Computer Science and Information Systems 16(1):289–311 https://doi.org/10.2298/CSIS180122013L

Improving Sentiment Analysis for Twitter Data by Handling Negation Rules in the Serbian Language

Adela Ljajić[1] and Ulfeta Marovac[1]

[1]State University of Novi Pazar, Vuka Karadžića bb,
36300 Novi Pazar, Serbia
{acrnisanin, umarovac}@np.ac.rs

Abstract. The importance of determining sentiment for short text increases with the rise in the number of comments on social networks. The presence of negation in these texts affects their sentiment, because it has a greater range of action in proportion to the length of the text. In this paper, we examine how the treatment of negation impacts the sentiment of tweets in the Serbian language. The grammatical rules that influence the change of polarity are processed. We performed an analysis of the effect of the negation treatment on the overall process of sentiment analysis. A statistically significant relative improvement was obtained (up to 31.16% or up to 2.65%) when the negation was processed using our rules with the lexicon-based approach or machine learning methods. By applying machine learning methods, an accuracy of 68.84% was achieved on a set of positive, negative and neutral tweets, and an accuracy of as much as 91.13% when applied to the set of positive and negative tweets.

Keywords: Sentiment Analysis, Serbian Language, Twitter, Negation Detection, Negation Rules, Machine Learning.

1. Introduction

Sentiment Analysis (SA) belongs to one of the sub-fields of Natural Language Processing (NLP). NLP is a technology that deals with ubiquitous human language that appears on websites, product descriptions, newspaper articles, social media, and scientific articles, all in thousands of languages and linguistic variations. SA belongs to the field of language technology which is on the rise, and with the appearance of machine processing of natural language and machine learning algorithms for classification, greater progress in this area occurred. The opinions of people expressed and written in the given language are not easy to analyze and it is not easy to determine their sentiment, given the complexity of the linguistic expressions and the different way of expressing oneself in a particular language. Some other problems in SA are the ambiguity of words or syntagms, non-standard writing, use of slang, neologisms, segmentation problems, irony, metaphor, and negation. Regardless of the inability of a computer to automatically determine sentiment, under certain assumptions the complexity of such a task can be simplified. The simplest systems determine the sentiment based on the occurrence of a word in the positive and negative sentiment lexicon. A few more advanced systems take into account POS tags, apply the rules of negation and detect irony. Although the determination of polarity, as one of the tasks of

SA, is an area which has been studied to some extent, some aspects that are linguistically specific (negation, irony, metaphor) still pose challenges and areas where improvements are expected. In this paper, we deal with the processing of grammatical rules of negation in the Serbian language. The main contribution of this paper is to show whether the treatment of these rules of negation and their integration into the method of classification of the sentiment can improve the accuracy of the prediction of short texts - in our case, tweets. The rules that are processed represent a subset of grammatical rules for the processing of negation in the Serbian language, i.e. the rules that are most commonly encountered in this type of short text. The applied rules and their influence on the polarity classification will be tested by an unsupervised lexicon-based method and a supervised machine learning method. It will be shown that, by integrating the rules of negation, there is a statistically significant improvement in the application of the method for the prediction of sentiment.

The rest of the paper is organized as follows. We begin with a description of related work in SA with negation in Section 2. Section 3 contains information about the dataset structure, sentiment lexicon, negation signals lexicon and other lexicons used. Section 4 describes the proposed method. The experiment is included in Section 5. Section 6 contains an evaluation and analysis of the results. Finally, we conclude and present directions for future work in Section 7.

2. Related work

In recent years there has been great interest in the research done in the field of SA. When determining text sentiment, researchers usually assume that a speaker has some attitude or sentiment about the subject, object or person, that the sentiment has a fixed value (good or bad) and that the sentiment in the text is represented by a word or combination of a small number of words. However, in the literature, there is little discussion of what a “sentiment” or “opinion/attitude” actually is [1]. In some review papers [2,3], the authors give an overview of the different approaches of determining sentiment. Sentiment can be analyzed at the document level, at the sentence level, or at the aspect level. Since SA is used to determine the subjective attitude about a phenomenon (object, person), some systems for SA include, in addition to positive and negative, neutral (objective) attitudes. Therefore, SA mainly involves research within three classes: positive, negative, and neutral. In [4] Pang et al. define the baseline method for sentiment analysis based on the sentiment lexicon. Sentiment lexicons contain words that express a positive or negative sentiment. By counting words that carry sentiment and using them as features for the machine learning methods, the sentiment of the sentence is determined. The extension of this method depends on the type of text, as well as on the specificity of the language on which this method is applied. Most papers deal with SA in English. Multilingual SA was introduced in [34] where authors compare their own implementation of existing approaches. The emphasis of the work is on the methods used for the Serbian language, and we will deal with them in more detail later on. The popularity of Twitter grows with the number of tweets, so an increasing number of authors have decided to test sentiment of this type of short text. An overview of the methods for analyzing sentiment data in Twitter is given in [19]. Most papers use supervised learning methods to determine sentiment, although not an insignificant number of approaches provide analysis by methods of unsupervised or combined semi-supervised learning.

2.1. Sentiment Analysis in Serbian

SA in the Serbian language was mostly dealt with by authors in [23] and [21]. Mladenović et al. in [23] used morphological dictionary rules for inflection expansion to build sentiment lexicons and a stop word list that contains lemmas and inflectional forms of the lemma in order to reduce feature space for machine learning classification. They used the Maximum entropy classifier method for the prediction of text sentiment in the Serbian language using a set of newspaper articles for the application of a 10-fold CV. To confirm the obtained results, the method has been tested for two more special sets: a set of newspaper articles and a set of movie reviews. The authors used rich lexical resources and worked with two classes (positive and negative). They used the attribute reduction method for attributes (words) found in sentiment lexicons and mapped them to a new set of attributes. They received an accuracy for a 10-fold CV of up to 95.6%, of up to 78.3% for movie reviews and of up to 79, 2% for the news test set. In their work, negation is not particularly addressed.

Batanović et al. in [21] created the Serbian movie review dataset – SerbMR using a dataset balancing algorithm that minimizes the sample selection bias by eliminating irrelevant systematic differences between the sentiment classes. The authors gave a SA using various combinations of n-grams, stems, lemmatizers, and different types of normalization of the attribute. Finally, applying optimal attribute settings over the NBSVM classifier, they achieved an accuracy of up to 85.55% for two, and 62.69% for three classes. Some authors, for example in [20] examine the influence of 2 modes of morphological normalization of the text (using a stemmer and lemmatizer for the Serbian language) on the influence of the sentiment classification on a data set of movie reviews. Using a combination of unigrams and bigrams, they achieved a statistically significant improvement compared to the baseline method for two classes (accuracy 86.11%) and for three classes (accuracy 63.02%). A comparative overview of the method of analyzing sentiment for the Croatian language is given in [22]. Due to the great similarity of the Croatian and the Serbian language in the complex-morphological sense and in other characteristics, we can consider that the results achieved are similar to those in the Serbian language. The authors used the “word embedding” method and “string kernels” comparison over three sets of short texts (game reviews, tweets from a specific domain and general-topic tweets). They showed that the “word embedding” technique and “string kernel” technique achieved an improvement over the bag of words (BOW) baseline method.

