

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/314424838>

# Negation Handling in Sentiment Analysis at Sentence Level

**Article** in Journal of Computers · January 2016

CITATIONS  
0

READS  
517

5 authors, including:



**Umar Farooq**  
Université Lumière Lyon 2  
5 PUBLICATIONS 5 CITATIONS

SEE PROFILE



**Antoine Nongaillard**  
Université des Sciences et Technologies de Lille 1  
38 PUBLICATIONS 75 CITATIONS

SEE PROFILE



**Yacine Ouzrout**  
Université Lumière Lyon 2  
171 PUBLICATIONS 452 CITATIONS

SEE PROFILE



**Muhammad Abdul Qadir**  
Capital University of Science & Technology  
130 PUBLICATIONS 497 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:



Standardisation of CPS for Smart Manufacturing [View project](#)



Product Lifecycle Management [View project](#)

# Negation Handling in Sentiment Analysis at Sentence Level

Umar Farooq<sup>1\*</sup>, Hasan Mansoor<sup>2</sup>, Antoine Nongaillard<sup>3</sup>, Yacine Ouzrout<sup>4</sup>, Muhammad Abdul Qadir<sup>5</sup>

<sup>1,4</sup> DISP Laboratory, University Lumiere Lyon 2, Lyon, France.

<sup>2</sup> Department of Informatics and Mathematics, University Joseph Fourier Grenoble.

<sup>3</sup> Univ. Lille, CNRS, Centrale Lille, UMR 9189 CRISTAL (SMAC), F-59000 Lille, France.

<sup>5</sup> Department of Computer Science, Capital University of Science and Technology, Islamabad, Pakistan.

\* Corresponding author. Tel.: 0033602533403; email: Umar.Farooq@univ-lyon2.fr

Manuscript submitted March 20, 2016; accepted March 31, 2016.

doi: ???

---

**Abstract:** Sentiment analysis is an automatic way to determine that whether opinions of people about a subject are favorable or unfavorable. One of the most important sub tasks in sentiment analysis is to determine the sequence of words affected by negation. Most of the existing sentiment analysis systems used traditional methods based on static window and punctuation marks to determine the scope of negation. However, these methods do not offer satisfactory performance due to variability in the negation scope, inability to deal with linguistic features and improper word sense disambiguation. In this paper, we investigate the problem of identifying the scope of negation while determining the polarity of a sentence. We propose a negation handling method based on linguistic features which determine the effect of different types of negation. Experiment results show that the proposed method improves the accuracy of both negation scope identification and overall sentiment analysis.

**Key words:** Sentiment Analysis, Opinion Mining, Negation Scope, Negation Handling

---

## 1. Introduction

Sentiment analysis also known as opinion mining refers to the process of determining opinions or emotions expressed in text about a subject. Sentiment analysis is a wide research area at the intersection of different domains of *natural language processing*, *computational linguistics* and *text mining*. Sentiment analysis can be performed at sentence level or document level. The sentence level determines that whether opinion expressed in a sentence about a subject is positive or negative whereas at document level the whole document is classified as positive or negative. In this paper we are focusing on a subtask called negation handling which is part of a larger research effort to improve the performance of sentence level sentiment analysis.

Negations are very important linguistics because they affect the polarities of other words. Negations include words such as no, not, shouldn't etc. When a negation appears in a sentence it is important to determine the sequence of words which are affected by this term. The scope of negation may be limited only to the next word after a negation or may be extended up to other words following negation. For example, in sentence "this mobile is not nice but it is working properly" the scope of negation is only limited to the next word after negation. On the other hand, in another sentence "the battery does not work for a long time" the scope of negation is until the end of the sentence. These examples show that the scope of negation is not fixed and varies based on different linguistics features such as conjunctions, punctuation marks and part of speech (POS) of negation etc. In addition, the presence of a negation term does not mean that all words in the sentence which carry polarities will be inverted [1]. However, most of the existing sentiment analysis systems used traditional methods (such as

static window [2], [3] and method based on punctuation marks [4]) to deal with this complex linguistic phenomenon. In static window a fixed number of words (i.e. 1, 2, 3, 4 or 5) after negation are considered as the scope of negation and polarities of only these words are inverted. On the other hand, the method based on punctuation marks inverts the polarities of words between negation and next punctuation mark.

