Sentiment Processing of Socio-political Discourse and Public Speeches

Gulmira Bekmanova[1,2(✉)], Banu Yergesh[1,2], Aru Ukenova[1], Assel Omarbekova[1,2], Assel Mukanova[3], and Yerkyn Ongarbayev[2]

[1] Faculty of Information Technologies, L.N. Gumilyov Eurasian National University, 010000 Astana, Kazakhstan
gulmira-r@yandex.kz
[2] L.N. Gumilyov Eurasian National University, 010000 Astana, Kazakhstan
[3] Higher School of Information Technology and Engineering, Astana International University, 010000 Astana, Kazakhstan

Abstract. The article deals with the development of an ontological model of words in public political discourse and texts of public speeches in the Kazakh language. The article presents an ontological model of the subject area of elections, a referendum, examples of processing queries from the knowledge base are given. A sentimental analysis of political discourse in social networks in the Kazakh language was carried out in order to determine the mood of the discussion in these sources.

Keywords: Intelligent Synonymizer · Political discourse · Database · Kazakh language · Ontology

1 Introduction

The concept of sentiment analysis has many different names, interpretations, tasks, such as, for example, sentiment analysis, intellectual analysis of opinions, search for opinions, search for subjectivity, sentiment analysis, etc. Therefore, in this work, the concept of sentiment analysis was adopted and, in connection with it, the following concepts were clarified.

Sentiment - the emotional attitude of the author, who expressed an opinion on certain objects (product, organization, personality, etc.), events (elections, uprising, war, etc.), phenomena (eclipse, flood, etc.), process (education, service, etc.) or its properties described in the text.

Opinion is a concept about something, belief, judgment, conclusion, conclusion, point of view or statement on a topic in which it is impossible to achieve complete objectivity, based on the interpretation of facts and an emotional attitude towards them.

Sentiment analysis of text is a class of content analysis methods in computational linguistics designed for automated detection of emotionally colored vocabulary in texts and emotional evaluation of authors (opinions) in relation to objects, phenomena, events, processes or their properties.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023
O. Gervasi et al. (Eds.): ICCSA 2023 Workshops, LNCS 14108, pp. 191–205, 2023.
https://doi.org/10.1007/978-3-031-37117-2_15

Sentiment of an opinion is a sign off, which indicates that the opinion will be positive, neutral or negative. It is also called opinion polarity, sentiment direction, or semantic direction.

Many researchers are engaged in sentiment analysis, so there are many different methods and algorithms that are used in research. On the one hand, in the applied research of sentiment analysis, machine learning methods, lexicon-based methods or linguistic methods can be applied.

On the other hand, the classification of sentiment analysis methods may depend on the level of their classification, such as document, sentence, or aspect level. Sentiment analysis at the document level classifies the entire document as positive or negative. In this case, it is considered that the document describes one object. Sentence-level sentiment analysis categorizes each sentence in a document as subjective or objective, and categorizes subjective opinions as positive or negative. And at the level of the object and aspect, the relation of the object to a particular aspect is determined, because the user can leave different opinions on several aspects of one object in one review.

Sentiment analysis methods can be used on various types of data such as news, reviews, blogs, or social media posts.

Each type of data has its own characteristics that must be taken into account when collecting, preparing, pre-processing data and describing objects.

The topic of political discourse on the Internet, automatic detection of hotbeds of political discourse in social networks is an urgent task of modern researchers. The next important task is to determine the sentiment of discussions in the centers of political discourse. The prerequisites for this study are several works by the authors related to the identification of terrorist threats in social networks [1], the work on the sentiment analysis of texts [2–6], as well as the creation of a morphological analyzer [7, 8].

An important element of the political discourse processing system is the development and further use of an intelligent synonymizer, which significantly increases the efficiency of the system that solves the problem of socio-political discourse processing sentiment. The development of an intelligent synonymizer application for socio-political discourse and texts of public speeches is an important task in the field of natural language processing. The socio-political discourse and texts of public speeches are often characterized by the use of specialized terminology and jargon, as well as by the repetition of certain words or phrases [9–13]. This can lead to a lack of clarity and precision in communication, as well as a loss of audience attention and engagement.

An intelligent synonymizer application can help resolve these issues by providing users with a variety of alternative words and phrases that convey the same meaning as the original word or phrase. The application can recommend synonyms based on the context and meaning of the original word or phrase by combining rule-based and machine learning approaches [14–19]. Data gathering, preprocessing, feature extraction, model training, and evaluation are all steps in the development of such an application. The data collection step entails gathering a large corpus of socio-political discourse and public speech texts that can be used to train and evaluate the synonymizer model [20–23]. In the preprocessing stage, the data is cleaned and tokenized, and stop words and other noise are removed [24]. The feature extraction stage involves identifying relevant features of the text, such as word frequency, co-occurrence patterns, and syntactic structures. The

model training stage involves selecting and implementing a machine learning algorithm, such as a neural network or a decision tree, and using the extracted features to train the model on the corpus of text. The evaluation stage involves testing the performance of the model on a separate test set of text, using metrics such as accuracy, precision, and recall [25–27].

Overall, the development of an intelligent synonymizer application for socio-political discourse and texts of public speeches has the potential to improve the quality of communication in these important domains, and to enhance audience engagement and understanding.

