Analytical Review of Methods for Identifying Emotions in Text Data*

Anastasia Dvoynikova
anastasia.dvoynikova@gmail.com

Oxana Verkholyak
overkholyak@gmail.com

Alexey Karpov
karpov@iias.spb.su

St.Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences (SPIIRAS),
St.Petersburg, Russia

Abstract

The sentiment analysis of text is one of the important tasks in the field of natural language processing. It is used in different areas. Despite the variety of existing methods, the systems of sentiment analysis of Russian-language texts give low accuracy compared to English-language ones. The article discusses basic methods for identifying emotions in text data and methods of text vectorization. The existing achievements in the field of computer sentiment analysis are analyzed. At the moment, there are many unsolved problems in the field of automatic sentiment analysis.

Keywords: *Sentiment-Analysis, Tonality of Text, Vectorize of Text, Machine Learning.*

1 Intoduction

Nowadays, a huge stream of information passes through the Internet, including the communication of network users. There are many open text sources that display data on people's opinions on various issues. To get more complete statistics of opinions, an analysis of the tonality of the text is necessary.

Sentiment-analysis is a field of computer linguistics and intellectual analysis of the text, focused on extracting subjective opinions and emotions from it [Minakov, 2013]. Sentiment-analysis finds practical application in many areas: assessing the quality of goods and services based on customer reviews on the Internet, analyzing negative emotions in messages, forecasting stock markets, political situations based on news feeds [Romanov et al, 2018]. Also, sentiment-analysis is necessary in automated systems in which a person communicates with a machine in natural language, for example, to analyze message histories [Bondareva and Lagerev, 2018]. In order to analyze such a volume of information in recent years, various methods have been proposed for automatically determining the tonality of the text, which will be considered in this article.

2 Methods of sentiment-analysis

There are several basic methods for determining the tonality of a text [Zvereva, 2014]. All of them can be divided into several categories (Figure 1).

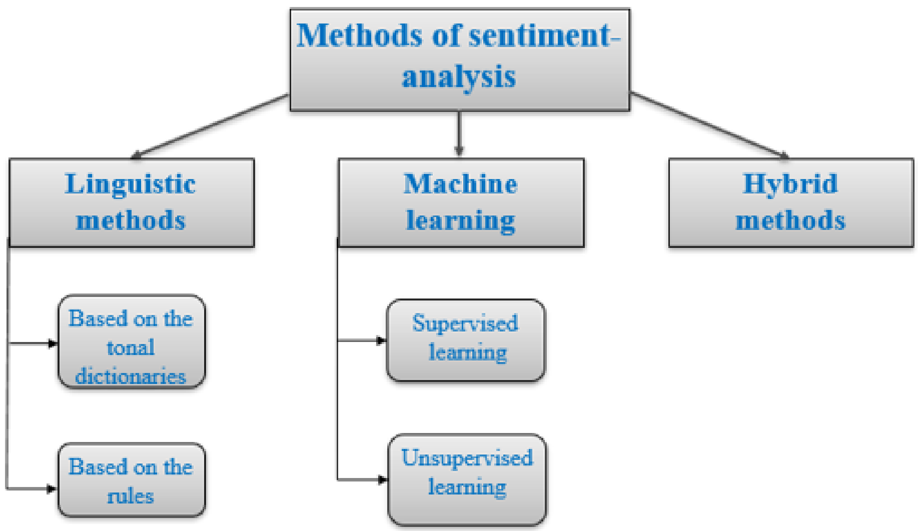

Figure 1: Classification of methods of sentiment-analysis

A description of each method is provided below.

1. Linguistic methods.
   (a) Method based on the tonal dictionaries.

   A tonal dictionary is a set of words or bigrams that are given a certain probability (weight) of belonging to a positive (positive weight) or negative (negative weight) class. The range of weights may be different depending on the dictionary. When analyzing a text, each word is searched in this dictionary and its weight is recorded. If the word is not in the dictionary, then its class is considered neutral and the weight will be zero. After all the weights are obtained, the membership of the given text in a certain tonality class is calculated. For this, the arithmetic mean of weights is most often calculated, and in rare cases, the sum of the weights is used, or artificial neural networks are used. Method based on the tonal dictionaries was used in the works [Tutubalina et al, 2015] [Posevkin and Bessmertniy, 2015].

   (b) Method based on the rules.

   This method requires a large set of rules of construction “if - then" rules. For example, if the particle “not" stands before an adjective of positive coloring, then this construction can be classified as negative. This method also implies the use of tonal dictionaries, in which words belong to a certain class (positive, negative, neutral, etc.). The sentiment analysis problem was solved using the method based on the rules in the works [Kan, 2011] [Panicheva, 2013].

