

## REVIEW

of official reviewer for the dissertation work of

**Karyukin Vladislav Igorevich on the topic “The research and development of a module for an intelligent system for analyzing and evaluating the social mood of society in the media space of the Republic of Kazakhstan”, granted for the degree of Doctor of Philosophy (PhD) in the specialty “6D070300 – Information Systems”.**

№ p/p	Criteria	Compliance with the criteria (you must mark one of the answer options)	Substantiation of the position of the official reviewer
1.	The dissertation's topic (as of its approval date) corresponds to the directions of the development of science and/or state programs.	<p>1.1 Compliance with priority areas of science development or state programs: the topic of the dissertation work, “The research and development of a module for an intelligent system for analyzing and evaluating the social mood of society in the media space of the Republic of Kazakhstan” corresponds to priority areas of science, such as intelligent systems, artificial intelligence, social research, and information technologies.</p> <p><b>1) The dissertation was completed within the framework of a project or target program financed from the state budget (indicate the name and number of the project or program)</b></p> <p>2) The dissertation was completed within the framework of another state program (indicate the name of the program)</p> <p>3) The dissertation corresponds to the priority direction of the development of science, approved by the Higher Scientific and Technical Commission under the Government of the Republic of Kazakhstan (indicate the direction)</p>	<p>The development of this topic requires the use of artificial intelligence methods, in particular natural language processing (NLP), machine learning, and deep learning models, which are the most important areas of science at this time. Also, this topic includes the study of social sentiment (SA), which requires the use of social sciences to analyze socially oriented programs and policies of the state. This work involves the development of an integrated system, including Big Data collection, processing, and analysis.</p> <p>The dissertation was carried out within the framework of the project for the commercialization of the results of scientific and technical activities “Opinion Monitoring Information System OMSystem (Opinion monitor system),” 0101-18-GK. (The main role of the PhD student was to develop a module for analyzing and evaluating social mood, machine learning models, and neural networks, conducting an experiment on analyzing social mood on the topic of vaccination against Covid-19, and developing an electronic Social Mood module).</p>
2.	Importance for science	The work <b>makes</b> / does not make a significant contribution to science, and its importance is well <b>disclosed</b> / not disclosed.	<p>This work is important because it describes a newly developed method for analyzing social mood using a lexical approach, machine, deep learning models, and marketing indicators of social mood.</p> <p>Experimental results on the analysis of the socio-political topic of vaccination against Covid -19 confirm the importance of the work.</p>

3.	The principle of independence	level: 1) <b>High;</b> 2) Medium; 3) Low; 4) There is no independence	In this work, the PhD student showed a high level of independence. He classified text data using more than ten machine and deep learning algorithms and evaluated their results using the metrics of accuracy, precision, recall, and F1-score. The trained models were then used to label new texts and to calculate social sentiment scores using marketing metrics. The doctoral student also experimented with analyzing the topic of vaccination against Covid-19 and developed the electronic Social Mood (eSM) module for a more detailed data analysis. The PhD student received an authorship certificate for this module.
4.	The principle of internal unity	4.1 Rationale for the relevance of the dissertation: 1) <b>Justified;</b> 2) Partially justified; 3) Not justified.	This topic is especially relevant in connection with the development of the Internet and the expansion of the media space over the past decades, in connection with which people have become more active in expressing their opinions online. It has led to the accumulation of a massive amount of data that needs to be analyzed to understand public sentiment. This study is also significant for understanding the political and social processes in the Republic of Kazakhstan, where the country's leadership seeks to develop policies and programs that meet the population's interests.
		4.2 The content of the dissertation reflects the topic of the dissertation: 1) <b>Reflects;</b> 2) Partially reflects; 3) Does not reflect	The content of the dissertation fully reflects the topic itself. It includes four main sections: theoretical analysis of social media monitoring methods and models, development of the data processing and analysis module of the OMSystem intellectual system, experimental research on data analysis methods and algorithms, and development of the eSM social mood analysis module.
		4.3. The purpose and objectives correspond to the topic of the dissertation: 1) <b>Correspond;</b> 2) Partially correspond; 3) Do not match	Yes, the goal and objectives are fully consistent with the dissertation topic. The work aims to develop a method for assessing the social mood of society in the media space of the Republic of Kazakhstan. To achieve this goal, the main tasks were to develop a module for analyzing and evaluating social mood using machine learning models, neural networks, and marketing indicators, as well as to evaluate the developed