2.2. Negation in Sentiment Analysis

The most commonly used method for processing negation is to add the suffix _NEG to a sentiment word that occurs in the negated scope and to treat the negation from the negation signal to the first punctuation mark [4, 5, 6], or to the first positive or negative sentiment word found [7]. Although most negation processing systems take into account all types of sentiment words, some take only adjectives [8] or adjectives and adverbs [9]. In [6], the authors examined the effect of the negation signal, as one form of valence shifters, on the change in the polarity of the word in the negated part. Most authors simply change the polarity of the negated word when processing negation (from n to -n and vice versa) [7, 10]. Sophisticated negation systems specifically treat sentiment words that appear in the scope of negation and show that negation with different intensities changes the polarities of the positive and negative sentiments of the word [11]. The authors who advocate this approach have shown in [12] that the production of such specific vocabulary for words that occur in the negation scope gives a significant improvement in the system for predicting sentiment. In addition to works dealing with negation using the negation signal, in [14] phenomena reminiscent of negation is processed because they change the polarity, but is not negation in the classical sense, as they do not contain words that are negation signals (no, not, did not ...). In [15], the authors suggested the detection of the scope of negation using the reinforcement learning method. Evaluation of the effects of the detected negation on the SA was done by measuring the correlation between the sentiment sentences and the corresponding “daily stock market return” as the measure of the prediction, and a correlation of 10.63% was obtained between the sentiment value and the stock market return. In [16], the authors detected negation signals and the scope of negation. They applied detected negation to the SA classifier for the entire set and the set containing negations. Their sophisticated method for predicting sentiment (when using negation treated attributes) in the whole set achieved an improvement in the accuracy of 0.01. Although most of the methods for evaluating the impact of negation on the sentiment of the text are based on supervised learning, the authors [10, 16] proposed to calculate the intensity of negation and its impact on polarity by applying its methods of calculating the intensity of negation and polarity of the text. In [10] they dealt with the analysis of negations in the Spanish language and their influence on SA. A corpus of tweets in Spanish was used along with an unsupervised method for predicting sentiment that, in addition to other resources, includes a processed negation. They obtained results that showed that by treating the negation, they achieved a significant improvement in accuracy in determining the polarity of tweet. In [18], the rules for identifying negation and calculating its intensity are described. The rules were created in order to improve the sentiment text analysis. The method that predicts sentiment is an unsupervised method that calculates the polarity and intensity of the words and phrases. They showed that there is a positive correlation between the polarity determined by the system and the one assigned by the 5 participants in the experiment.

It is difficult to say that there is a versatile, best classifier of text sentiment. Systems that achieve good results for long texts can be inappropriate for short ones. An example is the work done by Cerezo-Costas er al. [17], whose method used on a dataset containing sarcasm (Tweet Sarcasm 2014) took first place, while for the SemEval 2015 group (Analysis Task 10) it took 16th place. A review paper [13] provides an overview of most of these approaches in processing negation from the earliest papers published on the topic; represented ways of processing negation, the scope of negation, the selection of training attributes, and the boundaries in the negation model for SA.

For Serbian, there are no studies dealing with the detection of grammatical rules of negation in a text, and therefore the influence of such processing on SA. Batanović et al. applied a classical method of processing negation, by changing the polarity of words that appear after the negation signal [21]. By dealing with negation on a set of movie reviews, they achieved the following improvements in relation to the initial method [4]: for three classes they achieved the greatest improvement by means of the MNB method of 0.94% (accuracy of 54.72 to 55.66) marking two words after the negation; for two classes they achieved the greatest improvement with the SVM method of 0.66% (75.98 to 76.64) by marking only the first word after the negation.

3. Dataset and lexicons

This section presents the preparation of dataset and describes specialized lexicons for the SA and negation analysis. SA requires the use of a lexicon with the annotated sentiment of a word (sentiment lexicon). Negation signals, negative quantifiers, and particle intensifiers are needed to process negation as an indispensable part of the SA. The use of a stop words lexicon is widespread in text analysis and must be adapted to negation analysis so as not to lose the negation signals.

3.1. Dataset preparation and structure

So far, Serbian language sentiment analyses have been carried out on datasets of long texts (newspaper articles, movie reviews). In short texts, especially tweets, the sentiment is expressed more concisely than in long texts, but it is more difficult to determine, given the small number of words, ambiguity, informal writing, and often the presence of irony and sarcasm that significantly distort the classification. There is no annotated corpus of short texts for the analysis of sentiment in the Serbian language. For the purposes of this research, we collected 9059 tweets and annotated them. The tweets were collected using twitter streaming API from profiles of 15 official mass medias, 10 musicians, 5 public figures from the entertainment world, 5 athletes and 7 actors. The sentiments are grouped, according to sentiment, into three classes: negative, neutral and positive. The dataset contains 4334 negative, 2507 neutral and 822 positive tweets. In addition to positive, negative and neutral, irrelevant tweets were detected. These are tweets that do not contain clear information, contain only a link, or contain an informal phrase that cannot be attributed even to a neutral tweet. Irrelevant tweets were discarded. The data was annotated by 3 independent people, two men, and one woman. One person holds a master’s degree in electrical engineering, one is doctor of medicine, and one is a student of Serbian language and literature. Only tweets which all three annotators agreed on (7663 tweets) were taken into account. In 15.41% of the cases (1396 tweets) , the annotators did not agree. An imbalanced number of tweets by class are expected because the topic of the tweets we have collected is more negative than positive. Regardless of the variety of media and personalities, tweets that carry a negative sentiment prevail, while the number of tweets with a positive sentiment is far smaller.

3.2. Lexicons used

Sentiment lexicon

The most general sentiment lexicons are for the English language. The “General Inquirer” sentiment lexicon dates back to 1966 [27] and consists of 1915 positive and 2291 negative words. LIWC [28] contains 2300 words and 70 classes, such as negative emotions (hatred, anger, problematic issues...) and positive emotions (love, beauty, kindness ...). The MPQA lexicon [29] contains 6885 words from 8221 lemmas, of which 2718 are positive and 4912 are negative. The Bing Lui [30] Opinion Lexicon consists of 6786 words, of which 2006 are positive and 4783 negative. The SentiWordNet [31] is a lexicon where every word has an associated coefficient that determines how positive, negative, or objective it is. To the best of our knowledge, there is no publicly available sentiment lexicon for Serbian. For the needs of our analysis, a sentiment lexicon was created; the translation of the Opinion Lexicon [30] was taken as the starting version of our sentiment lexicon.

