

Detecting Relationship between Features and Sentiment Words Using Hybrid of Typed Dependency Relations Layer and POS Tagging (TDR Layer POS Tags) Algorithm

Siti Rohaidah Ahmad[#], Mohd Ridzwan Yaakub^{*}, Azuraliza Abu Bakar^{*}

[#]*Department of Computer Science, Faculty of Defence Science & Technology, Universiti Pertahanan Nasional Malaysia, 57000, Kuala Lumpur, Malaysia*
E-mail: sitiroidah@upnm.edu.my

^{*}*Data Mining and Optimization Research Group, Centre for Artificial Intelligence Technology, Faculty of Information Science & Technology, Universiti Kebangsaan Malaysia, 46000, Bangi, Selangor, Malaysia*
E-mail: ridzwanyaakub@ukm.edu.my, azuraliza@ukm.edu.my

Abstract— Through online product reviews, consumers share their opinions, criticisms, and satisfactions on the products they have purchased. However, the abundance of product reviews may be confusing and time-consuming for prospective customers as they read and analyze differing views before buying a product. The unstructured format of product reviews needs a sentiment mining approach in analyzing customers' comments on a product and its features. In this paper, the researchers explore and analyze the hybrid role of typed dependency relations (TDR) and part-of-speech tagging (POST) in detecting the relation between features and sentiment words. The researchers have also created a list of combination rules using TDR and POST to serve as a guide in identifying the relation between features and sentiment words in sentences. Results have shown that the hybrid algorithm could assist in identifying such a relationship and improve performance.

Keywords— typed dependency relations; part-of-speech tags; feature; sentiment word

I. INTRODUCTION

Online purchases are becoming more popularized due to recent advancements in information technology. Sellers normally provide feedback sections for users to comment, review or express their opinions on products they have purchased and used. The user comments are becoming useful sources of shared information; and are helpful to buyers, sellers, and product manufacturers alike. For buyers or consumers, the information from the comment sections enables them to assess and decide on whether or not to buy a product. For sellers and manufacturers, the customer feedbacks are valuable in helping to improve product development, marketing strategies, and customer relations. Opinions are sometimes written on websites using non-standard language and in unstructured forms. Therefore, a methodology to analyze the contents of customers' opinions, and to produce output is required to assist sellers, manufacturers, and consumers in decision-making. Sentiment analysis technology functions to analyze reviews or opinions from users regarding products, politics, services,

individuals, current issues, etc. [1]. This phenomenon has resulted in a heightened awareness among different parties and has currently become the focus of researchers in sentiment analysis technology.

In sentiment analysis, analyzing opinions on a large scale document is a challenging task [2]. Analysts need to identify three main constituents of a user comment: feature word, sentiment word, and the relation between these two words [3]. Failure to identify these three basic constituents could affect the results of sentiment classification. Hence, this study examined and developed an algorithm that uses the functions of typed dependency relation and parts-of-speech tags in identifying feature words, sentiment words, and the relation between both of them in sentences.

Ultimately, the aim of this research is to develop a method for identifying the relation between features and sentiment words. There are a number of approaches in finding the relations between features and sentiment words in a sentence. Previous literature mostly focused on product reviews. However, there are also studies that are inclined towards analyzing movie reviews. Studies by [4], [5] were the earliest to have extracted product features and sentiment

words from customer review data. They used Apriori algorithm to extract features with the highest frequency. The concept of feature pruning, namely: compactness pruning and redundancy pruning, has been used to eliminate incorrect features. Additionally, [6] used phrase dependency parsing in their research to identify product features, expressions of sentiments, and the relation between them. In their study, they proposed a new tree kernel function to model phrase dependency trees. In another research, [7] proposed six syntactic relationships based on typed dependency relations. They used dependency analysis to identify and extract features, sentiment words, and the relation between them.

