

Analyzing the Effect of Negation in Sentiment Polarity of Facebook Dialectal Arabic Text

KADDOURA, Sanaa <<http://orcid.org/0000-0002-4384-4364>>, ITANI, Maher <<http://orcid.org/0000-0002-0502-5130>> and ROAST, Chris <<http://orcid.org/0000-0002-6931-6252>>

Available from Sheffield Hallam University Research Archive (SHURA) at:

<https://shura.shu.ac.uk/28685/>

This document is the Published Version [VoR]

Citation:

KADDOURA, Sanaa, ITANI, Maher and ROAST, Chris (2021). Analyzing the Effect of Negation in Sentiment Polarity of Facebook Dialectal Arabic Text. *Applied Sciences*, 11 (11). [Article]

Copyright and re-use policy

See <http://shura.shu.ac.uk/information.html>

Article

Analyzing the Effect of Negation in Sentiment Polarity of Facebook Dialectal Arabic Text

Sanaa Kaddoura ^{1,*}, Maher Itani ^{2,†} and Chris Roast ^{3,†}¹ Computing and Applied Technology Department, College of Technological Innovation, Zayed University, 144534 Abu Dhabi, United Arab Emirates² Academic Development Division, Computing Department, Sabis Educational Services, 1200 Adma, Lebanon; mitani@sabis.net³ Department of Computing, Sheffield Hallam University, Sheffield S1 1WB, UK; C.R.Roast@shu.ac.uk

* Correspondence: sanaa.kaddoura@zu.ac.ae

† These authors contributed equally to this work.

Abstract: With the increase in the number of users on social networks, sentiment analysis has been gaining attention. Sentiment analysis establishes the aggregation of these opinions to inform researchers about attitudes towards products or topics. Social network data commonly contain authors' opinions about specific subjects, such as people's opinions towards steps taken to manage the COVID-19 pandemic. Usually, people use dialectal language in their posts on social networks. Dialectal language has obstacles that make opinion analysis a challenging process compared to working with standard language. For the Arabic language, Modern Standard Arabic tools (MSA) cannot be employed with social network data that contain dialectal language. Another challenge of the dialectal Arabic language is the polarity of opinionated words affected by inverters, such as negation, that tend to change the word's polarity from positive to negative and vice versa. This work analyzes the effect of inverters on sentiment analysis of social network dialectal Arabic posts. It discusses the different reasons that hinder the trivial resolution of inverters. An experiment is conducted on a corpus of data collected from Facebook. However, the same work can be applied to other social network posts. The results show the impact that resolution of negation may have on the classification accuracy. The results show that the F1 score increases by 20% if negation is treated in the text.



Citation: Kaddoura, S.; Itani, M.; Roast, C. Analyzing the Effect of Negation in Sentiment Polarity of Facebook Dialectal Arabic Text. *Appl. Sci.* **2021**, *11*, 4768. <https://doi.org/10.3390/app11114768>

Academic Editors: Barbara Guidi, Carlos A. Iglesias, Giulio Rossetti and Kevin Koidl

Received: 20 April 2021

Accepted: 18 May 2021

Published: 22 May 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: social networks; sentiment analysis; Arabic language; negation

1. Introduction

Sentiment Analysis can be defined as using natural language processing techniques to extract opinions from the text. This domain is very demanded in social networks where users express their opinions about particular subjects (such as products, policies, people, and events). Nowadays, we can witness an increased interest in people's opinions extraction, and classification [1–4]. Such opinion extraction benefits sellers as they indirectly gain insight into customers' attitudes towards products, informing their product design and advertising. Governments similarly take advantage of the ability to measure people's opinions and thoughts concerning laws, policies, and activities. The amount of data present makes manual extraction and classification opinions intractable and hence the need for automated aggregation. Moreover, users' posts tend to be short and use informal language to express their opinions. In the Arabic community, social network members use Dialectal Arabic (DA) because they use it in their conversations with friends or peers. From the user perspective, social networks are the forum for "written" conversations where they can express themselves freely without writing syntactically correct sentences that are relatively elaborate for what they need to say. This characteristic of DA posts is an obstacle for Arabic text sentiment analysis [5]: there is a limitation in the range of suitable analyses and corpora

for tools to use [6]; furthermore, common informal use of sarcasm and negation makes accurate opinion identification very hard.

Many research articles have explored the need for sentiment analysis in the Arabic language due to its popularity, particularly in most social media platforms. For example, in [7], the authors focused on studying the Arabic sentiment rather than English sentiments because the Arabic language has fewer resources than those available for the English language. Therefore, they have proposed an Arabic dataset collected from many tweets in different Arabic dialects. Their 151,000-opinion dataset was labeled as either positive or negative opinions. Their dataset was then used by many machine learning techniques showing that the ridge regression outperforms all other techniques with an accuracy of 99.9%. Similarly, the authors in [8] proposed a new approach to generate sentiment models using different machine learning techniques. Their approach was elaborated on using different social media platforms. In addition, they have handled some Arabic slang vocabulary and shorthand writings. They classified the sentiments into five categories from highly positive to highly negative, with 'No sentiment' as the middle one. The results of their approach showed high classification accuracy.