There are many forms of negation such as explicit (with explicit clues such as not, no etc.), implicit, diminishers and other restrained linguistic patterns. At highest structural level negations may appear in two forms, i.e. morphological and syntactic negations [5]. Morphological negations follow a specific structure where the root word is modified by either prefix (e.g. ir-, non-, un- etc. ) or suffix (e.g. -less ). Syntactic negations are those where explicit negation cues are used to invert a single word or a sequence of words. In complex situations negations may appear implicitly without any presence of the explicit words in the sentence such as in irony [6]. On other hand, diminishers also called reducers usually diminish the polarities of other words which they effect instead of inverting the polarity completely. Some of the existing studies such as [1] and [2] did not make a clear distinction between different types of negations and proposed a single method to determine the scope of negation for all or some types of negations. A clear distinction need to be made while determining the scope of different types of negation effectively because they affect polarities in different ways.

The objective of this paper is to improve the performance of sentence level sentiment analysis in the presence of negations. A negation handling method is proposed which determines the negation scope of different types of negations. Three linguistic features are used to determine the sequence of words affected by syntactic negation while considering some exceptions. A set of prefixes and suffix are used in order to handle morphological negations and to improve lexicon hits. Furthermore, the words affected by diminishers are also determined in order to reduce the polarities strengths of these words. This paper is organized as follows. In Section II, related work about negation scope identification is presented. Section III explains the proposed negation handling method. Finally, Section IV evaluates the performance of the proposed method.

## 2. Related Work

In literature several methods have been proposed to identify the scope of negation while determining the polarity of a sentence. Most of these methods are either based on static window or punctuation marks. In early study, the polarity of the whole sentence is inverted if a negation word appears in a sentence [7]. Some authors used the punctuation marks to identify the scope of negation. These methods invert the polarities of those words which come between the negation word and the next punctuation mark [4]. On the other hand, some studies also used static window of words following or around a negation in order to handle negation. The simplest way is to invert the polarity of the sentiment bearing word directly following the negation word [8]. In [9] the negation word is searched in a window from three to six words before an opinionated word; if negation is found then the polarity of words within this window is inverted. Similarly, in [10] and [11] the polarities of five next words to the negation are inverted. However, the effect of negation may be finish even after one word especially when conjunctions are used in a sentence. Another issue is that some of these methods only considered specific types of words such as adjectives [11] or adverbs and adjectives [12]. The polarities of only these types of words are inverted. In fact, other types of words such as verbs and nouns can also carry polarities and need to be considered if affected by a negation. In [2] a comparison of methods with different static windows is performed. The results show that the method which considered two words after negation performed the best. However, the overall performance of all these methods is not satisfactory and hence the author suggested that only static window is not enough to deal with such a complex natural language problem. All these methods work well in simple sentence. However in compound sentence these methods do not work well because not only the punctuation marks and position of the words from negation but also the conjugation words limit the scope of negation. Furthermore, in some situations even the punctuation mark or conjunctions do not limit the scope of negation. Similarly, sometimes the negation words do not invert the polarity such as in the case of word "not only". Therefore, a negation handling method must need to identify all these situations so that the polarities of

only those words are inverted which are actually affected by a negation.

A prominent method to determine the scope of negation is proposed in [1]. This method used three concepts: static delimiters, dynamic delimiters and heuristic rules based on word polarities to determine the scope of negation. The clause where negation appeared is determined using the Static delimiters and called the candidate scope which is further processed using dynamic delimiters and heuristic rules (based on words polarities) to eliminate some words from the candidate scope. However, this method also has some issues. The scope of the negation is determined by only considering the clause in which a negation appeared. However, the scope may be extend to another clause in some situations. Another issue is that sometimes heuristic rules do not work well and also need proper disambiguation of words according to context [6].

A survey on the role of negation in sentiment analysis is presented in [6] which described issues in existing methods. Two main issues in existing techniques are highlighted. First, an appropriate negation handling method need the awareness of polar expressions. Secondly, all polar words may not be in lexicon dictionary and hence some polar words may not be recognized. We can conclude that an efficient word sense disambiguation method and a lexicon dictionary which contains words with all senses are prerequisites for an effective negation handling method.