The authors [8] introduced the concept of pattern knowledge to extract features, which comprised of nouns or noun phrases, from review data. They used extracted features to identify sentiment words, namely adjectives or adverbs, which are closest to the features. The pattern knowledge produced is capable of identifying features, sentiment words, and the relation between features and sentiment words. However, this method still utilizes the concept of distance, which is executed by finding nearby features of sentiment words present in sentences. Problems occur when the distance between a sentiment word(s) and the feature is far apart. As a consequent, the process of identifying sentiment word(s) may be hindered. Meanwhile [9], suggested a hybrid pattern, which is a combined pattern based on noun phrases (cBNP). The hybrid pattern is a combination of four different patterns. The pattern is based on dependency relation between the sentiment terms represented by the adjectives with the product features represented by the nouns. The process to extract features and sentiment words is based on this hybrid pattern. On the other hand, [10] used three types of typed dependency relation, namely: ACOMP, XCOMP, and ADVMOD, to determine whether or not a review sentence is a sentiment sentence. Additionally, he also used typed dependency relation to identify product features. The approach used in this study is different from all the studies above. This study proposed a combination of typed dependency relation concept with the functions of part-of-speech tags to find feature words, identify sentiment words, and examine the relation between these two words.

From the discussions above, it has been identified that previous approaches in identifying relationships face constraints such as:

- 1) Existing dependency relationship techniques are unable to identify the relations between features and sentiment words in complex sentences.
- 2) The presence of more than one features and sentiment words in the same sentence hinders relationship matching process.
- 3) The use of distance in detecting relationships between features and word sentiment is flawed when more than one features and sentiment words are present in the same sentence.

- 4) Only three parameters of dependency relationships that exist in the Stanford Parser has been used in identifying the relationships between features and word sentiments.

From the analysis of previous literature, dependency relationship has been identified to be capable of identifying the type of word and the relationship between one word and another in a sentence. All parameters of typed dependency relationships in the Stanford Parser should be concurrently used since the sentence structures in customer reviews are primarily different. Consequently, the possibility of acquiring the combination of multiple types of TDR, especially for complex sentences, can be achieved.

II. MATERIAL AND METHOD

A. Methodology

This study was specifically designed to identify the relations between features and sentiment words. The main components of the methodology, which consist of 5 phases, are shown in Fig. 1:

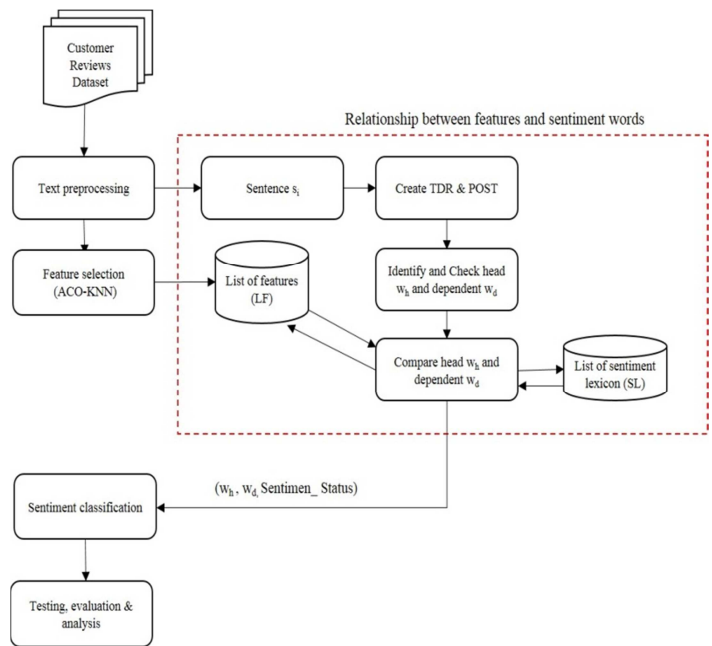


Fig. 1 Processes in identifying the relations between features and sentiment words

Phase 1: Text preprocessing: The researchers carried out the experiment on five data sets which were accessed from the Amazon website; covering five different types of electronic products. The same type of data sets were used in [4] and [5]. First, the researchers eliminated noisy data from selected customer reviews before proceeding to subsequent phases. Since most users do not have full mastery over the language, incorrect spellings, and ungrammatical sentence structures are identifiable within the data. Additionally, reviews which are written in short form words, with punctuation errors, with colloquial spellings, words, and structures, without correct capitalization of words, etc., are

also commonly found. Next, researchers performed POS tagging using a Standard Parser to identify nouns with tagging nouns, adjectives, and verbs; whereby these words are later on extracted from sentences.