2. Machine learning methods.

   (a) Supervised learning.

   The method is based on training the classifier on pre-annotated training text data [Kormalev, 2004] [Kotel'nikov and Klekovkina, 2012]. The most common methods in the field of tonal analysis are the naive Bayesian classifier and the support vector machine method.

   The naive Bayesian classifier (NB) is a probabilistic classifier based on the application of Bayes theorem with the assumption of class independence. It is used by the authors of the work [Lewis,1998]. Support Vector Machines (SVM) is a linear classifier. The main idea of the method is to construct a hyperplane separating the sample objects in the most optimal way. The algorithm works under the assumption that the greater the distance (gap) between the separating hyperplane and the objects of the common class, the smaller the average error of the classifier will be. The authors in [Zainuddin et al, 2014] used the support vector method to determine the tonality of the text.

   Neural networks are a set of structured computational elements that mimic the functioning of the human brain [Barskij, 2004]. Neural networks allow modelling the relationships between input and output data. In the field of sentimental analysis, the most common neural networks such as convolutional neural network (CNN) [Spiros et al, 2018] and recurrent neural networks (RNN) such as a neural network with a long short-term memory (LSTM) [Ben Amar et al, 2018] and Gated recurrent unit (GRU) [Aken et al, 2018].

   (b) Unsupervised learning.

   In contrast to the above method, the unsupervised learning method determines the relationship and patterns between objects without labeled data [Voronina and Goncharov, 2015]. Such methods include the Gaussian mixture model and k-nearest neighbors.

   K-means - the algorithm is based on the search of k training samples, the distance to which from the given sample is minimal. The most encountered class among k objects will be the class of the object of interest. This method was used in the article [Prabin Lama, 2013].

   The Gaussian mixture model is a probabilistic model that assumes that all data points are generated from a mixture of a finite number of Gaussian distributions with unknown parameters. The authors in [Pribill et al, 2014] used the GMM model in their work.

3. Hybrid methods Methods combining several different methods described above [Pazel'skaya and Solov'ev, 2011], [Krasnikov and Nikulichev, 2013]. In the article [Konig and Brill, 2006] a hybrid method was used for the problem of text classification, which includes a method based on tonal dictionaries and a method of support vectors. The authors achieved 72% accuracy using this method.

3 Vector representation of documents

Before using a machine classifier, it is necessary to present the text in numerical form (feature extraction). There are several ways to vectorize text [Hassaan Saeed, 2018], [Dhingra et al, 2017], they are presented in figure 2.

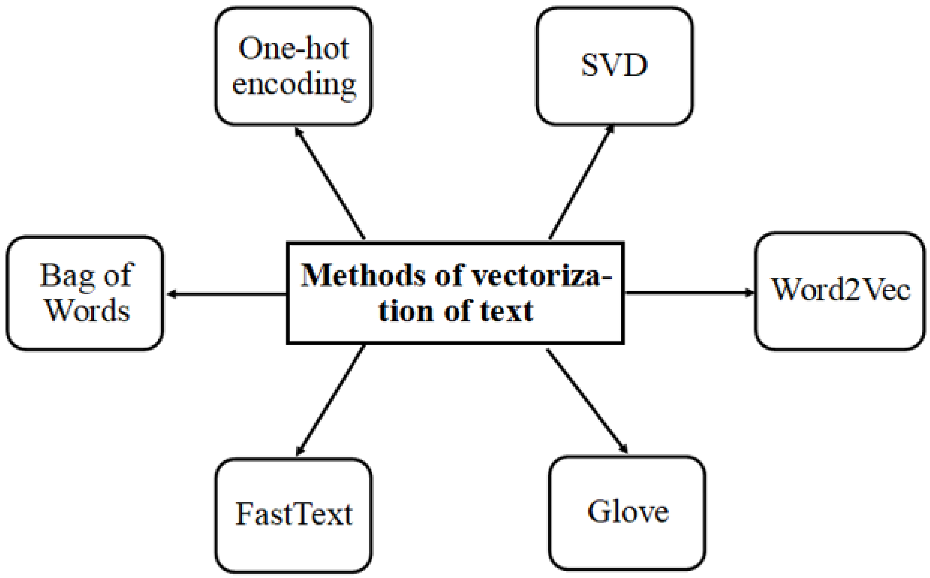

Figure 2: Methods of vectorization of text

- Bag of Words (BoW) – a model that represents a text as an unordered set of words [Soumya George et al, 2014]. Each word is assigned its own weight. Often used is TF-IDF (Term Frequency - Inverse Document Frequency)-word weight, reflecting the ratio of the word frequency in the document to the word frequency in all documents [Qaise and Ramsha Ali, 2018].

