



Methods for Analyzing Polarity of the Kazakh Texts Related to the Terrorist Threats

Gulmira Bekmanova^(✉), Gaziza Yelibayeva, Saltanat Aubakirova,
Nurgul Dyussupova, Altynbek Sharipbay, and Rozamgul Nyazova

L.N. Gumilyov Eurasian National University, Astana, Kazakhstan
gulmira-r@yandex.kz, gaziza_y@mail.ru,
saltanat759@mail.ru, sharalt@mail.ru,
nurgulya_1991@bk.ru, rozamgul@list.ru

Abstract. In this work we described the rule-based method, using dictionary for sentiment analysis of texts in the Kazakh language related to the terrorist threats. It provides an overview of the methods for analyzing polarity, parser, which analyzes the pages on the content of keywords from the database, morphological, syntactic and sentiment analysis of the texts in the Kazakh language.

Keywords: Sentiment analysis · Kazakh language · Classification · Production rules · Rule-based method · Morphological rules · Syntax rules · Forbidden content · Parser

1 Introduction

Sentiment analysis or opinion mining in natural languages is one of the fast growing natural language processing technologies. Sentiment analysis, also called opinion analysis, is a field of research that analyzes people's opinions, suggestions, assessments, attitudes and emotions in relation to such subjects as products, services, organizations, individuals, problems, events, topics and their attributes [1].

Many resources and systems have been developed for sentiment analysis of texts by now for English [1, 2]. A number of researches are conducting on sentiment analysis for Russian [3, 4], Turkish [5–7], Spanish [8], Arabic [9, 10] and other languages [8]. One approach was proposed for Spanish language to the subjectivity detection on Twitter micro texts that explores the uses of the structured information of the social network framework. For Arabic, there is a semantic approach to discover user attitudes and business insights from Arabic social media by building an Arabic Sentiment Ontology that contains groups of words which express different sentiments in different dialects [13]. Currently, there are emerging semantic models that recognize emotions and words and their expressions in different languages.

The work [11], which presents SEMO, a semantic model for recognizing emotions, which allows users to identify and quantify the emotional load associated with basic emotions hidden in short, emotionally saturated sentences (for example, news headlines, tweets, signatures). The idea, assessing semantic similarity of concepts by considering the occurrences and coincidences of the terms describing them on the pages

indexed by the search system, can be directly extended to emotions and words expressing them in different languages.

Here are some works on the analysis of attitudes for dual languages, Kazakh and Russian [12, 13]. The work [12] describes modern approaches for solving the problem of analyzing the opinions of news articles in the Kazakh and Russian languages using deep recurrent neural networks. Thus, studies show that good results can be achieved even without knowing the linguistic features of a particular language. It also proposes a model of deep neural network, which uses two-lingual embedding of words, to solve effectively the problem of mood classification for a given pair of languages. They apply this approach to two corpora of two different language pairs: English-Russian and Russian-Kazakh. It shows how to train a classifier in one language and predict it in another one. This approach ensures an accuracy of 73% for English and 74% for Russian languages. There are methods, proposed to analyze the mood of the texts in Kazakh, such as a baseline method, which reaches an accuracy of 60% and a method for studying bilingual embedding from a large unlabeled corpus using bilingual word pairs [13]. The analysis of the mood of the texts written in Kazakh is not sufficiently studied. The studies were conducted in [14–16], for the Kazakh language.

Computers are beginning to acquire the ability to recognize emotions. In 1995, Picard [17] reported on the key issues of “affective computing”, calculations that are associated with emotions, arise from them or affect them. Since then, many studies have been conducted. Many studies are related to the recognition of emotions from texts. The work [18] proposes an approach for recognizing emotions using network similarities. It also proposes a model for ranking emotions, based on semantic indicators of proximity, for example, Confidence, PMI, PMING.

There are many mobile devices such as smartphones, tablets, cameras and PCs in the world today. In addition, many applications for audio, video, chatting are being introduced day by day. Accordingly, the amount of text, audio and video information is increased. Therefore, the task of extracting emotions from textual, graphic, audio and video information becomes an important task.

Emotions are extracted not only from texts, but also from audio and video content [19, 20], from images [21]. Such applications can be used as marketing in social media, brand positioning, elections and financial forecasting.