### 3. Negation Handling

Negations are those words which affect the sentiment orientation of other words in a sentence. Examples of negation words include not, no, never, cannot, shouldn't, wouldn't, etc. Negation handling is an automatic way of determining the scope of negation and inverting the polarities of opinionated words that are actually affected by a negation. The portion of the sentence that negation affects is called the vicinity or scope of negation. Usually, in simple sentence (i.e. sentence contains only one clause) a negation may invert the polarities of all words in a sentence. However, in compound sentence (in which there are multiple clauses) a negation usually inverts the polarities of some words in a sentence and the number of words inverted varies based on different linguistic features. Therefore, it is little bit challengable to handle negation in compound sentence. In order to determine the negation scope, we define a list of negations which serve as an indicator of the presence of a negation. All the negation words are categorized in three classes, i.e. syntactic, diminisher and morphological negations. Table 1 shows the complete list of negations which are further extended by including the common misspellings. The syntactic negation class includes all those negations which completely invert the polarities of other words whereas the diminisher class includes all those negation words which reduce the polarities instead of inverting. On the other hand, the morphological class includes all those prefixes and suffix which can be used to form a morphological negation. These prefixes and suffix are also used as an primary indicator of the presence of a morphological negation. Based on this negation classification and issues highlighted in literature review, we propose a separate method for each class to determine the negation scope, which not only works well in simple but also in compound sentences.

Table 1. List of Negations

Negation Class	Negations
Syntactic	no, not, rather, couldn't, wasn't, didn't, wouldn't, shouldn't, weren't, don't, doesn't, haven't, hasn't, won't, wont, hadn't, never, none, nobody, nothing, neither, nor, nowhere, isn't, can't, cannot, mustn't, mightn't, shan't, without, needn't,
Diminisher	hardly, less, little, rarely, scarcely, seldom
Morphological	<i>Prefixes:</i> de-, dis-, il-, im-, in-, ir-, mis-, non-, un- , <i>Suffix:</i> -less

#### 3.1. Scope of Syntactic Negations

Syntactic negation is the well-known and most common negation type in user generated text. Most of the existing methods only determine the scope of syntactic negations. In order to determine negation scope in syntactic negations we proposed a method which uses three linguistic features with static window and also considers some expectations. The linguistic features include conjunction analysis, punctuation marks and

heuristics based on POS of negation term. The following section explains each linguistic feature, use of static window and exceptions in more details.

1) *Conjunction Analysis*: Conjunctions are those words which link clauses with each other in a sentence. Examples of conjunctions include: and, or, but, whereas, etc. Conjunction analysis is the important part of our technique because it is very handy to identify the scope of negation in compound sentence. For example in sentence "the camera is not great but it is OK for video call" the negation "not" is used in the left clause of the first sentence which inverts the polarity of a single opinionated word "great". The conjunction "but" does not allow the effect of negation to extended to the next clause of this sentence. Most of the conjunctions limit the scope of negation to individual clause. However, sometimes some conjunctions such as "or" and "and" allow the scope to be extended to the next clause. For example, in sentences "I cannot make a call or text" and "the screen is not bright and clear" same parts of speech patterns are used on either side of conjunctions. In first sentence nouns (i.e. call and text) and in second sentence adjectives (i.e. bright and clears) are used on either side of conjunctions "or" and "and". In both sentences the scope of negation is extended to the next clause. In order to perform conjunction analysis we are using the set of conjunctions defined in our previous study [13].

2) *Punctuation Marks*: Punctuation marks such as comma ",", semicolon ";", colon ":", exclamation mark "!", braces "(", ")", full stop ".", question mark "?" and single and double quotation marks also limit the scope of negation. A negation handling method should consider all these punctuation marks in order to determine the scope of negation. One exception associated with punctuation mark comma "," is that the scope of negation extends to the next clause whenever a comma is used either for "or" or "and" and the same parts of speech pattern is followed on either side.

3) *Heuristics based on Part of Speech (POS) of Negation*: The syntactic negation terms appear in different POS in a sentence such as adverb, noun, contraction, determiner, interjection, exclamation etc. We propose heuristic rules wherever a syntactic negation word appears as an adverb RB. A negation term with an adverb POS forms one of the following three possible semantic relationships. 1) The negation word RB may modify an adjective JJ which further describes a noun NN (i.e. RB, JJ, NN). 2). Similarly, the negation may modify a verb VB which further describes a noun (i.e. RB, VB, NN). 3) The negation term may modify an adverb which further modifies an adjective or a verb. In this case we have two different relationships (i.e. (RB, RB, JJ, NN) and (RB, RB, VB, NN)). First the semantic relation that a negation forms is determined using the proposed method in our previous study [14] and then all opinionated words involved in such a relationship are considered as the scope of negation.

