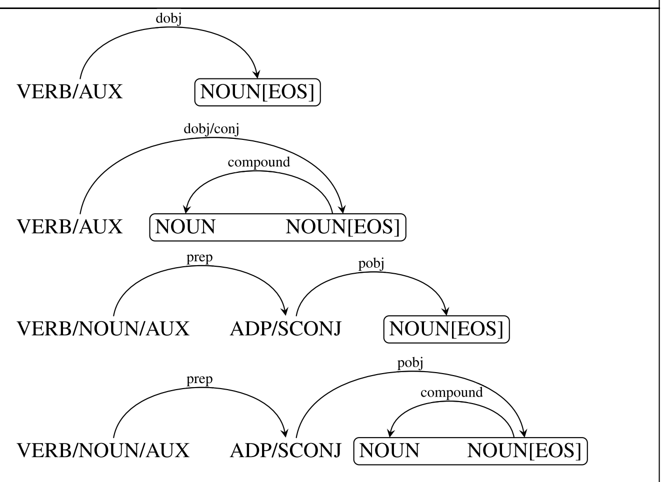
		
Rule 27	ADJ	
Rule 28	VERB	
Rule 29	ADJ	
Rule 30	—	

TABLE 8. ***(Continued.)*** **List of Type-01 and Type-02 Rules.**

		VERB/AUX ADP/SCONJ $\boxed{\text{ADJ }(\notin O_w)\ \text{ ADJ }(\notin O_w)\ \text{ NOUN[EOS]}}$ VERB/NOUN/AUX ADP/SCONJ $\boxed{\text{NOUN[EOS]}}$
Rule 31	—	ADJ PART $\boxed{\text{VERB}(\in AV_w)\ \text{ ADP}}$ ADJ PART $\boxed{\text{VERB}(\in AV_w)}$ ADJ PART $\boxed{\text{VERB}}$ CCONJ $\boxed{\text{VERB}(\in AV_w)\ \text{ ADP}}$

REFERENCES

[1] Y. Liu, J.-W. Bi, and Z.-P. Fan, "Ranking products through online reviews: A method based on sentiment analysis technique and intuitionistic fuzzy set theory," *Inf. Fusion*, vol. 36, pp. 149–161, Jul. 2017.

[2] K. Zhao, A. C. Stylianou, and Y. Zheng, "Sources and impacts of social influence from online anonymous user reviews," *Inf. Manage.*, vol. 55, no. 1, pp. 16–30, Jan. 2018.

[3] K. Akalamkam and J. K. Mitra, "Consumer pre-purchase search in online shopping: Role of offline and online information sources," *Bus. Perspect. Res.*, vol. 6, no. 1, pp. 42–60, Jan. 2018.

[4] D. Law, R. Gruss, and A. S. Abrahams, "Automated defect discovery for dishwasher appliances from online consumer reviews," *Exp. Syst. Appl.*, vol. 67, pp. 84–94, Jan. 2017.

[5] E. Cambria, S. Poria, D. Hazarika, and K. S. Kwok, "Discovering conceptual primitives for sentiment analysis by means of context embeddings," in *Proc. AAAI Conf. Artif. Intell.*, Apr. 2018, vol. 32, no. 1, pp. 1–11.

[6] T. A. Rana and Y.-N. Cheah, "Aspect extraction in sentiment analysis: Comparative analysis and survey," *Artif. Intell. Rev.*, vol. 46, no. 4, pp. 459–483, Dec. 2016.

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

[8] W. Jin, H. H. Ho, and R. K. Srihari, "A novel lexicalized HMM-based learning framework for web opinion mining," in *Proc. 26th Annu. Int. Conf. Mach. Learn.*, vol. 10, Jun. 2009, pp. 1–8.

[9] Y. Choi and C. Cardie, "Hierarchical sequential learning for extracting opinions and their attributes," in *Proc. ACL Conf. Short Papers*, Jul. 2010, pp. 269–274.

[10] P. Jiang, C. Zhang, H. Fu, Z. Niu, and Q. Yang, "An approach based on tree kernels for opinion mining of online product reviews," in *Proc. IEEE Int. Conf. Data Mining*, Dec. 2010, pp. 256–265.

[11] L. Chen, L. Qi, and F. Wang, "Comparison of feature-level learning methods for mining online consumer reviews," *Exp. Syst. Appl.*, vol. 39, no. 10, pp. 9588–9601, Aug. 2012.

[12] S. Li, R. Wang, and G. Zhou, "Opinion target extraction using a shallow semantic parsing framework," in *Proc. AAAI Conf. Artif. Intell.*, vol. 26, no. 1, 2012, pp. 1671–1677.

[13] F. L. Cruz, J. A. Troyano, F. Enríquez, F. J. Ortega, and C. G. Vallejo, "'Long autonomy or long delay?' the importance of domain in opinion mining," *Exp. Syst. Appl.*, vol. 40, no. 8, pp. 3174–3184, Jun. 2013.

