DETERMINATION THE TOTALITY OF NON-GOVERNMENTAL BUSINESS SUPPORT ORGANIZATIONS OF UKRAINE USING MACHINE LEARNING METHODS

Dariia Smolych[1]

DOI: https://doi.org/10.30525/978-9934-26-531-0-32

Abstract. Over the past ten years, despite the invasion and hostilities, Ukraine has seen a positive trend in the development of civil society, demonstrating the growing role of non-governmental organizations in the socio-political life of the country and its economic development. Scientists usually study the entire set of civil society organizations as a whole, there are separate works that cover certain areas: youth, environmental, human rights organizations, and others, without analyzing their number and development trends. *Subject of study* is non-governmental organizations in Ukraine, whose activities are aimed at supporting the development of entrepreneurship. Non-governmental business support organizations play the role of an effective independent institutional intermediary between business and the government sector, in order to protect the interests of business entities, establish appropriate communication that will promote business development, and as a result, enable the economic development of the country as a whole. Such organizations are non-profit, independent of the state, support and stimulate business development, including by implementing demanded services. In Ukraine, non-governmental business support organizations can be registered in the following organizational and legal forms: public organizations; public unions; employers' organizations; chambers of commerce and industry; associations; other associations of legal entities (unions, unions, etc.); other organizational and legal forms. *The purpose of this study* is to determine the general population of non-governmental organizations in Ukraine, whose activities are aimed at supporting the development of entrepreneurship. *Research methodology.*

[1] Candidate of Economic Sciences, Associate Professor,
Associate Professor of the Department of Management,
Lutsk National Technical University, Ukraine

The main stages of the study included: pre-processing of texts in the Ukrainian language, in particular, normalization, removal of stop words and the use of special stemming algorithms for morphological processing. This allowed to significantly improve the quality of training data. The BERT model was used for classification, which is able to work effectively with short texts due to its bidirectional architecture. The use of loss functions with class weights and the use of balancing methods allowed to equalize the influence of different classes on learning, which significantly increased the accuracy of predictions. The resulting model demonstrated high results in all main metrics - accuracy, completeness, prediction accuracy and F1-measure. This indicates the effectiveness of the model in solving the problem of classifying short texts. The model was used to classify registered non-governmental organizations in all regions of Ukraine. *Research conclusion.* As a result of the research, a short text classification model based on the BERT architecture was developed and successfully implemented to determine the number of non-governmental organizations that promote business in Ukraine. Natural language processing methods were applied for automatic classification of text data, which significantly increased the efficiency of analyzing large volumes of unstructured information.

1. Introduction

Non-governmental organizations protect and represent the interests of the public, participate in solving social problems, and can also provide services. There are a large number of non-governmental organizations in Ukraine, including youth, sports, environmental, human rights, and others, the direction of their activities is determined by statutory tasks.

The positive dynamics of the increase in non-governmental organizations in Ukraine as a whole over the past decade is associated, firstly, with the intensification of public activity and the growth of public consciousness after the Revolution of Dignity of 2013-2014, and secondly, as a result of the occupation of the Autonomous Republic of Crimea and the outbreak of war in eastern Ukraine during this period, the first internally displaced persons appeared who needed assistance. The growth in the number of non-governmental organizations in the period 2016-2019 is due to the implementation of many technical assistance projects and the positive impact of the implemented relevant activities.

During 2022-2024, the full-scale invasion was the main factor that contributed to the strengthening of public cohesion in order to counter the challenges of the war, civil society organizations, in addition to a specific area of their activity, focused on assisting the victims, the Armed Forces of Ukraine.

Within the framework of this study, we will consider the possibilities of using machine learning methods to establish a set of non-governmental organizations in Ukraine that provide incentives for business development and take an active part in the processes of liberalization of the conditions for conducting economic activity from among the general set of non-governmental organizations. The activities of non-governmental organizations supporting business in Ukraine are very important today, because companies and entrepreneurs, despite the losses and destruction associated with the war, continue their economic activities, need support, pay taxes, fill the country's budget, provide the population with jobs, financially support the Armed Forces of Ukraine and in the future will ensure the reconstruction of Ukraine.