2 Related Works

Sentiment processing of socio-political discourse and public speeches is a subfield of natural language processing (NLP) that involves the identification and analysis of emotions, attitudes, and opinions expressed in text data related to socio-political issues and public speeches.

The work [28] explores the nature of sentiment analysis in political texts, outlines the challenges faced by analysts of political texts, and discusses how these challenges influence the reliability of findings. It demonstrates ongoing research that is creating an algorithm that analyzes stylistic devices (such as sarcasm, irony, and hyperbole) and offers choices for figuring out the sentiment of texts in sentences that contain these stylistic devices.

Social media opinions and comments that incite racial tensions have been seen as posing a significant danger to social, political, and cultural stability as well as the peace in various nations. As a result, social media should be watched closely, and offensive comments should be quickly identified and blocked. The research [29–31] uses sentiment analysis to identify tweets that contain racist text. Adopting the sentiment analysis model [32] also showed that the emotions associated with all of the identified themes are primarily negative sentiments expressed toward socio-political problems. Moreover, this study [33] examines the sentiment analysis of social media data for the purpose of detecting anxiety or despair using a variety of artificial intelligence methods. A multimodal hate speech detection algorithm tailored for Greek social media is presented by the authors in [34]. The research focuses on Greek-language tweets that criticize immigrants and refugees, particularly those that use racist and xenophobic language. Similar to this, [35] suggests an advanced machine learning-based method for the automatic detection of hate speech in Arabic social media networks. In order to analyze the data, a distinct set of features are used and various emotional states are recorded. The research employs four distinct machine learning techniques, including Naive Bayes (NB), DT, SVM, and RF, along with features from the TF-IDF, profile, and mood.

Another paper [36] offers PolSentiLex, a sentiment lexicon for the Russian language that is openly available and created to identify sentiment in user-generated content about social and political problems. The next method proposed in [37] for evaluating news using blockchain technology, smart contracts, and NLP has a lot of potential. Its system assesses news without prejudice or interference from sociopolitical factors using a combination of human and machine opinion. In order to gauge public sentiment regarding

the spread of the COVID-19 virus and its potential societal repercussions, a sentiment analysis methodology was created in [38]. This was accomplished by extracting features using the TF-IDF and Lexical techniques, and classifying the data using Naive Bayes.

It was discovered that sentiment affects the way content is shared on social media, with emotionally charged Tweets getting more retweets and news content with high sentiment being more well-liked and getting more notice [39, 40]. Additionally, it has been demonstrated that including sentiment in recommendation systems raises the quality of recommendations by making them more specifically tailored to the participant's preferences [41]. However, the predominance of negative emotion in a collection of news stories may reduce the likelihood of meeting opposing views and getting a balanced media diet [42], as high levels of negative emotion indicate a bias in the news story [43].

Depending on the frequency and length of social unrest, there may be long-term consequences that raise risks in the financial, economic, and political spheres. The work [44] has annotated a dataset based on Nigeria's #EndSARS campaign to forecast the occurrence of societal unrest. A support vector classifier has also been used to model how the general public feels about the demonstration.

One of the potential challenges in developing an intelligent synonymizer in Kazakh language is the complexity and variability of the language used in socio-political discourse and public speeches. The use of specialized terminology, jargon, and metaphors can make it difficult to identify appropriate synonyms that convey the same meaning and nuance as the original text. In order to analyze texts in the Kazakh language that contain terrorist threats, this study [1] described a rule-based approach that used a dictionary. Additionally, the authors have published a number of works on sentiment analysis of texts [2–6] as well as morphological analysis [7, 8] for Kazakh language.

3 Development of a Database of Synonyms and an Ontological Model of Words in Public Political Discourse and Texts of Public Speeches

3.1 Review and Coordination of Sources of Political Discourse

While the term "discourse" comes from the Latin word discursus, meaning "conscious thought (reasonable thinking)", over time, this concept came to mean "scientific dispute", in some works discourse comes from the Latin "discurro – to wander", "discurrere" – to discuss, negotiate. Discourse is an attribute of the communicative-social field, it occurs where the statement has social consequences. Discourse is a concept that analyzes the social background of a thought, presented through an ambiguous conversation.

Currently, in connection with the events taking place in the world and the country, the study of political institutions, political processes, and political technologies is an urgent problem. Therefore, it is necessary to study the socio-political vocabulary, political texts written in the Kazakh language, to study them on their basis. Changes in the society that form the political cognitive consciousness of the society and the linguistic image of politics form the concept of political discourse in the relationship of humanity. For this reason, it is important to study political discourse to the fullest.

Political discourse is presented in many communication tools such as treaties, reports, election campaigns, editorials, newspaper comments, interviews and conferences. Politicians usually use regular language mixed with direct and colloquial language. They can also use proverbs and idioms. Through them, politicians make their language more informal, making it more understandable to the audience for many listeners who listen to them. Thus politicians often use two styles: for example, colloquial, and rhetorical style, which may also include political language, due to special terminology and speech structure.

Today, in addition to oral communication and dialogue, written texts in social networks and correspondence are also a form of social and political activity. The implementation of political action through text and conversation is the trend of the new time. They have political functions and consequences, so they are a source of political discourse.

The vast majority of political discourse studies are devoted to the texts and speeches of professional politicians or political institutions such as presidents and prime ministers, as well as governments, parliaments or political parties at the local, national and international levels. Therefore, political discourse is a tool that contributes to mass consciousness.