Negation signals lexicon

We named the words involved in creating a syntactic negation “negation signals”. The negation signal lexicon is derived from a list of all the word forms in the Serbian language that are involved in creating a syntactic negation. Negation terms (“ne” (no) and “ni” (no)) are included in the negation signal lexicon. In the Serbian language, the verbs “biti” (to be), “hteti” (to want) and “imati” (to have) have an irregular negation form, and the negation is written merged with the verb. The forms of these verbs are: “ne jesam”(to not be) in the present tense: “nisam” (I am not), “nisi” (you are not), “nije” (he/she/it is not), “nismo” (we are not), “niste” (you are not), “nisu” (they are not);“ne biti” (to not be) in the imperative mood: “nemoj” (do not), “nemojmo” (let’s not), “nemojte” (do not); “ne hteti” (do not want) in the present tense: “neću” (I do not want), “nećeš” (you do not want), “neće” (he does not want), “nećemo” (we do not want), “nećete” (you do not want), “neće” (they do not want); “ne imati” (do not have) in the present tense: “nemam” (I do not have), “nemaš” (you do not have), “nema” (he does not have), “nemamo” (we do not have), “nemate” (you do not have), “nemaju” (they do not have). These verbs in the negative form (negation) mingle with negatives and become special negative signals. All forms of negations of the verbs “jesam” (to be), “hteti” (to want) and “imati” (to have) are also included in negation signals lexicon. The total number of negation signals in this lexicon is 25.

Negative quantifiers lexicon

The negative quantifiers in the rules of negation in the Serbian language play a completely different role than negation signals, so it would be wrong to equate them. Thus, we have singled out negative quantifiers in a separate lexicon. Universal negative quantifiers are morphological “ni-” (no) negations of negative pronouns ((niko, ništa, nikakav... (nobody, nothing, no ...)), negative adverbs (nikad, nigde, nikako... (never, nowhere, noway ...)) and a small number of negative pronominal adverbials (nimalo, nijedanput... (not at all, never ...)) [32]. The negative quantifiers always stand with negation signal (almost always in front of it) and confirm the negation that lies with them. For example, in the sentence “Nikad nije lagao.” (“He has never lied.”), “Nikad” (“Never”) is a negative quantifier that agrees with the negation signal “nije” (“is not”) that changes the polarity of the word “lagao” (“lied”) from negative to positive. The agreement of negative quantifiers with negation means either the confirmation or intensification of the negating part.

Particle Intensifiers Lexicon

In the Serbian language, there are types of words that are called particles, and their scope is such that they express their relation only to the content of one term in a sentence [32]. Particles that affect negation are considered by intensifying or extending the scope of negation validity. We used particle intensifiers of negation by extending the scope of the validity of the negation even after the punctuation mark if the particle intensifier is behind it. The particle intensifier lexicon contains: “ni” (“no”), “nit” (“neither”) and “niti” (“neither”).

3.3. Stop Words Lexicon

The role of stop words is to remove those words (potential attributes) that do not bring meaning to the text. A stop words lexicon can be general (universal) or specific to the field. Specific lexicons are created mainly by extracting words that have the highest TF or the smallest TF/IDF. The authors in [23] experimentally confirmed that a universal lexicon gives approximately the same results as a specific lexicon. In the paper, we used the general stop words prepared in [24], which mostly consist of adverbs, prepositions, conjunctions. The stop word lexicon does not contain negation signals and words that participate in any rule of syntax negation.

4. The proposed method

4.1. Motivation

In sentences that have a certain construction in which a negation occurs, there is sometimes a neutralization of negation, sometimes there is an intensification and sometimes a change in polarity. The different ways in which negation can act are introduced in [25]. The idea is to process the rules determining when the negation changes the polarity of the word, and in which part of the sentence, and when the negation is actually a pseudo-negation and does not simply change the polarity but expresses the affirmative. In this paper, we want to show that by taking into account the specific rules of negation, by detecting the scope to which the signals of negation act, or by ignoring the negation signal, it is possible to improve the quality of the classification of tweets based on sentiment.

4.2. Normalization

Languages that are morphologically rich and highly inflected require the application of linguistic rules for the processing and classification of texts. In order for a text to be classified, it must first be normalized. Each text classification requires a specific normalization of the text. Normalization of tweets was done using:

- the tokenizer adapted to the specific structure of tweets,
- converting text from one alphabet to another – to Latin
- removing the stop word
- a Serbian language stemmer.

Tokenization and parsing of tweets

Tokenization is the first task in the preprocessing (normalization) of tweets. Tokenization was done by tokenizer built into the Python programming language using the *nltk.tokenize* module. The RegexpTokenizer was used to split the tweet into tokens and to process dates, various number formats and hashtags. Terms used by Twitter which are not relevant to our analysis (via, RT ...) were removed using the *re* python module and the regular expression. Since the Serbian language has two official alphabets (Cyrillic and Latin), it was necessary to process tweets written in Cyrillic and Latin, but separately. In order to avoid duplicate analyses, tweets written in Cyrillic were converted into Latin using the *cyrtranslit 0.4* python package. Following conversion into the Latin alphabet, stemming was performed.

Stemming

In the study we used a stemmer that encodes special Serbian Latin characters 'č', 'ć', 'š', 'đ', 'ž' to 'cx', 'cy', 'sx', 'dx', 'zx' [26]. It creates stems longer than 2 characters and has a dictionary of irregular verbs and their flections. Other words are stemmed using generalized stemmer rules. The tasks that the stemmer does not deal with are incorrect flections, sound changes, and short words. In addition to the above deficiencies, the accuracy of the stemmer is 90%, so we decided that it is sufficient to use it in the normalization of our text. An accuracy of 90% was obtained by machine stemming, after which the human read the stems and compared them. The main problem was noted in the words that have sound changes. The rules are designed to cover as many words as possible, and as sound changes occur in a fewer number of words, it is not possible to make an algorithmic rule that supports both words with a sound change and words without it. The other problem that came up with short words consisted of 3 or 4 characters, and which after stemming had only two, and the original meaning was in many cases difficult to extract. Adjusting the stem for the experiment was possible because the code is publicly available in the Python programming language. We tested stemming from a stem longer than 2 letters, irrespective of the length of the word, and stemming in a stem longer than 2 letters, but only for words longer than 4 letters. The experiment showed that stemming in a stem longer than 2 letters, but only for words longer than 4 letters, gives a better result. The Serbian language is morphologically rich; therefore, the use of a stemmer is expected to improve the classification of sentiment. By sticking together, different forms of the same word are reduced to the same stem, so the number of occurrences of that stems increases. The authors in [21] showed that using a stem on their data set increases the average number of occurrences of each word by approximately 50% and decreases the size of sentiment lexicons for each class by approximately 30-35%. After stemming, our dataset is ready to apply the rules for processing negation.

4.3. Negation rules in the Serbian language

Negation is an area studied both by logic and linguistics. The linguistic analysis of negation must include the logic rules of negation [32]. The question to be answered in the negation is to determine the scope to which logical negation will be applied in accordance with linguistic rules. The Serbian language generally has complex rules. Negation can be lexical, syntactic and morphological, depending on the type of language unit in which it appears. Depending on the range of the unit whose content is negated, negation can be partial (the content of the part of the sentence is negated) or total (the entire sentence content is negated).

Morphological negation is accomplished by using prefixes such as “ne-” (non-), “bez-” (no-), “ni-” (not-), “a-” (a-), “dis-” (dis-) and “in-” (in-). Morphological negation is exhausted only at the lexical level. It has an effect on sentence negation only if the lexemes appear in the role of sentence members (as negative words, i.e. negative lexemes). The morphological negation in this paper was processed in a way that negated lexemes were inserted into the sentiment lexicon. Lexical negation is related to the use of a word whose meaning has a negative component (“sumnja” (doubt) and “nedostatak” (a lack of)). Lexical negation was not processed in this paper.

