Министерство науки и высшего образования Российской Федерации
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ ИТМО
ITMO University

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА/GRADUATION THESIS

Токсичный твиттер: сравнительный анализ использования слова "токсичный" в социальных сетях / Toxic Twitter: the comparative analysis of the word "toxic" usage in social networks

Автор/ Author
Селина Мария Андриановна
Направленность (профиль) образовательной программы/Major
Анализ культурных данных и визуализация/ Data, Culture and Visualization 2019
Квалификация/ Degree level
Магистр
Руководитель ВКР/ Thesis supervisor
Хлопотов Максим Валерьевич, кандидат технических наук, Университет ИТМО, факультет инфокоммуникационных технологий, доцент (квалификационная категория "ординарный доцент")
Группа/Group
D42423
Факультет/институт/кластер/ Faculty/Institute/Cluster
институт международного развития и партнерства
Направление подготовки/ Subject area
45.04.04 Интеллектуальные системы в гуманитарной сфере

Обучающийся/Student	Документ подписан Селина Мария Андриановна 31.05.2021	Селина Мария Андриановна
	(эл. подпись/ signature)	(Фамилия И.О./ name and surname)
Руководитель ВКР/ Thesis supervisor	Документ подписан Хлопотов Максим Валерьевич 01.06.2021	Хлопотов Максим Валерьевич
	(эл. подпись/ signature)	(Фамилия И.О./ name and surname)

Министерство науки и высшего образования Российской Федерации
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ ИТМО
ITMO University

ЗАДАНИЕ НА ВЫПУСКНУЮ КВАЛИФИКАЦИОННУЮ РАБОТУ / OBJECTIVES FOR A GRADUATION THESIS

Обучающийся / Student Селина Мария Андриановна
Группа/Group D42423
Факультет/институт/кластер/ Faculty/Institute/Cluster институт международного развития и партнерства
Квалификация/ Degree level Магистр
Направление подготовки/ Subject area 45.04.04 Интеллектуальные системы в гуманитарной сфере
Направленность (профиль) образовательной программы/Major Анализ культурных данных и визуализация/ Data, Culture and Visualization 2019
Специализация/ Specialization Методы исследования культурных данных / Cultural Data Approaches
Тема ВКР/ Thesis topic Токсичный твиттер: сравнительный анализ использования слова "токсичный" в социальных сетях / Toxic Twitter: the comparative analysis of the word "toxic" usage in social networks
Руководитель ВКР/ Thesis supervisor Хлопотов Максим Валерьевич, кандидат технических наук, Университет ИТМО, факультет инфокоммуникационных технологий, доцент (квалификационная категория "ординарный доцент")

Срок сдачи студентом законченной работы до / Deadline for submission of complete thesis 01.06.2021

Техническое задание и исходные данные к работе/ Requirements and premise for the thesis
Determine the usage of the word 'toxicity' in a new social sense compared to the classical scientific one. Measure and visualization of the frequency of usage of this word in dependance of events and situations that happens around us.

Содержание выпускной квалификационной работы (перечень подлежащих разработке вопросов)/ Content of the thesis (list of key issues)
Toxicity in twitter, Most toxic social network, Critical discourse analysis theory, Analysis of usage of word 'toxic' by users, events and occasions

Перечень графического материала (с указанием обязательного материала) / List of graphic materials (with a list of required material)

Исходные материалы и пособия / Source materials and publications
1. Ashley A. Anderson, Dominique Brossard, Dietram A. Scheufele, Michael A. Xenos, Peter

Ladwig, The “Nasty Effect:” Online Incivility and Risk Perceptions of Emerging Technologies, Journal of Computer-Mediated Communication, Volume 19, Issue 3, 1 April 2014, Pages 373–387.
2. Georgakopoulos S.V., Tasoulis S.K., Vrahatis A.G., and Plagianakos V.P. 2018. Convolutional Neural Networks for Toxic Comment Classification. In Proceedings of the 10th Hellenic Conference on Artificial Intelligence (SETN ’18). Association for Computing Machinery, New York, NY, USA, Article 35, 1–6.
3. Chandrasekharan E, Samory M, Srinivasan A, Gilbert E. The Bag of Communities: Identifying Abusive Behavior Online with Preexisting Internet Data. Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. New York, NY, USA: ACM; 2017. pp. 3175–3187.
5. Irfan, R., King, C., Grages, D., Ewen, S., Khan, S., Madani, S., . . . Li, H. (2015). A survey on text mining in social networks. The Knowledge Engineering Review, 30(2), 157-170. doi:10.1017/S0269888914000277
6. Aggarwal C.C. (2011) An Introduction to Social Network Data Analytics. In: Aggarwal C. (eds) Social Network Data Analytics. Springer, Boston, MA. https://doi.org/10.1007/978-1-4419-8462-3_1

Дата выдачи задания/ Objectives issued on 14.05.2021

СОГЛАСОВАНО / AGREED:

Руководитель ВКР/ Thesis supervisor	Документ подписан Хлопотов Максим Валерьевич 14.05.2021 (эл. подпись)	Хлопотов Максим Валерьевич
Задание принял к исполнению/ Objectives assumed by	Документ подписан Селина Мария Андриановна 26.05.2021 (эл. подпись)	Селина Мария Андриановна
Руководитель ОП/ Head of educational program	Документ подписан Пучковская Антонина Алексеевна 26.05.2021 (эл. подпись)	Пучковская Антонина Алексеевна

Министерство науки и высшего образования Российской Федерации
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ ИТМО
ITMO University

АННОТАЦИЯ ВЫПУСКНОЙ КВАЛИФИКАЦИОННОЙ РАБОТЫ / SUMMARY OF A GRADUATION THESIS

Обучающийся/ Student
Селина Мария Андриановна
Наименование темы ВКР / Title of the thesis
Токсичный твиттер: сравнительный анализ использования слова "токсичный" в социальных сетях / Toxic Twitter: the comparative analysis of the word "toxic" usage in social networks
Наименование организации, где выполнена ВКР/ Name of organization
Университет ИТМО