Phase 2: Feature selection: The researchers chose a combination of ant colony optimization (ACO) and k-nearest neighbour (KNN) as the feature selection technique. Based on the researches by [11] and [12], ACO could potentially be used as a feature selection technique in text classification. Among the advantages of ACO are:

- ACO has fast ability in the convergence process.
- ACO has good discovery process ability in the problem space.
- ACO is proficient in outcome minimum subset feature.

Additionally, KNN functions as classifier algorithm that evaluates a subset of the features in the feature selection process by the ACO. The classifier algorithm (KNN) performance is used to evaluate the subset features.

Phase 3: Relationship between features and sentiment words: The researchers employed the combination of Typed Dependency Relations and POS tags to identify features, sentiment words, and the relation between them. The red dash dot box in Fig. 1 is the focus of this study. A detailed discussion on this subject will be available under The Hybrid of Typed Dependency Relations Layer and POS Tagging Algorithm.

Phase 4: Sentiment classification: The researchers classified the relation between features and sentiment words into groups of positive or negative sentiment words.

Phase 5: Testing, evaluation, and analysis: The researchers conducted manual verification by comparing the output acquired with a customer review data set that was already classified. The detailed discussion on testing, evaluation, and analysis of the proposed algorithm will be explained in later parts of this paper.

B. Concept of Typed Dependency Relations and POS Tags

In this study, the researchers employed stanford typed dependencies (STD) to find the relation between words in a sentence [13]. The representation of STD provides a simple description of the grammatical relationships in a sentence. Note that, there are 50 grammatical relations in STD. Results of the analysis revealed that various TDR types such as NSUBJ, NMOD, AMOD, XCOMP, DOBJ and others, connect words in sentences of the customer review data set. Table 1 shows the definitions of the TDRs based on grammatical relation categories.

TABLE I
SOME OF DEFINITIONS OF TDRS BASED ON GRAMMATICAL RELATION CATEGORIES

Categories	Typed dependency relations	Description (adopted from [13])
1	NSUBJ	A nominal subject is a noun phrase which is the syntactic subject of a clause.
2	AMOD	An adjectival modifier of an NP is any adjectival that serves to modify the meaning of the NP.
3	NMOD	A nominal modifier relation holding between the noun and the adjective.
4	DOBJ	The direct object of a VP is the noun phrase which is the object of the verb.
5	XCOMP	An open clausal complement (XCOMP) of a VP or an ADJP is a clausal complement without its own subject, whose reference is determined by an external subject.

In this research, the researchers used the Penn Treebank English POS tag set [14] because the customer review data for this study is in English. Table 2 lists some of these POS tags.

TABLE II
PENN TREEBANK POS TAGS

POS Tags	Description
NN	Noun
NNS	Noun plural
NNP	Proper noun, singular
NNPS	Proper noun, plural
JJ	Adjective
JJR	Adjective, comparative
JJS	Adjective, superlative
VB	Verb, base form
RB	Adverb

Meanwhile, every word in the sentences has its own POS tags. For that matter, a list of combination rules was derived based on TDR types and POS tags; with reference to the results of the analysis of customer review data sets. These rules are important as guidelines in finding the relations between features and sentiment words. Table 3 shows the some of the rules for typed dependency relations and POS tags for NSUBJ.

TABLE III
SOME OF THE RULES FOR TYPED DEPENDENCY RELATIONS AND POS TAGS
FOR NSUBJ

TDR	POS Tags	(Feature, sentiment word)
NSUBJ	(JJ/VBN→NN/NNS)	(camera, perfect)
(NSUBJ←)(→AD VMOD)	(NN←(VBZ/JJ)→RB)	(autofocus, well)
(NSUBJ)→(AMO D)	(JJ→(NN)→JJ/VBN)	(manual mode, rich)
(NSUBJ)→(NMO D)	(JJ→(NN)→NN)	(picture quality, excellent)

In this study, the relation between a word and another in a sentence is defined as 'one layer TDR'. Additionally, the relation between more than one TDR layers is known as 'multiple layer TDR'. Both one layer TDR and multiple layer TDR are used in finding the relation between features and sentiment words in a sentence. The algorithm is capable of combining five TDR layers in identifying features and sentiment words relation, based on the analyzed results of the proposed algorithm. Fig. 2 and 3 exemplify the sentences containing the relation between the feature, picture and sentiment word, perfect.