Syntactic (sentence) negation is accomplished by negation signals “ne” (not) and “ni” (not). The verbs which have an irregular negation form behave like negation signals. In this paper, we analyzed the syntactic (sentence) negation, i.e. the effect of the negation signal on a part of sentence or on the whole sentence.

The created negation rules

The rules of syntactic negation in the Serbian language are shown in [32]. The author presents the large numbers of negation and noted the rules that rarely appear. By analyzing the dataset of 3000 unannotated tweets, the rules of negation that most commonly appear in tweets were processed. For the initial scope of the negation, a set of words was taken from the first negation signal to the first punctuation mark. In some cases, it was necessary to change the scope of the negation. The following rules of negation improve the detection of the scope to which the negation relates and also improves overall SA. The rules of negation are:

1. The treatment of the negation type “Nije samo …. nego/već” (Not only .... but/instead) – the word that is the negation signal, which stands in front of the word “only”, is omitted. The scope of the negation is empty.
2. The treatment of the negation type “Nije…. nego/već” (Not .... but/instead) – in this case, the scope of the negation ends with the word “nego/već” (but/ instead). The negated part of the sentence has a smaller impact on SA than the following part of the sentence, so it should be omitted from SA.
3. The treatment of the negation in the questions of the type “Zar nije lepo?” (“Is it not nice?”) - in this case, the negation is neutralized because the sentiment after the negation signal does not change the polarity of the sentiment words. The scope of the negation is empty.
4. Treatment of the negation in the presence of an intensifier - If the intensifier appears after the negated word, then the next part of the sentence is negated, too. The scope of the negation is extended even after intensifier to the next punctuation mark.
5. Intensification of negation by negative quantifiers - agreement between negative quantifiers and the negation signal. Negative quantifiers increase the intensity of the negation.

The rules are applied in the order they are given. The correct order of the application of the rules is important because the Rule1 is, in fact, a pseudo-negation (the negation signal does not affect the following part which remains affirmative), and it excludes the application of the Rule2. Rule3 refers to a special question type that contains the negation signal preceded by the word “zar” (isn’t, aren't) or the word “jel” (isn’t, aren't). “Jel” is not grammatically correct, but we included it because it is used in informal speech and on Twitter for the same purpose as “zar”. These words belong to the group of negation neutralizers in questions. Rule4 is more general and can be applied to determine the scope of the negation independently of the language. Its use necessitates a change in the polarity of the negated word. Rule5 identifies negative quantifiers which agree with the negation (they confirm the negation) and in this way intensify it. When determining the sentiment of one tweet, the negation is processed in such a way that sentiment words from the scope of negation change the polarity and the number of negative quantifiers that affect the negated part assigned to that tweet. The rules are shown on the examples of tweets in Table 1.

Table 1. Tweets with the negation rules examples (highlighted - the initial scope of the negation, underlined - negation signal, bolded - the words that participate in the rule, framed - the final scope of negation)

Rule	The tweet containing the rule
1	**Ne** **samo** bezobrazluk, **već** i gaženje i ismevanje naroda. / **Not** **only** disgrace **but** also the violating and the mocking of the people.
2	Ovo se dešava kad se novac **ne** investira **nego** se troši. / This happens when money is **not** invested **but** consumed.
3	**Zar** vas nije sramota? / **Are** **not** you embarrassed?
4	Nije vredan, pametan **niti** lep. / He is not hardworking, smart **nor** nice.
5	***Nikad niko*** nije loše pričao o njemu. / ***Nobody*** ever (***never***) talked (did not talk) about him badly.

In the last rule (Rule 5), with the translation of Serbian into English, the double negation disappears because it does not exist in English (for clarification, word/words that would exist in the tweet if English had the same rule as Serbian are given in brackets).

4.4. Baseline methods

Method0 – This one performs SA without processing the negation, only based on the number of positive and negative words from the sentiment lexicon.

Method1 – This one processes negation by a simple change in polarity of the sentiment of the first word following the negation. The attributes are the following ones:

- the number of negation signals
- the number of positive words
- the number of negative words
- the number of negated positive words - one word after negation
- the number of negated negative words - one word after negation.

The negation can be processed simply by a change in the polarity of the word following the negation, for the words before a punctuation mark [4]. The authors in [21] analyzed the influence of negation on the words following it, and obtained the best results when there was a change in the polarity of the first word following the negation. This is why Method1 was taken as the baseline method of comparison which processes the negation, and which changes the polarity of the sentiment of only the first word following the negation.

4.5. The method

The method we propose (Method2) needs to determine the sentiment of the tweet, taking into account the rules of negation that we have processed. The method uses previously described resources such as the sentiment lexicon, the negation signal lexicon, the particle intensifier lexicon, and the lexicon of the negative quantifiers.

After normalizing the text, it was determined whether the text contains negation and if it does, which group of rules of negation the negation belongs to. After processing the rules of negation, attributes important for the determination of the sentiment were extracted. The attributes are the following ones:

- the number of negation signals
- the number of negative quantifiers
- the number of positive words
- the number of negative words
- the number of negated positive words - one word after negation
- the number of negated negative words - one word after negation
- the number of negated positive words – in the negation scope
- the number of negated negative words - in the negation scope

By using the selected attributes, prediction of the sentiments of tweets was performed using:

– an unsupervised classifier lexicon-based method (LBM)
– supervised machine learning classifiers (MLM).

In the experiment, these attributes are used for the application of LMB or MLM in relation to the two baseline methods (Fig. 1).

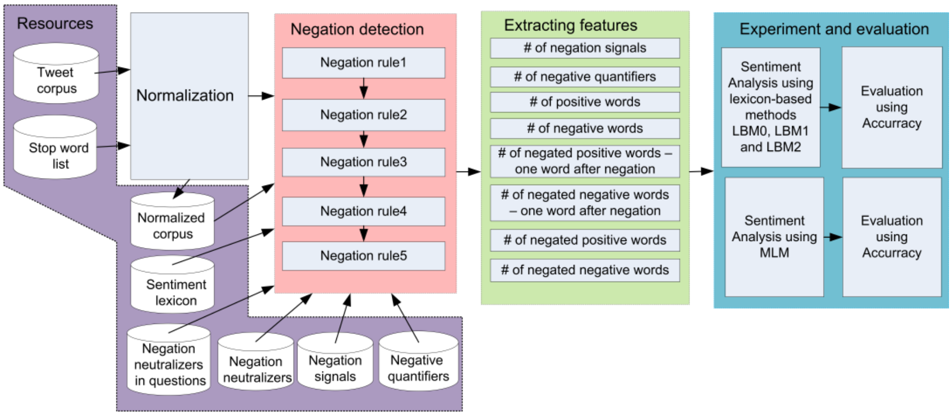

Fig 1. The architecture of the system

5. The experiment

Even though they are not structured texts, and are written with some informality, tweets can still be processed, and by implementing specific rules of negation the determination of the sentiment of these texts can be improved.