ХАРАКТЕРИСТИКА ВЫПУСКНОЙ КВАЛИФИКАЦИОННОЙ РАБОТЫ/ DESCRIPTION OF THE GRADUATION THESIS

1. Цель исследования / Research objective
Analysis of tweets, containing the word 'toxicity' and generating of text classification model for this data
2. Задачи, решаемые в ВКР / Research tasks
Studying of toxicity of social network Twitter and the consequences that can be caused by it, analyzing the critical discourse analysis theory and modeling technics that was used in a methods part of research. Manual classification of parsed data with tweets, and analysis of it. Creating a topic modeling for each group of tweets and evaluating of a precise text classification machine learning algorithm based on manually marked data.
3. Краткая характеристика полученных результатов / Short summary of results/conclusions
The topic generation modeling works very accurate and show the same patterns as in previous data analysis methods in a case of science usage of term 'toxicity'. But in the second group with social usage, LDA modeling wasn't very precise, and evaluate only a couple of comprehensible topics. The text classification algorithm shows very precise results. Depending on a volume of input data in each group the results vary, higher number of tweets in group gives better accuracy.
4. Наличие публикаций по теме выпускной работы/ Have you produced any publications on the topic of the thesis

1 Селина М.А. ТОКСИЧНЫЙ ТВИТТЕР: СРАВНИТЕЛЬНЫЙ АНАЛИЗ ИСПОЛЬЗОВАНИЯ СЛОВА ТОКСИЧНОСТЬ В СОЦИАЛЬНЫХ СЕТЯХ // Альманах научных работ молодых ученых Университета ИТМО. -2021. - С. 5 (Статья)

5. Наличие выступлений на конференциях по теме выпускной работы/ Have you produced any conference reports on the topic of the thesis

1 Пятидесятая научная и учебно-методическая конференция Университета ИТМО, 01.02.2021 - 04.02.2021 (Конференция, статус - университетский)

6. Полученные гранты, при выполнении работы/ Grants received while working on the thesis

7. Дополнительные сведения/ Additional information

Обучающийся/Student	Документ подписан Селина Мария Андриановна 31.05.2021 (эл. подпись/ signature)		Селина Мария Андриановна (Фамилия И.О./ name and surname)
Руководитель ВКР/ Thesis supervisor	Документ подписан Хлопотов Максим Валерьевич 01.06.2021 (эл. подпись/ signature)		Хлопотов Максим Валерьевич (Фамилия И.О./ name and surname)

Table of contents

Introduction

With the global popularization of the internet, social media is becoming available to more and more people. It is no longer just a tool to send free messages and calls, but an entire ecosystem with blogs, news publications from different branches of our lives, video feeds, and online radio conversations in which people share their interests and experiences. There is no longer any need to conduct additional polls to find out the mood of people, you can simply look at the statistics for mentions or hashtags. Social media these days is a tool to identify society's problems and, in some cases, even a mechanism to control the masses [1].

My attention has been drawn to the phenomenon of toxicity, increasingly appearing in my newsfeed. Toxic in science is defined as an object that is toxic or polluting, and dangerous to others. In the social area, the meaning of the word is roughly equivalent. Users have reversed the original use of the word and begun to use it in a social sense, referring to people, actions, or situations that are poisonous to their mental health [2]. In social media, a spike in the use of the word came in 2018. At the end of this year, this noun was selected as the word of the year in the famous English language publication 'Oxford Languages'.

The idea for my research was the question of how often is the word 'toxicity' used in a new social sense compared to the classical scientific one? Another point that I measure in my research is the frequency of usage of this word in correlation to events and situations that happens around us. In my project, I analyzed one of the most popular social media platforms, Twitter, where users share their thoughts and news through posts, called tweets.

As a result of this project, the social network Twitter was analyzed for frequency and variations in the use of the word ‘toxicity’. This research identified the frequency of use of words used with the term 'toxicity' and investigated the frequency of tweets with this word by date and their correlation to different events.

Chapter 1. Toxicity in Twitter

1.1 Most toxic social network

As the world becomes more connected online, our reliance on social media, such as Twitter, also becomes very important. Over the past 13 years, Twitter has become a place, which people use both professionally and personally. From human rights activists coordinating protests on Twitter to people in marginalized communities building solidarity networks, or politicians interacting with voters outside of traditional work hours, Twitter offers its users around the world the ability to connect across industries and regions at a speed that was almost unthinkable a decade ago.

In fact, Twitter provides the ‘up-to-the-minute reactions’ means that public figures are able to bypass traditional media outlets and engage directly with their audiences. However, the open and public nature of these interactions also means that the platform is vulnerable to being used to send violent and abusive content.

The radical commitment to free speech, which is the main rule of Twitter that differs it from other social networks, although one of the reasons for the popularity of the service, has set the stage for bullying. Unlike Facebook and Instagram, which have never positioned themselves as platforms where free speech is allowed and offensive content is removed, Twitter has made protecting the most obnoxious users its ideology. It has become home especially for those who like to bully people who use the platform for work, writes Warzel [5].

The service has become as toxic as it used to be revolutionary. As it turns out, Twitter chief Dick Costolo has been secretly editing negative comments on Barack Obama's posts. So, if you're not a celebrity, much less the President of America, you don't seem to be getting any help.

- Charlie Worzel, author of BuzzFeed News

In April 2019 there was a TED conference with Twitter founder Jack Dorsey, where he was mostly asked about harassment and bullying restrictions that need apply in this social network. Moreover, he has reported that the level of violence and misinformation on his social network has increased significantly.

We have experienced violence, harassment, manipulation, coordinated attacks and the spread of misinformation. This is a dynamic we did not expect in thirteen years. - the businessman admitted [6].

He also stressed that this problem cannot be solved by machine learning and algorithms. In addition, according to Dorsey, women of African descent are most likely to be targeted by Twitter harassment.
Later, in October 2019 the social network Twitter has decided to test new experimental responses to posts to reduce toxic behavior and general negativity in comments.
David Gasca, senior director of project management at Twitter, highlighted that they have serious ambitions in their desire to truly change the way social network users communicate with each other. It is believed that experiments will also take place during 2020. He also admitted that other changes could be more fundamental. As an example, Gasca mentioned the abandonment of a simple retweet in favor of a retweet with a mandatory comment.

1.2 Women abuse on Twitter

While talking about the toxicity on Twitter, the group of people who are the most harassed by the users can’t be unmentioned. This group is women, especially

women of color that are more frequently targeted according to research “Troll Petrol Findings” of Amnesty International Organization [9].
Amnesty International released its report on 18 December 2018. Together with startup Element AI, the organization analyzed 288,000 tweets addressed to 778 women working in politics or journalism. Using machine learning algorithms, the analysts found that women received 1.1 million abusive replays on Twitter in 2017. According to this research, a woman on Twitter takes abuse every 30 seconds [8].

After the quantitative and qualitative computer analysis, the organization proved these results with live interviews and tried to solve advise some solutions to prevent such harassment.