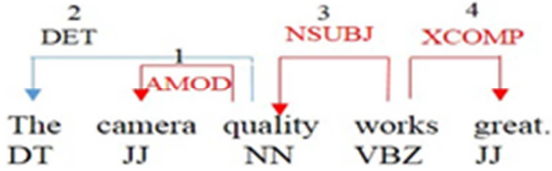


Fig. 2 The dependency relation for the sentence: This picture is perfect

The above example represents one layer TDR that is NSUBJ (perfect/JJ→picture/NN).

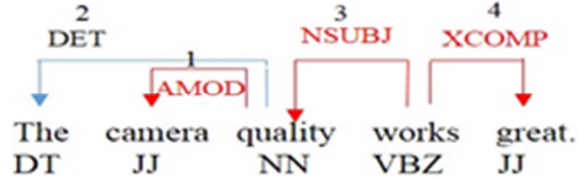


Fig. 3 The dependency relation for the sentence: The camera quality works great

However, in Fig. 3, three layers of TDR, namely AMOD-NSUBJ-XCOMP, need to be passed in order to connect the feature, camera quality with the sentiment word, great. The combination of three TDR layers in Fig. 3 is thus, categorized under 'multiple layer TDR' (camera/JJ←quality/NN←works/VBZ←great/JJ). In this case, POS tagging for the feature is the combination of adjective and noun (JJ-NN); while POS tagging for sentiment word is an adjective (JJ).

C. The Hybrid of Typed Dependency Relations Layer and POS Tagging Algorithm

The pseudocode of the algorithm in detecting the relation between features and sentiment words is shown in Fig. 4. The initial values of the proposed algorithm are labeled as sentences s_i , a list of sentiment lexicon (SL); and a list of product features (LF). The algorithm adopts a sentence from

a customer review as the point to start its analysis. For each sentence, the system generates TDRs and POS tags. It then identifies a feature for every review using POS tags and compares it with the (LF). Next, it checks and compares the head and the dependent for every TDR with the (LF) and (SL). If no new feature or sentiment word is found in the sentence s_i , the algorithm stops its analysis of the current review and begins to analyze the next review.

Input A:
Sentence s_i
List of feature (LF)
List of sentiment lexicon (SL)
Step 1: Input A
Step 2: Create typed dependency relations (TDR) for each s_i
Step 3: Create POS tags (POST) for each s_i
Step 4: Check and get POST for each TDR in s_i
Step 5: Check head w_h and dependent w_d for TDR then
Step 6: Compare head w_h and dependent w_d == LF;
Step 7: Compare head w_h and dependent w_d == SL
Step 8: If exists, set Feature_Status equal True and set Type of Sentiment_Status equal to P N.
Step 9: Add (w_h , w_d , Sentiment_Status) to list (w_h , w_d);
Step 10: Repeat the process for all the sentences.
END

Fig. 4 The Hybrid of Typed Dependency Relations Layer and POS Tagging (TDR Layer POS Tags) algorithm

III. RESULT AND DISCUSSION

Experimental results for the proposed algorithms are discussed in this section. In this study, the researchers focus on identifying the relation between features and sentiment words in customers' comments. The researchers are particularly interested in sentences that contain opinions on the product features. All the reviews were checked by the proposed algorithm. For the evaluation of the algorithm's performance, the researchers used precision (P), recall (R), and F-score (F) to measure its effectiveness in identifying the relations between the features and sentiment words. The precision, recall, and f-score are calculated by the following formula [15]:

$$Precision = \frac{TP}{TP+FP} \quad (1)$$

$$Recall = \frac{TP}{TP+FN} \quad (2)$$

$$F-score = \frac{2*(Precision*Recall)}{Precision+Recall} \quad (3)$$

whereby true positive (TP) is the number of comments from which the algorithm appropriately extracted the exact relations between features and sentiment words; false positive (FP) is the number of comments from which the algorithm incorrectly extracted the relations between features and sentiment words; and false negative (FN) is the number of comments that the algorithm has unsuccessfully attempted to identify the relations between features and sentiment words. Three series of experiments were conducted to test the effectiveness of the proposed algorithm, namely:

A. Experiment 1: Analysis of Precision, Recall, and F-score for Five Datasets

To verify the effectiveness of the proposed algorithm, the researchers evaluated the algorithm and compared the results with the hybrid pattern [8] and the pattern knowledge [9].

TABLE IV
PERFORMANCE (PRECISION, RECALL, AND F-SCORE) OF PROPOSED ALGORITHM AND BASELINE ALGORITHMS ON CUSTOMER REVIEW DATASETS

Dataset	TDR Layer POS Tags			Hybrid pattern			Pattern knowledge		
	P	R	F	P	R	F	P	R	F
Nikon	79.9	94.4	86.5	79.3	74.3	76.6	71.2	81.2	75.9
Nokia	83.3	94	88.3	81.3	74.5	77.8	73.6	82.1	77.6
Apex	88.7	82.5	85.4	81	72.9	76.8	78.2	97	86.6
Canon	83.1	89.4	86.1	76.3	70.5	73.3	73.9	92.1	82
Creative	89.6	91.8	90.6	76.9	66.5	71.3	69.6	76.2	72.8
Average	84.9	90.3	87.4	78.9	71.8	75.2	73.3	85.7	79

Table 4 shows the experimental results of the algorithm in terms of precision, recall, and F-score, in comparison to those of the hybrid pattern and the pattern knowledge algorithms; when applied to five different datasets. It is observable that the proposed TDR Layer POS Tags algorithm outperforms other algorithms from all three aspects: precision, recall, and F-score. The five customer review datasets are taken from Nikon, Nokia, Apex, Canon, and Creative datasets.



Fig. 5 Comparison of average precision, recall, and f-score for TDR layer POS tags algorithm, hybrid pattern, and pattern knowledge

Fig. 5 shows the average percentages of precision, recall, and F-score for all three algorithms: TDR Layer POS Tags, hybrid pattern, and pattern knowledge. The graph clearly exhibits that the TDR Layer POS Tags algorithm is capable of identifying the relation between features and sentiment words in sentences. Moreover, it increased the average values of precision from 5.93% to 11.61%, recall from 4.73% to 18.66%, and f-score from 8.41% to 12.22%. The increment indicates that the proposed algorithm is effective in extracting relations and is superior to the existing approaches for two main reasons. First, TDR is a relation between two words. Therefore, it simplifies the process of finding the relations between features and sentiment words. Second, the hybrid pattern and pattern knowledge

approaches could not possibly identify the relations between features and sentiment words that are present in complex sentences. Thus, it is conclusive that the TDR Layer POS Tags algorithm can be used in more practical settings. As such, the capability of the proposed algorithm is better than the existing hybrid pattern and pattern knowledge.

B. Experiment 2: Analysis TDR Layer POS Tags Algorithm with ACO-KNN, IG-GA and IG-RSAR As Feature Selection Techniques

The purposes of this experiment are to validate the combined ACO-KNN as feature selection techniques and demonstrate the appropriateness of the hybrid relationship between typed dependency relations layer and POS tagging algorithm in sentiment analysis. The experiment was done to trace the pairing of features and sentiment words. For baseline algorithms, the combination of information gain (IG) and genetic algorithm (GA); and the combination between information gain and rough set attribute reduction (RSAR), are used as feature selection techniques. Additionally, TDR layer POS tags algorithm is used to identify the relationship between features and sentiment words. We choose IG-GA because this combination has been proven to be an effective feature selection technique for sentiment analysis as exemplified by [16]. The RSAR was combined with IG as a feature selection technique for sentiment analysis, as done by [17].