[14] M. Tubishat, N. Idris, and M. A. Abushariah, "Implicit aspect extraction in sentiment analysis: Review, taxonomy, opportunities, and open challenges," *Inf. Process. Manage.*, vol. 54, no. 4, pp. 545–563, Jul. 2018.

[15] J. Yu, Z. J. Zha, M. Wang, and T. S. Chua, "Aspect ranking: Identifying important product aspects from online consumer reviews," in *Proc. 49th Annu. Meeting Assoc. Comput. Linguistics, Hum. Lang. Technol.*, Jun. 2011, pp. 1496–1505.

[16] Y. Wu, Q. Zhang, X. Huang, and L. Wu, "Phrase dependency parsing for opinion mining," in *Proc. Conf. Empirical Methods Natural Lang. Process.*, 2009, pp. 1533–1541.

[17] C. P. Wei, Y. M. Chen, C. S. Yang, and C. C. Yang, "Understanding what concerns consumers: A semantic approach to product feature extraction from consumer reviews," *Inf. Syst. E-Bus. Manage.*, vol. 8, pp. 67–149, Mar. 2010.

[18] G. Qiu, B. Liu, J. Bu, and C. Chen, "Opinion word expansion and target extraction through double propagation," *Comput. Linguistics*, vol. 37, no. 1, pp. 9–27, Mar. 2011.

[19] K. Liu, L. Xu, and J. Zhao, "Syntactic patterns versus word alignment: Extracting opinion targets from online reviews," in *Proc. 51st Annu. Meeting Assoc. Comput. Linguistics*, vol. 1, Aug. 2013, pp. 1754–1763.

[20] A. K. Samha, Y. Li, and J. Zhang, "Aspect-based opinion extraction from customer reviews," 2014, *arXiv:1404.1982*.

[21] Z. Yan, M. Xing, D. Zhang, and B. Ma, "EXPRS: An extended pagerank method for product feature extraction from online consumer reviews," *Inf. Manage.*, vol. 52, no. 7, pp. 850–858, Nov. 2015.

[22] A. M. Popescu and O. Etzioni, "Extracting product features and opinions from reviews," in *Natural Language Processing and Text Mining*. London, U.K.: Springer, 2007, pp. 9–28.

[23] S. Moghaddam and M. Ester, "Opinion digger: An unsupervised opinion miner from unstructured product reviews," in *Proc. 19th ACM Int. Conf. Inf. Knowl. Manage.*, Oct. 2010, pp. 1825–1828.

[24] S. S. Htay and K. T. Lynn, "Extracting product features and opinion words using pattern knowledge in customer reviews," *Sci. World J.*, vol. 2013, pp. 1–5, Dec. 2013.

[25] Z. Hai, K. Chang, J. Kim, and C. C. Yang, "Identifying features in opinion mining via intrinsic and extrinsic domain relevance," *IEEE Trans. Knowl. Data Eng.*, vol. 26, no. 3, pp. 623–634, Mar. 2014.

[26] S. Poria, E. Cambria, L.-W. Ku, C. Gui, and A. Gelbukh, "A rule-based approach to aspect extraction from product reviews," in *Proc. 2nd Workshop Natural Lang. Process. Social Media (SocialNLP)*, 2014, pp. 28–37.

[27] Y. Kang and L. Zhou, "RubE: Rule-based methods for extracting product features from online consumer reviews," *Inf. Manage.*, vol. 54, no. 2, pp. 166–176, Mar. 2017.

[28] V. R. Kumar and K. Raghuveer, "Dependency driven semantic approach to product features extraction and summarization using customer reviews," in *Proc. Adv. Comput. Inf. Technol., 2nd Int. Conf. Adv. Comput. Inf. Technol. (ACITY)*, vol. 3, Jul. 2012, pp. 225–238.

[29] Q. Liu, Z. Gao, B. Liu, and Y. Zhang, "Automated rule selection for opinion target extraction," *Knowl.-Based Syst.*, vol. 104, pp. 74–88, Jul. 2016.

[30] W. Maharani, D. H. Widyantoro, and M. L. Khodra, "Aspect extraction in customer reviews using syntactic pattern," *Proc. Comput. Sci.*, vol. 59, pp. 244–253, Jan. 2015.

[31] M. Z. Asghar, A. Khan, S. R. Zahra, S. Ahmad, and F. M. Kundi, "Aspect-based opinion mining framework using heuristic patterns," *Cluster Comput.*, vol. 22, no. S3, pp. 7181–7199, May 2019.

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

[33] M. Hussain, "Multi layered rule-based technique for explicit aspect extraction from online reviews," Ph.D. dissertation, Dept. Comput. Sci., COMSATS Univ. Lahore, Lahore, Pakistan, 2021.

[34] J. Lafferty, A. McCallum, and F. C. Pereira, "Conditional random fields: Probabilistic models for segmenting and labeling sequence data," in *Proc. Int. Conf. Mach. Learn.* San Francisco, CA, USA: Morgan Kaufmann, 2001.