> *Amnesty International interviewed 86 women both individually and in groups in the UK and USA. We spoke to female politicians, journalists, activists, bloggers, writers, comedians, games developers as well as women who use the platform but do not have a large following. Amnesty International also spoke with dozens of experts in the United Kingdom and United States working in the field of women's rights, identity-based discrimination, technology, and digital rights about violence and abuse against women on social media platforms. Amnesty International consulted with multiple organizations and individuals, particularly in the UK, when developing our recommendations and solutions for Twitter. The research highlights the particular experiences of violence and abuse on Twitter against women of color, women from ethnic or religious minorities, lesbian, bisexual or transgender women - as well as non-binary individuals – and women with disabilities, to expose the intersectional nature of abuse on the platform. In November 2017, the organization commissioned Ipsos MORI to conduct an online poll in 8 countries, including the UK and US, about women's experiences of abuse or harassment on social media platforms*

more generally and its impact on women's freedom of expression online as well as the psychological impact of online abuse and harassment.
-Amnesty International

Later, after this research women from all over the world boycott Twitter for a day with solidarity that necessary restrictions and punishments should be enacted to reduce the abuse and harassment against women [10]. They use the hashtag #WomenBoycottTwitter to share toxic attitudes that they experienced. These tweets arose a lot of discussions because most of it was against Twitter community standards but anyway wasn't deleted after. Here can be seen the clearest example of moderator's paradoxes. This incident shows us that the radical freedom of speech on Twitter is so rooted that even if you purposefully write a tweet that would violate all the rules of that community it won't be deleted.

At first sight, all this information about abuse can seem like just an imaginary problem in the online world. But different research shows that it not only online, but the consequences can also be very dangerous both for a person or for a whole society.

1.3 Impact of online abuse

With the development of social media as an online anonymous forum, there has been an explosion of personal opinions in public forums. Although the idea was to help people socialize more easily, and online platforms were supposed to help lead healthy discussions and exchange opinions, things have turned out differently. People say terrible things to each other, and the anonymity of the forums and support from like-minded people made it commonplace. People are ruthlessly criticized for their views, and often the criticism is personal - picking on religion, race, class, caste, gender, and even physical appearance.

But according to different research [21][22], this is not only words. Online bullying has a serious impact on the health of people. It can be even worse than abuse that we can experience offline because on the web people are hidden by their screens. It means that now you can anonymously say bad things to others avoiding any punishment. Nowadays users can register unlimited fake accounts and do whatever they want online without any judgment by society. This option makes them feel invincible and brave in the threatening of others. That’s why this online part of our life is even more dangerous for mental health than the real one.

Cyberbullying had the impact of amplifying symptoms of depression and post-traumatic stress disorder in young people who were inpatients at an adolescent psychiatric hospital, according to a new study published in the Journal of Clinical Psychiatry. The study addressed both the prevalence and factors related to cyberbullying in adolescent inpatients.

- University of Miami Miller School of Medicine

Cyberbullying considers as a sneaky form of bullying. It can bring people to a deep mental health crisis, suicidal thoughts, and even cause suicide itself. As research [21] confirms the deleterious effects of online bullying, more sensible approaches of detecting harmful posts and more strict measures against abusers.

Alarmingly, 41% of women who had experienced online abuse or harassment said that on at least one occasion, these online experiences made them feel that their physical safety was threatened.

- Amnesty International

According to this sociological survey processed by Amnesty International Company about half of the women experience some physical abuse just by online harassment cases. This may be caused by women's high sensitivity between mental health and physical one or it can show the real threat that can be caused by online toxic attitude. For now, there are no such researches for men or for different age groups that can show us the real relation between physical threat and online abuse in society overall.

Chapter 2. Research theory and methods

2.1 Critical discourse analysis theory

The theory that I used in this work was Critical Discourse Analysis (CDA). It is understood as a research program proposed by T. van Dijk, N. Fairclough, and R. Vodak and became popular in the 1990s-2000s mainly in sociology and anthropology, rather than in political science [15]. However, other approaches should not be reduced to it, and therefore it is reasonable to rely on the more general category of "critical discourse studies". Critical discourse studies as a research area emerged at the turn of the millennium during the development of a narrower concept of critical discourse analysis. However, it goes a little further and includes within the first concept all studies that deal with critical discourse interpretation.

The concept of discourse historically belongs to the field of linguistics, when in the 1950s and 1960s researchers tried to go beyond the analysis of a single phrase, which constituted the limit of application of linguistic methods. The very "discourse" was understood as speech conditioned by the social context, a linguistic unit larger than a sentence [17] [18]. The emergence of discourse studies coincided with the "linguistic turn" in the social sciences, which raised the problem of the non-neutrality of language in the creation and representation of the social.

In the 1980s, the founding school of what is now known as critical discourse analysis was established. Its characteristic feature, especially at the first stage in the 1980s-1990s [16], was the use of linguistic tools at the micro-level to expose specific symbolic spaces, namely media materials or politicians' speeches (also usually mediated by media). The theory developed ended up being dependent on the type of source being analyzed.

Later, in the first book that describes only Critical Discourse Analysis as a complete theory [27] the author, that now considered as a founder of this theory, wrote this description of it:

“The kind of discourse analysis which aims to systematically explore often opaque relationships of causality and determination between (a) discursive practices, events and texts, and (b) wider social and cultural structures, relations and processes; to investigate how such practices, events and texts arise out of and are ideologically shaped by relations of power and struggles over power; and to explore how the opacity of these relationships between discourse and society is itself a factor securing power and hegemony”

-Norman Fairclough

The purpose of critical discourse analysis is to highlight the linguistic discursive dimension of social and cultural phenomena and the processes of change occurring in these areas [19]. Critical discourse analysis uses detailed textual analysis to examine from a linguistic perspective how discursive processes take place in individual texts. However, it is not enough to analyze the text on its own, it is necessary to clarify the relationship between the text and the structures (cultural, social, etc.) in which the discourse emerges.

In the case of this research the discourse, that I was interested in, was the tweets posted by users on the social network Twitter and relationships were built on news, events, and situations as a social structure.

2.2 Latent Dirichlet allocation modelling

One of the methods that were used in this research is a topic modeling LDA. After data processing through the python programming code that allows using such modeling, the main topics of all tweets are generated. This helps to divide tweets into groups and find out the frequency distribution of them in a timeline. Another data analysis that can be shown by this processing is the number of tweets for each topic, which shows what its users are more interested in.

Latent Dirichlet Allocation (LDA) is a generative model used in machine learning and information mining that allows us to explain the results of observations using implicit groups so that we can determine the causes of similarity between some parts of the dataset. For example, if the observations are words collected in documents, it is argued that each document is a mixture of a small number of topics and that the occurrence of each word is related to one of the document topics.

