

Identifying Reproductive Behavior Arguments in Social Media Content Users' Opinions through Natural Language Processing Techniques

Irina Kalabikhina¹, Ekaterina Zubova², Natalia Loukachevitch³,
Anthony Kolotusha¹, Zarina Kazbekova¹, Evgeny Banin⁴, German Klimenko¹

¹ Faculty of Economics, M. V. Lomonosov Moscow State University, Moscow, 119991, Russia

² Yale University, New Haven, CT, 06511, USA

³ Research Computing Center of Lomonosov Moscow State University, Moscow, 119899, Russia

⁴ Kurchatov Institute, Moscow, 123182, Russia

Received 2 November 2022 ♦ Accepted 10 May 2023 ♦ Published 30 June 2023

Citation: Kalabikhina I, Zubova E, Loukachevitch N, Kolotusha A, Kazbekova Z, Banin E, Klimenko G (2023) Identifying Reproductive Behavior Arguments in Social Media Content Users' Opinions through Natural Language Processing Techniques. *Population and Economics* 7(2):40-59. <https://doi.org/10.3897/popecon.7.e97064>

Abstract

Big data provides researchers with valuable sources of information for studying demographic behavior in the population. One such source is the texts posted by social network users on various demographic issues. This study utilizes methods for automatically extracting user opinions from the “VKontakte” social network. The extracted texts are then classified using the Conversational RuBERT neural network model to investigate opinions related to reproductive behavior in the population. The classification process addresses two consecutive problems. Firstly, it aims to identify whether a user’s comment contains argumentation. Secondly, if an argument is present, it seeks to determine its type within the context of the “personal-public” dichotomy. To search for arguments and classify their types, six experiments were conducted, varying the dataset and the number of classes. The method employed for automatic extraction and classification of user opinions on the “VKontakte” social network has demonstrated the ability to accurately classify users’ comments, identifying the presence of argumentation and categorizing the arguments within the “personal-public” dichotomy. This enables the identification of personal and social attitudes, values, stories, and opinions, thus facilitating the study of reproductive behavior.

Keywords

reproductive behavior, personal-public dichotomy, automatic opinion extraction, argumentation, VKontakte, Conversational RuBERT

JEL codes: C8, J1

Introduction

Big data offers researchers new and valuable sources of information for studying demographic behavior in the population. One of these sources is the statements made by social network users on various demographic issues. In order to examine the reproductive behavior of the population, we initially analyzed this data by determining the “demographic temperature” of two types: the emotional background of specific demographic groups with pronatalist and antinatalist persuasions (Kalabikhina et al. 2021a), as well as the emotional background of statements related to reproductive behavior and the assessment of demographic policies (Kalabikhina et al. 2021b; 2022). We identify positive or negative attitudes towards reproductive aspects such as large families, childlessness, and abortions; however, automatically determining the arguments behind these attitudes remains a challenge.

While natural language processing methods are advancing rapidly, there are still complex problems without straightforward solutions. One such challenge is the identification and reproduction of arguments through automated methods. However, the ability to analyze statements and their justifications is of great interest, particularly in the field of social sciences. Solving this problem would enable us to answer questions not only about what people express in their opinions, but also the underlying basis for those opinions. In this study, we aim to conduct an analysis of argumentation regarding issues related to reproductive behavior based on user comments on the VKontakte social network.

Our study relies on methods for automatically extracting user opinions from the VKontakte social network, followed by the classification of these opinions using the Conversational RuBERT neural network model. The classification of statements serves two purposes: firstly, to identify the presence of argumentation within a user’s comment, and secondly, to determine the type of argument within the context of the “personal-public dichotomy.”

We adopt this approach based on the traditional consideration of the population’s attitudes in the context of demographic behavior to predict behavior and make decisions in the field of demographic policy. For example, we examine attitudes towards the desired number of children individuals plan to have and their perceptions of the ideal number of children in a given society at a specific time.

To address these challenges, we conducted six experiments using different datasets and varying the number of classes (two or three). This paper analyzes the key quality metrics of the trained classification models, including accuracy and F-score, and concludes by selecting the most relevant models for both tasks.

Literature review

In recent years, researchers have increasingly utilized comments from social media platforms like Reddit and Twitter as a valuable tool for studying public opinion. Social media data is regularly employed to assess the emotional spectrum and sentiment analysis (Thelwall and Stuart 2019; Mittos et al. 2020; Liu et al. 2021; Al-Rawi et al. 2021).

Several studies have focused on extracting arguments from text data for various purposes and research areas, often utilizing machine translation systems. For instance, researchers created an argument-annotated corpus of the Russian language and evaluated the perfor-

mance of various classifiers for argumentation analysis, which also extended to the analysis of argumentation in COVID-19 datasets (Fishcheva and Kotelnikov 2019; Kotelnikov et al. 2022).

Regarding argumentation extraction, it is worth mentioning studies that have identified comments as a source for extracting arguments. Some authors have conducted such studies using comment corpora from YouTube videos (Sagredos and Nikolova 2022) and news posts (Ehret and Taboada 2020). However, Castellano Parra et al., while analyzing comments on news posts and social media pages of several Spanish newspapers, found that both the level of engagement (number of comments per reader) and the proportion of reasoned comments in social networks were higher compared to news portals (Castellano Parra et al. 2020). Additionally, researchers note that on both platforms, arguments and viewpoints tend to belong to a limited circle of active users, which can contribute to the monopolization of discourse by certain individuals and groups (Jensen 2016).

The analysis of arguments from social media comments has gained popularity in the research community, particularly due to the increased discussions surrounding COVID-19-related topics such as vaccination and quarantine (Melton et al. 2021; Wawrzuta et al. 2021, 2022; Karami and Anderson 2020). Furthermore, comments are utilized for monitoring and making operational decisions regarding COVID-19 and its consequences (Han et al. 2020; Oyeboode et al. 2021).

There are individual studies that focus on analyzing comments expressing attitudes towards abortion and its legalization (Hasan and Ng 2013; Graells-Garrido et al. 2019; Misra et al. 2017). It is also worth highlighting works that analyze comment content on other demography-related topics, including parenthood aspects (Mencarini et al. 2017), health problems (Shah et al. 2019), the impact of various factors on demographic processes such as natural disasters (Mandel et al. 2012) and the COVID-19 pandemic, including vaccination for children (Miao et al. 2020; Glandt et al. 2021; Liu and Liu 2021; Thorpe Huerta et al. 2021; Abosedra et al. 2021), sexual harassment or violence (Andalibi et al. 2016; Xue et al. 2019; Al-Rawi et al. 2021; Lin et al. 2022), and attitudes towards genetic tests (Mittos et al. 2020), among others.