			module using the analysis of the topic of vaccination against Covid-19.
		<p>4.4 All sections and provisions of the dissertation are logically interconnected:</p> <ol style="list-style-type: none"> <li>1) <b>Are fully interconnected;</b></li> <li>2) The relationship is partial;</li> <li>3) There is no relationship</li> </ol>	<p>Sections of the dissertation are logically interconnected and have clear transitions. The first section describes the main aspects and features of domestic and foreign information and analytical systems and methods for determining the sentiment of texts. The second section describes the architecture and functionality of the developed OMSystem platform. The development of models for binary and multiclass text classification using machine learning algorithms and neural networks is presented, as well as a description of marketing technologies for analyzing social mood. The third section fully describes the experiment conducted to analyze the population's reaction to socio-political measures to vaccinate the population against Covid-19. Finally, the fourth section presents the developed Python Django framework electronic Social Mood (eSM) module, which is an application that analyzes data obtained using the OMSystem platform.</p>
		<p>4.5 New solutions proposed by the author (principles, methods) are argued and evaluated in comparison with known solutions:</p> <ol style="list-style-type: none"> <li>1) <b>There is a critical analysis;</b></li> <li>2) Partial analysis;</li> <li>3) The analysis is not one's own opinions, but quotes from other authors</li> </ol>	<p>The results of the constructed classification models were evaluated in comparison with the accepted values of the classification accuracy of machine learning (80%-85%), deep learning (85%-90%), and transformer models (91% and higher). The developed machine learning models exceeded the normal values. The indicators, such as the level of interest in the topic in society, the level of activity of discussion of the topic, and the level of social mood, are used for the first time in the framework of well-known developed systems for monitoring and analyzing the social media space.</p>
5.	The principle of scientific novelty	<p>5.1 Are scientific results and statements new?</p> <ol style="list-style-type: none"> <li>1) <b>Completely new;</b></li> <li>2) Partially new (25-75% are new);</li> <li>3) Not new (less than 25% are new)</li> </ol>	<p>The main provisions for defense are a newly developed method for analyzing social mood, which combines machine learning models for marking texts and marketing indicators for assessing social mood. Also, the integrated model includes more than ten machine and deep learning models, including the BERT transformer model. Another novelty of the work is the development of a new sentiment dictionary for the Kazakh</p>

			language, which is used to label data for machine learning models.
		<p>5.2 Are the conclusions of the dissertation new?</p> <p><b>1) completely new ;</b>  2) partially new (25-75% are new);  3) not new (less than 25% are new)</p>	The conclusions of this dissertation are new since the social mood evaluation module of the OMSystem platform was developed using machine learning models and marketing technologies. The experimental results were carefully studied and analyzed. Conclusions on the analysis of social mood on the topic of vaccination against Covid -19 were made.
		<p>5.3 Technical, technological, economic, or managerial solutions are new and justified:</p> <p><b>1) completely new;</b>  2) partially new (25-75% are new);  3) not new (less than 25% are new)</p>	Technically, in this work, many experiments were carried out on multiclass sentiment classification of texts from the media space of Kazakhstan using algorithms such as Naive Bayes, Logistic Regression, Support Vector Machine, K-nearest neighbors, Decision Tree, Random Forest, XGBoost, Neural Networks: Deep Neural Network, Convolutional Neural Network, and Recurrent Neural Network, as well as BERT transformer models. As a result, the best accuracy values were obtained from 95% to 98%, which is excellent for classifying a large set of text data.
6.	Validity of the main conclusions	All the main conclusions <b>are based</b> / not based on scientifically sound evidence or are reasonably well substantiated (for qualitative research and areas of study in the arts and humanities)	All the results obtained are based on strong evidence, as the developed method for analyzing social mood using machine learning methods and marketing technologies was applied to analyze the particularly relevant topic of vaccination against Covid-19. The results of the experiments were carefully studied and analyzed to understand the reasons for the negative public sentiment. In general, the experiment conducted on vaccination against coronavirus infection makes it possible to understand the public's attitude and the government's activities by evaluating the comments' sentiment analysis and semantic content. As a result, this will make it possible to correctly conduct a research policy for the population, determine the style of presenting information material, accelerate the implementation of such large-scale government tasks, and ensure the preservation of public health. In addition, the OMSystem is used as a serious analytical tool for assessing the perception of social and economic life by users, which will allow to promptly provide explanations to the population,