- One-hot encoding (direct encoding) – a method that converts words into vectors. The size of each vector is equal to the volume of all words in the text. Each element in the word vector is binary and consists of 0 and 1. Before encoding, all words that are present in the text are arranged alphabetically [Kedar Potdar et al, 2017], [Chaubard, 2016].

- Singular Value Decomposition (SVD) – a method that converts text into a sparse matrix $A_{mn} = \{a_{i1}, a_{i2}, a_{i3}, \ldots, a_{in}\}$, where a_{ij} is a weighted column vector of the frequency of the sentence members i in the document. If the document contains only m terms and n sentences, then the output will be a matrix of dimension m n [Jezek and Steinberger, 2004].

- Word2Vec (Toolkit developed by Google) is a neural network that generates word vectors. It is trained on two algorithms: BoW (predicts the word given context) and Skip-gram (predicts the context given word). Word2Vec first builds a dictionary from a learning text corpus and analyzes the vector representations of each word. In addition, Word2Vec can calculate the cosine distance between each word [Ma and Zhang, 2015].

- Glove is a method developed at Stanford University (USA). It is based on a method of calculating the frequency of words in the text corpus. In fact, it consists of two main stages. The first stage is the construction of the adjacency matrix from the training corpus. The second stage is matrix factorization to obtain vectors [Pennington et al, 2014].

- FastText is a model that converts into vectors not only words, but also symbolic n-grams from which words are composed. Due to this, it seems possible to calculate vector representations of unknown words [Armand et al, 2016].

4 Evaluation of classification results

The classification procedure is followed by a quantitative assessment of the results, which can be carried out using a set of the following indicators:

- Accuracy
$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

- Precision
$$Precision = \frac{TP}{TP + FP} \tag{2}$$

- Recall
$$Recall = \frac{TP}{TP + FN} \tag{3}$$

- F-score
$$F - score = 2 * \frac{Precision * Recall}{Precision + Recall} \tag{4}$$

where TP – true-positive, FP – false-positive, TN – true-negative, FN – false-negative.

5 Experimental computer systems for detecting emotions in text

On the Internet platform Kaggle in 2018, international competitions were held to identify “toxic" (containing negative emotions) comments (Jigsaw Toxic Comment Classification Challenge) [Kaggle, 2018]. In total, about 4,500 teams participated in them. The contestants had to create a model that would detect different types of “toxicity" of the text, such as: threat, obscenity, insults and hatred.

The Google team introduced the Jigsaw text data set. The corpus consists of various English-language text comments from Wikipedia. Each text has one or more labels of six: toxic, severe toxic, obscene, threatening, insulting, and hateful to the individual. The data set is divided into training and test sets. The number of samples of the training and test data set – 159571 and 63978 comments, respectively [Kaggle Data, 2018].

Researchers in [Saif et al, 2018] tested on the Jigsaw database such models as logistic regression, as well as three models of neural networks (convolutional neural network-CNN, neural network with long short-term memory(LSTM) and combined CNN+LSTM). All comment texts were tokenized using the CountVectorizer from the scikit-learn Python library. The best result was shown by the combined neural network (CNN+LSTM), which consists of 2 layers of LSTM and 4 layers of CNN. The accuracy of classification for 6 classes was 96.45%.

In [Spiros et al, 2018], traditional methods (the Naive Bayes classifier, the method of k-nearest neighbors, the method of support vectors, and linear discriminant analysis) are compared with a neural network. A convolutional neural network trained on the word2vec model was used (the word2vec dictionary contains about 3 million words of English taken from Google news). The traditional support vector machine approach showed 81.1% accuracy, but CNN surpassed all traditional approaches and showed 91.2% accuracy.

The authors of the paper [Noever, 2018] suggested that the combination of several traditional methods to improve the accuracy of recognition of toxic comments. Thus, when testing a method based on a random forest on the Jigsaw database, the accuracy is 57.82%. If regression trees, the support vector method, and logistic regression are added to this method, the accuracy increases to 62.82%.

The article [Elnaggar et al, 2018] describes the idea of using a combined neural network model. The authors used a network consisting of a word embedding layer using the Glove method (which allows for each word in the text data to obtain a corresponding fixed-length vector using statistical information about this word), 2 layers of a recurrent neural network (GRU and LSTM) and 6 layers of a convolutional neural network. In this paper, the result of the work is estimated using the F-score = 79%.

In [Mai et al, 2018] the authors predicted toxicity of the comment in 2 stages: first, the comment was toxic or not, and then, if toxic, the type of toxicity was determined. The work used an ensemble of neural networks-CNN+LST+GRUB, which showed F-score=87.2%.