This method of topic modeling was proposed by David Blei, Andrew Ng, and Michael Jordan in 2003 [26]. LDA belongs to the family of generating probabilistic models in which topics are represented by the probabilities of occurrence of each word from a given set. Documents, in turn, can be represented as combinations of topics. A unique feature of LDA models is that topics need not be distinct, and words can occur in more than one topic; this gives a certain vagueness to given topics, which can be useful for solving the problem of language flexibility.

Blei and associates (2003) found that the Dirichlet distribution, a family of continuous distributions (a way of measuring grouping by distributions), is a convenient way to identify topics present in a corpus as well as appearing in various combinations in each document in the corpus [25]. In fact, the Latent Dirichlet Allocation (LDA) method gives us an observable word or lexeme from which we can attempt to determine a topic most accurately. Moreover, with it, the distribution of words in each topic, and the combination of topics in the document can be defined. To use topic modeling techniques in the application, a customizable pipeline that extrapolates topics from unstructured text data, and a way to preserve the best model should be created.

The theme modeling pipeline in this research includes these steps:

1. Loading the corpus.
2. Text preprocessing.

2.1 Removal of stop words

2.2 Punctuation removal

2.3 Lemmatization of words

3. Dictionary creation
4. Selecting the optimal number of topics
5. Visualizations

Corpus in this case is the tweets from the same group – in social or STEM usage of word 'toxicity'. On a step text preprocessing we firstly remove stop words with a standard pack of words, that we can download from this library [24].

```
import nltk
nltk.download('stopwords')
```

In punctuation removal, after the typical part, it was important to delete all links because on a first try without it, topics contain a lot of links which was a difficult to topic description analysis. This code I use for links removal:

```
import re
def clean_message(msg: str):
    msg = re.sub(r'http\S+','',msg)
    return msg.replace("\n", " ").lower()
df['text'] = df['text'].apply(clean_message)
```

In a part of topic lemmatization, the task is usually to convert words into their root word. But in the everyday life conversations, people often make some spelling and grammatical mistakes, this fact may lead to different types of uncertainties, such as lexical, syntactic, and semantic. Because of that analyzing information patterns from such data sets become more complex [13]. Lemmatization can make these uncertainties clearer, but on the stage of filtering data with mistakes artificial intelligence can’t make it very accurate, as a real human. Lemmatization can help, for example, to convert plural to single: ‘mice’ to ‘mouse’, convert the continuous form of the verb to present: ‘playing’ to ‘play’ and etc.

In the dictionary creation part, Gensim modeling creates an identification number for each word in all data. In this corpus the groups of numbers are generated, which means which word how many times occurs in each document. This list of numbers used as input in the LDA model processing.

```
# Creates DICTIONARY which is a mapping of word IDs to words.
words = corpora.Dictionary(doc_list)

# Turns each document into a bag of words.
corpus = [words.doc2bow(doc) for doc in doc_list]
```

In the next step, we just need to choose the number of topics that we want to generate. For this we need to find out the coherence score of each number of topics, for this, we build many LDA models with different values of the number of topics. After that, we need to analyze the data and pick the number of topics that give the highest coherence value. If we are choosing the number of topics at the end of a rapid growth of coherence value it usually gives meaningful and interpretable topics but picking a higher coherence score value topic can sometimes provide more granular sub-topics. In Figure 28 and Figure 29 coherence score is builds on a graph for each number of topics in each class.