The object of our study will be political discourse on electoral topics, since here everyone is a participant in political discourse: the president, deputies, members of the government, political parties and social movements, as well as citizens who have the right to vote. As a source of political discourse, one can take texts and articles on the topic of elections and referendums, speeches and appeals of leaders of political and public associations, as well as candidates for deputies. And all this will be covered through the media.

3.2 Selection and Development of Text Data

It is necessary to prepare corrective textual data based on socio-political discourse. The subject area of elections and referendums was chosen for the analysis. A study was made of 1200 sources applicable to this area.

It was revealed that according to the tone of messages in the articles, negative - 2.58%, neutral - 95.5%, positive - 1.92%. Of all the data, the sources of information in social networks accounted for 33.33%, the media of information - 66.7% (Fig. 1).

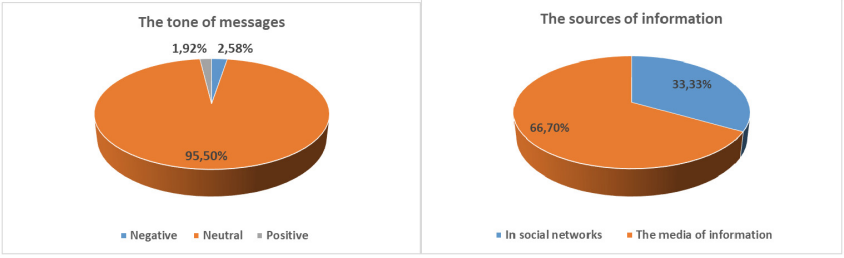

Fig. 1. Percentage of Sources

After analyzing the texts from these sources, more than 800 pages of textual data that met the requirements were collected and a text database was developed.

3.3 Ontological Model of the Subject Area of Elections, Referendum

To build an ontology there were collected, first of all, accumulated knowledge of the subject area and the basic concepts of selections from the collected texts. Their definitions, explanations, relationships to each other, subclasses, properties and axiomatic rules, their personalities were determined. The creation of the ontology was carried out in the Protégé environment.

A fragment of ontology classes "Electoral System of the Republic of Kazakhstan" is shown in Fig. 2.

To build an ontology, knowledge of the subject area and the basic concepts of samples from the collected texts were collected. Their definitions, explanations, relationships to each other, subclasses, properties and axiomatic rules, their personalities were determined. The creation of the ontology was carried out in the Protégé environment.

A fragment of ontology classes "Electoral System of the Republic of Kazakhstan" is shown in Fig. 2. Link to ontology https://clck.ru/33tTch.

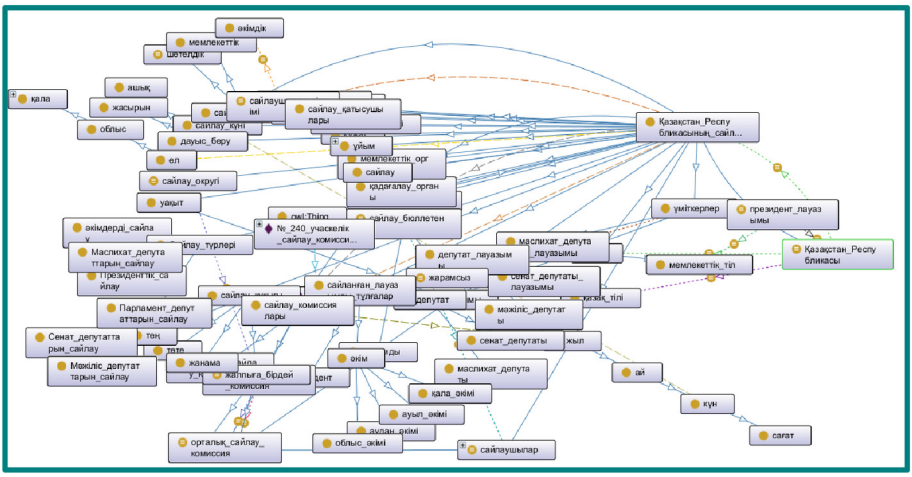

Fig. 2. A fragment of ontology classes "Electoral System of the Republic of Kazakhstan"

The figure shows the election commissions for the electoral system of the Republic of Kazakhstan, their types, definitions and terms of office (Fig. 3).

The figure reflects the central electoral commission, definition, term of office and members (Fig. 4). Since for each person it is indicated whether he is a member or the chairman of the central election commission, Reasoner automatically determines the chairman of the central election commission and the members of the central election commission.

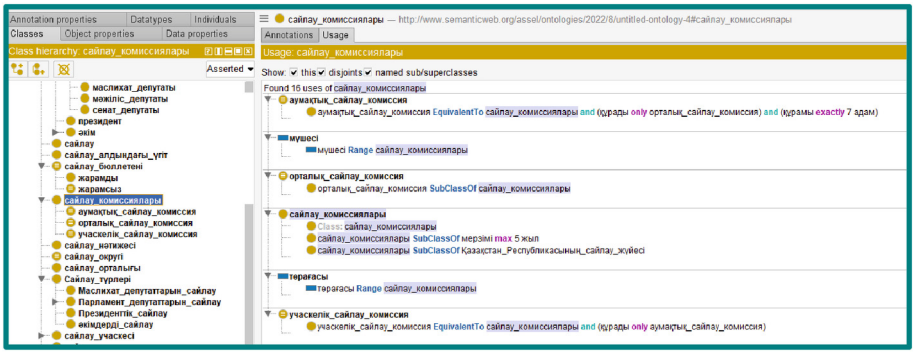

Fig. 3. Subclasses in the ontological model

The figure shows what qualities a candidate for the presidency should have. To do this, the following conditions must be met

$$\begin{aligned} &\textit{office of president} \equiv \exists\, \textit{state language proficiency} \sqcap \forall\, \textit{citizenship} \\ &(\textit{The Republic of Kazakhstan}) \sqcap \exists\, \textit{has Age} \geq 40 \sqcap \exists\, \textit{has higher education} \end{aligned} \tag{1}$$

here: office of president, state language, the Republic of Kazakhstan, higher education – classes of ontology, proficiency, age, has – denote relationships between classes.