The term “threat” usually refers to a specific unlawful act with the intent of one person to harm another one, which is expressed orally, in written or in another way. We will consider only written threats, which we will call prohibited content. In [22, 23], sentiment analysis is used to detect some prohibited content on the Internet. The determination of prohibited content by means of sentiment analysis is a new direction and works, that the authors of this article consider as useful for the further research, were published in the last ten years. The next article discusses the use of data mining techniques to detect prohibited terrorist activities on the network (Mahesh, Mahesh, Vinayababu 2010) [22]. The work [23] describes a system that detects forbidden statements with an accuracy of more than 80%, using natural language processing techniques and machine learning.

This work can be considered as an introduction and an attempt to apply the linguistic approach to identify texts containing prohibited content of a terrorist nature, which is written in the Kazakh language with keywords in English and Russian

languages. For this reason, this article describes the rule-based methods used in the analysis of feelings, and the approaches used to determine moods through the formalization of morphological and syntactic rules.

2 Method for Analyzing Text Polarity

There are three main methods for determining the text polarity:

1. Analysis of a text using vector analysis methods (often using n-gram models), comparing a text with the previously marked reference corpus on the chosen proximity measure and classifying (classification) it as negative or positive based on the result of the comparison.
2. Search for emotive vocabulary (lexical polarity) in the text according to pre-compiled polarity dictionaries (lists of patterns) using linguistic analysis. Based on the found emotive vocabulary, the text can be evaluated on a scale, which reflects the number of negative and positive vocabulary. This method can use both lists of patterns that are inserted into regular expressions, and rules for combining polarity vocabulary within a sentence.
3. Mixed method (combination of the first and the second approaches).

The analysis of the text polarity, which we are implementing now, consists of several stages (Fig. 1).

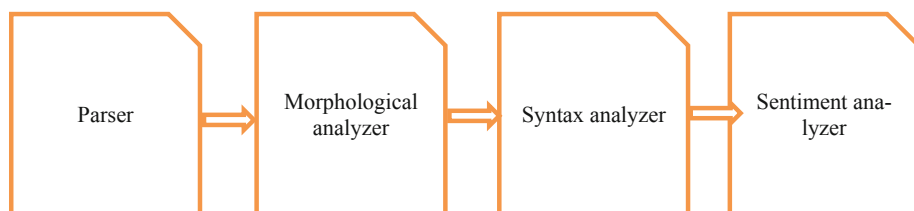


Fig. 1. Stages of the text polarity analysis

At the first stage, the processing of web pages is carried out by the special parser, which analyzes the pages for the content of keywords from the database (in our case, keywords of forbidden content).

At the second stage, the text of selected web pages is processed by a morphological analyzer to determine the parts of speech and the characteristics of each parts of speech.

At the third stage, a simplified syntax analyzer works: words and phrases are combined into polarity chains; the subject, predicate, and object are identified in the sentence.

At the fourth stage, a simplified sentiment analyzer works, it determines the polarity of the text.

3 Parser and Database

Parser is software for collecting and converting textual data into a structured format. The processing of web pages is carried out by the special parser, which analyzes the pages to determine if they contain the key words from the knowledge base (in our case, keywords of forbidden content). The operation of the parser is simple, and its effectiveness depends entirely on the knowledge base. The interface of the parser is shown in Fig. 2.