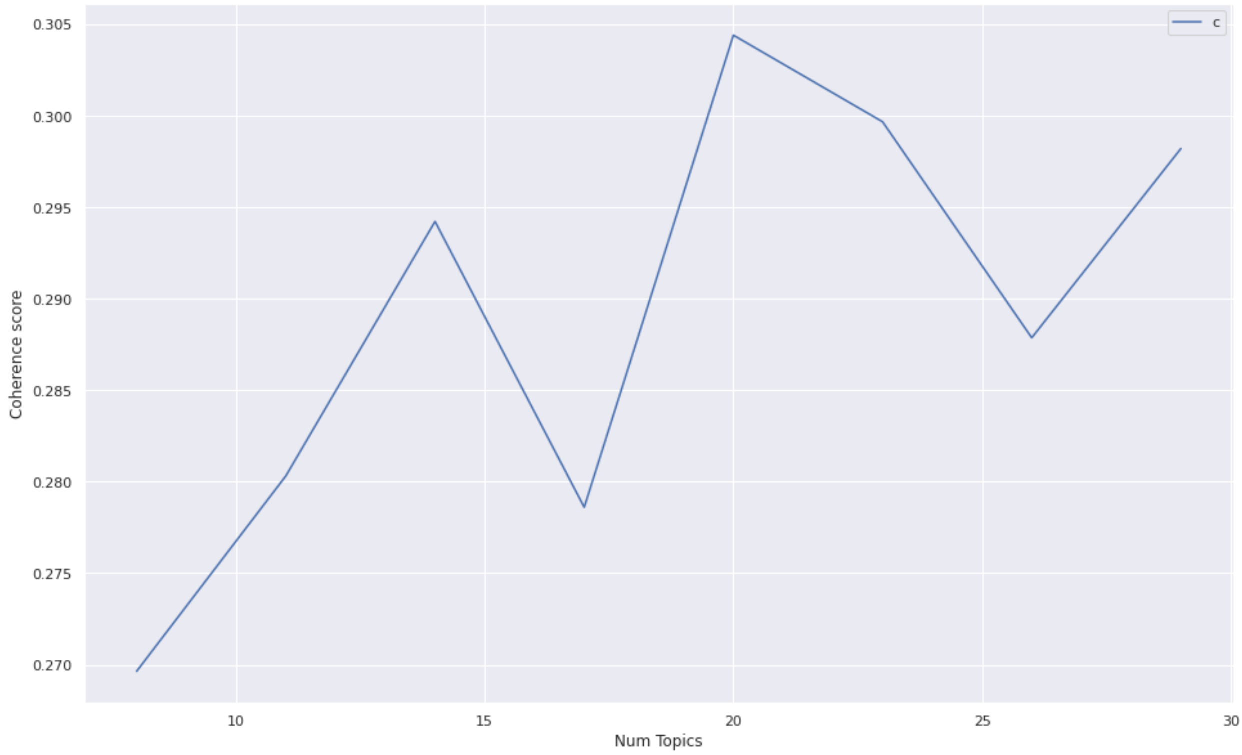

Figure 28. Coherence score for each number of topics in STEM toxicity class

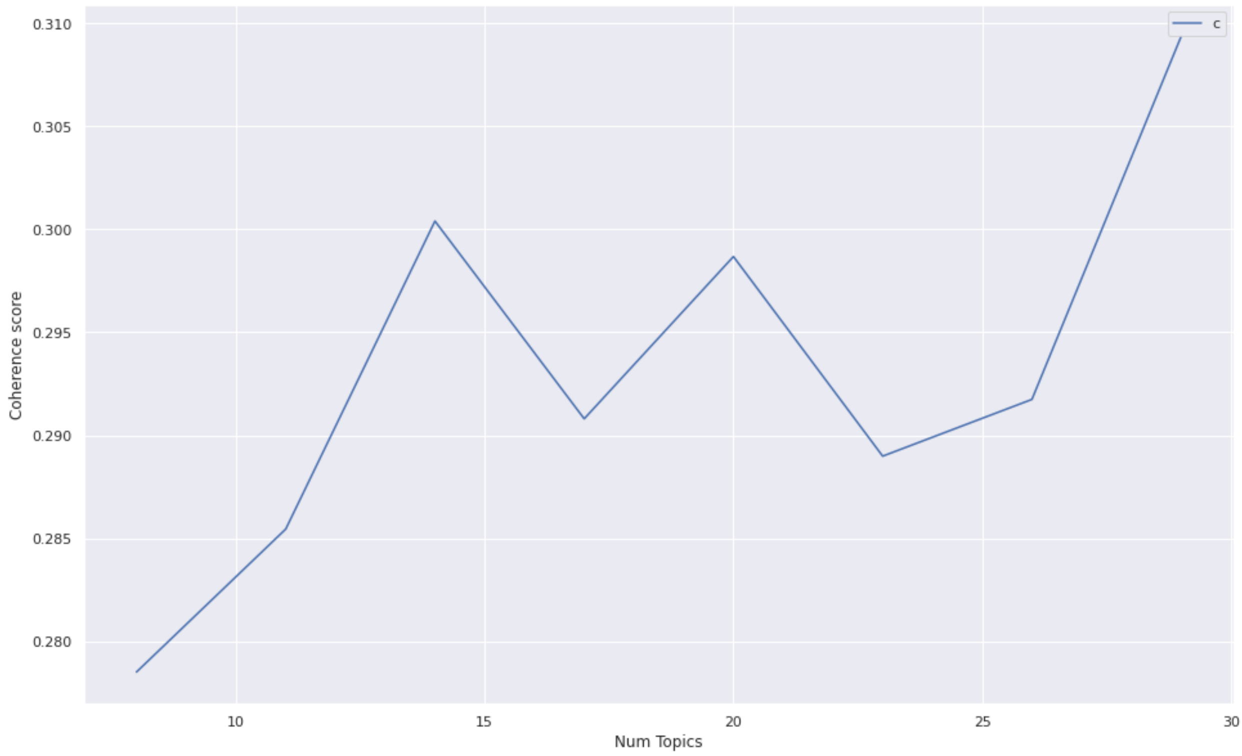

Figure 29. Coherence score for each number of topics in social toxicity class

In my case, I decided that 20 will be enough for a clear showing of the STEM theme, but at the same time not a lot to divide one real topic into two generated by a mistake.

And for the social toxicity topic, I chose to generate 14 topics, because this number has the highest affordable coherence value in Figure 29. After all this processing we can see the topic in such format:

```
(0,
  '0.097*"toxicity" + 0.092*"military" + 0.089*"basis" + 0.088*"sick" + '
  '0.087*"trail" + 0.011*"chemical" + 0.009*"guardian" + 0.008*"water" + '
  '0.008*"fragrance" + 0.007*"news"')
```

Firstly, the number of topics can be seen, then the weight of the 10 most popular words in the topic is shown. The weights reflect how important a keyword is to that topic. In this case, in the 0th topic, the word toxicity was the most frequent one, with a weight of 0.097. With these words, we can recognize the topic in most cases, if it generates right.

In the visualization part, we can use different types of bar charts and plots to show what is important. In this research, I visualize the timeline of topic distribution of each theme in an interactive chart that is shown in Figure 25. Another visualization that was done is the popularity of each topic, the number of tweets in each topic on a bar plot in Figure 26.

2.3 Scikit-Learn Multi-Class Text Classification

As it was mentioned before, the abuse is very common for nowadays, especially in the online world. Mostly, social networks clean their environment from harassment messages and posts by complaints of users. In that case, dangerous message can be only deleted after some reports, so firstly users need to read it. But the problem is that this abusive message will be seen by some people. Unequivocally, without reporting mechanism this harassment can be seen by much bigger number of people, but are we sure that abuse didn't reach its goal? The solution is to block dangerous messages and post before it can be seen by public, but for this the special machine learning mechanism is needed. In 2017 group of scientists tried to create such filter for unwanted publications called Bag of Communities (BoC) [11].

> *'On most sites today, moderation takes two primary forms: distributed social moderation [20, 30, 49, 50] and machine learning-based approaches [6, 10]. In the former, a site's users triage submissions via voting or reporting mechanisms—after which the site can take action. In the latter, online communities compile large datasets of example posts that have been moderated off-site, and thereafter train machine learning algorithms. The distributed social moderation approach is appealing because it can be deployed quickly and easily, and offloads the work of moderation to a large human workforce; yet, it requires vast amounts of human labor from the very people you would rather not see abusive posts (i.e., your users). Machine learning-based approaches can help by algorithmically triaging comments for a much smaller number of (perhaps paid) human moderators; yet they typically require vast amounts of labeled training data.' [11]*

In my research the text classification, based on the similar mechanism was done. My model was studied to classify the tweets by group of STEM or social meaning of the word 'toxicity'. This mechanism can be used further in the similar research, to save time of manual classifying of data. For algorithm learning I used the Scikit-Learn library.

The Scikit-Learn library is the most common choice for classical machine learning problems. It provides a wide choice of learning algorithms with and without a 'teacher'. Learning with a 'teacher' implies a marked dataset in which the value of the target attribute is known. While learning without a 'teacher' does not imply the presence of markup in the dataset - you need to learn how to extract useful information from arbitrary data. One major advantage of the library is that it builds on several common math libraries and integrates them easily with each other.

In the context of machine learning, classification refers to learning with a ‘teacher’. This type of learning implies that the data fed into the inputs of the system are already labeled, and that the important part of the features has already been divided into separate categories or classes. Therefore, the network already knows which part of the inputs are important, and which part can be independently checked. In case of this research the input was the marked data, and the checked output with this algorithm was the same data. This method was chosen to measure the productivity of model and to compare it with the manual classification.