Only a candidate who meets this axiom (rule) can run as a presidential candidate (Fig. 5).

The following figure shows that the precinct election commission is formed exclusively by the territorial election commission, as well as how many people are included in it (Fig. 6). The commission's term of office is 5 years.

As a rule, the composition of the precinct election commission can be of two types:

With the number of voters <= more than 2,000 people, the composition of the precinct election commission is 5–7 people.

With the number of voters => more than 2,000 people, the composition of the precinct election commission is 7–11 people.

In the created ontology, in accordance with the "Law of the Republic of Kazakhstan On Elections", a lot of provisions are created, as indicated above, and a definition of each concept is given.

With the help of ontology, we can get answers to various questions. To do this, we can create DL Query queries.

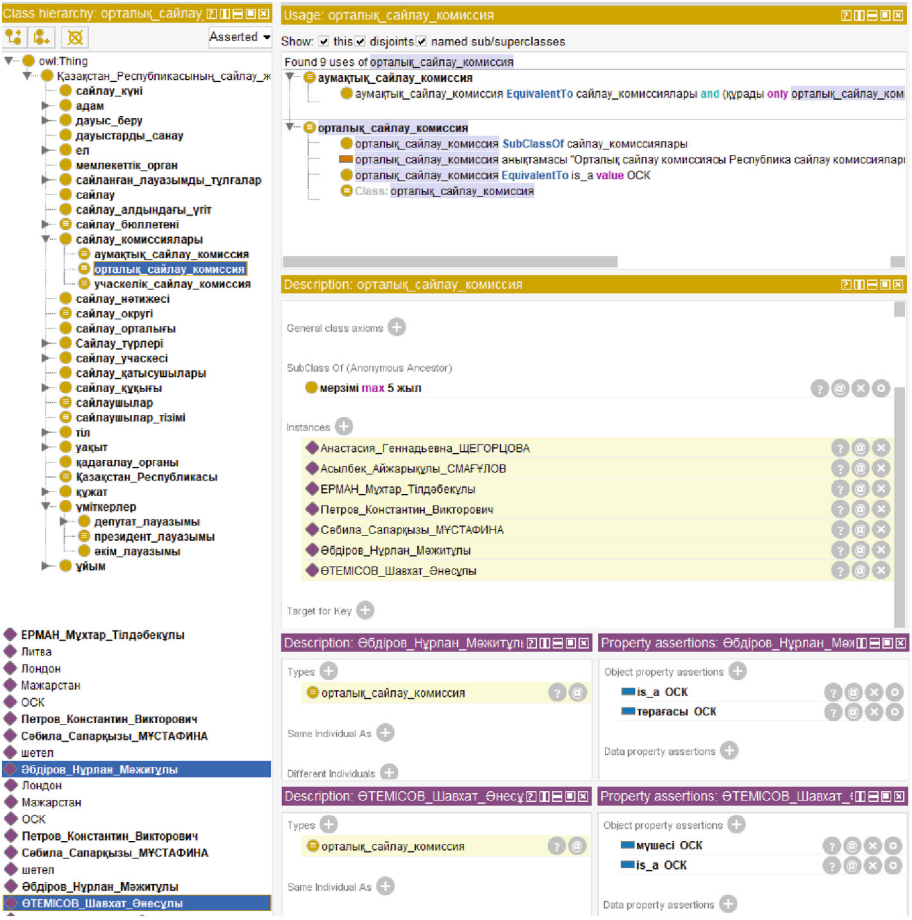

Fig. 4. Relationships and personalities in the ontological model

For example:

Question: “If there are 2 candidates marked on the ballot, is it valid or what ballot is it?” (Fig. 7).

Answer: Invalid

The question: “Who are the participants in the elections?” (Fig. 8).

Answer: election commission, voters, political parties, candidates (12).

Thus, the created ontology in the subject area “Electoral system of the Republic of Kazakhstan” includes rules, definitions, basic concepts and relationships that relate to the electoral system. Based on the rules, you can get new knowledge on demand from existing ones.