The analysis of social network comments within the context of demographic topics is not yet widely explored in the Russian academic literature. However, the use of user-generated content in demographic research, in general, has been increasing in recent years. There have been attempts to analyze the sentiment of Russian-language comments on social networks to determine the “demographic temperature” within pro-natalist and anti-natalist groups of users (Kalabikhina et al. 2021a). Furthermore, studies have focused on analyzing the opinions of social network users regarding reproductive behavior based on comments on VKontakte (Kalabikhina et al. 2021b; 2022). It has been demonstrated that comments can be a better source for analyzing the mood of statements compared to posts (Kalabikhina et al. 2021a).

Additionally, several datasets consisting of comments by pronatalists and antinatalists have been published, which serve as the basis for research data (Kalabikhina and Banin 2020; 2021).

Other researchers primarily concentrate on assessing the emotional background of posts (Donchenko et al. 2017) and comments (Sidorov and Slastnikov 2021; Smetanin and Komarov 2021), without specifically extracting the arguments underlying users' positions.

Methodological approach to the typology of arguments (reasons)

Indeed, the automated extraction of arguments from comments and determining their nature is a challenging task. Automating the entire process of extracting arguments, especially in the context of demographic topics, is highly complex. However, manually processing texts after automatically selecting comments with arguments can be time-consuming. As a result, the initial step towards automating argument extraction has involved identifying broader classes of arguments that can be applied to various demographic topics.

In this study, arguments are classified based on the “personal-public” dichotomy, depending on whether the commentator relies on personal experiences or references public ideas as supporting evidence. This approach is rooted in the division of opinions regarding individual and ideal aspects of reproductive behavior. For example, when assessing reproductive attitudes, questions often involve the expected and ideal number of children. The expected number refers to the children the respondent personally intends to have, considering their individual circumstances. On the other hand, the ideal number represents the respondent’s perception of the societal norm at that given time. Sometimes, questions may address the desired number of children, which is close to a normative question but focuses on individual plans irrespective of specific circumstances.

This approach is relevant for studying reproductive behavior and people’s attitudes towards demographic and family policies. By distinguishing between personal arguments/stories and statements about social norms, rules of demographic behavior, or motivations behind others’ actions, it becomes possible to monitor the balance between “personal” and “public” perspectives. Monitoring the increasing share of “personal” arguments can provide valuable insights and help address concerns related to demographic topics. Filtering personal arguments allows for a more focused analysis and identification of specific arguments within the dataset.

It is important to note that the division between personal and public arguments is not limited to reproductive issues alone. Similar divisions can be observed in various domains, such as marital decisions, migration, or self-preservation behaviors, where personal and social experiences shape individuals’ attitudes, relationships, and actions. For instance, arguments related to vaccination, including personal reasons against vaccination and discussions about vaccination in society, can also be classified within this framework.

A similar approach was found in the article by Kiesel et al. (2022). The authors aimed to automate the search for values by utilizing sets of values from various opinion polls such as Schwartz and the World Values Survey. The primary objective of their work was to establish connections between people’s statements (in natural language) and their value orientations, and to automate this process by training artificial intelligence to make these connections. It is noteworthy for our study that the authors’ described system of values included a distinction between personal and public values.

Data

In our study, we utilize two sets of demographic data. The primary dataset consists of comments extracted from VKontakte, a social network, focusing on topics related to childbirth (Kalabikhina and Banin 2020; Kalabikhina and Banin 2021). The additional dataset comprises statements made by VKontakte users discussing various aspects of the coronavirus infection (Kotelnikov et al. 2022).

Reproductive data

Data collection

The source of textual data on reproductive topics is thematic groups on the social network VKontakte (vk.com). In the initial stage, we gathered unique group identifiers in the form vk.com/<unique group identifier> by utilizing the built-in API (application programming interface) and searching for keywords such as “mother,” “mommies,” “children,” “child,” “baby,” “health,” “birth,” “pregnancy,” “parents,” “childfree,” “child hate,” and others. Based on the type of keyword, these thematic groups were categorized into two subgroups: anti-natalists (keywords: “childfree,” “child hate,” etc.) and pronatalists (keywords: “family,” “children,” “baby,” “pregnancy”). We collected approximately 1000-1500 unique group addresses along with participant count data.

During the second stage, we excluded groups associated with advertising and those with low user activity from our sample. Since advertising posts were present in all groups, we prioritized groups that contained meaningful texts related to motherhood, fatherhood, demographic policy, and similar topics. We employed an automated method to filter out the least active groups, considering factors such as a small number of subscribers, low publication activity, minimal mention of keywords, and low comment activity. These parameters were determined iteratively. For pronatalist groups, the cut-off threshold for the number of users was set at 10,000 subscribers, resulting in the collection of 341 target groups of pronatalists. The number of antinatalist groups was considerably smaller, but they exhibited higher user activity, so our filtering process accounted for this specificity. For antinatalists, the cut-off threshold corresponded to 500 subscribers (given the smaller size of such groups), resulting in the selection of 8 active antinatalist groups.

- Pronatalists: maximum number of subscribers – 1482303, minimum – 72570, average – 309 000 subscribers per group.
- Antinatalists: maximum number of subscribers – 61071, minimum – 619, average – 8,950 subscribers per group.
- After the completion of the final list of groups, text information (posts and comments on posts) was collected from those groups. Based on the collected information, a comment corpus was formed:
 - all words were reduced to lowercase,
 - stop words¹ were removed using functions from the nltk or gensim library,
 - punctuation was removed,
 - numerical data was excluded.

To reduce the volume of text data, we employed stemming (removal of word endings) or lemmatization (reducing words to their base forms using the MyStem lemmatizer). The sample structure and the list of core groups were presented in previous works (Kalabikhina and Banin 2020; Kalabikhina and Banin 2021). After collecting the textual data, we conducted a keyword search within the collected texts to identify the most relevant texts for further annotation. Table 1 provides a list of topics and keywords that were used to select the texts.