The process of creating multi-class text classification algorithm on a Scikit-Learn model contains the following steps:

1. Data preprocessing
2. Creating training sets
3. Creating a classifier
4. Training the classifier
5. Making predictions
6. Evaluating classifier performance
7. Adjusting parameters

First, I prepared the data set for the classifier: convert the data into a form correct for classification and handle any anomalies in the data. Missing values in the data or any other deviations may negatively affect the performance of the classifier. The next step is to divide the data into training and test sets. There is an excellent train_test_split function for this purpose in Scikit-Learn.

One of the most popular approaches how to extract some patterns from text is to use a model, called bag of words where for each document, a tweet text, in this case, the presence and frequency of words is taken into consideration, but the order in which they occur is ignored. To use this, I calculate a measure called ‘Term Frequency - Inverse Document Frequency (tf-idf). After it, each of the 10600 tweets that there is in the database is represented by 5996 features, representing the tf-idf score for different unigrams and bigrams. To train supervised classifiers, the text of tweets should be transformed into a vector of numbers. For this, vector representations tf-idf should be taken to train classifiers to check the unseen text of the tweet and predict for which group, STEM or social, it will rely on.

After import of all required libraries, I have chosen the Naive Bayes Classifier (NBA), because it outperforms many other classification algorithms. But Scikit-Learn library contains not only the NBA classification model. Overall, there are about 12 main Classificatory models. In my research, I compare the 4 most suitable: Logistic Regression, Naive Bayes (Multinomial), Linear Support Vector Machine, and Random Forest. After comparing, the NBA shows the most accurate result for my data model, and I decide to use it.

Naive Bayes Classifier is based on Bayes' theorem with the assumption of feature independence. NBA assumes that the presence of a character in a class is unrelated to the presence of any other feature. For example, fruit can be considered an apple if it is red, round, and about 8 centimeters in diameter. Even if these characteristics depend on each other or on others, in any case, they contribute independently to the probability that the fruit is an apple. Because of this assumption, the algorithm is called ‘naive’.

After fitting and training this algorithm, I tried to make some predictions just by asking the model to classify one random tweet. Then I compare the results with my marking, and it was the same. Continue with the NBA model, I made the confusion matrix and check the discrepancies between actual and predicted labels. The results were sufficient and then I made the classification report for each class, which gives about 0.96 accuracies f1-score for the whole model. The full code, that was written for this research dataset can be found on a Google Colab [43].

Chapter 3. Data analysis

3.1 Data mining

Data Mining is a method of data analysis designed to find previously unknown patterns in large amounts of information. These patterns make it possible to make effective management decisions and optimize business processes [14].

> *'Social media shatters the boundaries between the real world and the virtual world. We can now integrate social theories with computational methods to study how individuals (also known as social atoms) interact and how communities (i.e., social molecules) form. The uniqueness of social media data calls for novel data mining techniques that can effectively handle user - generated content with rich social relations. The study and development of these new techniques are under the purview of social media mining, an emerging discipline under the umbrella of data mining.* ***Social Media Mining*** *is the process of representing, analyzing, and extracting actionable patterns from social media data.' [12]*

The first step in any Digital Humanities research should be the choice of data for analysis. In this research data is presented as tweets corpus, parsed from Twitter. For the analysis subject, it was decided to take the word 'toxicity'. The word 'toxic' was chosen in social discourse theory, as a cultural phenomenon for analysis. For example, in the social phrase "toxic relationship" toxic means unhealthy or abusive, "toxic person" means a person, who spread negativity as it is considered by the author.

Overall, there were 10600 tweets, containing the word 'toxicity'. Data set contain tweets from 13 days, mostly at the end of each month. Parsing was operated by the special program ' Optimized-modified-GetOldTweets3-OMGOT' [20] that was written by user ‘marquisvictor’ on the GitHub platform. To receive tweets that I was interested in the searching request was:

```
python GetOldTweets3.py --querysearch «#toxicity» --since 2019-01-31 --
                              until 2019-12-30--lang es --maxtweets 100000
```

That request collects all tweets with the ‘toxicity’ hashtag from 31 January to 30 December 2019. After parsing, the data was sorted that way, to receive tweets from only 13 certain days, for fewer data with more illustrative content.

The dataset received consisted of 10,600 tweets, all of which used the word 'toxicity' in some sense or another. The data was then marked manually, using a written by my colleague Python program [23], into three types of datasets:

- Use of the noun 'toxicity' as a social phenomenon
- Toxicity as a scientific term – STEM (science, technology, engineering, and math)
- Failed to identify a link to one type or the other (Unknown)

3.2 Primary manual analysis

After creating the necessary data sets, I moved on to analysis using the R programming language. Statistical processing of the data and graphics was done in the free and open-source software development environment RStudio [3]. Figure 1 shows the distribution of tweets into groups in the sorted arrays.