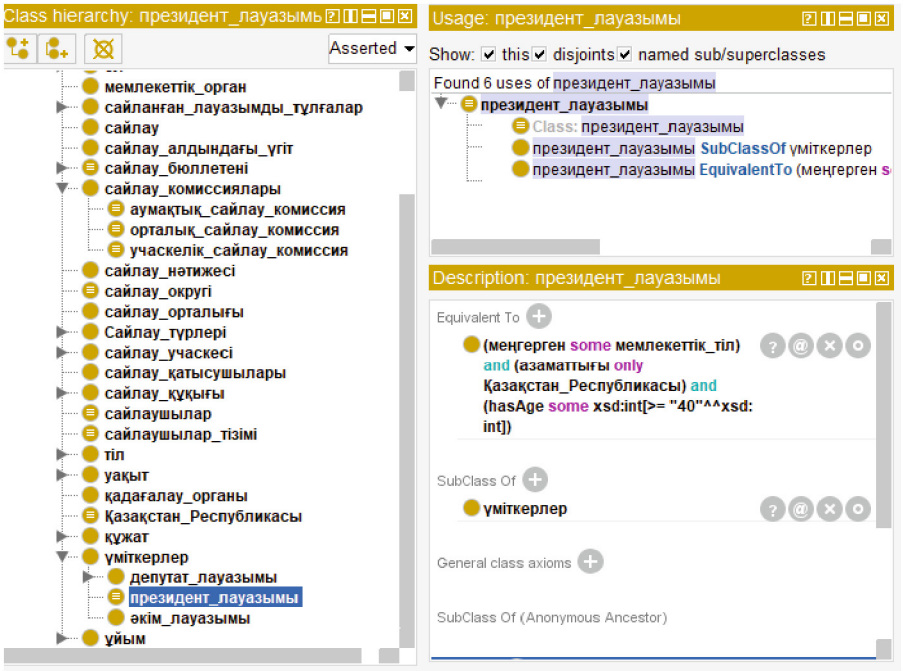

Fig. 5. Representation of axiomatic rules in the ontological model

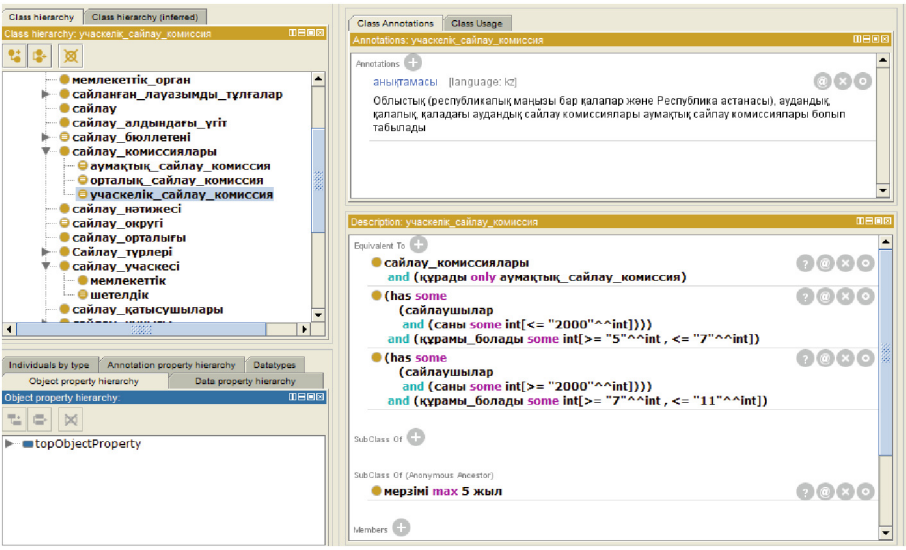

Fig. 6. Representation of axiomatic rules in the ontological model

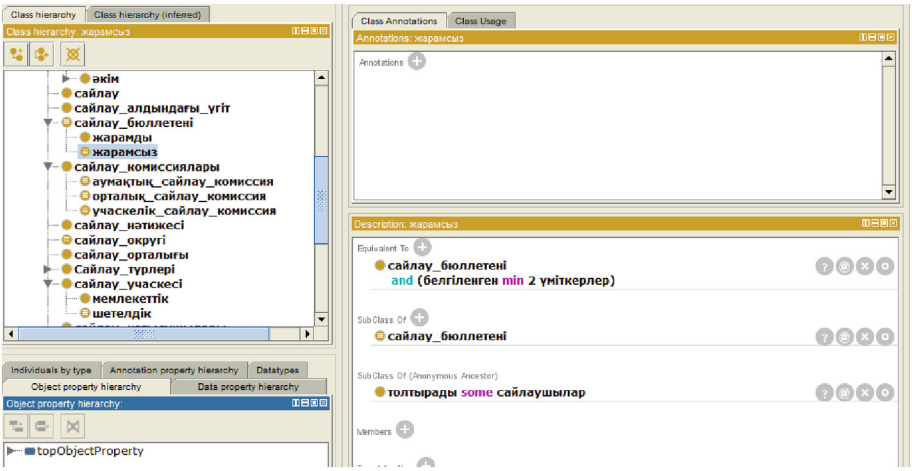

Fig. 7. Representation regarding the question "Is a ballot valid if 2 candidates are marked on the ballot or which ballot is it?"

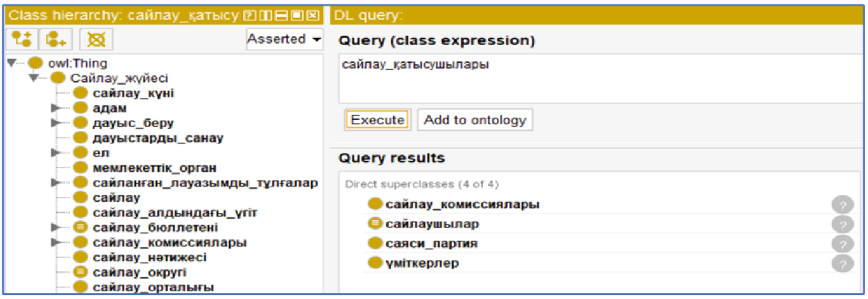

Fig. 8. Representation regarding the question "Who are the participants in the elections?"