The authors of the work [Ben Amar et al, 2018] conducted experiments not with the search for the best model, but with various methods of preprocessing texts. They used the LSTM neural network as a model. After the experiment, the researchers decided to focus on the best, in their opinion, methods of preprocessing the text. They removed stop words and links from the text, normalized only bad words to increase their toxicity weight, and included translation of bad words, in case the word was not in an English dictionary, it would be searched in dictionaries of other languages. This approach showed an accuracy of 97.72%.

Researchers [Aken et al, 2018] in their work compared different neural networks: convolutional and recurrent. The best accuracy on Jigsaw data was shown by a bidirectional neural network such as GRU. Moreover, using the vector representation of both Glove and FastText the accuracy was the same 98.3%.

The results of all studies in the Toxic Comment Classification Challenge are summarized in the diagram shown in figure 3.

6 Computer systems of sentiment analysis of Russian texts

Every year, the international conference "Dialogue" competitions of automatic processing systems of the Russian language – Dialogue Evaluation [Dialogue Evaluation] are held. One of the main topics of the competition was sentiment-analysis of the text.

So, in 2012, the ROMIP database [ROMIP] was provided for the competition, which included people's reviews of various films, books and digital cameras. In total, there are about

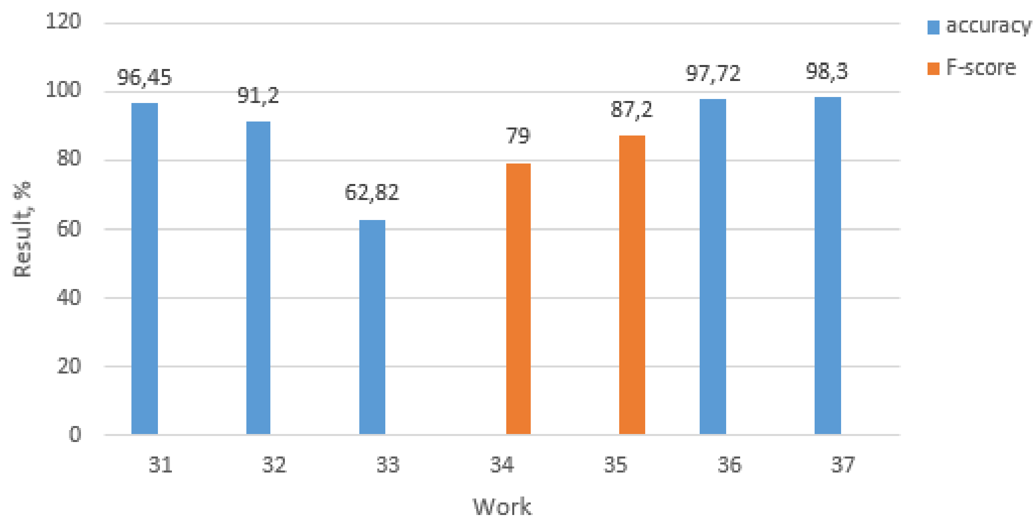

Figure 3: Result of Toxic Comment Classification Challenge

50 thousand text fragments containing people's opinions about the product. Tonality marking was carried out on a 5-point scale. Participants were asked to classify reviews into 2, 3 and 5 classes. The work [Pak and Paroubek, 2012] showed the best result of classification on 5 classes. The authors used the n-gram method of support vectors, used binary weights instead of the traditional TF-IDF, and trained the model on a combined database consisting of reviews of movies, books and cameras. Applying this approach on the ROMIP database, the average value of the F-score = 30.63% is obtained.

The competition in 2013 used the same database as in 2012, in addition to it, the organizers included a database of news feeds, which contained direct and indirect speech with an assessment of the sentiment of the statement [ROMIP]. The score can take one of 4 values: positive, negative, mixed score or no score. This database contains about 5 thousand news fragments. The authors of [Blinov et al, 2013] achieved the best value of F-score=65.9% in binary classification, and 35.36% in classification into 5 classes. When dividing the data into 2 classes, the authors used the method of maximum entropy, and when dividing into 5 classes-the method of support vector machine.

In 2015, Dialogue Evaluation was assigned a broader task. Participants were provided with the SentiRuEval-2015 database [SentiRuEval], which includes reviews of restaurants and cars. The volume of the database was about 18 thousand reviews. SentiRuEval-2015 contained the text of the review, selected target aspects, i.e. components, or characteristics of the assessed object. For restaurant themes such aspects are kitchen, interior, service, price. For cars, the list of aspects includes safety, comfort, reliability, appearance, prices, road quality. Participants had to perform several tasks: to identify the aspect terms (a set of terms in which the target aspect is expressed), to determine their tone and the immediate response itself. The best result of solving this problem is described in [Tarasov, 2015]. The author used recurrent neural networks and obtained the result of the F-score equal to 61.9% and 64.7% for restaurants and cars, respectively.