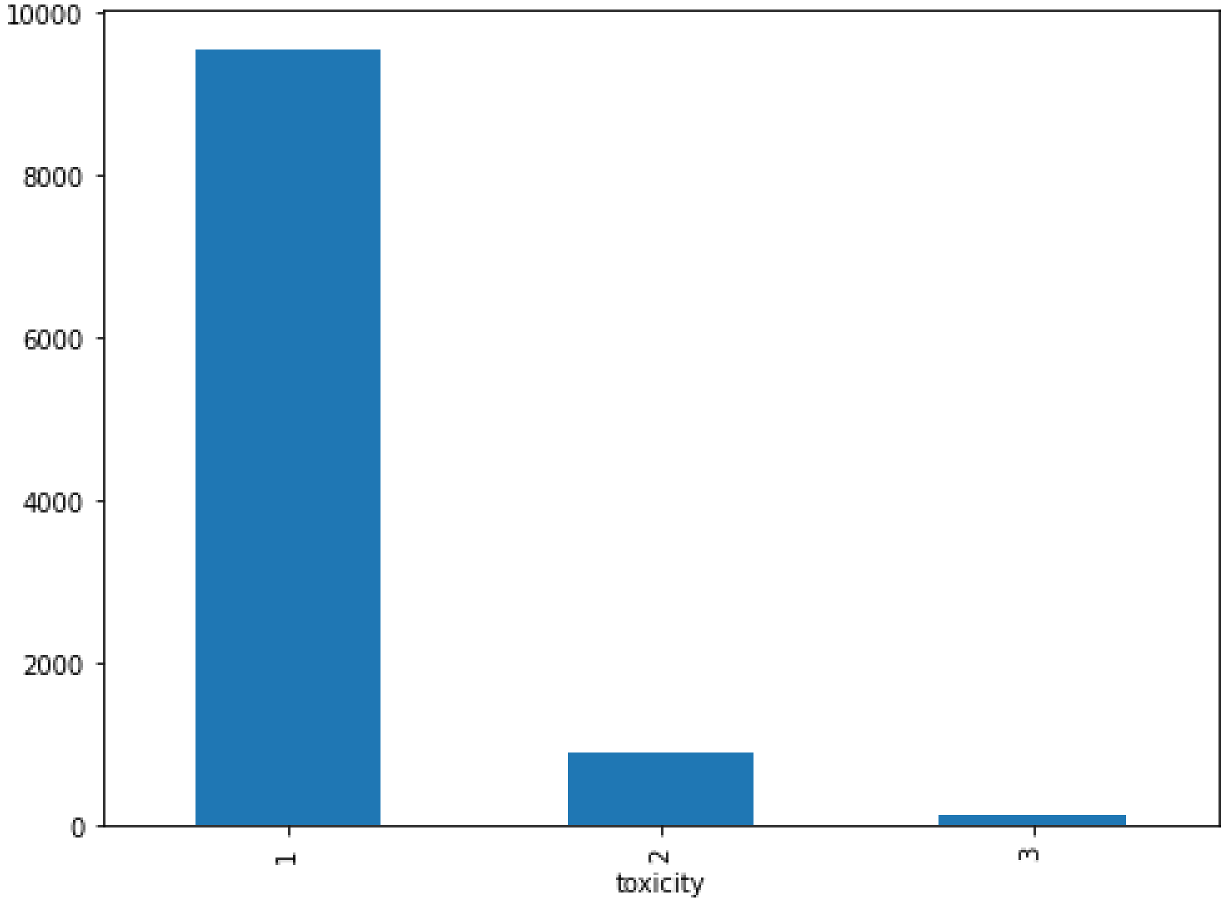

Figure 1. Compartment of tweets number in sorted groups

Then a chronological analysis of the data was performed, using RStudio to calculate on what dates tweets with this or that variant of the word 'toxicity' was written most often. The resulting infographics are shown in Figure 2, using the noun 'toxicity' as a social phenomenon, and Figure 3, using the word 'toxicity' as a scientific term. In both cases, the horizontal axis denotes the number of tweets on a given day marked on the ordinate axis.