[35] Y. Li, Z. Qin, W. Xu, and J. Guo, "A holistic model of mining product aspects and associated sentiments from online reviews," *Multimedia Tools Appl.*, vol. 74, no. 23, pp. 10177–10194, Dec. 2015.

[36] A. K. Samha and Y. Li, "Aspect-based opinion mining using dependency relations," *Int. J. Comput. Sci. Trends Technol. (IJCST)*, Feb. 2016, pp. 113–123.

[37] M. Tubishat, N. Idris, and M. Abushariah, "Explicit aspects extraction in sentiment analysis using optimal rules combination," *Future Gener. Comput. Syst.*, vol. 114, pp. 448–480, Jan. 2021.

[38] G. G. Qiu, B. Liu, J. Bu, and C. Chen, "Expanding domain sentiment lexicon through double propagation," in *Proc. 21st Int. Joint Conf. Artif. Intell.*, Jun. 2009, pp. 1–6.

[39] Q. Q. Liu, B. Liu, Y. Zhang, D. S. Kim, and Z. Gao, "Improving opinion aspect extraction using semantic similarity and aspect associations," in *Proc. AAAI Conf. Artif. Intell.*, Mar. 2016, vol. 30, no. 1, pp. 1–7.

[40] T. A. Rana and Y.-N. Cheah, "Sequential patterns rule-based approach for opinion target extraction from customer reviews," *J. Inf. Sci.*, vol. 45, no. 5, pp. 643–655, Oct. 2019.

[41] J. Feng, W. Yang, C. Gong, X. Li, and R. Bo, "Product feature extraction via topic model and synonym recognition approach," in *Big Data*. Singapore: Springer, Nov. 2019, pp. 73–88.

[42] S. Poria, E. Cambria, and A. Gelbukh, "Aspect extraction for opinion mining with a deep convolutional neural network," *Knowl.-Based Syst.*, vol. 108, pp. 42–49, Sep. 2016.

[43] A. Da'u, N. Salim, I. Rabiu, and A. Osman, "Weighted aspect-based opinion mining using deep learning for recommender system," *Exp. Syst. Appl.*, vol. 140, Feb. 2020, Art. no. 112871.

[44] A. Nawaz, S. Asghar, and S. H. A. Naqvi, "A segregational approach for determining aspect sentiments in social media analysis," *J. Supercomput.*, vol. 75, no. 5, pp. 2584–2602, May 2019.

[45] A. Bagheri, M. Saraee, and F. D. Jong, "An unsupervised aspect detection model for sentiment analysis of reviews," in *Natural Language Processing and Information Systems*. Berlin, Germany: Springer, Jun. 2013, pp. 140–151.

[46] M. Dragoni, M. Federici, and A. Rexha, "An unsupervised aspect extraction strategy for monitoring real-time reviews stream," *Inf. Process. Manage.*, vol. 56, no. 3, pp. 1103–1118, May 2019.

[47] P. Ray and A. Chakrabarti, "A mixed approach of deep learning method and rule-based method to improve aspect level sentiment analysis," *Appl. Comput. Informat.*, vol. 8, pp. 78–163, Mar. 2022.

[48] K. M. Karaoglan and O. Findik, "Extended rule-based opinion target extraction with a novel text pre-processing method and ensemble learning," *Appl. Soft Comput.*, vol. 118, Mar. 2022, Art. no. 108524.

PALLAVI MISHRA is currently pursuing the Ph.D. degree with the Department of Computer Science and Engineering, Faculty of Science and Technology (IcfaiTech), ICFAI Foundation for Higher Education (Deemed to be University), Hyderabad, Telangana, India. She has published two works with CRC Press (Taylor & Francis Group): her first article, "Big Data Digital Forensics and Cyber Security," was published, in 2020, and her second book chapter, "Collaborative Filtering Techniques-Algorithms and Advances," was published, in 2021. Her research interest includes natural language processing.

SANDEEP KUMAR PANDA (Member, IEEE) is currently an Associate Professor and the Head of the Department of Artificial Intelligence and Data Science, Faculty of Science and Technology (IcfaiTech), ICFAI Foundation for Higher Education (Deemed to be University), Hyderabad, Telangana, India. He has ten lakh seed money projects from IFHE. He has published 50 papers in international journals and international conferences and book chapters in repute. He has 17 Indian patents on his credit. He has four edited books named *Bitcoin and Blockchain: History and Current Applications* (CRC Press, USA), *Blockchain Technology: Applications and Challenges* (Springer ISRL), *AI and ML in Business Management: Concepts, Challenges, and Case Studies* (CRC Press, USA), and *The New Advanced Society: Artificial Intelligence and Industrial Internet of Things Paradigm* (Wiley Press, USA), in his credit. His research interests include blockchain technology, the Internet of Things, AI, and cloud computing. He is a member of ACM and a Life Member of IAENG. He received the Research and Innovation of the Year Award 2020 from MSME, Government of India, and DST, Government of India, New Delhi, in 2020. He received the "Research Excellence Award" from Brand Honchos, in 2022.