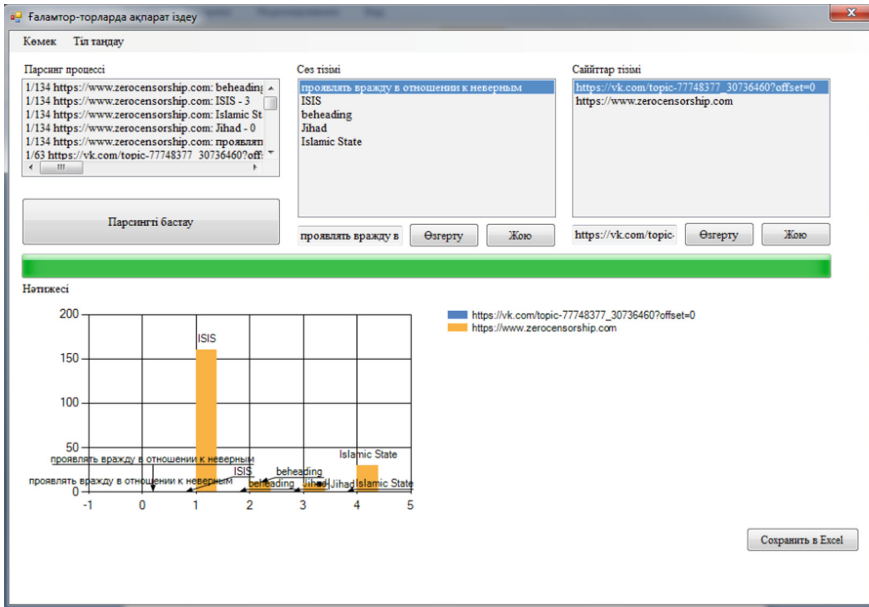


Fig. 2. The interface of the parser

WordNet [24] was used as a prototype of the database structure for the parser, SENTIWORDNET [25] was used at the stage of sentiment analysis. The database was filled completely symmetrically in three languages: Kazakh, English and Russian. This is because, historically, Kazakhstanis often use a mixed language, for example, they speak Kazakh, using terms or stable expressions from Russian and/or English.

WordNet is a large lexical database of the English language. Nouns, verbs, adjectives and adverbs are grouped into sets of cognitive synonyms (synsets), each of them expresses a distinct concept. Synsets are interlinked by means of conceptual-semantic and lexical relations. The resulting network of meaningfully related words and concepts can be navigated with the browser. WordNet is also freely and publicly available for download. WordNet’s structure makes useful tool for computational linguistics and natural language processing. WordNet superficially resembles a thesaurus, in that it groups words together based on their meanings. However, there are some important distinctions. First, WordNet interlinks not just word forms—strings of

letters—but specific senses of words. As a result, words that are found in close proximity to one another in the network are semantically disambiguated. Second, WordNet labels the semantic relations among words, whereas the groupings of words in a thesaurus does not follow any explicit pattern other than meaning similarity [24].

The work [25] presented SENTI WORDNET 3.0, an enhanced lexical resource explicitly devised for supporting sentiment classification and opinion mining applications (Table 1).

Table 1. Fragment of the table in the WordNet format

Noun	Adjective	Verb	Definition	Hyperonym	Hyponym	Holonim	...
Wahhabism			Religious - political trend in Sunni Islam, the middle of the XVIII century. It is based on the teachings of Muhammad al-Wahhab, Salafis	Islam, a radical religious and political movement		Islam	
Gazavat			War for faith is one of the jihad aspects	Holy War, Islam		Islam	
Jihad			The struggle for faith is used by the organizers of the terrorist activities for involving Muslims in it	Holy War, Islam		Islam	
Islamism			World outlook and practical activities of fanatical adherents of Islam, whose faith its institutions (dogmas) in culture, in public, political and economic life, applying extreme measures to achieve goals	Islam, a radical religious and political movement		Islam	
Pan-Islamism			Religious and political current in Islam, preaching the idea of the unity of Muslims around the world and the need for their unity in a single Muslim state. Often it is used as a cover for policy intervention in the affairs of other states and support for various extremist and terrorist structures	Islam, religious and political ideology		Islam	
Religious extremism			Intolerance towards people who have other religions, accompanied by calls for committing or committing antisocial, often unlawful, including violent, acts against citizens	Islam, a radical religious and political movement		Extremism	

(continued)

Table 1. (continued)

Noun	Adjective	Verb	Definition	Hyperonym	Hyponym	Holonim	...
Salafism			A direction in Islam that unites Muslim religious figures who, at different periods of the history of Islam, made calls to orient themselves toward the way of life and faith of the early Muslim community, to righteous ancestors (as-salaf as-salihun). Salafia means “understanding of religion in the form in which it was understood by the prophet and his companions”	Islam, a radical religious and political movement		Tawhid (monotheism)	
...							