4 Sentiment Analysis

The tasks of sentiment analysis are mainly implemented for English, Italian, Russian, Chinese and other languages, and research related to the sentiment analysis of texts in the Kazakh language has recently begun. Thus, there are very few ready-made tools, mathematical models, linguistic resources for the sentimental analysis of the Kazakh language.

To solve the problem of sentiment analysis, it is necessary to perform syntactic analysis after morphological analysis. For this, the syntactic rules of simple sentences of the Kazakh language were formalized and their constituent trees were built. The formalization of the syntactic rules of sentences was carried out using Chomsky's context-free (CS) grammar. During syntactic analysis, the components of the sentence are determined. More details about the sentiment analysis method can be found in the article [4].

Below is a model based on formalized production rules for determining the tonality of phrases in the text in the Kazakh language:

1) If the phrase consists of a noun with a positive coloring and a verb of a positive form with a neutral assessment following it, then the tone of this phrase is positive.

Displayed equations are centered and set on a separate line.

$$\frac{\omega \in L,\ \omega = \zeta \cdot \alpha \cdot \beta \cdot \xi,\ \alpha \in N,\ sent(\alpha) = 1,\ \beta = V_Post,\ sent(\beta) = 0}{sent(\omega) = 1} \tag{2}$$

here and below ζ, ξ – any chains of words, including empty ones. For example, “ инаугурация болды” (‘the inauguration was”), “ сыйыақы алды” (“received reward”).

2) If a phrase consists of a noun with a negative connotation followed by a positive verb with a neutral evaluation, then the tone of this phrase is negative.

$$\frac{\omega \in L,\ \omega = \zeta \cdot \alpha \cdot \beta \cdot \xi,\ \alpha \in N,\ sent(\alpha) = -1,\ \beta = V_Post,\ sent(\beta) = 0}{sent(\omega) = -1} \tag{3}$$

For instance, “ бұрмалау тіркелді” (“tampering has been reported”).

3) If the phrase consists a noun with a positive color and a verb of the positive form with a neutral color following it, and if after the verb there are words of negation (not/no), then the tonality of the phrase is negative.

$$\frac{\omega \in L,\ \omega = \zeta \cdot \alpha \cdot \beta \cdot \xi,\ \alpha \in N,\ sent(\alpha) = 1,\ \beta = V_{Post},\ sent(\beta) = 0,\ \gamma = @}{sent(\omega) = -1} \tag{4}$$

As an example, “ бостандық берген жоқ” (“did not give freedom”).

4) If a phrase consists a very positive adjective followed by a neutral noun, the tone of the phrase is very positive.

$$\frac{\omega \in L,\ \omega = \zeta \cdot \alpha \cdot \beta \cdot \xi,\ \alpha \in Adj,\ sent(\alpha) = 2,\ \beta = N,\ sent(\beta) = 0}{sent(\omega) = 2} \tag{5}$$

In this instance, “ құрметті депутат” (“honorable deputy”), “ көрнекті саясаткер” (“prominent politician”).

5) If a phrase consists an adjective with a very negative connotation followed by a noun with a neutral connotation, then the tone of this phrase is very negative.

$$\frac{\omega \in L,\ \omega = \zeta \cdot \alpha \cdot \beta \cdot \xi,\ \alpha \in A,\ sent(\alpha) = -2,\ \beta = N,\ sent(\beta) = 0}{sent(\omega) = -2} \tag{6}$$

For example, “ жемқор әкім” (“corrupt mayor”).

6) If a phrase consists an adjective with a very positive coloring, a negative particle (emes) and a noun with a positive coloring following it, then the phrase has a negative coloring.

$$\frac{\omega \in L,\ \omega = \zeta \cdot \alpha \cdot \beta \cdot \gamma \cdot \xi,\ \alpha \in Adj,\ sent(\alpha) = 2,\ \beta = @,\ \gamma \in N,\ sent(\gamma) = 1}{sent(\omega) = -1} \tag{7}$$

For instance, “ биік емес мәртебе” (“not a high status”).

7) If a phrase consists a positive adjective followed by a neutral positive verb, then the tone of the phrase is positive.

$$\frac{\omega \in L,\ \omega = \zeta \cdot \alpha \cdot \beta \cdot \xi,\ \alpha \in Adj,\ sent(\alpha) = 1,\ \beta \in V_Post,\ sent(\beta) = 0}{sent(\omega) = 1} \quad (8)$$

As an example, “ дұрыс таңдайды” (“chooses correctly”).

8) If a phrase consists an intensifying adverb (“ өте” (“very”), “ ең” (“the most”)), followed by a positive adjective and then a noun with a neutral color, then this phrase has a very positive color.

$$\frac{\omega \in L,\ \omega = \zeta \cdot \alpha \cdot \beta \cdot \gamma \cdot \xi,\ \alpha \in AdvIntens,\ \beta = Adj,\ sent(\beta) = 1,\ \gamma = N,\ sent(\gamma) = 0}{sent(\omega) = 2} \quad (9)$$

In this instance, “ өте перспективті саясаткер” (“the most perspective politician”).