In [9], the authors focused on tracking individuals' sentiments towards news in the Arab world. They built a corpus that contains around 7200 tweets that are manually annotated. A tracking system is presented in their study, considering sentiment classification and named entity recognition techniques. After applying the dataset to different machine learning classifiers, the authors showed that tracking people's sentiments is well-performing while exploring important entities extracted from social networks in different Arab regions. Accordingly, their technique helps predict the future sentiment of new entities. The authors in [10] have presented another approach for sentiment analysis. They have created their ontology based on their country tourism. They worked on the Twitter platform to collect tweets using POS tagger. The obtained entities have been compared to the specific domain ontology. As a result, they extracted the essential feature from their collected data. The latter can then be combined with machine learning techniques to improve sentiment analysis of their study. In [11], the authors created a dataset to annotate the Saudi dialect sentiment analysis and built a lexicon for this dialect to develop a weighted lexicon-based algorithm. The latter will help detect and classify Arabic sentences as polar or non-polar, taking into consideration some additional linguistic features such as negation.

In a previous work [12], we showed that the presence of a pattern (part of the post consisting of one or two words) could help determine the polarity of the post as positive, negative, dual neutral, or spam. However, results showed that about 17% of errors were due to negation. Therefore, exploiting patterns in a lexicon-based classifier and disregarding that their sentiment may be affected by other words, such as inverters, may lead to incorrect classification.

On the other hand, these inverters have more variability in their influence. In other words, they do not always flip the polarity of sentiment in a post. As a consequence, the treatment of such words is not trivial. Besides, the complex nature and richness of the Arabic language complicate their analysis. For example, the exact words that can flip the sentiment of what comes after them may hold a meaning not related to negation, such as *ما أجمل*, which means "how beautiful" despite the use of the word *ما* which can act as an inverter in another context. Moreover, the exact words can appear as substrings of another word. When considering Dialectal Arabic (DA) with the absence of grammar and informal spelling, this problem becomes even more challenging.

This paper presents an approach to treat the negation issue and its effect on sentiment analysis. We modify the classification used in previous work [12] to consider negation. Our results show that treating negation has the potential to lead to an increase in classification performance. This performance increase relies on the negated positive frequency and the negative targets.

The rest of the paper is divided as follows: In Section 2, some essential background information is added. Section 3 illustrates the materials and methods used in this paper.

Section 4 presents the results of the presented work, followed by a discussion in Section 5. Finally, Section 6 concludes the paperwork and includes possible work extensions.

2. Background

2.1. Naïve Bayes-Based Approach

The previous work [12] that this study is based on compares and contracts the machine learning approach, Naïve Bayes, against the lexicon-based approach in classifying the sentiments of Facebook posts. The classification was done in five categories: positive, negative, dual (either positive or negative), neutral, and spam. In this approach, the authors assume that expression opinions using phrases has limitations and depends on the dialect. For example, condolence can be expressed by one verse from Qur'an. Sarcasm can be expressed by phrases taken from famous plays. Thus, to extract patterns from the corpus, manual tagging was used. The phrases used for tagging each post consist of one or two words that determine the polarity of the post. The corpus used in this work consists of posts that discuss politics, social events, and sports. The posts were written in informal Arabic language with the following dialects: Lebanese, Syrian, Iraqi, Egyptian, Libyan, Tunisian, Algerian, Saudi, and Sudanese. The classification was done using the Naïve Bayes probabilistic classifier that assumes that the features used for classification are independent. This classifier was used in two setups: The first setup uses a vector holding the following features: Positive phrase, negative phrase, positive emoticon, negative emoticon, spam, and manual tag. The manual tag can be positive, negative neutral, dual, or spam. The second setup uses a vector holding the following features: Positive phrase, negative phrase, positive emoticon, negative emoticon, and subjectivity. Subjectivity is a Boolean value that identifies whether the post is subjective or objective. The common fields in both vectors are Boolean values to indicate the presence or absence of a feature. The performance measure of setup one was 85.61%, whereas in setup two it was 75.17%.

2.2. Valence Shifters

Sentiment analysis is predicated upon the concept that words, words sequence, and words group can determine a sentiment value. Hence, some lexical items communicate either negative or positive valence or attitude [13]. For example, the verb "boost" gives a positive valence, whereas "discourage" gives a negative valence. Some words have a positive valence, and, however, when a valence shifter, such as negation, enters the sentence, it becomes negative. For example, "John is clever" versus "John is not clever". When a word with a positive valence is combined with a negation such as not, the valence changed from positive to negative. The author in [13] describes significant determiners of sentiment in terms of valence shifters, which are words that modify the polarity of a sentiment of another word. For example, valence shifters can intensify a sentiment such as "so" in "so strong". Valence shifters can also weaken a sentiment such as "slightly" in "slightly hard". Valence shifters can even invert sentiment such as "not" in "not easy". The English language has Valence shifters divided into two main categories that are either based on sentence or discourse.