The use of machine learning methods as one of the ways to refine the classification of non-governmental organizations by their areas of activity will contribute to research in this area, as these methods can provide a solution to the problem of access to relevant data that are not available from official statistical sources.

Classification of short texts is a critically important aspect of natural language processing in the modern digital world. The growing volumes of text information that are generated and consumed daily require effective methods of organizing and analyzing data. Text classification has become a powerful tool that allows researchers, specialists and organizations to systematize and structure this information, facilitating access to knowledge and improving the decision-making process.

Today, with the increase in the volume and accessibility of documents in digital form and the subsequent need to organize them, techniques for automated classification of texts into predefined categories have become particularly important.

Let's consider several examples of the application of such classification. Classification of customer reviews in business by categories, such as satisfaction or dissatisfaction, helps to analyze moods and trends. In the

media, automatic classification of news by topics (politics, economics, culture) contributes to better organization of content and facilitates the work of journalists. Social media use publication classification to filter content by topic and mood, which improves interaction with the audience. Scientists classify articles to organize literature by topic, which facilitates search and contributes to scientific achievements.

Short texts, which usually consist of one or two sentences, often lack context, which makes their classification difficult. As a rule, these short texts contain news headlines, social topics, product reviews, etc. These texts are unstructured and characterized by irregularity. Therefore, highlighting the features of short texts and their accurate classification have become one of the key tasks in the field of natural language processing (NLP).

Among the scientists conducting research in this area, we can note the works of foreign and domestic authors, in particular: Alper Kursat Uysal [1], Litofchenko Ya., Karner D., Mayer F. [2], Volosyuk Yu. [3], Holub T., Zelenyova I., Grushko S., Lutsenko N. [4].

Since computers work with numerical data, converting texts into digital vectors is a necessary step in their classification. One of the basic methods of text transformation was direct encoding, in which a single word is represented as a vector, where one dimension is set to 1 and all others are set to 0. However, such a representation faces the problems of high sparsity and a significant increase in data size, and also does not take into account the importance of words in the text.

The next approach TF-IDF estimates the importance of a word in a document or set of texts by calculating the frequency of the word, but it ignores the order of the words and does not reflect the sequence of information. Further work has focused on constructing distributed dense vectors of words with low dimensionality [5].

Word2Vec is a significant step forward, as it allows us to take into account the context of a word through the methods of core word prediction by its surroundings (CBOW) and core word prediction by its surroundings (Skip-gram), which significantly improves the semantic mapping compared to previous methods. However, Word2Vec is a static model that does not solve the problem of word polysemy, since each word is assigned a single vector, regardless of the context [6].

To solve this problem, contextual models have been developed, such as ELMo [7], which generates semantic vectors of words that change depending on the context, as well as GPT [8], which generates text based on previous words, and BERT [9], which analyzes text from both sides to understand the meaning of words in different situations.

The ELMo model uses bi-lateral recurrent neural networks (BiLSTM) [10] to derive context-dependent word meanings. The GPT and BERT models are based on transformer architectures and attention mechanisms, which allow them to more effectively process the context of words in text.

Before the advent of contextual models such as ELMo, GPT, and BERT, text classification was based on convolutional neural networks (CNN) [11] and recurrent neural networks (RNN) [12], which performed well on sequential data. However, these models had limitations, particularly in handling long-term dependencies, which gave rise to the development of the transformer architecture.

Transformers are neural network architectures that are organized around an attention mechanism, which allows the model to effectively focus on different parts of the input data when processing sequences. They consist of encoding and decoding blocks that work together in concert to perform tasks such as translation, text classification, or speech generation.

2. BERT Model

BERT (Bidirectional Encoder Representations from Transformers) is based on the transformer architecture, but uses only the encoder part of this architecture. This allows the model to effectively process the context on both sides of each word in a sentence, the model consists of several layers of self-attention and volumetric fully connected layers. The main characteristic of BERT is its ability to understand bidirectional context by processing text from both left to right and right to left simultaneously.