4) *Static Window*: Static windows of words with different sizes are used in early studies to determine the scope of negation, which already yields a reasonable performance. However, only static window of words does not offer satisfactory performance especially in compound sentences. We are using static window of five words within a clause instead of a sentence (as in existing methods) with other linguistic features when the following conditions hold true. 1) When the negation appears in a POS other than adverb (as we have heuristic rules for adverb). 2) When there are more than five words after the negation term within the same clause.

5) *Exceptions in Handling Syntactic Negations*: We already discussed some linguistic features to determine the scope of negation. This section explains two exceptions associated with syntactic negation, i.e. when the negation does not have any scope at all and when the negation word inverts the polarity of the whole clause/sentence without affecting any opinionated word. Some exceptions related to particular linguistic feature are already discussed in respective sections. In some situations syntactic negations do not have any scope when used with other words such as "not just", "not only", "no wonder", "not to mention". Our method uses these exceptional words to determine such situations.

Sometimes, especially in product reviews there is no opinionated word within the scope of negation and the negation word negates a non opinionated word (which is usually a product feature). In such case the polarity of non opinionated word (which is actually 0) is inverted to negative. The strength of the negative polarity can be decided on the bases of popularity of the product feature which is negated. This situation occurs frequently when the users want to express opinion about a product in the absence of a particular product feature. For example, in

sentences "*there is no CD ROM in this laptop*" the negation negate the features "*CD ROM*" which do not carry any opinion. This exceptional case occurs when there is no opinionated word within the negation scope and the negation word (which is either determiner (TD) POS or negation term 'without') negate a product feature.

### 3.2. Scope of Diminishers and Morphological Negations

The diminisher negations are different from the syntactic negations because they usually reduce the polarities of other words instead of completing inverting the polarities. In addition, the words that diminishers affect may not follow the negation term but may be anywhere in the sentence unlikely in syntactic negation. For example, in sentences "*this mobile hardly lags*" and "*the application crashes rarely*" the diminishers (i.e. hardly and rarely) reduce the strengths of negative polarities of words "*lags*" and "*crashes*". In first sentence, the negation affected the word following but in second it affected the word before the negation term. Based on these facts, we can say that the method proposed for syntactic negations will not give accurate results if used for diminishers. For example, if we use the proposed scope identification method for syntactic negation for the first sentence then the polarity of the word lags will be completely inverted and the polarity of this sentence and the sentence "*this mobile do not lags*" will be equivalent. However, the polarities of both of these sentences should not be equivalent because one meant that the application never lags, however another conveyed that it lags but rarely. Therefore, the diminishers should be deal differently than syntactic negations. Indeed, the diminishers appear in a sentence as an adverb and affect the polarity of either an adjective or a verb. In order to determine the scope of diminishers we uses the diminishers list down in Table 1 to indicate the presence of such type of negation in a sentence. The affected adjective or verb is determined using the following two heuristics. 1) Adverb is usually used immediately before or after an adjective or a verb, which it modifies. 2) If neither an adjective nor a verb is immediately before or after the diminishers then it likely to affect the nearby adjective or verb within the same clause. After determining the affected adjective/verb the polarity of this word is diminished by using a reducing factor of 0.2. For example, if the polarity of lags is -0.75 and it is affected by diminisher 'hardly' as in above example then the polarity strength of word lags is reduced to -0.15 using the reducing factor of 0.2 (i.e.  $-0.75 \times 0.2 = -0.15$ ).

In some cases the negation term and the negated opinionated word are combined in a single word, e.g. in words such as end-less, impolite, dishonest, non-cooperative, etc. This type of negation is called morphological negation which can be formed by using either one of the nine prefixes (i.e. de-, dis-, il-, im-, in-,ir-, mis-, non-, un-) or one suffix (i.e. -less) with a root word. We are using these prefixes and suffix in order to indicate the presence of a morphological negation. One method to handle such negations is to first make an effort to lexicalized all these types of words. If such negation appears in a sentence then polarity is obtained from the lexicon directly without any scope identification and polarity inversion. However, uncommon words may not necessarily be in the lexicon dictionary which may lead to wrong classification. Therefore, we proposed a simple method to handle such situations. In fact, the negation scope of morphological negation is limited only to the root word. The morphological negation word is decomposed in order to determine the root word. The polarity of the root word is obtained from the lexicon which is then inverted. If the root word is also not lexicalized then the polarity of the synonym is obtained. In this way by decomposing the morphological negation and by using the polarity of synonyms set the lexicon hits can be increased.