The aim of this paper was to determine the extent to which the processing of negation by means of the proposed rules can improve the SA of short texts, or in our case, selected tweets. Thus, a comparison was carried out of the obtained accuracy of the SA of the text, which processes negation by means of the proposed method (Method2) and the accuracy of the baseline method (Method0 and Method1).

For all three cases of negation processing by means of Method0, Method1, and the proposed Method 2, two techniques of sentiment prediction were implemented:

– LBM
– MLM.

In addition, both techniques of sentiment prediction (LBM, MLM) were performed on:

– the entire set (ALL)
– the set which includes only negation (OnlyNeg)
– the set which includes only negations which were included by the rules (OnlyRuleNeg).

The analysis which included the complete set, the set with negation signals and the one in which the rules were applied is the same methodology that was used by Jiménez-Zafra et al. [10]. Statistical analysis by means the unsupervised lexicon-based method was used on a set of three classes. A set of three classes (3-class) consists of positive, negative and neutral tweets, and a set of two classes (2-class) consists of positive and negative tweets. Methods of machine learning were also implemented on 3-class and 2-class. Table 2 shows the number of tweets for 2-class and 3-class for ALL and OnlyNeg and OnlyRuleNeg.

Table 2. The number of tweets in the entire set, the set with negation signals and the set of negations included by the rules

	3-class	2-class
ALL	7664	5156
OnlyNeg	3747	2726
OnlyRuleNeg	2313	1733

6. Evaluation and analysis of the results

6.1. Statistical analysis by means of the lexicon-based method

In order to justify the application of the method which includes the processed rules of negation, the quality of classification of the tweets by sentiment was analyzed by applying LBM. The polarity of the tweet was determined according to the following formula (1):

$$Tweet\ sentiment = \begin{cases} positive\ if\ sumPos > sumNeg \\ neutral\ if\ sumPos = sumNeg \\ negative\ if\ sumPos < sumNeg \end{cases} \quad (1)$$

This method was applied in three ways, depending on the selected method for processing the negation (Method0, Method1 and Method2). The first method (LBM0) classifies words from the sentiment lexicon only based on the positive and negative sentiment. The second method (LBM1) classifies tweets by including the negation of only the first word following the negation signal. The third method (LBM2) includes everything included in the second and first, in addition to the rules which we included for detecting and processing negation.

In LBM1 and LBM2, the number of positive and negative words is corrected by the number of words that changed polarity by processing the negation. The analysis was performed for all three methods and in three cases: for the whole set (ALL), for the set within which negation occurs at least once (OnlyNeg), and for the set with a negation which is included by the rules which we processed (OnlyRuleNeg). An analysis was given in the case of 3-class. The absolute improvements and relative improvements of

LBM1 and LBM2 were calculated in relation to LBM0 (baseline) and are calculated according to formula (1) and (2), respectively.

$$AbsChange\ in\ accuracy = \ compared\ method - baseline \tag{2}$$

$$RelChange\ in\ accuracy = \frac{compared\ method - baseline}{baseline} * 100 \tag{3}$$

In the sequel, the improvement will be considered as the relative improvement. The results are shown in Table 3.

Table 3. The result of applying SA on different datasets using lexicon-based methods

		LBM0	LBM1	LBM2
ALL	Accuracy	48.57%	51.07%	53.73%
	Improv.		2.50%	5.16%
	Rel. improv.		5.15%	10.62%
OnlyNeg	Accuracy	39.66%	44.76%	50.23%
	Improv.		5.09%	10.56%
	Rel. improv.		12.86%	26.63%
OnlyRuleNeg	Accuracy	38.86%	46.89%	50.97%
	Improv.		8.03%	12.11%
	Rel. improv.		20.66%	31.16%

Table 4.The statistical analysis

Data set	Analyzed method	Compared method	P(3 classes)
ALL	LBM2	LBM0	<0.0001
ALL	LBM2	LBM1	<0.0001
OnlyNeg	LBM2	LBM0	<0.0001
OnlyNeg	LBM2	LBM1	<0.0001
OnlyRuleNeg	LBM2	LBM0	<0.0001
OnlyRuleNeg	LBM2	LBM1	<0.0001

Table 3 indicates a significant improvement following the application of the method which includes our rules of negation. These results are encouraging and justify the application of the MLM for training and predicting sentiment, which will be shown in subsection 6.4. What is of great significance is the large improvement in the method LBM2 for the analysis of the set which consists only of negation, and even more importantly, when the set analyzed is the one which contains only negations included by the rules we have processed. Greater accuracy was expected and achieved for the entire set (ALL) than it was for the set which consists only of negations (OnlyNeg) and the set which consists only of the processed negations (OnlyRuleNeg). This is an indicator of the bad influence of the presence of negation on the prediction of sentiment. However, the LBM1 method (the baseline method for processing negation), and especially our proposed LBM2 method which includes the processed negation, provide a greater improvement for the set OnlyNeg and OnlyRuleNeg.

6.2. Statistical justification of the results

In order to verify the statistical significance of the obtained improvement in the classification, we will test the hypothesis that the method of text classification based on sentiment is significantly better at classifying text if the LBM2 rules of negation are used compared to the methods LBM0 and LBM1. In order to prove it, we will apply the MC Nemar test. When applying this method, it is necessary to construct a 2x2 matrix which consists of the ratio of accurately and inaccurately classified tweets by implementing two different methods. In all these cases, by comparing the results obtained by the method LBM2 and the methods LBM1 and LBM0, a statistically significant improvement was determined (Table 4).

6.3. Testing the influence of various means of attribute selection on the accuracy of predicting sentiment

The previous chapter contains a lexicon-based classification of tweets by means of three methods, and it was proven that the LBM2 method creates a significant improvement. The next step is for the same set of attributes used in the lexicon-based methods to be used in the application of the ML method. Since ML methods do not work well with a small number of attributes, especially numerical attributes, the baseline set of attributes should be expanded. To the set of attributes listed in subsection 4.5, we added more attributes. Additional attributes (word attributes), were obtained by a transformation of the normalized text into a vector representation of words. Text to word vector transformation was performed in Weka software using the StringToWordVector method for transformation to word attributes and the AttributeSelection method for selecting attributes.

In order to test the influence of the word attributes, we used three methods of ML: Naïve Bayes, Logistic regression and SVM. ML methods were applied only at word attributes on a 3-class dataset. The reason behind the selection of these three methods is that they mutually differ quite significantly and have proven to provide good results in the classification of texts and the determination of sentiment. For each case of text transformation into vector representation of words, we also performed a reduction of attributes by the application of the technique of information gain. Information entropy (gain) takes values ranging from 0 to 1 depending on how much information the attribute brings. All of the attributes which have a value greater than 0 were used, that is, all of the ones which carry any kind of information will be found on the list of attributes. The results of the influence of various means of transforming text into vector representation of words and the selection (reduction) of attributes from the text on the accuracy of the prediction of sentiment are shown in the Table 5. A 5-fold cross validation was carried out by applying the Naïve Bayes, Logistic and SVM methods. A 5-fold cross-validation was performed because a large number of attributes were used, and a 10-fold would be time and memory limited.