The self-attention mechanism is the central element of BERT and other transformers. Its role is to model the relationships between words in the text regardless of their position. Self-attention allows the model to determine which words are most important for understanding the meaning of each individual word in the context.

During model training, the attention parameters are adjusted to better reflect the relationships between words. The model learns to pay attention

to those words that are most important for the context and the specific classification task. These parameters are updated during each training step, which improves the model's ability to make accurate predictions.

In many cases, the training data may be unevenly distributed across classes, so balancing techniques are used to improve the quality of the model. For example, you can resample the data for a minority class to ensure that all classes have an equal number of examples before training the model.

This helps to avoid situations where the model pays more attention to one class, which can reduce classification accuracy. In our case, we resampled the data for a smaller class using the resample function from the sklearn library. This allows us to equalize the number of examples in both classes before training, which improves the model's performance and avoids bias in predictions.

During model training, special numbers (weights) are calculated for the classes that are used to estimate the errors (loss function). This helps to balance the influence of classes on training.

In our case, the CrossEntropyLoss loss function is used, modified to take into account the class weights. This function ensures that the model weights are updated correctly during training, especially in conditions of class imbalance.

Formally, the loss function is calculated as:

$$L = -\frac{1}{N}\sum_{i=1}^{N}\omega_{y_i}\log\left(p_{y_i}\right), \tag{1}$$

N – number of examples,

ω_{y_i} – class weight y_i,

p_{y_i} – the probability of the correct class predicted by the model.

Training is performed using the AdamW optimizer, which is used to update the model weights. This optimizer includes regularization using weight reduction, which helps to avoid overtraining the model. During training, the model goes through several cycles during which losses are calculated and the model is gradually adjusted to improve the accuracy of predictions.

After the training process is complete, the model goes through a validation phase, where the main evaluation metrics include:

– Accuracy: shows the proportion of correct predictions among all examples.

– Recall: reflects the ability of the model to correctly classify all positive examples.

– Precision: estimates how many of the predicted positive examples are actually positive.

– F1-measure: a balanced measure between accuracy and completeness.

– ROC-AUC: a classification quality indicator that takes into account all possible decision thresholds.

BERT model is extremely effective when working with short texts due to its ability to fully take into account the bidirectional context even in small fragments of text. Using self-attention allows the model to focus on important words and discard less important parts of the text, making it suitable for tasks such as classifying organizations by their names or descriptions of activities.

3. Data preparation for training the BERT model taking into account the peculiarities of the Ukrainian language

Data preparation for training the BERT model consists of several important stages: pre-processing of the input text, extraction of the most key elements of the text and the formation of the training set. The quality of such a set largely depends on the quality of the pre-processing of the data.

The input text goes through several stages of processing. First, the text is normalized, which includes converting all words to lowercase and correcting errors. Then the text is cleaned of special characters, and unnecessary characters and punctuation are removed. One of the features of processing Ukrainian texts is the morphological complexity of the language. Words can have different forms depending on the context, including cases, genders and tenses. This complicates text processing, so stemming is used.

Stemming allows you to reduce words to their basic form, which helps reduce the variability of word forms and simplify further processing. In the process of creating data, a specially configured stemmer for the Ukrainian language was used, which takes into account case endings and verb suffixes.

Another important aspect is the removal of stop words. These words are not essential for the model, such as conjunctions and prepositions, and they

need to be removed to reduce noise in the data. At this stage, a dictionary of Ukrainian stop words was used, which contains 1892 words.

The key stage is to extract the most significant elements of the text that will be used for classification. The processed text remains structured features that allow for more efficient classification.

At the final stage of data preparation for training the BERT model, keywords that define the activities of non-governmental organizations that support business are compared with words that were obtained from the names of non-governmental organizations registered in the Volyn region (2322 organizations).