TABLE V
AVERAGE RESULTS ON DATA SETS OF DIFFERENT APPROACHES

Approach	Precision	Recall	F-score
Proposed Algorithm	85.76	90.01	87.70
IG-GA	83.02	83.69	83.17
IG-RSAR	81.18	84.58	82.70

Table 5 shows a comparison between the results of ACO-KNN, IG-GA and IG-RSAR as feature selection algorithms with TDR layer POS tags algorithm. The combination of ACO-KNN with the TDR layer POS tags algorithm was used to detect the relations between POS features and sentiment words. This combination is able to improve the accuracy of sentiment analysis; in comparison to the performance results from IG-GA and IG-RSAR algorithms.

C. Experiment 3: Significance t-test

A significance test is used to assess significant differences in the mean values between two algorithms, namely: ACO-KNN algorithm and; IG-GA and IG-RSAR algorithms. To be able to identify significant differences between the two algorithms, the significant level for ACO-KNN must be less by 0.05 than the IG-GA and IG-RSAR algorithms. The 'significance' columns display the comparative significant differences, (+) or (-), for both algorithms (see Table 6).

TABLE VI
SIGNIFICANCE TEST FOR ALGORITHM ACO-KNN WITH ALGORITHMS IG-GA
AND IG-RSAR

Dataset	ACO-KNN & IG-GA (Value p)	Significance	ACO-KNN & IG-RSAR (Value p)	Significance
Nikon	0.1434	-	0.0002	+
Nokia	0.0001	+	0.0233	+
Apex	0.0087	+	0.0253	+
Canon	0.4553	-	0.0069	+
Creative	0.1158	-	0.047	+

Statistical results from the significance test (t-test) indicate that the ACO-KNN algorithm is better than the IG-RSAR & IG-GA algorithms in terms of its significance; except for Nikon, Canon, and Creative datasets. The researchers have identified three reasons behind the good performance of the proposed approach:

1) The significant difference between this work and previous studies by [4], [5] and [18] is that the approach used in this research employs a more systematic feature selection process. The feature selection method (ACO-KNN) used is capable of producing optimum feature set. Based on this approach, texts have to be preprocessed before feature selection takes place. This step is important to ensure that the feature set is optimized with only the highest quality and most important features.

2) The proposed method in this study differs from previous methods as used by [6], [7], and [10]. This study employs an interactive method in finding the relationship between features and sentiment words, whereby a principle-based approach that combined typed dependency relation with POS tagging was used. Additionally, this method is also capable of identifying relationships between features and sentiment words for up to five layers TDR (multiple layers).

3) Most of the previous studies, the study by [4], [5] for example, only took into account nouns or noun phrases in identifying features in sentences. However, an analysis on selected data sets revealed that features may consist of a noun, a verb, or even the combinations of noun-adjective or noun-verb. Additionally, sentiment words in previous studies [4] and [19] are mainly identified as adverbs, verbs or adjectives; whereas sentiment words in this study may consist of verb, adverb, noun or the combinations of adverb-adjective, noun-verb, noun-adjective or adverb-adjective-noun.

IV. CONCLUSIONS

In this paper, the researchers have proposed a hybrid algorithm between TDR Layer and POS Tags in identifying the expressions in sentiment words, and the relations between features and sentiment words in a sentence. The proposed algorithm is able to identify multiple layers of type dependency relations between features and sentiment words in a sentence due to its utilization of both TDR layers and

POS tags. Additionally, the experimental evaluation has shown that the algorithm is able to outperform two baseline approaches. It is, therefore, conclusive to claim that the proposed algorithm can be used in practical settings. For prospective research, the researchers intend on further investigating, improving, and refining the algorithm for implicit feature detection and selection; implicit sentiment word detection and selection; and sentiment word orientation identification. The researchers also plan on conducting a study to investigate the use of pronoun features in customer review datasets in the future.

ACKNOWLEDGMENT

The authors gratefully acknowledge Universiti Pertahanan Nasional Malaysia and Universiti Kebangsaan Malaysia as well as the Geran Galakan Pensyarah Muda for supporting this research project through grant no. GGPM-2016-006.