Valence shifters that are Sentence-based can be categorized in terms of inverters, intensifiers, modals, pre-suppositional items, and irony:

- **Inverters:** These are words that change the polarity of a word from positive to negative and vice versa, for example, "not", "never", "none", "neither". Saying "not bad at all" changes the polarity from negative to positive; however, saying "not good" changes the polarity from positive to negative.
- **Intensifiers** can amplify the sentiment, for example: "badly" in "badly injured"; others make the meaning of the word weak, such as "slightly" in "slightly interested". Their impact is not consistent in the language context; their position can significantly influence their effect.
- **Modals:** words which may be used to assume future consequences, hence they convey a probability of an even happening. Words related to opinions may not give normal

behavior, such as “If Marwan were lazy, he would fail in their exams”. The words “lazy” and “fail” are negative; however, “would” and “was” affected the sentiment of the negative words. Thus, the sentence sentiment is that Marwan is not lazy not likely to fail in the exam.

- Pre-suppositional Items: these words can affect the valence when an event does not meet the confident expectation. For example, the word “almost” in “he almost passed” implies that the person did not pass. The word “passed” conveys a positive sentiment but “almost” changed the sentiment to a negative one. This also applies to “barely” in “the water was barely enough”. In both examples, the pre-suppositional items negatively shifted the neutral and positive sentiment. Some parts of speech may affect the sentiment in a similar way, such as saying “failed to pass” or “impossible to enjoy”.
- Irony is the last category of sentence valence shifters we consider here. Context plays a significant role in its recognition and impact, to the extent that readers may not consistently identify an invention of sentiment. The irony arises when some words are used to intensify the sentiment but flip the polarity. The word “genius” is a positive word that can give a negative sentiment such as “the genius student did not solve the exercise”.

Discourse-based valence shifters are concerned with sentence structure. In general, they show that the valence of a sentence cannot be assumed to be a simple composition of the valence of the clauses involved. Hence, in the case of DA, where grammar rules are not followed, their relevance to our analysis context is limited. Discourse-based valence shifters include the following:

- Particular conjunctions that combine clauses such as “but”, “although”, and “however” can affect words that show opinions within their range. For example, “mean” is a negative word, but when used in “Although he is mean, he treats birds in a nice way,” it has a positive meaning.
- Discourse structure: Sentences may consist of two parts: dominant and illustrative. The illustrative part supports the dominant part such that when the dominant part is opinionated, the illustrative part intensifies its sentiment. Consider the sentence “He is a great hunter. He caught five animals today”. The first part of the sentence contains the word “great,” so it is positive and opinionated. However, the second part of the sentence is neutral and intensifies the first sentence.
- Reported speech: some sentences like “he said that the episode is nice” do not mean that the author agrees on the positive sentiment. However, if it is “he said that the episode is nice and I totally agree” is supported by a phrase to give positive sentiment. Thus, “I agree” and “I do not agree” are keywords that can be used in reported speech and may change the sentiment of a sentence.

The fact that negation is a challenge to sentiment analysis in Arabic is widely recognized [14], and that few works mention the challenges of negation. In [15], the author reports that negation makes classification more challenging. In [16], the authors present three contributions related to negation: (1) presenting an approach that identifies the negation’s scope with the presence of inverter such as none, not, or barely, (2) presenting a methodology that identifies the polarity of a fragment of a sentence that has an inverter, and (3) the scope of negation concept is introduced in a system of opinion retrieval.

Similarly, in [17], the author highlights the inaccuracy of some detection techniques that neglect the presence of negation words in the Arabic language sentences. She claims that some classification techniques consider negation words as stop words; thus, they are removed. Consequently, this will lead to false detection of the real meaning of a sentence because negation reveals the truth of one’s opinion. Accordingly, the author proposed an approach in which the negation entailment is significantly showing a higher detection accuracy.

Lexicon-based approaches in [18,19] are based on the effect of negation using a list of negating terms and shift or transform valence in various ways. Based valence on accumulative values with negative words decrease the value (for instance, “not” has the impact of “−4”). This approach means that a negation would lower the valence but not fully invert it. For example, if “beautiful” is weighted +5, then “not awesome” will

not be weighted -5 . However, the result of adding -4 is 1. Thus, “not beautiful” still has a positive polarity albeit less than “beautiful”.

By contrast, in [6], the frequency of inverters in a post was used as one of the features in the classification; however, the behaviour of inverters and their exact effect on the classification process was not addressed. In [20], it was assumed that negating terms always precede the targets directly, and, hence, the targets’ polarity was inverted whenever preceded by an inverter. Clearly, word order and the scope of inverters are strongly linked to the form of a specific language. For instance, English inverters are commonly preceding their targets.