## 4. Results

In order to evaluate the performance of the proposed negation handling method we performed two sets of experiments. The first set of experiments is performed to measure the performance of negation scope identification. The second set of experiments is performed to evaluate that how much the performance of the sentence level sentiment analysis has been improved by the proposed negation handling method.

### 4.1. Evaluation of Negation Scope Identification

In order to assess the performance we implemented the proposed negation handling method. One thousand

compound sentences, each of which contains a negation are randomly sampled from review sites such as *Amazon*, *Ebay* and *Cnet*. Indeed, it is little bit more challengeable to determine the negation scope in compound sentences hence our data set only contains compound sentences. These sentences already had good English. The existing feature extraction method [11] is implemented in order to determine product features, as they are required in one negation exception. The proposed and existing negation handling methods are used to determine the negation scope for all these compound sentences. Fig. 1 compares the accuracy of the proposed method with exiting methods such as whole sentence based (WSB) [7], punctuation marks based (MP) [4], static windows (SW) of words with different lengths of K (i.e. SW-1, SW-2, SW-3, SW-4 and SW-5) [2], [3], [9], [11], adjective immediately following negation (Adj-1) [11] and adjectives within the three words following negation (Adj-3). Results show that our proposed negation scope identification methods outperformed the existing methods and the accuracy is 83.3%. However, the existing methods, i.e. WSB, MP, Adj and SW with different lengths did not offer satisfactory performance.

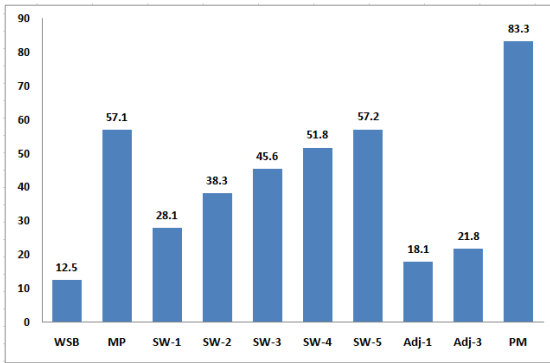


Fig. 1. Accuracy of Negation Scope.

Table 2. Performance of Sentence Sentiment Analysis

Methods	Precision	Recall	F_measure	Accuracy
WSB [7]	0.06	0.04	-	-
SW, K=1 [2]	-	-	0.51	51.3
SW, K=2 [2]	-	-	0.52	52.7
SW, K=3 [2]	-	-	0.51	52.0
SW, K=4 [2]	-	-	0.52	52.2
SW, K=5 [3]	0.18	0.77	-	-
SW, K=6 [9]	-	-	-	84.2
HB [1]	-	-	-	88.4/85.8
PM	0.93	0.91	0.92	91.3

## 4.2. Evaluation of Sentence Level Sentiment Analysis

In order to measure the contribution of the proposed negation handling method to the overall performance of sentence level sentiment analysis, the polarities of the sentences selected in Section 4.1. for previous experiment are determined. The negation scope within each sentence is already determined using the proposed negation handling method. In order to disambiguate the opinionated words according to the context the proposed method in our previous study is adopted [14]. The lexicon dictionary built in this study which support the word sense disambiguation is used to obtained polarities of different words. Subsequently, the sentiment analysis classification method proposed in our previous study [13] is used to determine the polarity of the whole sentence. Table 2 shows the performance of the sentence level sentiment analysis, which is measured in term of accuracy, precession, recall and F-measure. The results of the sentence level sentiment analysis using the proposed negation handling method are compared with existing methods such as WSB [7], SW with different length K (i.e. K=1, 2, 3, 4, 5 or 6) [2], [3], [9], [11] and heuristic based (HB) method [1]. The results show that the proposed method (PM) improved the performance of sentence level sentiment analysis considerably. In addition, it also offered the best recall and precision. The WSB method offered the worst accuracy and HB method offered reasonable performance at sentence level. On the other hand, the performance of static windows of different lengths offered more or less similar performance. The improvement in performance by the proposed method is because of three main reasons. First, the proposed method determines scope of different types of negations more effectively. Second, an appropriate word sense disambiguation method is adopted. Finally, all opinionated POS (i.e. adjective, verb, adverb and noun) are considered while determining the polarity.