After that, the training dataset is checked for compliance with the information specified in the charters of these organizations. In case of discrepancies in the lists of key terms, the necessary adjustments are made to ensure the accuracy of the data for the model. The final version was applied to the data for Rivne region (2470 NGOs) and Sumy region (2500 NGOs). At the output, we get a classified dataset, ready for further use in training the BERT model.

4. Training a BERT model for text classification

Training a BERT model for text classification involves several key steps: model initialization, training parameters, the actual training process, and validation of the results on validation data. Each of these steps is critical to the model's performance and ability to make accurate predictions.

Model Initialization

The first stage is to load a universal pre-trained BERT language model, which specializes in multilingual text processing. In our case, the model is used for binary classification, which involves determining whether an NGO supports a business. This process is achieved by tuning the model for specific tasks, where the final results are two possible classes: «supports» and «does not support». Tuning includes fine-tuning the model's weights, which allows improving its accuracy and adapting to the specifics of the data it works with.

Setting up training parameters

To train the model, optimal parameters are determined that ensure its effective functioning. One of the key parameters is the optimizer – an algorithm that is responsible for updating the model's weights during

training. For this process, the popular gradient descent method with weight correction is used, which allows minimizing model errors. In addition, it is important to set the learning rate, which determines how quickly or slowly the model parameters will change during the learning process.

Another important parameter is the use of weights for each class, which allows balancing the model since the input data is unbalanced (there are significantly fewer organizations that support business than those that do not).

The learning process

After initialization, the direct process of training the model begins. It consists in gradually improving the model's ability to recognize patterns in the texts of non-governmental organizations. The input texts are processed as sequences of tokens that are fed to the model's input. The model makes a prediction that is compared with real data, and based on this, a loss function (error function) is calculated. The loss function shows how much the prediction differs from the correct result. The model tries to minimize this function by gradually adjusting its parameters.

When training the model, it is important to monitor its performance on the training data set. This allows you to find out whether the model is able to learn the necessary information, and at the same time avoid overfitting, when the model adapts too much to a specific data set and loses the ability to work effectively on new examples.

Model evaluation on validation data

After each training stage, the model is evaluated on a separate validation data set that is not used for training. This allows us to assess its ability to generalize the acquired knowledge and make predictions on new texts. Metrics such as accuracy, recall, precision, and F1-measure are used to evaluate the results. These metrics show how well the model classifies texts and whether it takes into account all possible options.

Another important metric is ROC-AUC, which reflects the ratio between the sensitivity and specificity of the classifier. A high ROC-AUC value indicates a high quality classification.

After completing the training process and evaluating the model on the validation data, the following results were obtained.

Validation Accuracy: 0.9716. This means that the model correctly classifies 97.16% of the cases. This is a very high accuracy, indicating that the model is learning well on the presented data.

Validation Recall: 0.9752. The recall for class 1 (positive class) is 97.52%, which means that the model detects 97.52% of all true positive cases. High recall indicates that the model is good at detecting positive samples.

Precision: 0.9657. The recall for class 1 (positive class) is 96.57%, which means that of all the samples that the model classifies as positive, 96.57% are actually positive. High accuracy indicates that the model is rarely wrong when predicting a positive class.

Validation F1-Score: 0.9704. The F1-Score is the harmonic mean between precision and completeness. A value of 0.9704 indicates a balanced model performance in terms of both precision and completeness.

Validation ROC-AUC: 0.9717. The ROC-AUC (Area Under the Receiver Operating Characteristic Curve) is 0.9717, which indicates a very good ability of the model to separate positive samples from negative ones. A value close to 1 indicates an excellent discriminatory ability of the model.

Confusion Matrix – shows how many times the model correctly or incorrectly classified objects for each class:

[[426 14]
[10 394]]

426: Number of correct predictions for class 0 (negative class).

14: Number of false predictions for class 0 (model predicted positive class instead of negative).

10: Number of false predictions for class 1 (model predicted negative class instead of positive).

394: Number of correct predictions for class 1 (positive class).