In 2016, the competition was held on the SentiRuEval-2016 database. It includes feedback from Twitter (tweets) about banks and mobile operators. In addition to the text of the review itself, the database contains information about which object the review refers to and about the assessment of tonality in the range from -1 to 1 (negative, neutral and positive). The task of the participants was to determine the reputational attitude of the tweet in relation to a company. The authors of work [Arkhipenko et al, 2016] used a two-layer GRU neural network and fed the input vector in the reverse sequence. Using this method, the F-score = 55.17% and 55.94% for banks and mobile operators, respectively, was achieved.

Several works on sentiment-analysis of Russian-language texts outside the competition are known. Thus, in [Mirzayanova, 2019] the author collected his own database from the site kinopoisk.ru, containing people's reviews of various movies with scores ranging from 1 to 10. The volume of the database is about 1 million words. For experiment all reviews were divided first on 3, and then on 5 classes. Various methods were used for classification: the Naive Bayes classifier, the support vector method, as well as neural networks: multilayer perceptron and LSTM networks. The naive Bayes classifier showed the best results in classification. The average value of the F-score was 72% and 26.8% when classified into 3 and 5 classes, respectively.

The author of the work [Bob'yakova, 2017] used a combination of several approaches: supervised machine learning, namely a naive Bayes classifier, and a dictionary approach. The frequency dictionary initially included 100 words with the maximum frequency in the Russian language according to the "Frequency Dictionary of the Modern Russian Language". But since most of the words in the compiled dictionary turned out to be pronouns, conjunctions, prepositions, the dictionary was reduced to 29 words. The experiments were conducted on a database consisting of 100,000 texts from Twitter. The author used the removal of stop words, links, hashtags, as well as words with the maximum frequency of occurrence for the preprocessing of texts. Documents were presented in the form of bigrams and unigrams. With this approach, precision reached 86.6%, and recall - 89.1%.

In work [Barskij, 2004] for the sentiment-analysis the database from "posts" of a social network "Vkontakte" with the emotional coloring corresponding to them was used. The volume of the database was about 3000 texts. To classify texts, the author applied a naive Bayes classifier. This method showed an accuracy of 70%.

7 Discussion

Every year the number of conferences in the field of sentimental analysis increases, and the number of publications on the analysis of the text in Russian and in other foreign languages grows. According to the Google Academy [Google scholar] in 2018, about 700 works on sentimental analysis of Russian-language texts were published, while about 8,500 works on English-language texts were published. Also, based on the above material, it can be determined that the researchers obtained the accuracy of tonal analysis of Russian-language texts about 80%, and in English-language texts the accuracy reaches 96%.

8 Conclusion

The presence of numerous works on the topic of sentient-analysis suggests that the topic is relevant today and is in demand in many areas such as the economic market, politics, marketing, etc. But, as can be seen from the analytical review, the systems of sentiment-analysis of Russian-language texts are less developed than foreign texts. Also, Russian-language sentiment analysis gives a rather low accuracy compared to English. Therefore, now, the task of improving the accuracy of tonal analysis of Russian-language texts remains relevant.

Acknowledgements

This research is supported by RFBR projects No. 18-07-01407 and 19-29-09081.

References

[Minakov, 2013] Minakov I.A. (2013) *Sentiment-analisis of the text and its application to improve the quality of transitions on relevant ads.* Bulletin of Samara state technical University. Technical science, 1(37), 58-63, Samara, Russia (In Rus.) = Analis emotional'noy tonal'nosti texta i ego primenenie dlya povishenia kachestva perehodov po relevantnim obyavleniyam: Vestnic Samarskogo gosudarstvennogo universiteta: Tehnicheskie nauki, 1(37), 58-63, 2013.

[Romanov et al, 2018] Romanov A.S., Vasil'eva M.I., Kurtukova A.V. and Mesheryakov R.V.(2018) *Text tonality analysis using machine learning techniques.* Proceedings 2nd International Conference "R. PIOTROWSKI'S READINGS LE & '2017", 86-95, Saint-Petersburg, Russia (In Rus.) = Analiz tonal'nosti texta s ispol'zovaniem metodov mashinogo obuchenia. Proceedings 2nd International Conference "R. PIOTROWSKI'S READINGS LE & AL'2017", 86-95, 2018.