¹ Stop words are frequently used words that do not add any additional information to the text. For example, we removed conjunctions, pronouns and prepositions that do not carry a semantic load

Table 1. Keyword lists for extracting statements on reproductive topics

Theme	Characteristic words
Maternity capital / child benefits	Maternity capital, payments, benefits
Abortion	Abortion
Large family	Large family, many children
Childlessness	Childfree childless, no children
Parental vacation	Maternity leave
Individualism	in one’s own, selfish, responsibility, for oneself, personality, develop

Source: compiled by the authors.

Annotation of reproductive data according to the author’s position and argumentation

During the first stage, users’ statements on reproductive topics were categorized based on the author’s position towards the given topic, namely “for,” “against,” or “other.” As part of this categorization process, irrelevant statements were excluded.

In the second stage, additional markup was added to the data, indicating the presence of arguments within the marked statements.

For the categorization based on the author’s position, a random selection of sentences from the collected sample was assigned to annotators for markup. Each sentence was independently labeled by three annotators. Since each sentence could potentially involve multiple topics, annotators provided labels for all six topics. The final label for each sentence was determined through majority voting based on the evaluations of the annotators. The reproductive dataset’s statistics by topics and positions can be found in Table 2 of the study conducted by Kalabikhina and Banin (2021). The total number of labeled comments in the sample amounted to 5,412.

During the second stage, the dataset on reproductive topics was further annotated to indicate the presence of arguments within each statement. Annotators determined whether a statement contained an argument that could be used in a dispute. If an argument was

Table 2. Distribution of author’s markups on topics related to the birth of children.

Topic	Relevant	For	Against	Other
Abortion	1374	709	161	504
Having many children	341	75	153	113
Parental leaves	992	201	376	435
Individualism	739	464	119	156
Maternity capital / Childcare benefits	813	184	370	259
Childlessness	1422	853	134	435
Total	5681*	2486	1313	1902

Note: the total number of markups (*) in the table is greater than the total number of comments analyzed since some comments contained opinions on several topics.

Source: compiled by the authors.

identified, the statement was additionally categorized based on the type of argumentation: either public or personal.

Preliminary analysis of comments

Before automatically extracting arguments, we conducted a preliminary manual analysis of the comments to identify potential approaches for their automatic classification. Table 3 presents examples of comments on each of the analyzed topics, with the authors' spelling and punctuation retained.

Table 3. Examples of comments from the analyzed database on topics of interest

Topic	Comment examples
Abortion	<ul style="list-style-type: none"> • “in the 20-30s, industry was reviving at a rapid pace, it was not easy, but the Soviet people survived the war, and they are talking about abortion as a factor in population decline, nonsense”; • “and how to avoid it? move to another country to have an abortion? not everyone has this possibility.”
Maternity capital / Childcare benefits	<ul style="list-style-type: none"> • “maternity capital does not motivate it is useless”; • “matcapital serves well for those who already have their own corner, apartment + one room, the homeless (oh, renters) have nothing to count on there.”
Childlessness	<ul style="list-style-type: none"> • “no children no problems”; • “Without children, all the quarrels in the family flare up.”
Having many children	<ul style="list-style-type: none"> • “why have many children, I think one should be dressed, put, grown up, fed, watered and given education”; • “Many children = terrible mother.”
Parental leaves	<ul style="list-style-type: none"> • “recently I was at an interview, the hr-manager was a girl about my age, at the end of the interview she asked an epic question when I was planning a maternity leave”; • “after all, maternity leave is not sitting on your neck and does not eat from your pocket.”

Note: the texts of the comments are given in their original form.

Source: compiled by the authors.

Not all comments contain arguments, and there are various types of arguments for each topic, which makes it challenging to identify unified arguments that can be automatically extracted across topics. We attempted to find several typical arguments for our topics, which include reproductive behavior and attitudes towards pronatalist demographic policy. In our view, classifying arguments as positive or negative, as opposed to measuring the emotional tone of the texts themselves, has limited practical value, as such classification partially duplicates the assessment of the emotional tone of the texts and the arguments within them. Making a political decision based solely on the proportion of positive/negative arguments would require additional manual analysis of the content. Therefore, we decided to categorize arguments into more general types: **personal** arguments (statements about personal experiences, attitudes, values, stories) and **public** arguments (statements about ideal attitudes, values, norms, and recommendations on behavior).

Classification of arguments into personal and public

In demography, such a criterion is important for determining demographic behavior. For instance, in surveys on the number of children, there are different types of questions aimed at distinguishing between “public” attitudes and individual actions. Two examples of such questions are:

1. “What is the ideal number of children in your opinion?”
2. “How many children do you plan to have based on your capabilities?”

It has been demonstrated in numerous datasets from various countries that the response to the second question predicts the birth rate of corresponding generations very effectively.

What will monitoring personal and public arguments provide? While we still may not be able to make political decisions based solely on the change in the proportion of personal/public arguments without additional manual analysis of the content, it offers a new way to evaluate the situation. An increase in the share of personal arguments serves as a signal of an intensification of issues, indicating the need for careful manual examination of the argument texts during that period. If individuals shift from discussing matters in general terms to sharing personal stories, it is often associated with a rise in problems related to the selected topics.

Personal and public arguments were also classified in the topics of large families, childlessness, abortion, benefits, and holidays.

Personal cases typically involve narratives about the experiences or current situations of the authors or their families, or their personal opinions on the matter.

Public cases involve the authors’ reflections on how to (or not to) live and behave in a demographic context, encompassing all residents of a country, region, or social group.

Examples of comments with arguments in the categories “personal” and “public” are presented in Table 4.