REFERENCES

- [1] B. Liu, *Sentiment Analysis and Opinion Mining*. Morgan & Claypool Publishers, 2012.
- [2] B. Seerat and F. Azam, "Opinion Mining: Issues and Challenges (A Survey)," *International Journal of Computer Applications*, vol. 49, no. 9, 2012.
- [3] A. Bagheri, M. Saraee, and F. de Jong, "Care more about customers: Unsupervised domain-independent aspect detection for sentiment analysis of customer reviews," *Knowledge-Based Systems*, vol. 52, pp. 201–213, 2013.
- [4] M. Hu and B. Liu, "Mining Opinion Features in Customer Reviews," in *Proceeding AAAI'04 Proceedings of the 19th national conference on Artificial intelligence*, 2004, vol. 21, no. 2, pp. 755–760.
- [5] M. Hu and B. Liu, "Mining and Summarizing Customer Reviews," *Proceedings of the 2004 ACM SIGKDD international conference on Knowledge discovery and data mining KDD 04*, pp. 168–177, 2004.
- [6] Y. Wu, Q. Zhang, X. X. Huang, and L. Wu, "Phrase Dependency Parsing for Opinion Mining," in *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing Volume 3 EMNLP 09*, 2009, no. August, pp. 1533–1541.
- [7] G. Somprasertsri and P. Lalitrojwong, "Mining Features-Opinion in Online Customer Reviews for Opinion Summarization," *Journal of Universal Computer Science*, vol. 16, no. 6, pp. 938–955, 2010.
- [8] S. S. Htay and K. T. Lynn, "Extracting product features and opinion words using pattern knowledge in customer reviews," *The Scientific World Journal*, vol. 2013, 2013.
- [9] K. Khan, B. Baharudin, and A. Khan, "Identifying Product Features from Customer Reviews using Hybrid Patterns," *International Arab Journal of Information Technology*, vol. 11, no. 3, pp. 281–285, 2014.
- [10] A. Qadir, "Detecting opinion sentences specific to product features in customer reviews using typed dependency relations," in *Proceeding eETTs '09 Proceedings of the Workshop on Events in Emerging Text Types*, 2009.
- [11] M. R. Ahmad, Siti Rohaidah; Bakar, Azuraliza Abu; Yaakub, "Metaheuristic algorithms for feature selection in sentiment analysis," in *Science and Information Conference (SAI), 2015*, 2015, pp. 222–226.
- [12] M. H. Aghdam, N. Ghasem-Aghaee, and M. E. Basiri, "Text Feature Selection using Ant Colony Optimization," *Journal Expert Systems with Applications: An International Journal*, vol. 36, no. 3, pp. 6843–6853, 2009.
- [13] M. De Marneffe and C. D. Manning, "Stanford typed dependencies manual." *20090110 Httpnlp Stanford*, vol. 40, no. September, pp. 1–22, 2010.
- [14] A. Taylor, M. Marcus, and B. Santorini, "THE PENN TREEBANK: AN OVERVIEW," in *Treebanks*, 2003, pp. 5–22.
- [15] M. R. Yaakub, Y. Li, A. Algarni, and B. Peng, "Integration of Opinion into Customer Analysis Model," in *2012 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology*, 2012, pp. 164–168.

- [16] A. Abbasi, H. Chen, and A. Salem, "Sentiment Analysis in Multiple Languages: Feature Selection for Opinion Classification in Web forums," *ACM Transactions on Information Systems*, vol. 26, no. 3, pp. 1–34, 2008.
- [17] B. Agarwal and N. Mittal, "Sentiment Classification using Rough Set based Hybrid Feature Selection," in *Proceedings of the 4th Workshop on Computational Approaches to Subjectivity, Sentiment & Social Media Analysis (WASSA 2013)*, 2013, no. June, pp. 115–119.
- [18] H. Zhang, A. Sekhari, Y. Ouzrout, and A. Bouras, "Jointly identifying opinion mining elements and fuzzy measurement of opinion intensity to analyze product features," *Engineering Applications of Artificial Intelligence*, vol. 47, pp. 122–139, 2016.
- [19] G. Qiu, B. Liu, J. Bu, and C. Chen, "Opinion Word Expansion and Target Extraction through Double Propagation," *Journal Computational Linguistics*, vol. 37, no. 1, pp. 9–27, 2011.