Social networks “Vkontakte” and “Facebook”, and selected foreign Internet resources distributing terrorist content were selected for the research and replenishment of the database. To begin with, we directly reviewed the individual profiles of social network users, mentioning prohibited words from the list provided by the US Department of National Security in 2011, which is used to analyze networks for prohibited content [26].

4 Morphological Analysis

The Kazakh language is a typical Turkic language, retaining most of the features common to this group and possessing a number of typical Kypchak features. The structural-typological characteristic of the Kazakh language is mainly related to its membership in agglutinative languages. As a rule, a description of the agglutinative type applies a features set, which takes into account not only phonetic, but also morphological and syntactic features. The order of the endings in the Kazakh language is strictly defined. For example, for nouns, first the plural endings are added to the base of the word, then the possessive endings (meaning the object belongs to a person), then the case endings and then the conjugation form endings (added only to animate nouns). In general, it can be said that the Kazakh language is a well formalized language. We have developed a rule-based morphological analyzer, which is described in [27–29].

5 Syntax Analyzer

Visual analysis of prohibited content shows that it consists of only simple sentences, that is, they do not have complex compound sentences. It follows that to build a syntax analyzer of prohibited content, it is sufficient to formalize and implement

the syntax rules for simple sentences in the Kazakh language, that is, to build a simplified parser.

There are 20 types of simple sentences in the Kazakh language. The formalization of the syntax rules of simple sentences and the construction of the corresponding syntactic analyzer are given in [30]. Some examples of simple sentences containing forbidden fragments are provided below.

A simple sentence consisting of the subject will be represented as:

$$SS(Q(Q1(S))M(M1(N Adj Pron Adv))) \quad (1)$$

Where SS – simple sentence

Q – structure

Q1– structure with the first index

S – Subject

M – semantics

M1 – semantics with the first index

N – Noun

Adj – Adjective

Pron – Pronoun

Adv – Adverb

Example: Bomb. Explosion.

A simple sentence consisting of the subject and the predicate will be presented as:

$$SS(Q(Q2(S P))M(M2(MS (N Adv Num)MP (N V Adj Adv Num)))) \quad (2)$$

Where SS – simple sentence

Q – structure

Q2 – structure with the second index

S – Subject

P – Predicate

M – semantics

M2 – semantics with the second index

MS – semantics of subject

N – Noun

Adv – Adverb

Num – Numeral

MP – semantics of predicate

N – Noun

V – Verb

Adj – Adjective

Adv – Adverb

Num – Numeral

Example: The bomb exploded. The explosion occurred. Set fire. The explosion is terrible. The building was crushed. Four blew up. Infidels died.

A simple sentence consisting of the subject, the complement and the predicate will be presented as:

$$SS(Q(Q3(S OP)))M(M3(MS (N Pron Num Adj) MA (N Pron Num Adj) MP (N V)))) \quad (3)$$

- Where SS – simple sentence
- Q – structure
- Q3 – structure with the third index
- S – Subject
- O – Object
- P – Predicate
- M – semantics
- M3 – semantics with the third index
- MS – semantics of subject
- N – Noun
- Pron – Pronoun
- Num – Numeral
- Adj – Adjective
- MA – semantics of object
- N – Noun
- Pron – Pronoun
- Num – Numeral
- Adj – Adjective
- MP – semantics of predicate
- N – Noun
- V – Verb

Example: Terrorists crashed the city. Five soldiers got killed.

The complete list of formal syntactic rules for simple sentences is as follows:

- SS (Q(Q1(S)) M(M1(N Adj Pron Adv)))
- SS(Q(Q2(S P)) M(M2(MS (N Adv Num)MP (N V Adj Adv Num))))
- SS(Q(Q3(S O P)) M(M3(M_S (N Pron Num Adj) M_A(N Pron Num Adj) M_P (N V))))).
- SS(Q(Q4(S C P)) M(M4(M_S (N Pron Num Adj) M_C(N Adv Num Adj) M_P (V))))).
- SS(Q(Q5(S A C P)) M(M5(M_S (N Pron Num Adj) M_A(N Adj Num Pron) M_C (N Adv) M_P(V))))).
- SS(Q(Q6(S A C P)) M(M6(M_S (N Pron Num Adj) M_C(N Adv Pron) M_A (N Pron Adj Num) M_P(V))))).
- SS(Q(Q7(S D A P)) M(M7(M_S (N Pron Adj Num) M_D(Adv Pron Num Adv) M_A (N Adj Num) M_P(V))))).
- SS(Q(Q8(A D S P)) M(M8(M_A (N Pron Adj Num) M_D(Adj Adv Num) M_S (N Num Adj) M_P(V))))).