Table 4. Examples of comments with arguments in the categories “personal” and “public”

Type of argument (personal/ public)	A comment
Personal	<ul style="list-style-type: none"> • “I have been married for 5 years, no children! repeatedly changed jobs, and no one was interested in why we do not have children and when I will go on maternity leave! only work experience, education, characteristics and personal qualities”
Personal	<ul style="list-style-type: none"> • “And I also know that it’s a shame to feel like a beggar, that your peers poke a finger at you and say that you are from a large family of a rogue when you don’t have your own corner of the house, when the cheapest ones buy notebooks and there aren’t enough of them, therefore I can say first-hand that children are the responsibility of the parents and you need to think about what you will support them for, whether you can provide them not only with a future, but also just a normal existence».
Personal	<ul style="list-style-type: none"> • “Personally, my Children, this is my meaning of life, I love them very much, and I love them more than life!”
Personal	<ul style="list-style-type: none"> • “I have three children, too, they make a mess better than without children. There is no meaning to life without children”

Type of argument (personal/ public)	A comment
Personal	<ul style="list-style-type: none"> • “My mother gave birth to 12 children and she never smelled bad and she made masterpieces out of nothing ... girls, let’s not forget about ourselves, no matter how hard it is for us, for one simple reason, that we are women, and also for the fact that our children and our husbands would be proud of us”
Personal	<ul style="list-style-type: none"> • “Even without a calculator, I can perfectly understand that I won’t be able to keep my wife and child on maternity leave: I saw the prices for children a couple of times – my inner Jew told me in detail what he thinks about me in advance and where should I go if I suddenly decide to do your poor photocopy”
Public	<ul style="list-style-type: none"> • “in countries with a ban on abortion, the crime rate goes off scale”
Public	<ul style="list-style-type: none"> • “the author correctly said – no children until you get on your feet, and even better until you live for yourself to the fullest”
Public	<ul style="list-style-type: none"> • “society receives more benefits from childfree than from children, we don’t have a decree, we don’t have to take time off from work because of children’s problems and illnesses, we don’t take a place in schools, kindergartens and clinics, for all kinds of care payments, maternity capital”
Public	<ul style="list-style-type: none"> • “and then send both children to an orphanage? What will happen then? Have you already spent your maternity capital on improving living conditions, for example”
Public	<ul style="list-style-type: none"> • “this is an abnormal situation in the country, when salaries are beggarly, prices for everything are sky-high, child allowance is also pennies, and a woman on maternity leave falls out of life and becomes mega-dependent”
Public	<ul style="list-style-type: none"> • “so I’m personally for abortion and not for ruined lives”

Note: the texts of the comments are given in their original form.

Source: compiled by the authors.

It is worth highlighting that during the markup, there were some controversial cases: some comments were difficult to classify into one of two categories – personal or public. For example, some comments contained elements of both: “Here I am raising three children with child benefits, and everyone else must too!” This sentence, on one hand, describes the personal experience of the author, while on the other hand, it expresses ideas about how society should be organized. This comment is also noteworthy because it likely contains an element of irony. The author is probably quoting their opponents and ridiculing their opinion. There were a lot of comments with irony and citations in the sample, but they were not excluded at this stage since our current research tasks do not require it.

Manual markup of comments

Manual markup of comments was performed by six annotators, who are also the authors of this article. Each comment was independently evaluated by three annotators. The markup process involved two parameters: 1) the presence of an argument (whether an argument was present or not); 2) the type of argument (personal or public). In the first step, annotators determined whether a comment contained an argument. Arguments were defined as argumentative statements that could be used to persuade an opponent regarding a particular

viewpoint, regardless of whether the argument was in favor of or against that viewpoint. The statement was marked as an argument if it had the potential to be used as such in a dispute.

Here are examples illustrating the markup process:

Example 1 (no argument, despite a negative attitude towards abortion):

“Of course, they don’t care about fetuses and embryos themselves: – they never oppose selective abortions of girls (Caucasus, Central Asia, Arab countries, China, etc.)”

Example 2 (no argument, despite a positive attitude towards maternity leave):

“The fact that she is sitting at home on maternity leave does not necessarily mean that she is doing well and resting.”

Example 3 (argument present, expressing a negative attitude towards maternity leave due to the perceived loss of a woman’s professional skills):

“She will come out of maternity leave, come to work, and it turns out that she doesn’t know how to work – she doesn’t need a credit in her work, she needs a result.”

In the second stage, comments identified as containing arguments in the first stage were further marked to indicate whether the argument was of a “personal” or “public” nature. Thus, each comment in the reproductive dataset includes the independent markup results from three annotators.

On average, there was a satisfactory agreement (over 70%) among the annotators regarding the presence of an argument and its type. Please refer to Tables 5 and 6 for more details.

Table 5. Results of manual markup of the comment dataset. The number of matches by argument presence parameter between each two annotators

Annotators	Number of matches: argument / no argument					
	Anno-tators 1 and 3	Anno-tators 1 and 2	Anno-tators 2 and 3	Anno-tators 4 and 5	Anno-tators 4 and 6	Anno-tators 5 and 6
% of matches	86,4	84,2	78,6	73,8	73,3	71,2
Number of matches	2336	2274	2124	1997	1985	1928
Total Marked Comments	2705	2705	2705	2707	2707	2707

Source: compiled by the authors.

Table 6. Results of manual markup of the comment base. The number of matches by argument type parameter between each two annotators. Note: Only cases where both annotators marked the comment as containing an argument were considered

Annotators	Number of matches: personal / public *					
	Anno-tators 1 and 3	Anno-tators 1 and 2	Anno-tators 2 and 3	Anno-tators 4 and 5	Anno-tators 4 and 6	Anno-tators 5 and 6
% of matches	82,1	90,3	77,8	84,0	74,5	80,6
Number of matches	724	646	526	516	533	625
Total comments (which both markups marked as suggestive)	883	719	677	617	717	777

Source: compiled by the authors

Descriptive statistics of the markup results of each annotator are presented in Table 7.

Table 7. Descriptive statistics of the markup results of each annotator

Annotator	Arguments		of which: personal stories		Total Marked Comments
	abs.	%	abs.	%	abs.
1st group of comments					
Annotator 1	1034	38,2	381	36,8	2705
Annotator 2	835	30,9	303	36,3	2705
Annotator 3	1101	40,7	578	52,5	2705
2nd group of comments					
Annotator 4	953	35,2	335	35,2	2707
Annotator 5	1051	38,8	323	30,7	2707
Annotator 6	1280	47,3	516	40,3	2707

Source: compiled by the authors.

COVID-19 Data

The COVID-19 data utilized in this study consisted of comments from VKontakte users on news reports pertaining to COVID-19. The comments were selected based on keywords related to three aspects of the COVID-19 pandemic: “masks,” “quarantine,” and “vaccination” (Nugamanov et al. 2021).

In the first stage, the comments were annotated according to the author’s stance on the mentioned aspects, categorized as “for,” “against,” or “other.” In the second stage, the annotated statements were further analyzed to determine the presence of an argument substantiating the author’s position on a particular aspect (Kotelnikov et al. 2022), specifically focusing on the presence of arguments “for,” arguments “against,” or no argument. A total of 9,550 statements were annotated, with approximately 2,000 of them containing arguments. For this study, the presence of an argument, irrespective of the topic, was used as the dataset for pre-training the Conversational RuBERT model. By incorporating COVID-19 data into the reproductive dataset, the aim was to investigate the potential for enhancing the processing outcomes of the target collection through pre-training the model on data from a distinct but related topic encompassing population and demographic studies.