[Bondareva and Lagerev, 2018] Bondareva I.V. and Lagerev D.G. (2018) *Research of methods of vector representation of textual information for solving the problem of tonality analysis.* All-Russian scientific conference "Information technologies of intellectual decision support", 10-15, Ufa-Stavropol, Russia (In Rus.) = Issledovanie metodov vectornogo predstavleniya textovoi informacii dlya reshenia zadachi analiza tonal'nosti. "Informacionie tehnologii intelectual'noy podderzhki prinyatiya resheniy", 10-15, 2018.

[Zvereva, 2014] Zvereva P.P. (2014) *Sentimental analysis of the text (on the material of printed texts of the newspaper "The New York Times" about Russia and Russians).* // Bulletin of the Moscow state University. Series: Linguistics. 5, 32-37, Moscow, Russia (In Rus.) = Sentient analiz texta (na materiale pechatnih textov gazeti "The New York Times" o Rossii I rossianah). Vestnik MGOU, seria: Lingvostica, 5, 32-37, 2014.

[Tutubalina et al, 2015] Tutubalina E.V., Ivanov V.V., Zagulova M.A., Mingazov N.R., Alimova I.S. and Malih V.A. (2015) *Testing of methods of text tonality analysis based on dictionaries.* Electronic libraries, 18(3-4), 138-160 (In Rus.) = Testirovanie metodov analiza tonal'nosti texta osnovanih naslovaryah. Electronnie biblioteki, 18(3-4), 138-160, 2015.

[Posevkin and Bessmertniy, 2015] Posevkin R.V., Bessmertniy I.A.(2015) *Application of sentimental analysis of texts for evaluation of public opinion.* Scientific and technical Bulletin of information technologies, mechanics and optics, 15(1), 169-171 (In Rus.) = Prienenie sentiment analiza textov dlya ocenki obshestvennogo mnenia. Nauchno-tehnicheskiy vestnic inforacionnih tehnologiy mehaniki i optiki, 15(1), 169-171, 2015.

[Kan, 2011] Kan D. (2011) *Rule-based approach to sentiment analysis at ROMIP .* Avaible at http://www.dialog-21.ru/digests/dialog2012/materials/pdf/Kan.pdf

[Panicheva, 2013] Panicheva P. (2013) *The system of sentimental analysis ATEX, based on the rules, in the processing of texts of various subjects.* Conference "Dialogue", 2, 101-113 (In Rus.) = Sistema sentiment analiza ATEX osnovanaya na pravilah pri obrabotke textov razlichnih thematic. Komputernaya lingvistica i intelectual'nie tehnologii, 2, 101-113, 2013.

[Kormalev, 2004] Kormalev D.A. (2004) *Applications of machine learning methods in text analysis problems.* Software systems: theory and applications: proceedings of the International conference, Pereslavl-Zalessky, 2, 35-48 (In Rus.) = Prilozhenia metodov mashinnogo obuchenia v zadachax analiza texta. Trudi conferencii: programmnie sistemi: teoria I prilozhenia 2, 35-48, 2004.

[Kotel'nikov and Klekovkina, 2012] Kotel'nikov E.V., Klekovkina M.V. (2012) *Automatic text tonality analysis based on machine learning methods.* Conference "Dialogue", 11 (18), 27–36 (In Rus.) = Avtomaticheskij analiz tonal'nosti tekstov na osnove metodov mashinnogo obucheniya. Komp'yuternaya lingvistika i intellektual'nye tekhnologii, 11 (18), 27–36, 2012.

[Lewis,1998] Lewis D.D. (1998) *Naive (Bayes) at forty: The independence assumption in information retrieval.* Proceedings of 10th European Conference on Machine Learning, 4–15.

[Zainuddin et al, 2014] Zainuddin, Nurulhuda, Selamat(2014) *Sentiment Analysis Using Support Vector Machine.* International Conference on Computer, Communication and Control Technology, 333-337.

[Barskij, 2004] Barskij A.B. (2004) *Neural networks: recognition, management, decision-making.* Finance and statistics. 176 (In Rus.) = Nejronnye seti: raspoznavanie, upravlenie, prinyatie reshenij. Izd-vo Moskva: Finansy i statistika, 176, 2004

[Spiros et al, 2018] Georgakopoulos A. G., Tasoulis S.K., Vrahatis, Plagianakos V.(2018) *Convolutional neural networks for ment classification.* Proceedings of the 10th Hellenic Conference on Artificial Intelligence. Avaible at https://arxiv.org/abs/1802.09957

[Ben Amar et al, 2018] Ben Amar I. , Coppin A., Lecomte E. (2018) *Final Report for the Toxic Comment Classification Challenge* // The Toxic Comment Classification Challenge Avaible at

[Aken et al, 2018] Betty van Aken, Risch J., Krestel R., Löser A. (2018) *Challenges for Toxic Comment Classification: An In-Depth Error Analysis.* EMNLP, 33-42.