Our inverter scope supports this, finding that in 96.5% of the cases, the inverters had a scope of one, which means that they directly preceded the target, as expected according to the Arabic grammar rules [21].

This paper contributes to the above body of work by further analyzing valence shifter in sentiment classification for Arabic and DA. We focus on inverters. We follow the assumption mentioned in [6,20] in working with negation focusing upon inverters whose targets have scope one.

3. Materials and Methods

The majority of analyses of Arabic do not focus specifically upon DA. Within Modern Standard Arabic (MSA), negation is achieved by a well-defined set of words that can be used to invert the polarity of opinionated words, for example, as in *ليس جميل* (not beautiful). For sentiment analysis, when the spelling rules are appropriately followed, these words can be detected easily. However, in DA, inverters are less well defined. Different words originate from different dialects without reference to any agreed spelling rules. Thus, in addition to the challenges already discussed, DA sentiment analysis taking account of negation is particularly challenging.

Table 1 represents a table that shows some Arabic negation words, referred to as inverters, for both used in MSA and their corresponding word in DA. While MSA inverters can be used in DA, DA inverters are not conformant with MSA. It is even more complicated when the negation is part of the word itself such as the DA inverter, the Arabic letter M (م) acts as a prefix and negates the target to which it is attached, such as *محببتو* (I did not love him).

Table 1. Some Arabic Negation Words (Inverters).

MSA Inverters	DA Inverters	Translation
لا	مفي	does/do not
لم	مافي	does/do not
لما	منو	without
لات	ماكو	there is no
لن	مو	There will not be
بلا	بلاش	without
ليس	مش	No
من دون	مانو	without
بدون	م	without

The dataset [22] used in this paper consists of 1000 posts. The distribution of posts in the dataset is available in Table 2. It is based on two corpora. The first corpus is a corpus of new (CN) that contains posts retrieved from the Facebook page of AlArabiya News [23]. The other corpus is a corpus of arts, referred to as CA, contains posts retrieved from the Facebook page of The Voice [24], which is a page that covers the news of a singing competition.

Table 2. Distribution of Posts in the Dataset.

	Arts Corpus	News Corpus
Negative	224	230
Positive	233	236
Dual	151	161
Spam	197	193
Neutral	195	180

All posts in the dataset are classified into three sets: spam (containing patterns used to advertise a page or a product), negative (containing patterns used to express negative opinions), and positive (containing patterns used to express positive opinions). In [12], more details are provided about the classification procedure.

The subsequent analysis showed 17% of posts are incorrectly classified due to ignoring negation. These misclassified negation cases are analyzed in this work, and the following different scenario case identified: Inverters are expressed in many forms even within the same dialect such as *منو* spelled as “Manu” and *مانو* spelled as “Maanu”. These words can be used to say, for example, not beautiful as *مانو جميل* and *منو جميل*. The exact words used in negation may have other meanings that are not related to negation like *ما أروع*, spelled as “Ma Arwa’a” which means “how great”. The word *ما* can be used in negation to say not beautiful as *ما حلو*.

Strings representing inverters can occur as part of words not related to negation, such as the occurrence of *لا* (la) in *ملايكة*, spelled as “Maleykeh” which means angels. In this context, *لا* is part of the word *ملايكة*; however, in another context, it can be separate negation word such as *لا تروح*, spelled as “la truh” which means “do not go”.

Negation also takes place by using suffixes as in *ترحش* (truhish—do not go), prefixes as in *متروح* (Matruh—do not go) or as a separate word preceding the target as in *لا تروح* (la truh—do not go). The scope of negating terms is variable, the vast majority of cases, the scope was 1, i.e., the target came directly after the inverter. Not all inverters flip the polarity of an opinionated term. These characteristics correspond to cases in the first category of negatives mentioned in [16]. However, it is essential to note that these cases are the product of the analyzed corpora, and other cases exist in Arabic [16] but were not present.

The following cases discuss the specific obstacles to inverters in sentiment analysis that the sample used in this work are illustrated. They are different from the valence shifters represented in the literature:

1. **Tokenization:** When inverters with scope one are the adequate form some sentiment analysis. Our corpora contradicted this norm because spelling rules were not followed in DA on social media. Consider the negated pattern *ماتهرأ*; ideally, space should separate the subject from the negative pattern. As there is no space, this phrase will not be properly tokenized. Thus, it will not be treated in a proper way. Unluckily, improper tokenization is frequently available in DA. One solution would be to search for inverters prior to tokenization and thus split the negation from the subject. This approach, however, would raise other challenges at the tokenization stage. For example, legitimate cases where a negation may be falsely identified based on lexical matching. Consider the word *مشروعة* (legal): negation sensitive tokenization might result in a classifier detecting *روعة* as a positive pattern, preceded by an inverter *مش*, an incorrect inversion. This process is challenging when an inverter is a prefix. For example, the verb *عجبه* (liked him), can be negated by adding the letter *م* (M)

as a prefix to the word, leading to **معجبه**, but the exact word has another meaning of admirer or fan which is a positive meaning.