Experiments

In this study, two classification tasks were conducted: one to determine the presence¹ of an argument and another to identify the type of argument (public or personal).

A total of six experiments were performed, each addressing a specific problem. Experiments 1-3 focused on determining the presence of an argument, while Experiments 4-6

¹ Initially, the presence of an argument was attempted to be defined using conjunctions such as “because” or “thus.” However, after evaluating the results, it was decided to manually annotate the comments for the presence of an argument to be used later in machine learning.

aimed to classify the type of argument. The data used in Experiments 2 and 3 included both reproductive and COVID-19 data, while the remaining experiments utilized only reproductive data. The training, testing, and validation samples were divided in a ratio of approximately 80:10:10 to ensure an even distribution across the main topics of interest (as described in the “Data collection” section).

The empirical part of the study was conducted using the Python environment with the PyTorch, Transformers, and Scikit Learn libraries. The Conversational RuBERT model¹ served as the basis for all experiments. This model is a pre-trained Russian language model that has been additionally trained on social network data and user dialogues. The experiments employed specific parameters, including a learning rate of 0.0005 and a batch size of 64. The number of epochs, ranging from 4 to 6, was selected based on the best F-score obtained on the validation set. The model’s performance was evaluated using accuracy and F1-score metrics, and the results are summarized in Tables 8 and 9. More detailed information on the model’s quality can be found in the Appendix, specifically Tables A1-A6.

Results

Presence of an argument

For the task of detecting the presence of arguments the comments of VKontakte users were classified into two classes: Class “1” if at least two out of three annotators agreed that the comment contained an argument, and Class “0” if only one annotator or none of them indicated the presence of an argument. To solve this problem, three variants of the experiment were considered (metrics for comparison are presented in Table 8). The ratio of comments by class averaged 40:60 for comments with arguments versus those without arguments.

In Experiment 1, the reproductive data was used, consisting of comments from VKontakte users on reproductive behavior and the evaluation of demographic policy measures. The training, test, and validation sample comprised a total of 5,410 comments. Standard pre-processing techniques were applied, including removing capital letters, punctuation, stop words, and empty comments.

Table 8. Classification model by the presence of an argument (Class “0” – no argument, Class “1” – presence of an argument).

	F-score – Class “0”	F-score – Class “1”	Accuracy
Experiment 1: reproductive data only	0,81	0,61	0,75
Experiment 2: reproductive and COVID-19 data	0,82	0,49	0,73
Experiment 3: pre- training on COVID-19 data, fine-tuning on reproductive data	0,79	0,48	0,70

Source: created by the authors.

Notes: Experiment 1 used only reproductive data, Experiment 2 used reproductive and COVID-19 data in the training set while only demographic data was used in the test and validation sets, Experiment 3 used COVID-19 data to pre-train the model, while fine-tuning was conducted on reproductive data.

¹ <https://huggingface.co/DeepPavlov/rubert-base-cased-conversational>

In Experiment 2, COVID-19-related comments from VKontakte users, such as vaccination, mask-wearing, and quarantine restrictions, were added to the reproductive data. The new comments were included only in the training set, and the additional data consisted of 6,716 comments, forming the training sample for the COVID-19 data.

In Experiment 3, training was conducted in two stages: pre-training on the COVID-19 data and fine-tuning and evaluation of the model on the reproductive data.

The results in Table 8 indicate that the addition of COVID-19 data did not improve the results of argument extraction in the reproductive collection. The model trained on the reproductive data alone performed better in this task.

Type of Argument

In addition to identifying the presence of arguments, our main interest was to determine the content of arguments, specifically the distinction between personal and public arguments. This classification task was important for our study, even at a high level of generalization, to ensure that the model's performance was not sensitive to specific argument topics, which could affect its overall quality.

As described above, we classified arguments based on the “personal-public” context, i.e., whether the commentator used personal experiences or cited public ideas as an argument.

This task was more challenging than identifying the presence of an argument due to the great number of controversial cases where statements contained elements of both types, but annotators should have chosen only one class to assign to each comment.

To address this, we conducted two types of experiments. In Experiments 4 and 5, we focused on classifying comments with argumentation into two classes. The reproductive data used for training was reduced to 2,049 comments, including only those where at least two out of three annotators indicated the presence of an argument. In Experiment 5, Class “1” represented comments with personal arguments, as agreed upon by at least two annotators, while Class “0” included all other comments, including controversial cases. The ratio of observations between Class “0” and Class “1” was approximately 65:35. Similarly, Experiment 5 classified comments into public-type argumentation, with a ratio of observations between Class “0” and Class “1” at approximately 45:55.

In Experiment 6, we trained the model on the full set of reproductive data, which was divided into three classes: “0” for no argumentation, “1” for public-type argumentation, and “2” for personal-type argumentation. The number of comments in each class was approximately 70:15:15.

The classification results showed that Experiments 1 and 2 achieved the best performance in identifying the presence of an argument. This can be attributed to Experiment 1 training the model exclusively on reproductive data, which reduced the risk of misclassification due to topic differences between reproductive behavior and the pandemic. Moreover, Experiment 2 had a larger number of observations in the training set, which likely contributed to its higher quality compared to other experiments.

In terms of determining the type of argument on the comparison of the quality metric, Experiments 4 and 5 performed better than Experiment 6. This outcome was expected as Experiments 4 and 5 addressed a simpler problem with two classes instead of three. Additionally, these experiments benefited from more reliable data since the reduced sample automatically excluded many ambiguous cases. The distribution of observations across classes in Experiments 4 and 5 was also more balanced compared to Experiment 6.

Table 9. Classification model according to the type of argument (“personal” – based on personal experience, “public” – based on public perceptions).

	F-score – class “0”	F-score – class “1”	F-score – class “2”	Accuracy
Experiment 4: two classes with the emphasis on personal arguments	0,86	0,71	-	0,81
Experiment 5: two classes with the emphasis on public arguments	0,70	0,81	-	0,77
Experiment 6: three classes	0,78	0,52	0,34	0,67

Source: compiled by the authors.