The tonality of the entire text is defined as the arithmetic mean of the measurements of the tonality of lexical units (sentences) and the rules for their combinations.

$$sent(L) = \frac{\sum_{i=1}^{n} sent(\omega)}{n} \quad (10)$$

This emotional speech recognition method is used to determine the mood of the conversation in these sources by analyzing the tone of texts obtained from official and unofficial political information sources posted on social media in the Kazakh language.

5 Conclusion

In order to address the practical issue of identifying the sources of political discourse, the tone of discussion in these sources, identifying hotspots of inciting hatred, negativity, enmity, or vice versa, and highlighting the political events, phenomena, and discussions that are most highly regarded by society. The resolution of this problem will aid in the implementation of the Hearing State Concept, and the sophisticated artificial intelligence methods used to complete the tasks will enhance the Republic of Kazakhstan’s degree of technological and scientific advancement.

References

1. Bekmanova, G., Yelibayeva, G., Aubakirova, S., Dyussupova, N., Sharipbay, A., Nyazova, R.: Methods for analyzing polarity of the Kazakh texts related to the terrorist threats. In: Misra, S., et al. (eds.) ICCSA 2019. LNCS, vol. 11619, pp. 717–730. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-24289-3_53
2. Yergesh, B., Bekmanova, G., Sharipbay, A.: Sentiment analysis of Kazakh text and their polarity. Web Intell. **17**(1), 9–15 (2019). https://doi.org/10.3233/WEB-190396
3. Bekmanova, G., Yergesh, B., Sharipbay, A.: Sentiment analysis model based on the word structural representation. In: Mahmud, M., Kaiser, M.S., Vassanelli, S., Dai, Q., Zhong, N. (eds.) BI 2021. LNCS (LNAI), vol. 12960, pp. 170–178. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-86993-9_16

4. Bekmanova, G., Yergesh, B., Sharipbay, A., Mukanova, A.: Emotional speech recognition method based on word transcription. Sensors **22**(5) (2022). https://doi.org/10.3390/s22051937
5. Yergesh, B., Bekmanova, G., Sharipbay, A.: Sentiment analysis on the hotel reviews in the Kazakh language. In: Paper Presented at the 2nd International Conference on Computer Science and Engineering, UBMK 2017, pp. 790–794 (2017). https://doi.org/10.1109/UBMK.2017.8093531
6. Yergesh, B., Bekmanova, G., Sharipbay, A., Yergesh, M.: Ontology-based sentiment analysis of Kazakh sentences. In: Gervasi, O., et al. (eds.) ICCSA 2017. LNCS, vol. 10406, pp. 669–677. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-62398-6_47
7. 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**(2), 257–263 (2016)
8. Bekmanova, G., et al.: A uniform morphological analyzer for the Kazakh and Turkish languages. In: Paper Presented at the CEUR Workshop Proceedings, pp. 20–30 (2017)
9. Raxmatovna, B.N.: Specific features of political speech. Central Asian J. Lit. Philos. Cult. **3**(12), 80–87 (2022)
10. Tameryan, T.Yu., et al.: Political media communication: bilingual strategies in the pre-election campaign speeches. Online J. Commun. Media Technol. **9**(4), e201921 (2019)
11. Al Maani, B., et al.: The positive-self and negative-other representation in Bashar Al-Assad's first political speech after the Syrian uprising. Theory Pract. Lang. Stud. **12**(10), 2201–2210 (2022)
12. Sotvoldiyevna, U.D.: Political Euphemisms in English and Uzbek languages (A comparative analysis). Eurasian J. Learn. Acad. Teach. **9**, 92–96 (2022)
13. Dave, P.: Analysis of the political power speeches of Jr. Martin Luther King and Barrack Obama: in the light of critical discourse analysis as a literary research method. Vidhyayana-Int. Multi. Peer-Rev. E-Journal-ISSN **7**(5), 2454–8596 (2022)
14. Abdurashetona, A.M., Ismailovich, I.O.: Methods of tagging part of speech of Uzbek language. In: 2021 6th International Conference on Computer Science and Engineering (UBMK). IEEE (2021)
15. Fiorelli, M., et al.: Metadata-driven semantic coordination. In: Garoufallou, E., Fallucchi, F., William De Luca, E. (eds.) MTSR 2019. CCIS, vol. 1057, pp. 16–27. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-36599-8_2
16. Langer, A.M.: Analysis and Design of Next-Generation Software Architectures. Springer, New York (2020). https://doi.org/10.1007/978-3-030-36899-9
17. Lai, C.: Fast retrieval algorithm of English sentences based on artificial intelligence machine translation. In: Atiquzzaman, M., Yen, N., Xu, Z. (eds.) 2021 International Conference on Big Data Analytics for Cyber-Physical System in Smart City, vol. 102. Springer, Singapore (2022). https://doi.org/10.1007/978-981-16-7466-2_117
18. Abdurashetona, A.M., Mokhiyakon, U.: Software features and linguistic features of Uzbek Synonymizer. In: 2022 7th International Conference on Computer Science and Engineering (UBMK). IEEE (2022)
19. Bekmanova, G., et al.: Linguistic foundations of low-resource languages for speech synthesis on the example of the Kazakh language. In: Gervasi, O., Murgante, B., Misra, S., Rocha, A.M.A.C., Garau, C. (eds.) Computational Science and Its Applications–ICCSA 2022 Workshops: Malaga, Spain, 4–7 July 2022, Proceedings, Part III. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-10545-6_1
20. Ibrahim, M.: A corpus-based comparative analysis of assertive strategies in Pakistani democratic and dictatorial speeches. J. Appl. Linguist. TESOL **5**(4), 6–19 (2022)
21. Mohammed, T.A.S., Banda, F., Patel, M.: The Topoi of Mandela's death in the Arabic speaking media: a corpus-based political discourse analysis (2022)