SS(Q(Q₉(A C S P)) M(M₉(M_A (N Pron Adj Num) M_C(Adv Num) M_S (N Pron Adj Num) M_P(V)))).

SS(Q(Q₁₀(A S C P)) M(M₁₀(M_A (N Pron Adj Num) M_S(N Pron Num Adv Adj) M_C (Adv Num) M_P(V)))).

SS(Q(Q₁₁(C S A P)) M(M₁₁(M_C (N Num Adv) M_S(N Pron Num Adv Adj) M_A (Adv Num) M_P(V)))).

SS(Q(Q₁₂(D S A P)) M(M₁₂(M_D (Adj Num Adv Pron) M_S(N Adj) M_A (N Adj Adv Num) M_P(V)))).

SS(Q(Q₁₃(D S C P)) M(M₁₃(M_D (Adj Num Adv Pron) M_S(N Adj) M_C (N Adj Adv Num) M_P(V)))).

SS(Q(Q₁₄(S C D A P)) M(M₁₄(M_S (N Pron Adj Num) M_C(N Adv) M_D (N Adj Adv Num) M_A (N Adj Num) M_P(V)))).

SS(Q(Q₁₅(S D A C P)) M(M₁₅(M_S (N Pron Adj Num) M_D(Adj Adv Num) M_A (N Adj Num) M_C (N Adv Adj) M_P(V)))).

SS(Q(Q₁₆(C S D A P)) M(M₁₆(M_C (N Adv Num) M_S(N Pron Num Adj) M_D(Num Adj) M_A(N Num Adj) M_P(V)))).

SS(Q(Q₁₇(C A D S P)) M(M₁₇(M_C (N Adv Num) M_A(N Pron Num Adj) M_D(Num Adj) M_S(N Pron Num Adj) M_P(V)))).

SS(Q(Q₁₈(D S C A P)) M(M₁₈(M_D (Adj Adv Num) M_S(N Pron Num Adj) M_C(Adv N) M_A(N Pron Num Adj) M_P(V)))).

SS(Q(Q₁₉(D S A C P)) M(M₁₉(M_D (Adj Adv Num) M_S(N Pron Num Adj) M_A(N Pron Num Adj) M_C(Adv N) M_P(V)))).

SS(Q(Q₂₀(D A S C P)) M(M₂₀(M_D (Adj Adv Num) M_A(N Pron Num Adj) M_S(N Pron Num Adj) M_C(Adv N) M_P(V)))).

The input of the syntactic analyzer is the text containing elements of forbidden content with morphological tags.

At this stage, the initial syntax analysis is performed: words and phrases are combined into polarity chains, the subject, predicate and object are distinguished in the sentence. The result of the syntactic analyzer will be the text containing forbidden content marked by morphological and syntactic tags.

6 Sentiment Analysis

At this stage, the polarity object is highlighted. It is set by the user or determined automatically: in each sentence, a so-called named entity is searched, for example, a proper name, animate real, etc. For a given object, the polarity of the text is calculated [14–16, 31]. To work at the sentiment analysis stage, a database of words and phrases that are emotional in color and refer to prohibited content (hereinafter - the prohibited fragment) is used. In our case, we use database tables that have a SENTI WORDNET database structure (Table 2).