Note: Experiments 4 and 5 utilize restricted reproduction data, which was sampled based on the criterion that “at least 2 out of 3 annotators agreed that the comment contains an argument.” In Experiment 4, Class “1” comprises comments with personal arguments, as deemed by at least 2 out of 3 annotators. All other comments are classified as Class “0”. In Experiment 5, Class “1” consists of comments with public arguments, as determined by at least 2 out of 3 annotators. All other comments are categorized as Class “0”. In Experiment 6, the full reproductive data was utilized and divided into three classes: “0” indicating no argument, “1” representing public arguments, and “2” denoting personal arguments.

Conclusion

The annotation of user comments related to reproductive behavior in the VKontakte social network revealed that approximately 40% of the selected comments contained an argument expressing the author’s positive or negative attitude towards the demographic topic. Among these comments, around 40% were classified as personal arguments, while the remaining were classified as public arguments based on social attitudes, values, and norms. These findings indicate that there is a solid foundation for analyzing the arguments related to demographic viewpoints in social network user comments and identifying personal and public arguments.

The method employed in this study, which involved automatically extracting user opinions from VKontakte comments and classifying them using the Conversational RuBERT neural network model, demonstrated its effectiveness in accurately identifying the presence of arguments and determining their types in terms of the personal-public dichotomy.

Throughout the study, six experiments were conducted, varying the dataset and the number of classes, to address the challenges of argument retrieval and classification. The results obtained indicate that the developed model achieved a high level of accuracy in classifying user comments. This automated approach can be further utilized for data analysis in similar contexts. It has the potential to monitor social networks and promptly identify an increase in personal-type critical statements regarding demographic issues, contributing to the improvement of pronatalist socio-demographic policies.

Moreover, the study opens up new possibilities for applying the developed algorithm to diagnose other types of demographic behaviors, such as self-preservation, matrimonial, and migratory behaviors. Testing the algorithm in the analysis of these behaviors can help identify the scale of personal-type statements and track the dynamics of such statements over time.

In conclusion, the findings of this study support the effectiveness and accuracy of the developed model in classifying user comments. The proposed approach can enable automation in analyzing similar types of data, paving the way for future research and applications in the field of demographic behavior analysis (self-preservation, matrimonial or migratory).

Acknowledgements

The study was carried out within the framework of research work «The reproduction of the population in socio-economic development». The work of Natalia Loukachevitch on the COVID-2019 data and their use in the current reproductive behavior study was supported by the Russian Science Foundation (grant No. 21-71-30003).

References

- Abosedra S, Laopodis NT, Fakhri A (2021) Dynamics and asymmetries between consumer sentiment and consumption in pre-and during-COVID-19 time: Evidence from the US. *The Journal of Economic Asymmetries* 24: e00227. <https://doi.org/10.1016/j.jeca.2021.e00227>
- Al-Rawi A, Grepin K, Li X, Morgan R, Wenham C, Smith J (2021) Investigating public discourses around gender and COVID-19: a social media analysis of Twitter data. *Journal of Healthcare Informatics Research* 5: 249–69. <https://doi.org/10.1007/s41666-021-00102-x>
- Andalibi N, Haimson OL, De Choudhury M, Forte A (2016) Understanding social media disclosures of sexual abuse through the lenses of support seeking and anonymity. In: *Proceedings of the 2016 CHI conference on human factors in computing systems*, San Jose (USA), May 7–12. Association for Computing Machinery, New York, 3906–18. <https://doi.org/10.1145/2858036.2858096>
- Castellano Parra O, Meso Ayerdi K, Pena Fernandez S (2020) Behind the comments section: The ethics of digital native news discussions. *Digital Native News Media: Trends and Challenges*: 8(2). <https://doi.org/10.17645/mac.v8i2.2724>
- Donchenko D, Ovchar N, Sadovnikova N, Parygin D, Shabalina O, Ather D (2017) Analysis of comments of users of social networks to assess the level of social tension. *Procedia Computer Science* 119: 359–67. <https://doi.org/10.1016/j.procs.2017.11.195>
- Ehret K, Taboada M (2020) Are online news comments like face-to-face conversation? A multi-dimensional analysis of an emerging register. *Register Studies* 2(1): 1–36. <https://doi.org/10.1075/rs.19012.ehr>
- Fishcheva I, Kotelnikov E (2019) Cross-lingual argumentation mining for Russian texts. In: WMP van der Aalst et al. (eds) *Analysis of Images, Social Networks and Texts*. International Conference, Kazan (Russia), July 17–19. Springer, Cham: 134–44. https://doi.org/10.1007/978-3-030-37334-4_12
- Glandt K, Khanal S, Li Y, Caragea D, Caragea C (2021) Stance Detection in COVID-19 Tweets. In: *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint 46 Conference on Natural Language Processing (Volume 1: Long Papers)*. Stroudsburg (USA), August 1–6. Association for Computational Linguistics, 1596–611. <https://doi.org/10.18653/v1/2021.acl-long.127>
- Graells-Garrido E, Baeza-Yates R, Lalmas M (2019) How representative is an abortion debate on twitter? In: *Proceedings of the 10th ACM Conference on Web Science*, Boston (USA), 30 June–3 July. Association for Computing Machinery, New York, 133–4. <https://doi.org/10.1145/3292522.3326057>
- Han X, Wang J, Zhang M, Wang X (2020) Using social media to mine and analyze public opinion related to COVID-19 in China. *International Journal of Environmental Research and Public Health* 17(8): 2788. <https://doi.org/10.3390/ijerph17082788>
- Hasan KS, Ng V (2013) Stance classification of ideological debates: Data, models, features, and constraints. In: *Proceedings of the sixth international joint conference on natural language processing*, Nagoya (Japan), October 14–18. Asian Federation of Natural Language Processing, 1348–56. URL: <https://aclanthology.org/I13-1191.pdf>