From Table 5 we can conclude that the attributes used for training which make the greatest contribution are unigrams (U), especially in the case when only the presence or absence in the tweet is being detected (the greatest improvement to the NB-row 1 in Table 5). If the vector representation of words which contains the presence or absence of a word is replaced by the number of occurrences of a word, the accuracy of the prediction decreases, which can be explained by the fact that the text of the tweet is rather short and that the words rarely repeat several times, both in the same tweet and in the entire group. In addition, the normalization of the IDF brings no improvement – quite the contrary; the reason is the same as the one for the application of the TF normalization.

Table 5. Accuracy of the ML method in the determination of sentiment by transforming text into vector representations of words

	Feature type	Filter applied	# of features	NB	Log	SVM
1	U	word presence	408	57.54%	64.67%	63.26%
2	U	word count	404	56.36%	64.00%	62.83%
3	U	TF	404	54.54%	64.21%	63.01%
4	U	TF/IDF	404	54.54%	64.21%	63.00%
5	U	TF+ Normalized tweet length	265	39.26%	62.31%	60.92%
6	U	TF/IDF+ Normalized tweet length	255	28.77%	62.04%	60.82%
7	U+B+T	word presence	1023	57.40%	64.87%	63.88%
8	U+B+T	word count	1015	55.94%	64.52%	63.61%
9	U+B	word presence	756	57.49%	65.30%	64.31%
10	U+B	word count	750	56.21%	65.16%	63.97%

The normalization of the length of the tweet has a negative effect on the prediction of all three methods. The negative influence of normalization of the length of the tweet is especially clear in the application of the Naïve Bayes method, as a result of the assumption on the independence of the variables, which contributes to poor results. Unigrams (U) combined with bigrams (B) and trigrams (T) offer less improvement compared to unigrams combined with bigrams. Unigrams combined with bigrams (highlighted in Table 5, row 9) offer the second highest improvement (the greatest improvement for Log and SVM) and will also be tested during the evaluation, along with unigrams for Naïve Bayes (highlighted in Table 5. row 1).

The results presented in Table 5 do not have great values of accuracy, but the aim of this analysis was to determine which selection of attributes from the text is the most suitable as an addition to the basic attributes which we have described above.

6.4. Machine learning methods

In order to evaluate the proposed method for handling negation, the following machine learning methods were used: Naïve Bayes, Logistic Regression, SVM, and J48-decision tree method. The attributes used to train contain all the attributes from the lexicon-based methods (LBM0, LBM1, and LBM2), and additional textual attributes selected in the manner presented in sub-section 6.3. Additional attributes would certainly bring better results for each method, but our aim is not to present the best system of classification of tweets by means of a SA, but to determine whether the attributes obtained by the rules of detecting negation influence the improvement in the classification of tweets in the Serbian language. Table 6 shows the accuracy of the proposed method for NB, LOG, J48 and SVM for:

- the entire set: with unigrams (ALL U); with unigrams and bigrams (ALL U+B),
- the set which contains only negations: with unigrams (OnlyNeg-U); with unigrams and bigrams (OnlyNeg-U+B),
- the set which contains only negations included by the rules: with unigrams (OnlyRuleNeg-U); with unigrams and bigrams (OnlyRuleNeg-U+B)
- all this on the 3-class set.

Table 6. Accuracy of the proposed method for the various groups and the different ML methods, for 3-class

MLM2	ALL U	ALL U+B	OnlyNeg U	OnlyNeg U+B	OnlyRule Neg U	OnlyRule Neg U+B
NB	57.82%	57.80%	62.53%	61.97%	62.17%	62.77%
LOG	68.46%	68.84%	69.25%	69.76%	68.96%	69.69%
J48	61.88%	61.92%	61.57%	61.65%	61.95%	62.04%
SVM	67.45%	65.98%	65.86%	67.65%	67.83%	68.61%

Table 7 presents the same information as in Table 6, but for 2-class.

Table 7. Accuracy of the proposed method for various groups and different ML methods, for 2-class

MLM2	ALL U	ALL U+B	OnlyNeg U	OnlyNeg U+B	OnlyRule Neg U	OnlyRule Neg U+B
NB	86.49%	86.47%	84.63%	85.03%	85.46%	85.98%
LOG	91.15%	91.13%	90.46%	90.53%	90.82%	90.45%
J48	86.88%	86.80%	85.29%	84.92%	84.30%	83.32%
SVM	89.56%	89.36%	88.74%	88.92%	88.29%	88.75%

The results we have achieved by applying all three aforementioned methods (two baseline: MLM0 and MLM1 and the proposed MLM2) for the 3-class and 2-class groups are given in the following tables.

Table 8 shows the results of the application of the methods for the entire set, for all three classes. The table also indicates the accuracy for the "Only words" (instances when only the textual attributes of the vector representation of words are used). The improvements were calculated in relation to the baseline MLM0 method. From Table 8 it can be concluded that the proposed method which includes the rules of negation we have processed (MLM2) provides better results than all the three previous cases. In order to better present the effects of the application of the processed rules, in Table 9 we present the results for the set of tweets which obligatorily contain at least one negation.

Table 8. The results and improvement for the entire group, for 3-class

	ALL-U		ALL-U+B	
	Accuracy	Improvement	Accuracy	Improvement
Only words	64.6660%		65.3053%	
MLM0	67.4843%		68.3586%	
MLM1	68.2281%	1.1022%	68.7500%	0.5726%
MLM2	68.4629%	1.4501%	68.8413%	0.7061%

Table 9. The results and improvements for the set OnlyNeg, for 3-class

	OnlyNeg-U		OnlyNeg-U+B	
	Accuracy	Improvement	Accuracy	Improvement
Only words	67.0224%		68.7033%	
MLM0	67.4673%		68.6683%	
MLM1	68.5615%	1.6218%	69.7892%	1.6323%
MLM2	69.2554%	2.6503%	69.7625%	1.5935%

Table 10. The results and improvements for the set OnlyRuleNeg, for 3-class

	OnlyRuleNeg-U		OnlyRuleNeg-U+B	
	Accuracy	Improvement	Accuracy	Improvement
Only words	67.7182%		68.6690%	
MLM0	68.4825%		69.4336%	
MLM1	68.2663%	-0.3157%	68.9148%	-0.7472%
MLM2	68.9581%	0.6945%	69.6930%	0.3736%

Table 10 shows the accuracy of the prediction of sentiment of only those tweets which contain a negation, or to be precise, the type of negation which was processed by the proposed rules (OnlyRuleNeg). The negative improvement achieved with the MLM1 method indicates that the simple change in the polarity of the word following the negation signal in the case of our dataset leads to worse results than does the exclusion of this rule. The reason for this is that the set OnlyRuleNeg contains tweets which are the most difficult to process. The proposed method, MLM2, offers an improvement even in the case when we use only unigrams, and in the case when both unigrams and bigrams are used together. The results which were achieved for the set 2-class, for all the tweets, are given in Table 11.

Table 11. The results and improvements for the entire set, for 2-class

	ALL-U Accuracy	Improvement	ALL-U+B Accuracy	Improvement
Only words	90.1474%		90.8362%	
MLM0	90.4134%		90.9179%	
MLM1	91.0926%	0.7512%	91.1120%	0.2135%
MLM2	91.1508%	0.8156%	91.1314%	0.2348%

We can see that the accuracy achieved for the set 2-class is a lot greater compared to the one for 3-class. This indicates the problematic nature of neutral tweets.