## 5. Conclusion

One of the main reasons behind the errors in sentence level sentiment analysis is the inability to accurately determine the effect of negation on other words. In this paper, we focused on negation handling, which is part of



a larger research effort to improve the performance of sentence level sentiment analysis. We proposed a method which can handle different types of negation while determining the polarity of a sentence. This method efficiently determine the sequence of words affected by syntactic negation while considering the exceptions such as when the negation even don't have a scope and when the negation inverts the polarity of a clause even though no opinionated word is within the scope. In addition, the effect of diminishers is also incorporated to reduce the strengths of polarities of those words which are affected. Furthermore, the morphological negations are also handled in order to increase the lexicon hits and subsequently to improve the overall performance.

## References

- [1] Jia, L., Yu, C., & Meng, W. (2009). The effect of negation on sentiment analysis and retrieval effectiveness. *Proceedings of the 18th ACM conference on Information and knowledge management* (pp. 1827-1830).
- [2] Hogenboom, A., Van Iterson, P., Heerschop, B., Frasincar, F., & Kaymak, U. (2011, October). Determining negation scope and strength in sentiment analysis. *Proceedings of 2011 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. (pp. 2589-2594).
- [3] Wilson, T., Wiebe, J., & Hoffmann, P. (2009). Recognizing contextual polarity: An exploration of features for phrase-level sentiment analysis. *Computational linguistics*, 35(3), 399-433.
- [4] Pang, B., Lee, L., & Vaithyanathan, S. (2002, July). Thumbs up?: sentiment classification using machine learning techniques. *Proceedings of the ACL-02 conference on Empirical methods in natural language processing-Volume 10* (pp. 79-86)
- [5] Givón, T. (1993). *English grammar: A function-based introduction* (Vol. 2). John Benjamins Publishing.
- [6] Wiegand, M., Balahur, A., Roth, B., Klakow, D., & Montoyo, A. (2010, July). A survey on the role of negation in sentiment analysis. *Proceedings of the workshop on negation and speculation in natural language processing* (pp. 60-68). Association for Computational Linguistics.
- [7] Popescu, A. M., & Etzioni, O. (2007). Extracting product features and opinions from reviews. *Natural language processing and text mining* (pp. 9-28). Springer London.
- [8] Heerschop, B., van Iterson, P., Hogenboom, A., Frasincar, F., & Kaymak, U. (2011). Analyzing sentiment in a large set of web data while accounting for negation. *Advances in Intelligent Web Mastering-3* (pp. 195-205). Springer Berlin Heidelberg.
- [9] Narayanan, R., Liu, B., & Choudhary, A. (2009, August). Sentiment analysis of conditional sentences. *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing*. (pp. 180-189).
- [10] Grefenstette, G., Qu, Y., Shanahan, J. G., & Evans, D. A. (2004, April). Coupling niche browsers and affect analysis for an opinion mining application. *Coupling approaches, coupling media and coupling languages for information retrieval* (pp. 186-194).
- [11] Hu, M., & Liu, B. (2004, August). Mining and summarizing customer reviews. *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 168-177).
- [12] Benamara, F., Cesarano, C., Picariello, A., Recupero, D. R., & Subrahmanian, V. S. (2007, March). Sentiment Analysis: Adjectives and Adverbs are better than Adjectives Alone. *Proceedings of ICWSM*.
- [13] Farooq, U., Nongaillard, A., Ouzrout, Y., & Qadir, M. A. (2013). Product reputation evaluation: the impact of conjunction on sentiment analysis. *Proceedings of 7th Intl Conference on Software, knowledge, Information Management and Applications (SKIMA)*. (pp. 590-602).
- [14] Farooq, U., Dhamala, T. P., Nongaillard, A., Ouzrout, Y., & Qadir, M. A. (2015, December). A word sense disambiguation method for feature level sentiment analysis. *Proceedings of 2015 9th International Conference on Software, Knowledge, Information Management and Applications (SKIMA)* (pp. 1-8).