22. Liu, M.: Stancetaking in Hong Kong political discourse: a corpus-assisted discourse study. Chin. Lang. Discourse **13**(1), 79–98 (2022)
23. Afzaal, M.: "Review of Literature." A Corpus-Based Analysis of Discourses on the Belt and Road Initiative: Corpora and the Belt and Road Initiative, pp. 17–37. Springer, Singapore (2023)
24. Anand, S., Keefer, R.: From description to code: a method to predict maintenance codes from maintainer descriptions. Maintenance Reliab. Condition Monit. **2**(2), 35–44 (2022)
25. Ma, Y., et al.: An end-to-end dialogue state tracking system with machine reading comprehension and wide & deep classification. arXiv preprint arXiv:1912.09297 (2019)
26. Saravanan, S., Sudha, K.: GPT-3 powered system for content generation and transformation. In: 2022 Fifth International Conference on Computational Intelligence and Communication Technologies (CCICT). IEEE (2022)
27. Dmytriv, A., et al.: Comparative analysis of using different parts of speech in the Ukrainian texts based on stylistic approach. In: CEUR Workshop Proceedings, vol. 3171 (2022)
28. Tretyakov, E., et al.: Sentiment analysis of social networks messages. In: Klimov, V.V., Kelley, D.J. (eds.) Biologically Inspired Cognitive Architectures 2021: Proceedings of the 12th Annual Meeting of the BICA Society. Springer, Cham (2022). https://doi.org/10.1007/978-3-030-96993-6_61
29. Goswami, S., Hudnurkar, M., Ambekar, S.: Fake news and hate speech detection with machine learning and NLP. PalArch's J. Archaeol. Egypt/Egyptol. **17**(6), 4309–4322 (2020)
30. Lee, E., et al.: Racism detection by analyzing differential opinions through sentiment analysis of tweets using stacked ensemble GCR-NN model. IEEE Access **10**, 9717–9728 (2022)
31. Alshalan, R., Al-Khalifa, H.: A deep learning approach for automatic hate speech detection in the Saudi Twittersphere. Appl. Sci. **10**(23), 8614 (2020)
32. Chu, K.E., Keikhosrokiani, P., Asl, M.P.: A topic modeling and sentiment analysis model for detection and visualization of themes in literary texts. Pertanika J. Sci. Technol. **30**(4), 2535–2561 (2022)
33. Babu, N.V., Kanaga, E.G.M.: Sentiment analysis in social media data for depression detection using artificial intelligence: a review. SN Comput. Sci. **3**, 1–20 (2022)
34. Perifanos, K., Goutsos, D.: Multimodal hate speech detection in Greek social media. Multimodal Technol. Interact. **5**(7), 34 (2021)
35. Aljarah, I., et al.: Intelligent detection of hate speech in Arabic social network: a machine learning approach. J. Inf. Sci. **47**(4), 483–501 (2021)
36. Koltsova, O., et al.: PolSentiLex: sentiment detection in socio-political discussions on Russian social media. In: Filchenkov, A., Kauttonen, J., Pivovarova, L. (eds.) Artificial Intelligence and Natural Language: 9th Conference, AINL 2020, Helsinki, Finland, 7–9 October 2020, Proceedings. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-59082-6_1
37. Mahmud, Md.A.I., et al.: Toward news authenticity: synthesizing natural language processing and human expert opinion to evaluate news. IEEE Access **11**, 11405–11421 (2023)
38. Widodo, D.A., Iksan, N., Sunarko, B.: Sentiment analysis of Twitter media for public reaction identification on COVID-19 monitoring system using hybrid feature extraction method. Int. J. Intell. Syst. Appl. Eng. **11**(1), 92–99 (2023)
39. Holt, K., Ustad Figenschou, T., Frischlich, L.: Key dimensions of alternative news media. Digital Journalism **7**(7), 860–869 (2019). High-Choice Information Environments, vol. 25
40. Chang, W.-L., Tseng, H.-C.: The impact of sentiment on content post popularity through emoji and text on social platforms. In: Cyber Influence and Cognitive Threats, pp. 159–184. Academic Press (2020)
41. Dang, C.N., Moreno-García, M.N., De la Prieta, F.: An approach to integrating sentiment analysis into recommender systems. Sensors **21**(16), 5666 (2021)

42. Wu, C., et al.: SentiRec: sentiment diversity-aware neural news recommendation. In: Proceedings of the 1st Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 10th International Joint Conference on Natural Language Processing (2020)
43. Rozado, D., Al-Gharbi, M., Halberstadt, J.: Prevalence of prejudice-denoting words in news media discourse: a chronological analysis. Soc. Sci. Comput. Rev. 08944393211031452 (2021)
44. Oladele, T.M., Ayetiran, E.F.: Social unrest prediction through sentiment analysis on Twitter using support vector machine: experimental study on Nigeria's# EndSARS. Open Inf. Sci. **7**(1), 20220141 (2023)