Table 12. The results and improvements for the set OnlyNeg, for 2-class

	OnlyNeg-U Accuracy	Improvement	OnlyNeg-U+B Accuracy	Improvement
Only words	89.2268%		90.3994%	
MLM0	88.9949%		90.2421%	
MLM1	89.3984%	0.4534%	90.4622%	0.2439%
MLM2	90.4622%	1.6487%	90.5356%	0.3252%

For the set which includes only negation (OnlyNeg), in Table 12 we find a lower percentage of accuracy for all methods. The reason for this is that all the tweets which contain negation are more difficult to classify. However, the improvement which was achieved by using the method MLM2 is greater (1.648%) compared to the improvement for the same method achieved for the entire set (0.8156%).

Table 13. The results and improvements for the set OnlyRuleNeg, for 2-class

	OnlyRuleNeg-U Accuracy	Improvement	OnlyRuleNeg-U+B Accuracy	Improvement
Only words	88.9273%		89.5040%	
MLM0	89.6711%		89.4403%	
MLM1	90.3058%	0.7078%	89.5557%	0.1290%
MLM2	90.8252%	1.2870%	90.4512%	1.1303%

Table 13 shows improvements for the set of tweets that contain the rules of negation we have processed (OnlyRuleNeg). The set of neutral tweets significantly disrupts the quality of the classification of sentiments both for the entire set and for the set with negations. This can be seen by comparing the results of the classifications for the sets 3-class and 2-class. Neutral tweets contain words with sentiment and they determine the sentiment of that tweet. The incorrect impact of sentiment words in neutral tweets can be solved by the introduction of the degree of polarity, which might significantly influence the quality of the classification of the neutral tweets. The accuracy obtained for the set 2-class is satisfactory.

The lexicon-based methods (LBM0, LBM1, LBM2) offer lower accuracy than the ML method since in addition to the attributes which contain the number of occurrences of positive and negative terms, only the attributes which are a direct consequence of the processing of rules of negation were used. However, an improvement in Method1 and Method2 compared to Method0 are expressed better by a comparison of the lexicon-based methods (LBM0, LBM1, LBM2), precisely for the aforementioned reasons.

Using 10-fold cross validation, no qualitative data is obtained because there are no classification data for the individual element from the test set. Therefore, we cannot use the a MC Nemar test for a comparison of the methods, but in this case, we will use T-test. By applying the T-test we analyzed whether the results were obtained by various methods (Only words, MLM0, MLM1, MLM2) for all the sets (All, OnlyNeg and OnlyRuleNeg) in the case when we analyzed 3-class and in the case when we analyzed 2-class. (Table 14)

Table 14. The results of the T-test

Analyzed method	Compared method	P (3-class)	P (2-class)
Only words	MLM0	0.028954	0.250274
Only words	MLM1	0.015784	0.042767
Only words	MLM2	0.004331	0.008634
MLM0	MLM1	0.089519	0.006044
MLM0	MLM2	0.007086	0.004997
MLM1	MLM2	0.017721	0.033535

In all of the studied cases, except for the comparison of the MLM0 and MLM1 for 3-class and Only words and MLM0 for 2-class, a statistically significant improvement was proven. In all the other cases we obtained significance at the $p<0.05$ level. What is of special importance is that in the case of the comparison of the proposed MLM2 method with Only words and MLM0 methods (in the case of 3-class and in the case of 2-class) we obtain significance at the $p<0.01$ level, based on which we can confirm that the proposed MLM2 method can be used to achieve a statistically significant improvement compared to the other two methods. The improvement achieved by applying the MLM2 method is statistically significant compared to the MLM1 method ($p<0.05$).

7. Conclusion and future work

In this paper we have provided an overview of the group of rules for processing negation which occurs in tweets. It has been shown that tweets containing negation are more complicated for SA and that the application of the proposed rules contributes to improvement in determining sentiment. For a more detailed processing of the rules of negation, it is necessary to ensure the existence of an corpus with annotated negation scopes. The existence of this type of resource could enable an additional analysis of the phenomenon of negation and the determination of additional rules which might help the process of negation to be included more successfully into the system of determining the sentiment of the short texts it is contained in. The application of a morphological dictionary might greatly improve the results of the SA, but the aim of this paper was to prove that with the application of the rules of negation there is an improvement in the prediction of sentiment.

The next planned step is the creation of a special corpus which consists of annotated scopes of negation (the scope of the signal of negation in the part of the text). Furthermore, we plan to use the existing morphological dictionary and expand the dictionary of synonyms from the Serbian WordNet [33]. In addition, we also plan to analyze the influence of intensifiers on general sentiment. We also plan to create a special lexicon that will be based on the corpus and which will contain the coefficient of terms that occur in the presence of negation. Using these additional resources, we expect to obtain a significant improvement in training the negation itself, as well as an improvement in the general system for detecting the polarity of short texts.

Acknowledgment. This work was partially supported by the Ministry of Education, Science and Technological Development, Republic of Serbia, Project III-44007. The authors are grateful to Professor Milena Stanković, PhD, for her support in writing this article.

References

1. Hovy, E.H.: What are Sentiment, Affect, and Emotion? Applying the Methodology of Michael Zock to Sentiment Analysis. In: Gala N., Rapp R., Bel-Enguix G. (eds) Language Production, Cognition, and the Lexicon. Text, Speech and Language Technology, Springer, Cham, Vol. 48, 13-24. (2015)
2. Pang, B., Lee, L.: Opinion Mining and Sentiment Analysis. Found. Trends Inf. Retr., Vol. 2, No. 1-2, 1-135. (2008)
3. Liu, B., Zhang, L.: A Survey of Opinion Mining and Sentiment Analysis. In: Aggarwal C., Zhai C. (eds) Mining Text Data. Springer, Boston, MA. (2012)
4. Pang, B., Lee, L., Rd, H., Jose, S.: Thumbs up? Sentiment Classification using Machine Learning Techniques. EMNLP ’02 Proceedings of the ACL-02 conference on Empirical methods in natural language processing, Vol. 10, 79–86. (2002)
5. Das R., S., Chen Y., M.: Yahoo! For Amazon: Sentiment Extraction from Small Talk on the Web. Management Science, Vol. 53, No. 9, 1375–1388. JSTOR. (2007)
6. Polanyi, L., Zaenen, A.: Contextual Valence Shifters. In: Shanahan J.G., Qu Y., Wiebe J. (eds) Computing Attitude and Affect in Text: Theory and Applications. The Information Retrieval Series, Vol 20. Springer, Dordrecht. (2006)
7. Kennedy, A., Inkpen, D.: Sentiment classification of movie and product reviews using contextual valence shifters. In Proceedings of the Workshop on the Analysis of Informal and Formal Information Exchange during Negotiations, Ottawa, Ontario, Canada. (2005)
8. Hu, M., Liu, B.: Mining and summarizing customer reviews. In Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining (KDD '04). ACM, New York, NY, USA, 168-177. (2004)
9. Benamara, F., Cesarano, C., Picariello, A., Recupero, D.R., Subrahmanian, V.S.: Sentiment Analysis: Adjectives and Adverbs are Better than Adjectives Alone. ICWSM. (2007)
10. Jimenez-Zafra, S. M., Martin Valdivia, M. T., Martinez Camara, E., Urena-Lopez, L. A.: Studying the Scope of Negation for Spanish Sentiment Analysis on Twitter. IEEE Transactions on Affective Computing, IEEE, Vol. PP, Issue: 99. (2017)
11. Zhu, X., Guo, H., Mohammad, S., Kiritchenko, S.: An Empirical Study on the Effect of Negation Words on Sentiment. In ACL (1), 304-313. (2014).
12. Kiritchenko, S., Mohammad, S., Zhu, X.: Sentiment Analysis of Short Informal Texts. J. Artif. Intell. Res., Vol. 50, 723-762. (2014)
13. Wiegand, M., Balahur, A., Roth, B., Klakow, D., Montoyo, A.: A survey on the role of negation in sentiment analysis. In Proceedings of the Workshop on Negation and Speculation in Natural Language Processing (NeSp-NLP '10), Roser Morante and Caroline Sporleder (Eds.). Association for Computational Linguistics, Stroudsburg, PA, USA, 60-68. (2010)