- Kalabikhina IE, Banin EP (2020) Database “Pro-family (pronatalist) communities in the social network VKontakte”. *Population and Economics* 4(3): 98–103. <https://doi.org/10.3897/popecon.4.e60915>
- Kalabikhina IE, Banin EP (2021) Database “Childfree (antinatalist) communities in the social network VKontakte”. *Population and Economics* 5(2): 92–6. <https://doi.org/10.5281/zenodo.4612131>
- Kalabikhina IE, Banin EP, Abduselimova IA, Klimentko GA, Kolotusha AV (2021a) The measurement of demographic temperature using the sentiment analysis of data from the social network VKontakte. *Mathematics* 9(9): 987. <https://doi.org/10.3390/math9090987>
- Kalabikhina IE, Lukashevich NV, Banin EP, Alibayeva KV, Rebrey SM (2021b) Avtomaticheskoye izvlecheniye mneniy pol’zovateley sotsial’nykh setey po voprosam reproduktivnogo povedeniya [Automated extraction of social network users’ opinions on reproductive behaviour]. *Programmnyye sistemy: teoriya i prilozheniya* 4(51): 33–63. <https://doi.org/10.25209/2079-3316-2021-12-4-33-63> (in Russian)
- Kalabikhina IE, Lukashevich NV, Banin EP, Alibayeva KV (2022) Avtomaticheskii analiz reproduktivnykh tsnostey pol’zovateley seti vkontakte [Automated analysis of reproductive values of Vkontakte users]. *Intellektual’nyye sistemy. Teoriya i prilozheniya* 1(26): 90–6. (in Russian)
- Karami A, Anderson M (2020) Social media and COVID-19: Characterizing anti-quarantine comments on Twitter. *Proceedings of the Association for Information Science and Technology* 57(1): e349. <https://doi.org/10.1002/pra2.349>
- Kiesel J, Alshomary M, Handke N, Cai X, Wachsmuth H, Stein B (2022) Identifying the Human Values behind Arguments. In: *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, Dublin (Ireland), May 22–27. Association for Computational Linguistics, 4459–71. <https://doi.org/10.18653/v1/2022.acl-long.306>
- Kotelnikov E, Loukachevitch N, Nikishina I, Panchenko A (2022) RuArg-2022: Argument Mining Evaluation. In: *Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference “Dialogue 2022”, Moscow (Russia), June 15–18*, 333–47. <https://doi.org/10.28995/2075-7182-2022-21-333-348>
- Lin AE, Young JA, Guarino JE (2022) Mother-Daughter sexual abuse: An exploratory study of the experiences of survivors of MDSA using Reddit. *Children and Youth Services Review* 138: 106497. <https://doi.org/10.1016/j.childyouth.2022.106497>
- Liu S, Liu J (2021) Public attitudes toward COVID-19 vaccines on English-language Twitter: A sentiment analysis. *Vaccine* 39(39): 5499–505. <https://doi.org/10.1016/j.vaccine.2021.08.058>
- Liu S, Li J, Liu J (2021) Leveraging transfer learning to analyze opinions, attitudes, and behavioral intentions toward COVID-19 vaccines: social media content and temporal analysis. *Journal of Medical Internet Research* 23(8): e30251. <https://doi.org/10.2196/30251>
- Mandel B, Culotta A, Boulahanis J, Stark D, Lewis B, Rodrigue J (2012) A demographic analysis of online sentiment during hurricane Irene. In: *Proceedings of the 2012 Workshop on Language in Social Media*, Montréal (Canada), June 7. Association for Computational Linguistics, 27–36. URL: <https://aclanthology.org/W12-2104.pdf>
- Melton CA, Olusanya OA, Ammar N, Shaban-Nejad A (2021) Public sentiment analysis and topic modeling regarding COVID-19 vaccines on the Reddit social media platform: A call to action for strengthening vaccine confidence. *Journal of Infection and Public Health* 14(10): 1505–12. <https://doi.org/10.1016/j.jiph.2021.08.010>
- Mencarini L, Hernandez Farias DI, Lai M, Patti V, Sulis E, Vignoli D (2017) Happy parents’ tweets: An exploration of Italian Twitter data using sentiment analysis. *Demographic Research*: 40: 25. <https://doi.org/10.4054/DemRes.2019.40.25>
- Miao L, Last M, Litvak M (2020) Twitter data augmentation for monitoring public opinion on COVID-19 intervention measures. In: *Proceedings of the 1st Workshop on NLP for COVID-19 (Part 2) at EMNLP 2020*, Online. Association for Computational Linguistics. <https://doi.org/10.18653/v1/2020.nlpccovid19-2.19>

- Misra A, Oraby S, Tandon S, Ts S, Anand P, Walker M (2017) Summarizing dialogic arguments from social media. In: V.Petukhova, Y.Tian (eds.) Proceedings of the 21st Workshop on the Semantics and Pragmatics of Dialogue, Saarbrücken (Germany), August 15-17, 126-36 <https://doi.org/10.21437/SemDial.2017-14>
- Mittos A, Zannettou S, Blackburn J, Cristofaro ED (2020) Analyzing genetic testing discourse on the Web through the lens of Twitter, Reddit, and 4chan. *ACM Transactions on the Web (TWEB)* 14(4): 1–38. <https://doi.org/10.1145/3404994>
- Nugamanov E, Loukachevitch N, Dobrov B (2021) Extracting sentiments towards COVID-19 aspects. In: A. Pozanenko et al. (eds.) Supplementary Proceedings of the XXIII International Conference on Data Analytics and Management in Data Intensive Domains (DAMDID/RCDL 2021), Moscow (Russia), October 26-29, 299–312. URL: paper24.pdf (ceur-ws.org)
- Oyebode O, Ndulue C, Adib A, Mulchandani D, Suruliraj B, Orji FA, Orji R (2021) Health, psychosocial, and social issues emanating from the COVID-19 pandemic based on social media comments: text mining and thematic analysis approach. *JMIR medical informatics* 9(4): e22734. <https://doi.org/10.2196/22734>
- Sagredos C, Nikolova E (2022) 'Slut I hate you': A critical discourse analysis of gendered conflict on YouTube. *Journal of Language Aggression and Conflict* 10(1): 169–96. <https://doi.org/10.1075/jlac.00065.sag>
- Shah Z, Martin P, Coiera E, Mandl KD, Dunn AG (2019) Modeling spatiotemporal factors associated with sentiment on Twitter: synthesis and suggestions for improving the identification of localized deviations. *Journal of medical Internet research* 21(5): e12881. <https://doi.org/10.2196/12881>
- Sidorov N, Slastnikov S (2021) Some Features of Sentiment Analysis for Russian Language Posts and Comments from Social Networks. *Journal of Physics: Conference Series* 1740(1): 12–36. <https://doi.org/10.1088/1742-6596/1740/1/012036>
- Smetanin S, Komarov M (2021) Share of Toxic Comments among Different Topics: The Case of Russian Social Networks. In: 2021 IEEE 23rd Conference on Business Informatics (CBI), Bolzano (Italy), September 1-3, 65–70. <https://doi.org/10.1109/CBI52690.2021.10056>
- Thelwall M, Stuart E (2019) She's Reddit: A source of statistically significant gendered interest information? *Information processing & management* 56(4): 1543–58. <https://doi.org/10.1016/j.ipm.2018.10.007>
- Thorpe Huerta D, Hawkins JB, Brownstein JS, Hswen Y (2021) Exploring discussions of health and risk and public sentiment in Massachusetts during COVID-19 pandemic mandate implementation: A Twitter analysis. *SSM-Population Health* 15: 100851. <https://doi.org/10.1016/j.ssmph.2021.100851>
- Wawrzuta D, Jaworski M, Gotlib J, Panczyk M (2021) What arguments against COVID-19 vaccines run on Facebook in Poland: content analysis of comments. *Vaccines* 9(5): 481. <https://doi.org/10.3390/vaccines9050481>
- Wawrzuta D, Klejdysz J, Jaworski M, Gotlib J, Panczyk M (2022) Attitudes toward COVID-19 Vaccination on Social Media: A Cross-Platform Analysis. *Vaccines* 10(8): 1190. <https://doi.org/10.3390/vaccines10081190>
- Xue J, Macropol K, Jia Y, Zhu T, Gelles RJ (2019) Harnessing big data for social justice: An exploration of violence against women-related conversations on Twitter. *Human Behavior and Emerging Technologies* 1(3): 269–79. <https://doi.org/10.3390/10.1002/hbe2.160>