Table 1

The classification model evaluation matrix is presented

	precision	**recall**	**f1-score**	**support**
0	0.98	0.97	0.97	440
1	0.97	0.98	0.97	404
accuracy	0.97	0.97	0.97	844
macro avg	0.97	0.97	0.97	844
weighted avg	0.97	0.97	0.97	844

Note: Created by the author

Thus, the modeling process used a pre-trained BERT model, which was configured for binary classification of texts from non-governmental organizations.

Thanks to the self-attention mechanism and taking into account the context, the model effectively identified connections between words, which improved its performance. Data balancing and parameter tuning allowed to avoid overtraining problems and ensure high accuracy of predictions. Evaluation on validation data confirmed the ability of the model to make accurate and generalized predictions.

The resulting model is saved for further use in predictions.

5. Using the BERT model to identify business support NGOs

At the initial stage, the resulting BERT model is used together with text processing tools. The model analyzes texts, converting them into numerical sequences using tokenization, which allows you to «understand» the content of the text and its structure. This creates the basis for accurate further forecasting.

Next, information is extracted from a database containing data on Ukrainian NGOs. This is key information for determining whether a particular organization contributes to business development. The texts are prepared for analysis by normalizing them and limiting them to 128 tokens, which optimizes the model's operation and ensures a uniform input data format.

The prepared data is fed to the input of the BERT model, which analyzes each record and provides a prediction as to whether the organization supports business. The model uses a probabilistic approach to determine whether a non-governmental organization belongs to one of the categories: «supports business» or «does not support business». This allows you to automate the process of classifying a large amount of data.

After the classification is complete, the results are stored in the database. A new field is added to each organization indicating whether it supports business development. This simplifies the integration of machine learning results into the existing system for further analysis and use of the results.

To collect and evaluate quantitative data on non-governmental organizations supporting business in Ukraine, the paid software product YC Markets [13] of the YouControl analytical system was selected, which

allows searching and systematizing data on all existing organizations in Ukraine.

This innovative analytical tool allows you to conduct an independent multifactor analysis of Ukrainian organizations with the ability to form lists of companies according to specified criteria. In particular, it allows you to quickly find Ukrainian organizations by the selected KVED code, region, settlement, date of registration and subsequent ranking and generation of an electronic report.

YC Markets data sources are: Unified State Register of Legal Entities, Individual Entrepreneurs and Public Organizations; relevant data sets of the State Statistics Service of Ukraine and the State Tax Service of Ukraine; analytical data of the company YouControl.

Thus, using the paid analytical tool YC Markets [13], the potential number of active civil society organizations was determined according to the specified organizational and legal forms in which non-governmental business support organizations can be registered (public organization, public union, employers' organization, chamber of commerce and industry, association, other associations of legal entities), which amounted to 108,547 organizations that carry out activities in accordance with the classifier of types of economic activity [14] under the following codes: 94.11 – activities of organizations of industrialists and entrepreneurs; 94.12 – activities of professional public organizations; 94.99 – activities of other public organizations.

As previously noted, from the totality of civil society organizations, it is quite difficult to accurately identify those that are engaged in business support. For example, according to KVED 94.99, as of 2024, there were 89,845 public organizations in Ukraine, including youth, sports, cultural, environmental, entrepreneurship support, and others.

The results of identifying the number of non-governmental organizations specifically supporting business based on the use of the BERT model are presented in Table 2.

Data collection was conducted in 2024, excluding the temporarily occupied territory of the Autonomous Republic of Crimea and the city of Sevastopol. Due to the dynamic nature of the sector, data may differ in future periods.