Table 2. Fragment of the table in the SENTI WORDNET format

#POS	PosScore	NegScore	Terms	Gloss
n	0	0,5	Ваххабизм#1 Уаһхабизм#1 Вахабизм#1 Уаһабизм#1 (Wahhabism)	XVIII ғасырдың ортасында Мухаммад әл-Ваххаб оқытуларының негізінде пайда болған сүнниттік исламның радикалды-саяси ағын, салафиттер (a Sunni Islamic doctrine and radical-political movement founded by Muhammad al-Wahhab in the XVIII century, Salafi)
n	0	0,875	Газават#1 Жиһад#1 кәпірлерге_қарсы_соғыс#1 Аллаһтың_жолында_күресу#1 (Gazawat, war against infidels, war in God's path)	«Дін үшін соғыс», джихадтың бір бөлігі (War for religion, a part of jihad)
n	0	0,875	Джихад#1 аль_Жиһад#1 Жиһад#1 Жиһат#1 исламдық_джихад#1 (jihad, al-jihad, Islamic jihad)	Дін үшін соғыс, мұсылмандарды еліктіру үшін қолданылатын түсінік (war for religion, a concept is used to attract muslims)
n	0	0,5	Исламизм#1 (Islamism)	Исламның бірбеткей, ессіз ұстанушыларының көзқарастары мен тәжірибелік іс-әрекеттері, оның (ислам) белгілерінің бастылығын уағыздап жүрулері қоғамда, саясатта және экономикалық өмір кырларында, мақсаттарына жету үшін шекті іс-әрекеттерді орындау (views and practical actions of one-sided and unconscious Islam followers, preaching the unity of its signs, executing ultimate actions to meet their goals in society, politics and economic life)
n	0	0,5	Панисламизм#1 (Panislamism)	XIX ғасырдың соңында пайда болған, мұсылман реформаторы Джемал ад-Дина ал-Афгани (1839–1897) есімімен байланысты идеология және қозғалыс; жер бетіндегі барлық мұсылманқауымының бірлігі концепциясы мен олардың біртұтас ислам мемлекетіне бірігу қажеттілігіне негізделген діни-саяси идеология және тәжірибе. Көбінесе әртүрлі террористік және экстремистік топтарды ақшалай қолдау үшін “жаппа” ретінде қолданылады (ideology and movement related the name of Islamic reformer Jemal al-Dina al-Afgani (1839–1897) formed in the XIX century; religious-political ideology and practice which is based in the idea of necessity of association for all Islamic communities in the world; mostly used in different terrorist and extremism groups as “cover” to get financial support)
n	0	0,875	Діни экстремизм#1 (religious extremism)	Басқа дінді ұстайтын адардармен есептеспеушілік, қоғамға қарсы,

(continued)

Table 2. (continued)

#POS	PosScore	NegScore	Terms	Gloss
				құқыққа қарсы, зорлықтық іс-әрекеттерді жасауға үндеумен қатар жүреді. діттеген саяси мақсаттарына қол жеткізу үшін, дінді қалқан ете отырып әрекет етуші, дінге ешқандай қатысы жоқ іс-қимыл. Басқаша атағанда - терроризм (лаңкестік). Мұндай әрекеттер көбіне мемлекеттің құрамын күшпен өзгертуге немесе үкіметті басып алуға бағытталады. Қауіпті жағдайға айналған осындай әрекеттерді іске асыру, көбіне қарулы топтар құрып, өзара өшіктіру әдістерін қолдану арқылы дін және ұлттық қайшылықтарды қоздырумен, сондай-ақ адам құнын жаппай бұзумен қатар жүреді. (ignoring other religion followers, goes together with a call to do violent actions against society, rights; actions which are not related to religion but using it as a “cover”. In other words, it is terrorism. Usually, these actions are targeted to change a state composition or take it over by force; goes together with actions for creating armed groups, unleashing religious and national contradictions, violating human rights)
n	0	0,5	Сәләфия#1 салафия#1 (salafia)	Ислам дінінің негізі боп табылатын Құран мен Сүннетті «әс-сәләф әс-салихун», яғни Мұхаммед пайғамбардың (салл Аллаху ғалейһи уә сәлләм) сахабалары, олардың ізбасарлары мен Ислам үмметінің имамдары қалай түсінсе солай түсіну дегенді білдіред (means to understand Quran and Sunni, which are the basis of the Islam religion, as companions of the prophet Muhammad and followers and Islam community (ummah) understand)
...				

7 Evaluation of Results

Currently, methods for objective testing of textual markup systems have not yet been developed. Therefore, the testing method currently used by us is based on periodic subjective evaluations of small text collections by an expert [32].