Another example would be **حُب**, the word may mean lover or did not love at the same time depending on context. In MSA or with diacritized text, the problem can be resolved. However, since DA cannot be diacritized, other solutions are needed. Smarter tokenization may help eliminate such cases by splitting prefixes and checking for the legitimacy of the remaining string as a standalone Arabic word. For example, **مشروعة**, will be split into **مشروعة** but after checking the words before it, splitting the word may be disregarded if it turned out the **مشروعة** is an adjective to what comes before it. However, such “solutions” have their own problems of additional complexity and still legitimate cases by wrongly interpreted. For example, the letter **م** is an inverter, yet there are plenty of words that start with this letter without being inverted.

2. **Fake Inverters:** The strings representing inverters have other usages not related to negation. Consider the phrase **ما أحلاها** (how beautiful she is): the phrase consists of a string used usually for negation followed by a positive word. If MSA was used, searching for patterns that belong to the morphological group used usually for comparison **أفعل التفضيل** would resolve the ambiguity. However, since such tools do not exist for DA, and if we assumed that they are available, the spelling irregularities would hinder the usage of such tools. For example, **ما نحلى** (how beautiful) is a positive pattern that is usually used to praise the beauty of an object, yet this pattern is not written **أ** (A) as it should be. What complicates the problem is when the exact “fake” negation cases appear in legitimate negation scenarios. Consider the phrase **بلا أحلى صوووووووت بلا نيله** at (not a beautiful voice at all). This is a positive pattern that is preceded by an inverter that makes it negative.

Another example is **مش أحلى صوت** (not the most beautiful voice). The positive sentence is preceded by an inverter that makes it negative. There is an important observation that many targets of the fake inverters start with the letter **أ** (A). However, this alone cannot be considered as a rule because, in other cases, a real inverter changes the polarity of a sentence starting with the same letter. One way to reduce the number of misleading cases is to filter targets consisting of 4 letters (when pronouns are not used as suffixes such as **أجملها**) and patterns consisting of all spelling variants of the letter **أ** such as **إ، آ، أ، ء، ئ،** since these are considered as fake targets. The patterns that fail to satisfy these rules may not be targets of fake inverters. However, those who satisfy the rules can be fake or legitimate or inverter’s targets.

3. **Odd Negation:** Although real inverters usually flip the polarity of sentimental targets, there are cases when this is not true. Consider the phrase **ما تسب** (do not curse). Although the negative word curse comes after a real inverter, the phrase is still negative after negation. The target in these cases has the exact properties as other patterns when negation is valid, i.e., flipping polarity. For example, the phrase **ما تزعل** (do not be sad) has the same POS-features as the previous example because the two verbs are in the present tense). Furthermore, they have the same semantic features because both are negative patterns. Furthermore, they have the same syntactic features because both patterns come after the same inverter. Finally, they are representing orders; however, the overall outcome is still different. These cases make the aforementioned modified algorithm is prone to error. Work on this issue is still ongoing.
4. **Implicit Negation:** The sentiment of a negated pattern can be reversed by a dependent clause. Consider the phrase: **ما به عيب سوى عبادة الاصنام** (he would have flawless if he did not worship statues). In other words, the pattern “flawless” is implicitly negated because the overall sentence implies that “he has flaws”. The first segment of the sentence

is positive, but adding neutral phrases converts it to negative. MSA can easily detect these cases because few words imply exclusion, i.e., words used to show how something would change under a certain condition. In English, we can say, “It would have been flawless if it was red”, which means that an object has flaws and can become flawless if a particular condition is met. In MSA three common words are used for this purpose *لو، لولا، سوى، إلا، إنما*, the example below is illustrated:

- *لو أنه سمع النصيحة، لكان من السعداء* (if he had listened to the advice, he would have been happy now)
- *لولا التعب، لكان العمل ممتعا* (work would have been fun if we don't get tired)
- *لن ينجح سوى المجتهد* (only the hard worker will succeed)
- *لن ينجح إلا المجتهد* (only the hard worker will succeed)
- *إنما طلبت الأخضر* (I only asked for the green one)

However, not all these cases are applicable in DA, especially where spelling variants of these words can be used, use them without proper tokenization, or the same pattern can be used to express different meaning, such as *شو يدك بهاشغلة ولو*, in this phrase, the word *لو* is not used for exclusion, such cases are challenging.

5. **Neutral Targets:** In addition to their ability to flip the polarity of sentimental targets, inverters may act on neutral targets to produce a sentimental phrase. Consider the example *لا صوت ولا صورة* (no voice, no picture) the two patterns voice and picture are neutral, but when the inverter *لا* (La), comes before it, negative sentiment will be given to the negated phrase. It is complex to detect these cases because usually when a neutral target is negated, the output is a neutral phrase. The neutral patterns mentioned earlier cannot be used by themselves to express a positive sentiment, i.e., saying *صوت وصورة* is not used a positive phrase. Another example would be *ناقص* (missing) is considered neutral in DA since it does not express a sentiment as a stand-alone pattern; however, when preceded by the inverter, the phrase will express as a negative sentiment. Moreover, the same pattern *ناقص* can be used as a negative pattern in MSA as in *ناقص العقل*, and if an inverter comes before it in MSA, the overall sentiment becomes positive.