[Voronina and Goncharov, 2015] Voronina I.E., Goncharov V.A.(2015) *Analysis of emotional coloring of messages in social networks (on the example of the network "Vkontakte").* Vestnik VSU series: system analysis and information technologies, 4, 151-158(In Rus.) = Analiz emocional'noj okraski soobshchenij v social'nyh setyah (na primere seti "Vkontakte"). Vestnik VGU, seriya: sistemnyj analiz i informacionnye tekhnologii, 4, 151-158, 2015.

[Prabin Lama, 2013] Prabin Lama (2013) *Clustering system based on text mining using the k-means algorithm.* Avaible at https://core.ac.uk/download/pdf/38099883.pdf

[Pribill et al, 2014] Pribill Ju., Pribilova A., Matousek J. (2014) *GMM Classification of text-to-speech synthesis: identification of original speaker's voice* Avaible at https://www.researchgate.net/publication/265215872_GMM_Classification_of_Text-to-Speech_Synthesis_Identification_of_Original_Speaker's_Voice

[Pazel'skaya and Solov'ev, 2011] Pazel'skaya A.G., Solov'ev A.N. (2011) *The method of definition of emotions in Russian texts.* // Computational linguistics and intellectual technologies: collection of scientific articles, 10(17), 510-522 (In Rus.) = Metod opredeleniya emocij v tekstah na russkom yazyke // Komp'yuternaya lingvistika i intellektual'nye tekhnologii. Sbornik nauchnyh statej, 10(17), 510-522, 2011.

[Krasnikov and Nikulichev, 2013] Krasnikov I.A., Nikulichev N.N. (2013) *Hybrid algorithm of classification of text documents based on analysis of internal connectedness of the text.* //Engineering Bulletin of Don. 3 (In Rus.) = Gibridnyj algoritm klassifikacii tekstovyh dokumentov na osnove analiza vnutrennej svyaznosti teksta.//Inzhenernyj vestnik Dona. 3, 2013. Avaible at http://www.ivdon.ru/ru/magazine/archive/n3y2013/1773

[Konig and Brill, 2006] Konig A.C., Brill E. (2006) *Reducing the human overhead in text categorization.* //In Proceedings of the 12th ACM SIGKDD conference on knowledge discovery and data mining, 596-603.

[Hassaan Saeed, 2018] Hafiz Hassaan Saeed, Khurram Shahzad, Faisal Kamiran (2018) *Overlapping Toxic Sentiment Classification Using Deep Neural Architectures.* Proceedings of the 2018 IEEE International Conference on Data Mining Workshops, 1361-1366.

[Dhingra et al, 2017] Bhuwan Dhingra, Hanxiao Liu, Ruslan Salakhutdinov, William W. Cohen A. (2017) *Comparative Study of Word Embeddings for Reading Comprehension.* Avaible at https://arxiv.org/pdf/1703.00993.pdf

[Soumya George et al, 2014] Soumya George K, Shibily Joseph (2014) *Text Classification by Augmenting Bag of Words (BOW) Representation with Co-occurrence Feature.* //IOSR Journal of Computer Engineering, 16(1), 34-38.

[Qaise and Ramsha Ali, 2018] Shahzad Qaise, Ramsha Ali (2018) *Text Mining: Use of TF-IDF to Examine the Relevance of Words to Documents.* //International Journal of Computer Applications, 181(1), 25-29.

[Kedar Potdar et al, 2017] Kedar Potdar, Taher S. Pardawala, Chinmay D. Pai (2017) *A Comparative Study of Categorical Variable Encoding Techniques for Neural Network Classifiers.* //International Journal of Computer Applications,175(4), 7-9.

[Chaubard, 2016] Francois Chaubard, Rohit Mundra, Richard Socher (2016) *Deep Learning for NLP.* //Lecture Notes: Part I. Avaible at https://tensorflowkorea.files.wordpress.com/2017/03/cs224n-2017winter-notes-all.pdf

[Jezek and Steinberger, 2004] Karel Jezek, Josef Steinberger (2004) *Text Summarization and Singular Value Decomposition.* //International Conference on Advances in Information Systems, 245-254.

[Ma and Zhang, 2015] Long Ma, Yanqing Zhang (2015) *Using Word2Vec to Process Big Text Data.* //IEEE International Conference. Avaible at https://www.researchgate.net/publication/291153115_Using_Word2Vec_to_process _big_text_data

[Pennington et al, 2014] Jeffrey Pennington, Richard Socher, Christopher D.(2014) *Glove: Global vectors for word representation.* EMNLP, 1532-1543.