Table 2

Results of identifying the set of non-governmental organizations supporting business in Ukraine based on the use of the BERT model

Region	Number of potential non-governmental business support organizations	%	Number of non-governmental business support organizations	%
Vinnytsia	3514	3,24	140	3,53
Volyn	2322	2,14	112	2,82
Dnipropetrovsk	6333	5,83	175	4,41
Donetsk	5519	5,08	173	4,36
Zhytomyr	2558	2,36	129	3,25
Zakarpattia	2945	2,71	143	3,60
Zaporizhia	3724	3,43	149	3,75
Ivano-Frankivsk	3331	3,07	115	2,90
Kyiv	28036	25,83	1099	27,68
Kirovohrad	2093	1,93	117	2,95
Luhansk	2921	2,69	100	2,52
Lviv	7729	7,12	305	7,68
Mykolaiv	3263	3,01	106	2,67
Odesa	7112	6,55	277	6,98
Poltava	3266	3,01	116	2,92
Rivne	2470	2,28	77	1,94
Sumy	2500	2,30	71	1,79
Ternopil	2153	1,98	53	1,34
Kharkiv	5177	4,77	147	3,70
Kherson	2217	2,04	67	1,69
Khmelnytskyi	2543	2,34	61	1,54
Cherkasy	2767	2,55	94	2,37
Chernivtsi	2048	1,89	61	1,54
Chernihiv	2006	1,85	83	2,09
ARC Crimea	-	-	-	-
Total	108547	100	3970	100

Note: Calculated by the author based on data [13]

Figure 1 presents the percentage distribution of non-governmental business support organizations in the regions of Ukraine.

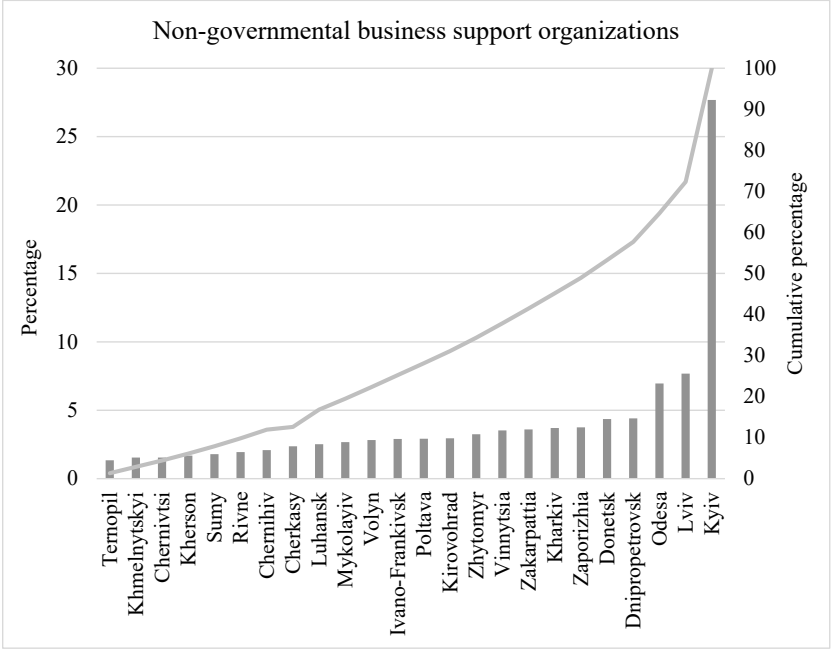

Figure 1 Percentage distribution of non-governmental business support organizations in the regions of Ukraine

Note: Created by the author

As can be seen from the data in the figure, the largest number of non-governmental business support organizations is in Kyiv region (including the city of Kyiv), Lviv region, Odessa region, Dnipropetrovsk region, Donetsk region, Zaporizhia region and Kharkiv region. This is largely due to the fact that the regional centers of these regions are the largest cities of Ukraine, which are home to a large percentage of the population and business entities that are the target audience of non-profit organizations, including non-governmental business support organizations. The smallest number is in Ternopil region, Khmelnytskyi region and Chernivtsi region.

6. Conclusions

The history of the development of civil society in the conditions of independent Ukraine has proven the ability and justification of the main mission of non-governmental organizations – to be a public voice for solving urgent issues. With the beginning of the full-scale invasion of Ukraine, Ukrainian business found itself in an extremely critical situation, despite the destruction and significant losses, Ukrainian companies continue their activities and adapt to new conditions, needing support. The growing economic presence of non-governmental organizations supporting business, their broad involvement in solving social problems and the contribution they make to the development of entrepreneurship, public good determines the relevance of the chosen topic of scientific research.