4. Results

After analyzing the corpora and the exceptional negations described above, an approach to addressing them was developed to address the challenges that negations invoke.

The approach developed was to augment the positive and negative patterns using the word operated on by an inverter. The result parent sets are termed Inversion Enhanced Sets. Formally, given a clause with an inverter “...inv trgt...” in such cases, if trgt was deemed positive, it was added to the positive patterns and to the negative patterns if negative. Extracted patterns that contain inverters were filtered. Then, the polarity of targets was checked then added to the set they belong to it. For example, if a negated phrase has a positive target, it will be added to the positive patterns set.

This process has two significant advantages. First, splitting the negated phrases will increase the number of patterns without losing the original one. For example, when we split the pattern “beautiful” from “not beautiful”, we gained one positive pattern. At the same time, we did not lose “not beautiful” because the algorithm checks if the sentimental pattern comes after an inverter before deciding the sentiment of the pattern, i.e., beautiful, will be considered negative if it comes after an inverter.

Moreover, the negated patterns in SA (or SN) may appear preceded by the different inverter in CN (or CA). This led to higher recall since the algorithm considers that the same pattern may appear preceded by different inverters.

When there is negation, the target cannot have sentiment by itself. For example, لا صوت ولا صورة (translated as there is no sound and there is no picture) is inverting two neutral targets by negation. Thus, when negation is applied, neutral nouns become negative and mean that someone is ugly and cannot sing. These patterns were not split.

The classification algorithm was modified to consider negation: when a positive (or negative) pattern is found, the previous word was checked to see if it is an inverter, and then if an inverter was found, the polarity of the pattern was inverted. The pseudocode, in Algorithm 1, below represents a modified algorithm that considers negation.

Algorithm 1: The pseudocode of the negation algorithm.

```

Result: Pattern polarity inverted for each inverter.
Input: The Unclassified Corpus P;
Output: The Classified Corpus P;
SpamCount = 0; NegCount = 0; PosCount = 0;
while Post P do
    Search P for Patterns;
    if P has a Spam Patter then
        | SpamCount++;
    else
        if P has Positive and Negative Patterns then
            | if Positive Pattern is preceded by an inverter then
            | | NegCount++;
            else
            | | PosCount++;
            end
        else
            if Negative Pattern is preceded by an inverter then
            | | PosCount++;
            else
            | | NegCount++;
            end
        end
    if P has Positive Patterns then
        | if Positive Pattern is preceded by an inverter then
        | | NegCount++;
        else
        | | PosCount++;
        end
    else
        if Negative Pattern is preceded by an inverter then
        | | PosCount++;
        else
        | | NegCount++;
        end
    end
    if P has Negative Patterns then
        | if Negative Pattern is preceded by an inverter then
        | | PosCount++;
        else
        | | NegCount++;
        end
    else
        if Positive Pattern is preceded by an inverter then
        | | NegCount++;
        else
        | | PosCount++;
        end
    end
end
end

```

Algorithm 1: Cont.

```

if SpamCount > 0 then
  | AutoClass = "Spam";
else
  | if PosCount > 0 && NegCount > 0 then
  | | AutoClass = "Dual";
  | else
  | | if PostCount > 0 then
  | | | AutoClass = "Pos";
  | | else
  | | | if NegCount > 0 then
  | | | | AutoClass = "Neg";
  | | | else
  | | | | AutoClass = "Neu";
  | |
  |
if AutoClass == ManualClass then
  | Classification = True;
else
  |

```

To study the effect of inverters from a different perspective, we ran the algorithm above on CA using SN and on CN using SA. To measure the performance of our approach, we use the classical F1-score that exploits precision and recall. If there is an exact match between automatic and manual classification, then a post is correctly classified, such as:

- tp: correct result and classified as being correct.
- fn: correct result missed by the classifier.
- fp: negative result that is incorrectly classified by the classifier as positive.

The following equations are used to calculate the precision, recall, and F1-score:

$$precision = \frac{tp}{tp + fn} \quad (1)$$

$$recall = \frac{tp}{tp + fp} \quad (2)$$

$$F1 - score = \frac{2 \times precision}{precision + recall} \quad (3)$$

Table 3 represents the results achieved before and after treating negation.

Table 3. Effect of negation on classification F1-score.