8 Conclusion

In this paper, we have developed a method for analyzing the polarity of the Kazakh texts related to terrorist threats. The database for the parser and sentiment analysis was developed. The syntax rules of simple sentences in the Kazakh language, which are sufficient for the presentation of texts related to terrorist threats, have been formalized. The stages of morphological, syntactic and sentiment analysis of texts in the Kazakh language are described.

The development of this work and the replenishment of the database and knowledge base of the prohibited content will allow detecting sites leading to terrorist propaganda. Currently, the database contains 1200 entries, which allowed us to detect more than 50 similar sites. Content analysis of sites allowed expanding the database with new keywords. Previously obtained results of determining the polarity of texts made it possible to automatically determine the list of sites that have extremely negative content. We hope that this work will be developed by virtue of its practical significance for society.

References

1. Liu, B.: *Sentiment Analysis and Opinion Mining*. Morgan & Claypool Publishers (2012)
2. Pang, B., Lee, L.: *Opinion mining and sentiment analysis*. Foundations and Trends® in Information Retrieval. Now Publishers (2008)
3. Loukachevitch, N.V., Chetviorkin, I.I.: Evaluating Sentiment Analysis Systems in Russian. *Artif. Intell. Decis. Making* **1**, 25–33 (2014). (In Russian)
4. Chetvirokin, I., Loukachevitch, N.: Sentiment analysis track at ROMIP 2012. In: *Proceedings of International Conference Dialog-2013*, vol. 2, pp. 40–50 (2013)
5. Akba, F., Uçan, A., Sezer, E.A., Sever, H.: Assessment of feature selection metrics for sentiment analyses: Turkish movie reviews. In: *Proceedings of the 8th European Conference on Data Mining*, pp. 180–184 (2014)
6. Yıldırım, E., Çetin, F., Eryiğit, G., Temel, T.: The impact of NLP on Turkish sentiment analysis. In: *Proceedings of the TURKLANG 2014 International Conference on Turkic Language Processing*, Istanbul (2014)
7. Eryiğit, G., Çetin, F., Yanık, M., Temel, T., Çiçekli, I.: TURKSENT: a sentiment annotation tool for social media. In: *Proceedings of the 7th Linguistic Annotation Workshop & Interoperability with Discourse, ACL 2013*, Sofia, Bulgaria (2013)
8. Sixto, J., Almeida, A., López-de-Ipiña, D.: An approach to subjectivity detection on twitter using the structured information. In: Nguyen, N.-T., Manolopoulos, Y., Iliadis, L., Trawiński, B. (eds.) *ICCCI 2016. LNCS (LNAI)*, vol. 9875, pp. 121–130. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-45243-2_11