[Armand et al, 2016] Joulin Armand, Grave Edouard, Bojanowski Piotr and Mikolov Tomas.(2016) *Bag of Tricks for Efficient Text Classification.* Avaible at https://www.researchgate.net/publication/319770220_Bag_of_Tricks_for_Efficient _Text_Classification

[Kaggle, 2018] *Kaggle: Toxic Comment Classification Challenge.* (2018). Avaible at https://www.kaggle.com/c/jigsaw-toxic-comment-classification-challenge/

[Kaggle Data, 2018] *Kaggle. Toxic Comment Classification Challenge. Data.* (2018). Avaible at https://www.kaggle.com/c/jigsaw-toxic-comment-classification-challenge/data

[Saif et al, 2018] Mujahed Saif, Alexander Medvedev, Maxim Medvedev, Todorka Atanasova (2018) *Classification of online toxic comments using the logistic regression and neural networks models.* //Proceedings of the 44th international conference "applications of mathematics in engineering and economics", 2048(1).

[Noever, 2018] Noever D.A. (2018) *Machine Learning Suites for Online Toxicity Detection.* Avaible at https://www.researchgate.net/publication/328146083_Machine_Learning_Suites _for_Online_Toxicity_Detection

[Elnaggar et al, 2018] Ahmed Elnaggar, Bernhard Waltl, Ingo Glaser, Jorg Landthaler, Elena Scepankova, Florian Matthes (2018) *Stop Illegal Comments: A Multi-Task Deep Learning Approach.* Proceedings of the 2018 Artificial Intelligence and Cloud Computing Conference, 41-47.

[Mai et al, 2018] Mai Ibrahim, Marwan Torki, Nagwa El-Makky(2018) *Imbalanced Toxic Comments Classification Using Data Augmentation and Deep Learning.* IEEE International Conference on Machine Learning and Applications, 875-878.

[Dialogue Evaluation] *Competitions Dialogue Evaluation.* Avaible at http://www.dialog-21.ru/evaluation/

[Pak and Paroubek, 2012] Pak A., Paroubek P. (2012) *Language Independent Approach to Sentiment Analysis (LIMSI Participation in ROMIP'11).* Conference "Dialogue", 11(18), 37-50.

[ROMIP] ROMIP *A collection of quotes from the news stream with markup on the estimated tonality.* Avaible at http://romip.ru/ru/collections/sentiment-news-collection-2012.html

[Blinov et al, 2013] Blinov P.D, Klekovkina M.V., Kotelnikov E.V., Pestov O.A.(2013) *Research of lexical approach and machine learning methods for sentiment analysis.* Proceedings of International Conference on Computational linguistics and intellectual technologies, 2(12), 48-58.

[SentiRuEval] *SentiRuEval-2015 database.* Avaible at https://drive.google.com/drive/folders /0B7y8Oyhu03y_fjNIeEo3UFZObTVDQXBrSkNxOVlPaVAxNTJPR1Rpd2U1WEktUV Nkcjd3Wms

[Tarasov, 2015] Tarasov D. (2015) *Deep recurrent neural networks for aspect-oriented analysis of user feedback tonality in different languages.* International conference "Dialogue", V.2, 14(21), 53-64 (In Rus.) = Glubokie rekurrentnye nejronnye seti dlya aspektno-orientirovannogo analiza tonal'nosti otzyvov pol'zovatelej na razlichnyh yazykah. Komp'yuternaya lingvistika i intellektual'nye tekhnologii: po materialam ezhegodnoj mezhdunarodnoj konferencii "Dialog", V.2, 14(21), 53-64, 2015.

[Arkhipenko et al, 2016] Arkhipenko K., Kozlov I., Trofimovich J., Skorniakov K., Gomzin A., Turdakov D. (2016) *Comparison of neural network architectures for sentiment analysis of Russian tweets.* Proceedings of International Conference on Computational linguistics and intellectual technologies. Avaible at http://www.dialog-21.ru/media/3380/arkhipenkoetal.pdf

[Mirzayanova, 2019] Mirzayanova S.V. *Sentient-analysis of text.* (2019). Final qualifying work ITMO University, SPb, Russia, p.45 (In Rus.) = Analiz tonal'nosti teksta, VKR, Universitet ITMO, 45 p.

[Bob'yakova, 2017] Bob'yakova D.A. (2017) *Development of an application for analyzing the tonality of texts from social networks.* Final qualifying work ITMO University, SPb, Russia, 74 p. (In Rus.) = Razrabotka prilozheniya dlya analiza tonal'nosti tekstov iz social'nyh setej, VKR, Universitet ITMO, 74 p., 2017.

[Google scholar] *Google scholar* Avaible at https://scholar.google.ru/