Other sources of information

- Jensen E (2016) NPR website to get rid of comments // National Public Radio, August 17. URL: <https://www.npr.org/sections/publiceditor/2016/08/17/489516952/npr-website-to-get-rid-of-comments>

Appendix

Table A1. Only reproductive data was used, categorized into two classes: “0” representing no argument and “1” indicating the presence of an argument

	Precision	Recall	F1-score	Number of observations in the validation set
Class “0”	0,73	0,91	0,81	324
Class “1”	0,79	0,50	0,61	218
Accuracy			0,75	542
Macro avg	0,76	0,70	0,71	542
Weighted avg	0,76	0,75	0,73	542

Source: compiled by the authors.

Table A2. Reproductive and COVID-19 data were exclusively used in the training sample, which consisted of two classes: “0” representing no argument and “1” indicating the presence of an argument

	Precision	Recall	F1-score	Number of observations in the validation set
Class “0”	0,72	0,94	0,82	346
Class “1”	0,77	0,36	0,49	195
Accuracy			0,73	541
Macro avg	0,75	0,65	0,66	541
Weighted avg	0,74	0,73	0,70	541

Source: compiled by the authors.

Table A3. The pre-training was conducted first on COVID-19 data and then on reproductive data, with a focus on two classes

	Precision	Recall	F1-score	Number of observations in the validation set
Class “0”	0,68	0,93	0,79	324
Class “1”	0,77	0,35	0,48	218
Accuracy			0,70	542
Macro avg	0,73	0,64	0,64	542
Weighted avg	0,72	0,70	0,66	542

Source: compiled by the authors.

Table A4. Only reproductive data was used in Experiments 4 and 5, but in a restricted format. The data was selected according to the following criteria: if at least 2 out of 3 annotators agreed that an argument was present, the comment was included. The data was then divided into two classes: “1” indicating personal according to at least 2 out of 3 annotators, and “0” representing all other cases

	Precision	Recall	F1-score	Number of observations in the validation set
Class “0”	0,94	0,79	0,86	202
Class “1”	0,61	0,87	0,71	76
Accuracy			0,81	278
Macro avg	0,77	0,83	0,79	278
Weighted avg	0,85	0,81	0,82	278

Source: compiled by the authors.

Table A5. The data selection process followed the following criteria: only comments in which at least 2 out of 3 annotators agreed that an argument was present were included. These comments were then categorized into two classes: “1” indicating that the comment was deemed public by at least 2 out of 3 annotators, and “0” representing all other cases

	Precision	Recall	F1-score	Number of observations in the validation set
Class “0”	0,62	0,80	0,70	94
Class “1”	0,88	0,76	0,81	184
Accuracy			0,77	278
Macro avg	0,75	0,78	0,76	278
Weighted avg	0,79	0,77	0,78	278

Source: compiled by the authors.

Table A6. Only reproductive data was used, categorized into three classes. The classification is as follows: “0” indicates no argument, “1” represents public arguments, and “2” signifies personal arguments

	Precision	Recall	F1-score	Number of observations in the validation set
Class “0”	0,72	0,85	0,78	324
Class “1”	0,57	0,48	0,52	123
Class “2”	0,46	0,27	0,34	95
Accuracy			0,67	542
Macro avg	0,59	0,54	0,55	542
Weighted avg	0,64	0,67	0,65	542

Source: compiled by the authors.

Information about the authors

- Kalabikhina Irina – Hab.Doctor (Economics), Professor, Head of Population Department, Faculty of Economics, Lomonosov Moscow State University, Moscow, 119991, Russia. E-mail: kalabikhina@econ.msu.ru
- Zubova Ekaterina – Fox Fellow at Yale University, New Haven, CT, 06511, USA. E-mail: ekaterina.zubova@yale.edu
- Loukachevitch Nataila – Hab.Doctor (Technical Sciences), leading researcher of Research Computing Center of Lomonosov Moscow State University. Moscow, 119899, Russia. E-mail: louk_nat@mail.ru
- Kolotusha Anthony – Candidate of Economic Sciences (Ph.D.), 2nd category Programmer, Faculty of Economics, Lomonosov Moscow State University, Moscow, 119991, Russia. E-mail: koav13a@econ.msu.ru
- Kazbekova Zarina – researcher of Population Department and Digital Economy Research Laboratory, Faculty of Economics, Lomonosov Moscow State University, Moscow, 119991, Russia. E-mail: kazbekova@econ.msu.ru
- Banin Evgeniy – Candidate of Technical Sciences (Ph.D.), researcher of Kurchatov Institute, Moscow, 123182, Russia. E-mail: evg.banin@gmail.com
- Klimenko German – postgraduate student, Faculty of Economics, Lomonosov Moscow State University, Moscow, 119991, Russia. E-mail: german89000@mail.ru