			identify alarming factors of the population, and evaluate the social mood.
7.	Basic provisions for defense	<p>The following questions need to be answered for each position separately:</p> <p>7.1 Is the position proven?  <b>1) proven;</b>  2) rather proven;  3) rather unproven;  4) not proven</p> <p>7.2 Is it trivial?  1) yes;  <b>2) no</b></p> <p>7.3 Is it new?  <b>1) yes;</b>  2) no</p> <p>7.4 Level to apply:  1) narrow;  2) medium;</p> <p><b>3) wide</b></p> <p>7.5 Is it proven in the article?  <b>1) yes;</b>  2) no</p>	<p>7.1 The developed social mood analysis method applies to various socio-political topics. This paper used it to analyze the topic of vaccination against Covid-19. The integrated machine learning model showed high values of accuracy in determining the sentiment of positive, negative, and neutral texts. The experiment results showed the population's reaction to the ongoing government vaccination measures, from which positive and negative sides are visible.</p> <p>7.2 The developed methods and models are not trivial.</p> <p>7.3 The developed method of social mood analysis is new and has not been used in other works.</p> <p>7.4 The method and models can be applied to analyze a wide range of socio-political topics</p> <p>7.5 The results of the thesis were presented in 12 scientific papers, of which two articles and one chapter in the book were published in journals (Q1, Q2) and a book series peer-reviewed in the Scopus database, four articles in journals recommended by the Committee for quality assurance in the field of education and science of the Ministry of Education and science of the Republic of Kazakhstan, and two articles in scientific conferences, peer-reviewed in the Scopus database, and three articles in the materials of international conferences.</p>
8.	The principle of certainty Reliability of sources and information provided	<p>8.1 Choice of methodology - justified or methodology in sufficient detail described  <b>1) yes;</b>  2) no</p>	<p>The methods used are described in detail in the second and third sections of the thesis. The second section describes the architecture and functionality of the developed OMSystem platform and the social mood analysis module. The development of binary and multiclass text classification models using machine learning algorithms and neural networks is presented, and a description of marketing technologies for social mood analysis is given. The third section fully describes the experiment conducted to analyze the population's reaction to socio-political measures to vaccinate the population against Covid-19.</p>

		<p>8.2 The results of the dissertation work were obtained using modern methods of scientific research and methods of processing and interpreting data using computer technologies:  <u>1) yes;</u>  2) no</p>	<p>This work includes a large number of software implementations. The data analysis and processing module was created in Python using many NLP libraries, especially NLTK and machine learning libraries. NLTK is one of the leading libraries for creating NLP programs. The OMSystem platform and the eSM social mood analysis module are developed on the Django framework Python.</p>
		<p>8.3 Theoretical conclusions, models, identified relationships, and patterns are proven and confirmed by experimental research ( for areas of training in pedagogical sciences, the results are proven based on a pedagogical experiment):  <u>1) yes;</u>  2) no</p>	<p>The paper widely presents experimental data on the development of machine learning models and the analysis of social mood on the topic of vaccination against Covid-19.</p>
		<p>8.4 Important statements <b>are supported</b> / partially confirmed / not supported by references to the relevant and reliable scientific literature</p>	<p>In this paper, more than 100 references to relevant sources of information are used, among which are the works of foreign and domestic scientists, such as Bekmanova G.T., Sharipbay A.A., Yergesh B.Zh., Tukeyev U.A., Karibayeva A., etc.</p>
		<p>8.5 Used literature sources <b>are sufficient</b> / not sufficient for a literature review</p>	<p>The number of literature sources is sufficient for a complete literature review of the dissertation.</p>
9	Principle of practical value	<p>9.1 The dissertation has a theoretical value:  1) <u>yes;</u>  2) no</p>	<p>The dissertation work is of great theoretical importance since, in the course of the work, the architecture and functionality of the intelligent OMSystem, the main algorithms for machine learning and neural networks, as well as methods for preprocessing, vectorization, and class balancing were considered. Moreover, a description of the calculation of marketing indicators of the level of interest in the topic in society, the level of activity in the discussion of the topic, and the level of social mood was given.</p>
		<p>9.2 The dissertation is of practical importance, and there is a high probability of applying the results obtained in practice :  <u>1) yes;</u>  2) no</p>	<p>This dissertation reflects the practical application of the developed models and the method of analyzing social mood. As part of the work, a practical experiment has already been carried out on the analysis of mood on the topic of vaccination against Covid-19. The paper states that in the future, it is planned to conduct more experiments in this area, expand the range of analysis of topics and strengthen the role and</p>



			significance of social analytics as one of the most important areas of natural language processing.
		9.3 Are the suggestions for practice new? <b>1) completely new;</b> 2) partially new (25-75% are new); 3) not new (less than 25% are new)	The proposed method for analyzing social mood is a new method that was used in this work and published in the scientific article “ On the development of an information system for monitoring user opinion and its role for the public ” in the highly rated Journal of Big Data (Q1, 99 percentile).
10.	Quality of writing and design	Quality of academic writing: <b>1) high;</b> 2) average; 3) below average; 4) low.	The text of the dissertation work has a clear scientific presentation. Grammar, stylistic, and punctuation errors were checked manually and with the use of an automatic text-checking and editing system.

The dissertation work of Karyukin Vladislav Igorevich on the topic “**The research and development of a module for an intelligent system for analyzing and evaluating the social mood of society in the media space of the Republic of Kazakhstan**” complies with the Rules for awarding the degree of Doctor of Philosophy (PhD). Its author Karyukin Vladislav Igorevich deserves to be awarded the Doctor of Philosophy (PhD) degree in the specialty “6D070300 – Information Systems”.

**Official reviewer:**  
**PhD, Senior Lecturer**  
**Department of Artificial Intelligence Technologies**  
**Faculty of Information Technologies**  
**L.N. Gumilyov Eurasian National University**

**Barlybayev A.B.**