	Before Treating Negation	After Treating Negation
CA—SN	0.73	0.93
CN—SA	0.73	0.73

5. Discussion

There are two main reasons behind the significant increase in performance of CA-SN and no improvement in CN-SA:

1. **Frequency of negation in posts:** In CA-SN, the number of positive and negative targets affected by inverters is 229% more than those of CN-SA, this means that as the frequency of presence of inverters in posts increases, the number of incorrectly classified posts will increase and therefore a solution like ours will be needed to resolve such cases. In other words, the performance improvement will be appreciated when there is a significant part of posts containing inverters.
2. **Increase in the number of patterns:** as mentioned earlier, splitting inverters from their targets led to increasing the set of patterns without losing the original one. The number of patterns in SN increased by 4% while that of SA decreased by 3%. The increase in the number of patterns in SN led to higher performance when classifying SN. In con-

trast, a decrease in the number of patterns of SA did not lead to a decrease in performance since, as we mentioned earlier, when splitting negated phrase and adding the target to its corresponding set, the algorithm will check the presence of any inverter before each pattern, and therefore the effect of the original negated phrase will be conserved. However, the decrease in the number of patterns was because the newly found patterns (after splitting) were already part of SA, so the set lost the negated phrases as a pattern and did not gain new ones. This did not affect performance.

When talking about the Arabic language, there is no standard structure, especially with negation. The Arabic language has various structures for different parts of speech. Until now, the research in Arabic natural language processing is considered in its beginning stage. More work should be done. It is normal to have rule-based methods that are performing better than machine learning when dealing with negation. This depends on the feature extraction process and the size of the available corpus, and the variety of valence shifters. In this corpus, only negation was available. Negation was not discussed in many research that deal with machine learning.

Although this research was applied to a corpus of posts collected from Facebook, the same approach can also be applied to any other social network. Twitter has an API that allows researchers to collect data in the form of tweets. Thus, the same work can be applied to the tweets corpus.

The proposed work in this paper can add value to recommender systems and building user-interest profiles as in [25]. The authors in [25] were trying to discover the popularity of posts on Instagram. Negation can add value to this work and help enhance the results of popularity ranking. Before considering the popularity, Instagram posts should be classified. Negation can help classification techniques and allows the model they suggest to understand the semantics of the popular posts. Popular post can then be predicted once the post is posted on Instagram. Instagram can expect whether the post will go viral or not. In [26], the authors classify tweets into positive and negative tweets to use the classification in building a user-interest profile. The authors do not consider negation in their work. Thus, many lost tweets were either incorrectly classified or unclassified. Negation and valence shifters can enhance the results of the work done in [26].

6. Conclusions

Negating Words have the tendency to invert the polarity of a word and to negate (or express the absence) of the term that comes after them. In this paper, we used Facebook posts to study the effect of negating words on sentiment polarity of a post. We highlighted few challenges that hinder using negating terms directly as classification features such as tokenization, odd negation, etc. We also show that treating negation may lead to increase in classification performance as we as their usage in patterns discovery. Highest improvement achieved after treating negation was 20%.

This work can be more expanded in the future work that can include resolving the following three issues:

1. **Odd negation:** More investigation is needed to find rules that can determine which patterns will flip their polarity when preceded by inverters.
2. **Fake inverters:** There are cases when applying the inverter without considering the nature of the targets will lead to incorrect classification. Such cases should be excluded when inverting the polarity of targets preceded by inverters.
3. **Complex Negation:** The Arabic language is complex by its nature, so there are many cases where negation is not trivially treated. Such cases include using exclusion words mentioned earlier.

This work can also be applied to any corpus of data collected from any social network. Twitter has an API that allows researchers to collect posts. So, This work will also be applied to tweets from Twitter, and a comparison will be made with this work. Another possible extension of this work is to apply it to different languages, such as the French language, because it is the closest language to Arabic.

Author Contributions: Literature, S.K., M.I. and C.R.; Conceptualization, S.K., M.I. and C.R.; Analysis, S.K., M.I. and C.R.; Writing, S.K., M.I. and C.R. All authors have read and agreed to the published version of the manuscript