14. Sharif, W., Samsudin, N. A., Deris, M. M., Naseem, R.: Effect of negation in sentiment analysis. 2016 Sixth International Conference on Innovative Computing Technology (INTECH), 718-723. IEEE. (2016)
15. Pröllochs, N., Feuerriegel, S., Neumann, D.: Negation scope detection in sentiment analysis, Decis. Support Syst., Vol. 88, No. C, 67-75. (2016)
16. Reitan, J., Faret, J., Gambäck, B., Bungum, L.: Negation Scope Detection for Twitter Sentiment Analysis. Association for Computational Linguistics. (2015)
17. Cerezo-Costas, H., Celix-Salgado, D.: Gradiant-analytics: training polarity shifters with CRFs for message level polarity detection. In Proceedings of the 9th international workshop on semantic evaluation (SemEval-2015), 539–544. Denver, Colorado: Association for Computational Linguistics. (2015, June).
18. Asmi, A., Ishaya, T.: Negation identification and calculation in sentiment analysis. In The Second International Conference on Advances in Information Mining and Management, 1-7. (2012)
19. Giachanou, A., Crestani, F.: Like it or not: A survey of Twitter sentiment analysis methods, ACM Computing Surveys (CSUR), Vol. 49, No. 2, 28:1–28:41. (2016)
20. Batanović, V., Nikolić, B.: Sentiment classification of documents in Serbian: The effects of morphological normalization. In Proceedings of the 24th Telecommunications Forum (TELFOR), Belgrade, Serbia, 1-4. (2016)
21. Batanović, V., Nikolić, B., Milosavljević, M.: Reliable Baselines for Sentiment Analysis in Resource-Limited Languages: The Serbian Movie Review Dataset. In Proceedings of the 10th International Conference on Language Resources and Evaluation (LREC 2016), Portorož, Slovenia, 2688-2696. (2016)
22. Rotim, L., Šnajder, J.: Comparison of Short-Text Sentiment Analysis Methods for Croatian. In Proceedings of the 6th Workshop on Balto-Slavic Natural Language Processing. Association for Computational Linguistics, 69-75. (2017)
23. Mladenović, M., Mitrović, J., Krstev, C., Vitas, D.: Hybrid Sentiment Analysis Framework for A Morphologically Rich Language. Journal of Intelligent Information Systems, Vol. 46:3, 599–620. (2016)
24. Marovac, U., Pljaskovic, A., Crnisanin, A., Kajan, E.: N-gram analysis of text documents in Serbian language, In Proceedings of the 20th Telecommunications Forum (TELFOR), Belgrade, Serbia, 1385-1388. (2012)
25. Jimenez-Zafra, S. M., Taule, M., Martin Valdivia, M. T., Urena-Lopez, L. A., Marti M., A.: SFU ReviewSP-NEG: a Spanish corpus annotated with negation for sentiment analysis. A typology of negation patterns. Language Resources and Evaluation, Springer, Vol. 52, No. 2, 533-569. (2018)
26. Milosevic, N.: Stemmer for Serbian language. arXiv preprint arXiv:1209.4471. (2012)
27. Stone, P. J., Dunphy, D. C., Smith, M. S.: The General Inquirer: A Computer Approach to Content Analysis. M.I.T. press. (1966)
28. Pennebaker, J. W., Chung, C. K., Ireland, M., Gonzales, A., Booth, R. J.: The Development and Psychometric Properties of LIWC2007. The University of Texas at Austin, 1–22. (2007)
29. Wilson, T., Wiebe, J., Hoffmann, P.: Recognizing Contextual Polarity in Phrase-Level Sentiment Analysis. In Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing, 347–354. (2005)
30. Hu, M., Liu, B.: Mining and Summarizing CustomerReviews. In Proceedings of the Tenth SIGKDD International Conference on Knowledge Discovery and Data Mining, ser. KDD ’04. New York, NY, USA: ACM, 168–177. (2004)
31. Baccianella, S., Esuli, A. Sebastiani, F.: SENTI WORD NET 3.0 : An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining. In Proceedings of the Seventh International Conference on Language Resources and Evaluation, Vol. 0, 2200–2204. (2008)
32. Kovačević, M.: Sintaksička negacija u srpskome jeziku. Izdavačka jedinica Univerziteta u Nišu. (2002)
33. Krstev, C., Pavlović-Lažetić, G., Vitas, D., Obradović, I.: Using Textual and Lexical Resources in Developing Serbian Wordnet. Romanian Journal of Information Science and Technology, Vol.7, No. 1-2, 147—161. (2004)

34. Dashtipour, K., Poria, S., Hussain, A., Cambria, E., Y. A. Hawalah, A., Gelbukh, A., Zhou, Q.: Multilingual Sentiment Analysis: State of the Art and Independent Comparison of Techniques. Cognitive Computation, Springer, Vol. 8, 757-771. (2016)

Adela Ljajić is a teaching assistant at the State University of Novi Pazar (Serbia). She received an M.S. degree in Information Systems and Technologies from the University of Belgrade (Serbia). She is currently a Ph.D. student in Electrical Engineering and Computer Science at the University of Niš (Serbia). Her current research interests are related to Natural Language Processing, Text Mining, Machine Learning and Sentiment Analysis with a special focus on treatment of negation. She has authored/co-authored a significant number of publications in international conferences in the field of NLP and Sentiment Analysis.

Ulfeta Marovac is a teaching assistant with PhD at the State University of Novi Pazar, Serbia. She holds a PhD in Computer Science from the University of Belgrade, Serbia. Her thesis “Mining sequential patterns for determination of protein characteristics” proposed the model for classification of proteins in COG categories based on amino acid n-grams (sequences of n length). Her research interests focus on, biological sequences analysis, text mining, natural language processing, logic and its applications in computer science. She is authored and coauthored research publications in Journal of Multiple-Valued Logic and Soft Computing, Match Communications in Mathematical And In Computer Chemistry, Publications de'lInstitut Mathematique.

Received: January 22, 2018; Accepted: June 26, 2018