9. Mohammad, S., Salameh, M., Kiritchenko, S.: Sentiment lexicons for Arabic social media. In: Proceedings of the 10th Edition of the Language Resources and Evaluation Conference, Portorož (Slovenia) (2016)
10. Samir, T., Abdul-Nabi, I.: Semantic sentiment analysis in Arabic social media. *J. King Saud Univ. Comput. Inf. Sci.* **29**(2), 229–233 (2016)
11. Franzoni, V., Milani, A., Biondi, G.: SEMO: a semantic model for emotion recognition in web objects. In: Proceedings IEEE/WIC/ACM International Conference on Web Intelligence, WI (2017)
12. Sakenovich, N.S., Zharmagambetov, A.S.: On one approach of solving sentiment analysis task for Kazakh and Russian languages using deep learning. In: Nguyen, N.-T., Manolopoulos, Y., Iliadis, L., Trawiński, B. (eds.) ICCCI 2016. LNCS (LNAI), vol. 9876, pp. 537–545. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-45246-3_51
13. Abdullin, Y.B., Ivanov, V.V.: Deep learning model for bilingual sentiment classification of short texts. *Sci. Tech. J. Inf. Technol. Mech. Opt.* **17**(1), 129–136 (2017)
14. Yergesh, B., Bekmanova, G., Sharipbay, A.: Sentiment analysis of Kazakh text and their polarity. In: *Web Intelligence* (2019)
15. Yergesh, B., Bekmanova, G., Sharipbay, A.: Sentiment analysis on the hotel reviews in the Kazakh language. In: 2nd International Conference on Computer Science and Engineering, UBMK (2017)
16. Yergesh, B., Bekmanova, G., Sharipbay, A., Yergesh, M.: Ontology-based sentiment analysis of Kazakh sentences. In: Gervasi, O., Murgante, B., Misra, S., Borruso, G., Torre, C.M., Rocha, A.M.A.C., Taniar, D., Apduhan, B.O., Stankova, E., Cuzzocrea, A. (eds.) ICCSA 2017. LNCS, vol. 10406, pp. 669–677. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-62398-6_47
17. Picard, R.W.: *Affective Computing*. MIT Media Laboratory Perceptual Computing Section Technical Report No. 321. Media Lab. Massachusetts Institute of Technology, Cambridge University (1995)
18. Biondi, G., Franzoni, V., Li, Y., Milani, A.: Web-based similarity for emotion recognition in web objects. In: Proceedings - 9th IEEE/ACM International Conference on Utility and Cloud Computing, UCC 2016, pp. 327–332 (2016)
19. Poria, S., Chaturvedi, I., Cambria, E., Hussain, A.: Convolutional MKL based multimodal emotion recognition and sentiment analysis. In: Proceedings - IEEE International Conference on Data Mining, ICDM, Art. no. 7837868, pp. 439–448 (2017)
20. Arunnehr, J., Kalaiselvi Geetha, M.: Automatic human emotion recognition in surveillance video. *Stud. Comput. Intell.* **660**, 321–342 (2017)
21. Jiang, R., Ho, A.T.S., Cheheb, I., Al-Maadeed, N., Al-Maadeed, S., Bouridane, A.: Emotion recognition from scrambled facial images via many graph embedding. *Pattern Recogn.* **67**, 245–251 (2017)
22. Mahesh, S., Mahesh, T.R.: Vinayababu, M: Using data mining techniques for detecting terror-related activities on the WEB. *J. Theor. Appl. Inf. Technol.* **16**(2), 99–104 (2010)
23. De Smedt, T., De Pauw, G., Van Ostaeyen, P.: Automatic Detection of Online Jihadist Hate Speech: CLiPS Technical Report Series (2018)
24. <https://wordnet.princeton.edu/>
25. Esuli A., Baccianella S., Sebastiani F.: SentiWordNet 3.0: an enhanced lexical resource for sentiment analysis and opinion mining. In: Proceedings of the International Conference on Language Resources and Evaluation, Valletta, Malta (2010)
26. Analyst’s Desktop Binder, U.S. Department of Homeland Security [Электронный ресурс]. <https://epic.org/foia/epic-v-dhs-media-monitoring/Analyst-Desktop-Binder-REDACTED.pdf>

27. Sharipbayev, A., Bekmanova, G., Buribayeva, A., Mukanova, A., Kaliyev, A.: Semantic neural network model of morphological rules of the agglutinative languages. In: The 6th International Conference on Soft Computing and Intelligent Systems. The 13th International Symposium on Advanced Intelligent Systems, Kobe, Japan, pp. 1094–1099 (2012)
28. Yergesh, B., Mukanova, A., Sharipbay, A., Bekmanova, G., Razakhova, B.: Semantic hyper-graph based representation of nouns in the Kazakh language. *Computacion y Sistemas* **18**(3), 627–635 (2014)
29. Zhetkenbay, L., Sharipbay, A., Bekmanova, G., Kamanur, U.: Ontological modeling of morphological rules for the adjectives in Kazakh and Turkish languages. *J. Theor. Appl. Inf. Technol.* **91**, 257–263 (2016)
30. Razakhova, B.Sh., Sharipbaev, A.A.: Formalization of syntactic rules of the Kazakh language *Вестник. Special issue.- Astana: L.N. Gumilyov ENU*, pp. 42–50 (2012)
31. Yergesh, B., Sharipbay, A., Bekmanova, G., Lipnitskii, S.: Sentiment analysis of Kazakh phrases based on morphological rules. *Theor. Appl. Sci. Tech. J.* **2**(38), 39–42 (2016). Kyrgyz State Technical University named after I. Razzakov
32. Pazelskaya, A., Solovyev, A.: A method of sentiment analysis in Russian texts. In: *Dialog* (2011)