Non-governmental support organizations are independent of the state, pre-legal associations of business entities and/or individuals and stakeholders, whose activities are non-profit and aimed at ensuring a favorable business environment and expanding business development opportunities at the local, national, international levels by providing organizational, resource or other types of necessary assistance to business structures, while satisfying the interests of members and the public.

Non-governmental business support organizations focus on the following areas of activity: protection and representation of business interests in relations with authorities; legal support; information support and training; establishing relationships and communication support; resource support.

Using the paid analytical tool YC Markets, the potential number of active civil society organizations was determined by the analyzed organizational and legal forms in which non-governmental business support organizations can be registered (public organization, public union, employers' organization, chamber of commerce and industry, association, other associations of legal entities), which as of 2024 amounted to 108,547 organizations.

As a result of the research, a short text classification model based on the BERT architecture was developed and successfully implemented to determine the number of non-governmental organizations that promote business development in Ukraine, the total number of which as of 2024 was 3970 organizations.

Natural language processing methods were applied for automatic classification of text data, which significantly increased the efficiency of

analyzing large volumes of unstructured information. The main stages of the research included: pre-processing of texts in Ukrainian, in particular, normalization, removal of stop words and application of special stemming algorithms for morphological processing. This significantly improved the quality of training data.

The BERT model was used for classification, which is able to work effectively with short texts due to its bidirectional architecture. The model was adapted for binary classification – determining the number of non-governmental organizations that contribute to business development. The use of loss functions with class weights and the use of balancing methods allowed to equalize the influence of different classes on learning, which significantly increased the accuracy of predictions.

The resulting model demonstrated high results in all main metrics – accuracy, completeness, prediction accuracy and F1-measure. This indicates the effectiveness of the model in solving the problem of classifying short texts. The model was used to classify non-governmental organizations in all regions of Ukraine. The largest number of non-governmental business support organizations is in Kyiv region (including the city of Kyiv), Lviv region, Odessa region, Dnipropetrovsk region, Donetsk region, Zaporizhia region and Kharkiv region. This is largely due to the fact that the regional centers of these regions are the largest cities of Ukraine, which are home to a large percentage of the population and business entities that are the target audience of non-profit organizations, including non-governmental business support organizations. The smallest number is in Ternopil region, Khmelnytskyi region and Chernivtsi region.

References:

1. Alper Kursat Uysal. An improved global feature selection scheme for text classification. *Expert Systems with Applications*. 2016. Vol. 43. P. 82–92.

2. Litofcenko J., Karner D., Maier F. Methods for Classifying Nonprofit Organizations According to their Field of Activity: A Report on Semi-automated Methods Based on Text. *VOLUNTAS: International Journal of Voluntary and Nonprofit Organizations.* 2020. P. 227–237.

3. Волосюк Ю. В. Методи класифікації текстових документів в задачах Text mining. *Наукові записки Українського науково-дослідного інституту зв'язку*. 2014. № 6 (34). С. 76–81.

4. Голуб Т.В., Зеленьова І.Я., Грушко С.С., Луценко Н.В. Програмна реалізація автоматичного класифікатора текстів на основі уточненого методу

формування простору ознак категорій. *Телекомунікаційні та інформаційні технології*. 2020. № 1 (66). С. 227–237.

5. Yu, C. T., & Salton, G. (1976). Precision weighting – an effective automatic indexing method. *Journal of the Association for Computing Machinery*, 23(1), 76–88. DOI: https://doi.org/10.1145/321921.321930

6. Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. Paper presented at the Advances in neural information processing systems.

7. Peters, M. E., Neumann, M., Iyyer, M., Gardner, M., Clark, C., Lee, K., & Zettlemoyer, L. (2018). Deep contextualized word representations. arXiv preprint arXiv:1802.05365

8. Radford, A., Narasimhan, K., Salimans, T., & Sutskever, I. (2018). Improving language understanding by generative pre-training. Academic Press.

9. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding / J. Devlin, M. W. Chang, K. Lee, K. Toutanova // Google research. URL: https://github.com/google-research/bert

10. Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural Computation, 9(8), 1735–1780.

11. K. Yoon, "Convolutional neural networks for sentence classi6cation," in Proceedings of the Conference on Empirical Methods in Natural Language Processing, pp. 1746–1751, Doha, Qatar, October 2014.

12. M. Tomas, K. Martin, B. Lukas, C. Jan, and K. Sanjeev, "Recurrent neural network based language model," in Proceedings of the 11th Annual Conference of the International Speech Communication Association, pp. 1045–1048, Makuhari, Chiba, Japan, September 2010.

13. Програмний продукт YC Markets аналітичної системи YouControl. URL: https://youcontrol.market/

14. Національний класифікатор видів економічної діяльності ДК 009:2010. Наказ Держспоживстандарту України від 11 жовтня 2010 р. № 457. URL: https://zakon.rada.gov.ua/rada/show/vb457609-10#Text

References:

1. Alper Kursat Uysal. An improved global feature selection scheme for text classification. *Expert Systems with Applications*. 2016. Vol. 43. P. 82–92.

2. Litofcenko J., Karner D., Maier F. Methods for Classifying Nonprofit Organizations According to their Field of Activity: A Report on Semi-automated Methods Based on Text. *VOLUNTAS: International Journal of Voluntary and Nonprofit Organizations.* 2020. P. 227–237.

3. 3.Volosiuk Yu.V. (2014). Metody klasyfikatsii tekstovykh dokumentiv v zadachakh Text mining [Methods of classification of text documents in Text mining tasks]. *Naukovi zapysky Ukrainskoho naukovo-doslidnoho instytutu zviazku*. № 6 (34). S. 76–81.

4. Holub T.V., Zelenova I.Ia., Hrushko S.S., Lutsenko N.V. (2020). Prohramna realizatsiia avtomatychnoho klasyfikatora tekstiv na osnovi utochnenoho metodu

formuvannia prostoru oznak katehorii [Software implementation of an automatic text classifier based on a refined method of forming the space of category features]. *Telekomunikatsiini ta informatsiini tekhnolohii.* № 1 (66). S. 227–237.

5. Yu, C. T., & Salton, G. (1976). Precision weighting – an effective automatic indexing method. *Journal of the Association for Computing Machinery*, 23(1), 76–88. DOI: https://doi.org/10.1145/321921.321930

6. Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. Paper presented at the Advances in neural information processing systems.

7. Peters, M. E., Neumann, M., Iyyer, M., Gardner, M., Clark, C., Lee, K., & Zettlemoyer, L. (2018). Deep contextualized word representations. arXiv preprint arXiv:1802.05365.

8. Radford, A., Narasimhan, K., Salimans, T., & Sutskever, I. (2018). Improving language understanding by generative pre-training. Academic Press.

9. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding / J. Devlin, M. W. Chang, K. Lee, K. Toutanova // Google research. URL: https://github.com/google-research/bert

10. Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural Computation, 9(8), 1735–1780.

11. K. Yoon, "Convolutional neural networks for sentence classi6cation," in Proceedings of the Conference on Empirical Methods in Natural Language Processing, pp. 1746–1751, Doha, Qatar, October 2014.

12. M. Tomas, K. Martin, B. Lukas, C. Jan, and K. Sanjeev, "Recurrent neural network based language model," in Proceedings of the 11th Annual Conference of the International Speech Communication Association, pp. 1045–1048, Makuhari, Chiba, Japan, September 2010.

13. Prohramnyi produkt YC Markets analitychnoi systemy YouControl. Available at: https://youcontrol.market/

14. Natsionalnyi klasyfikator vydiv ekonomichnoi diialnosti DK 009:2010. Nakaz Derzhspozhyvstandartu Ukrainy vid 11 zhovtnia 2010 r. № 457. Available at: https://zakon.rada.gov.ua/rada/show/vb457609-10#Text