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data is available on this link: <https://shurda.shu.ac.uk/id/eprint/59/> (access date 1 January 2021).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Kumar, K.N.; Uma, V. Intelligent sentiment-based lexicon for context-aware sentiment analysis: Optimized neural network for sentiment classification on social media. *J. Supercomput.* **2021**, 1–25, doi:10.1007/s11227-021-03709-4.
2. Jindal, K.; Aron, R. A systematic study of sentiment analysis for social media data. *Mater. Today Proc.* **2021**, in press.
3. Abd El-Jawad, M.H.; Hodhod, R.; Omar, Y.M.K. Sentiment analysis of social media networks using machine learning. In Proceedings of the 14th International Computer Engineering Conference (ICENCO), Cairo, Egypt, 29–30 December 2018; pp. 174–176.
4. Sangam S.; Shinde S. Sentiment classification of social media reviews using an ensemble classifier. *Indones. J. Electr. Eng. Comput. Sci. IJEECS* **2019**, *16*, 355–363.
5. Oueslati, O.; Cambria, E.; HajHmida, M.B.; Ounelli H. A review of sentiment analysis research in Arabic language. *Future Gener. Comput. Syst.* **2020**, *112*, 408–430.
6. Hamouda, A.A.; El-taher, F.Z. Sentiment analyzer for arabic comments system. *Int. J. Adv. Comput. Sci. Appl.* **2013**, *4*, doi:10.14569/IJACSA.2013.040317.
7. Gamal, D.; Alfonse, M.; El-Horbaty E.S.; Salem A.B. Twitter benchmark dataset for Arabic sentiment analysis. *Int. J. Mod. Educ. Comput. Sci.* **2019**, *11*, 33.
8. Yu L.S.; Al Baadani S. A sentiment analysis approach based on Arabic social media platforms. *DEStech Trans. Eng. Technol. Res. ICMEIT* **2018**, doi:10.12783/dtetr/icmeit2018/23467.
9. Al-Laith, A.; Shahbaz, M. Tracking sentiment towards news entities from Arabic news on social media. *Future Gener. Comput. Syst.* **2021**, *118*, 467–484.
10. Ramanathan, V.; Meyyappan, T. Twitter text mining for sentiment analysis on people's feedback about Oman tourism. In Proceedings of the 4th MEC International Conference on Big Data and Smart City (ICBDSC), Muscat, Oman, 15–16 January 2019; pp. 1–5.
11. Assiri, A.; Emam, A.; Al-Dossari, H. Towards enhancement of a lexicon-based approach for Saudi dialect sentiment analysis. *J. Inf. Sci.* **2018**, *44*, 184–202.
12. Itani, M.; Hamandi, L.; Zantout, R.; Elkabani, I. Classifying sentiment in Arabic social networks: Naïve Search versus Naïve Bayes. In Proceedings of the 2nd International Conference on Advances in Computational Tools for Engineering Applications (ACTEA 2012), Beirut, Lebanon, 12–15 December 2012.
13. Polanyi, L.; Zaenen, A. Contextual valence shifters. In *Computing Attitude and Affect in Text: Theory and Applications*; Springer: Dordrecht, The Netherlands, 2005; pp. 1–10.
14. Zahidi, Y.; Younoussi, Y.E.; Yassine A.A. Arabic sentiment analysis problems and challenges. In Proceedings of the X International Conference on Virtual Campus (JICV), Tetouan, Morocco, 3–5 December 2020; pp. 1–4.
15. El-Halees, A. Arabic opinion mining using combined classification approach. In Proceedings of the International Arab Conference on Information Technology (ACIT), Riyadh, Saudi Arabia, 11–14 December 2011.
16. Jia, L.; Yu, C.; Meng, W. The effect of negation on sentiment analysis and retrieval effectiveness. In Proceedings of the 8th International Conference on Information and Knowledge Management, Hong Kong, China, 10–11 July 2009.
17. Al-Khawaldeh, F.T. A study of the effect of resolving negation and sentiment analysis in recognizing text entailment for Arabic. *arXiv* **2019**, arXiv:1907.03871.
18. Maynard, D.; Bontcheva, K.; Rout, D. Challenges in developing opinion mining tools for social media. In Proceedings of the @NLP Can u Tag# Usergeneratedcontent, Istanbul, Turkey, 26 May 2012; pp. 15–22.
19. Taboada, M.; Brooke, J.; Tofiloski, M.; Voll, K.; Stede, M. Lexicon-based methods for sentiment analysis. *Comput. Linguist.* **2011**, *37*, 267–307.
20. Farra, N.; Challita, E.; Abou Assi, R.; Hajj, H. Sentence-level and document-level sentiment mining for arabic texts. In Proceedings of the IEEE international Conference on Data Mining Workshops, Sydney, NSW, Australia, 13–17 December 2010; pp. 1114–1119.
21. Shoukry, A.M. Arabic Sentence Level Sentiment Analysis. Master's Thesis, The American University in Cairo, Cairo, Egypt, 2013.
22. Itani, M.; Chris R.; Al-Khayatt, S. Corpora for sentiment analysis of Arabic text in social media. In Proceedings of the 8th International Conference on Information and Communication Systems (ICICS), Irbid, Jordan, 4–6 April 2017.
23. Al Arabiya. Available online: <http://www.facebook.com/AlArabiya> (accessed on 1 January 2012).
24. MBC The Voice. Available online: <http://www.facebook.com/MBCTheVoice> (accessed on 1 January 2012).

-
25. Carta, S.; Podda, A.S.; Recupero, D.R.; Saia, R.; Usai, G. Popularity Prediction of Instagram Posts. *Information* **2020**, *11*, 453.
 26. Khattak, A.M.; Batool, R.; Satti, F.A.; Hussain, J.; Khan, W.A.; Khan, A.M.; Hayat, B. Tweets Classification and Sentiment Analysis for Personalized Tweets Recommendation. *Complexity* **2020**, *2020*, 8892552.