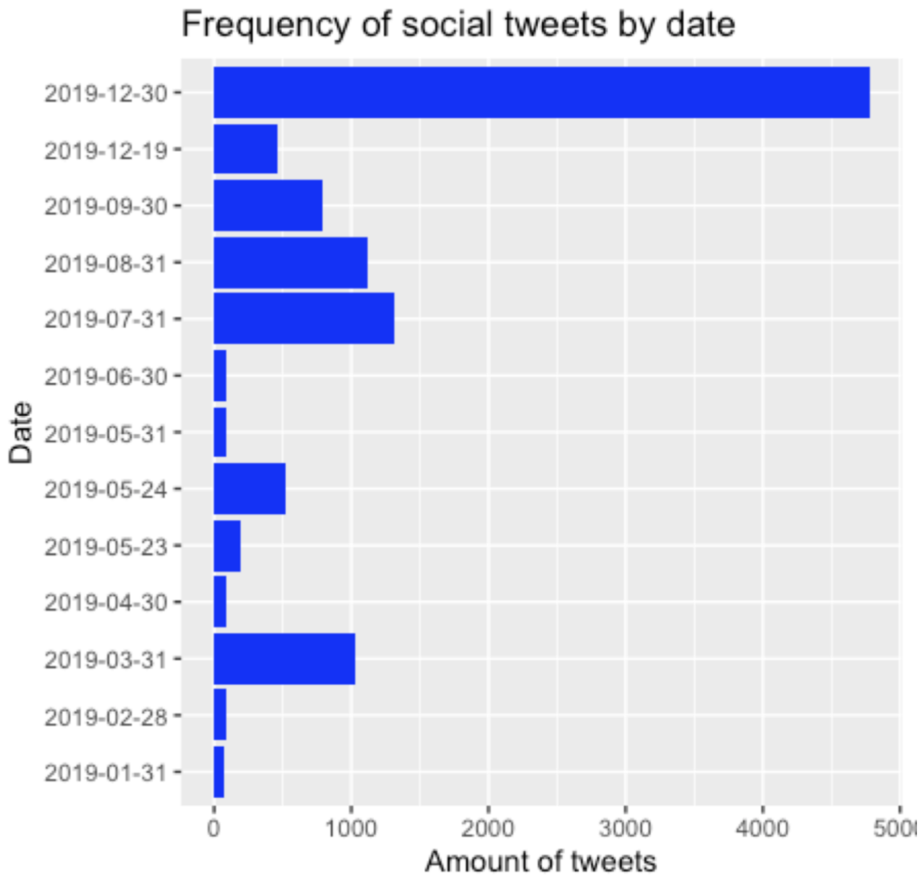

Figure 2. Frequency of social tweet posting on timeline

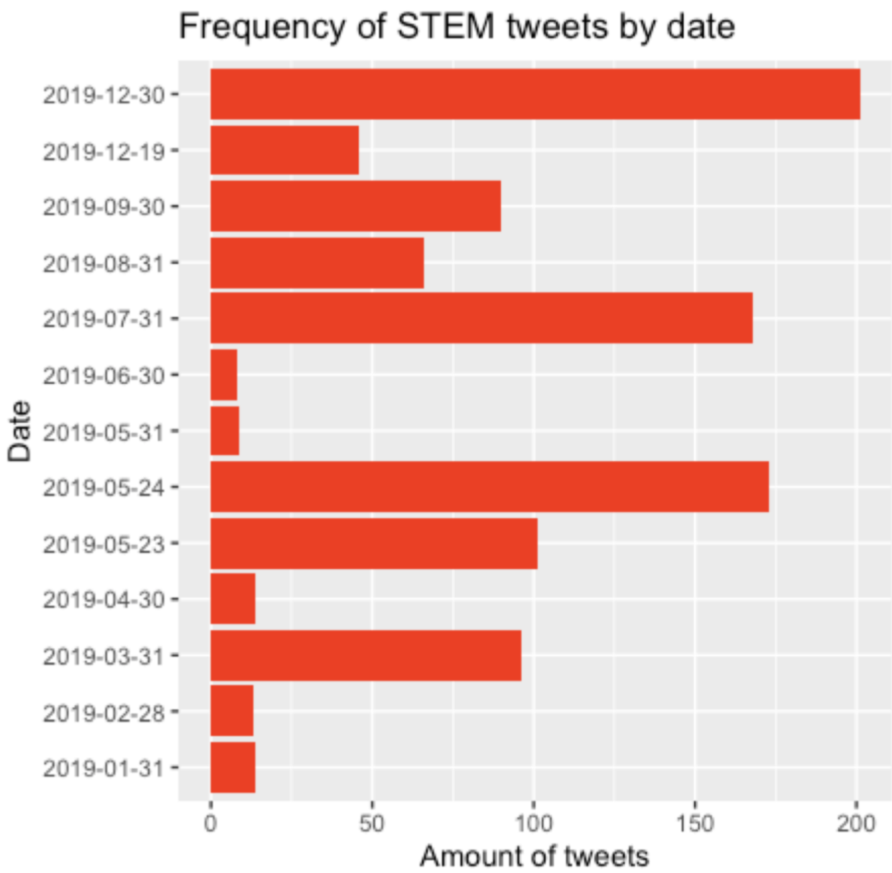

Figure 3. Frequency of STEM tweet posting on timeline

It is clear from Figure 2 that on December 30 the number of tweets using the word 'toxicity' to describe social phenomena increased significantly compared to the other dates examined. After examining the content of the tweets on that day, it turned out that most of them contained wishes for the upcoming year. That is, the surge in the use of the word falls at the very end of the year due to users taking stock of the outgoing 'toxic' year and intending to end their 'toxic' habits in the new year.

We can observe from Figure 3 that in this case, the distribution of tweets by date is about the same. Examining the content of tweets from the dates with the largest spike in the chart, I found the date of 24 May the most interesting. On that day, most of the tweets are related to the news, which narrates about toxic military bases that cause illness in people.

Further, my research provides a word cloud that shows the most frequently used words in tweets with the social use of the word 'toxicity' Figure 4 or with the scientific use of the term Figure 5. The word 'toxicity' and all of its cognates, as well as all prepositions and particles, have been removed from the list of the most commonly used words. The font size indicates the frequency of use of the word, the larger the font, the more frequently it is used.

Figure 4. Word cloud of most frequently used words in social group of tweets

Figure 5. Word cloud of most frequently used words in STEM group of tweets

From Figure 4 and Figure 5, there is a natural correlation between the most frequently used words and the meaning of the use of the word 'toxicity'. In the case of the social version of the term, the most frequently used words are 'people', 'relationships', 'life', 'family', i.e., all the social institutions that surround and concern society. In Figure 5 we can observe the most used words such as 'chemicals', 'water', 'waste', 'disease', 'air', 'health', which is rather obvious given the specific scientific field of the word 'toxicity'. But in the word cloud, we can also notice the words 'bases' and 'people' in large print, which do not quite correspond to the scientific domain. This discrepancy is due to the fact that these words are very often used in the news from 24 May, which caused a stir in Figure 3, in the context of military bases causing illness in people by their toxicity.

Chapter 4. Model analysis

4.1 Topic modelling with Latent Dirichlet Allocation (LDA) in case of STEM usage of word toxicity

Firstly, I make a topic modeling of tweets connected with STEM usage of toxicity words. LDA modeling calculates 20 different topics of all Scopus of data in this group. These topics are rather recognizable, and they can be easily connected with tweets. In this part, I will describe it and make some examples of tweets that relate to them.

Topic description

```
(0,
'0.097*"toxicity" + 0.092*"military" + 0.089*"basis" + 0.088*"sick" + '
'0.087*"trail" + 0.011*"chemical" + 0.009*"guardian" + 0.008*"water" + '
'0.008*"fragrance" + 0.007*"news"')
```

The First generated topic describes the resonance event that was reported in many news channels. This incident is about research that shows how military bases in the US chemically pollutes the environment [28]. The news article was posted on 24.05.2019, especially in this period the main burst of activity has happened in Twitter society as it was generated by this topic on the timeline Figure 6.

Figure 6. Time series of tweets posted in 0^{th} topic

```
(1,
  '0.034*"free" + 0.030*"bag" + 0.017*"chemical" + 0.017*"product" + '
```

```
'0.016*"plastic" + 0.016*"harmful" + 0.015*"use" + 0.015*"friendly" + '
'0.015*"app" + 0.015*"open"')
```

This topic, as it can be seen from main words, connected with raising people's awareness about how bad and harmful plastic bags are for our environment. Another tweet that influences this topic is the advertisement of a dry-cleaning company that announces its opening on 1 of April 2019 on this date, the maximum number of tweets on this topic is shown on timeline Figure 7.

"We are finally open. We use no harmful or toxic chemicals, no plastic bags and all of our products are eco-friendly such as our free garment bag. Download our app for free pick-up and delivery by texting the word "GreenO" to the number 555-888."

-@greenorganicdry

Figure 7. Time series of tweets posted in first topic

```
(2,
'0.020*"toxicity" + 0.011*"jar" + 0.011*"pot" + 0.008*"chemical" + '
'0.006*"use" + 0.006*"product" + 0.006*"study" + 0.006*"new" +
0.006*"honey" '
'+ 0.006*"glass"')
```

In the third case, generated words of this topic show us some description of eco-friendly candles [29]. Mostly, this group is formed because of many reposts to promote this product. This tweet was posted on 31 of June and the first greatest number of tweets on timeline Figure 8 in this group dated on the same day. Later, in January the second-largest point is shown, most probably that is the result of grows of this brand.

"súki's Exclusive New Honey Pot soy candles come in a mouth-blown glass jar - These glass jars are completely toxic free and dishwasher safe. They are designed to recycle - re-use your jar as a honey pot, coffee bean pot, vase, money jar, sugar pot"

- súki candles

Figure 8. Time series of tweets posted in second topic

```
(3,
  '0.034*"toxicity" + 0.014*"product" + 0.008*"include" + 0.006*"death" + '
  '0.006*"analysis" + 0.005*"body" + 0.005*"find" + 0.005*"link" + '
  '0.005*"algae" + 0.005*"energy"')
```

This topic tells us about the dangers for the health of blue-green algae. In this group, I found a couple of related types of tweets, one of them was the news about dog deaths because of this toxic plant [30], which was posted on 29.08.19, on timeline Figure 8 the increased number of tweets related to this article happened at the beginning of September. Another one was telling just about algae toxicity especially in water in some regions [31]. The second article was published on 30 of April, related to this, the maximum number of tweets that were posted in this topic group Figure 9 accounts for the beginning of June.

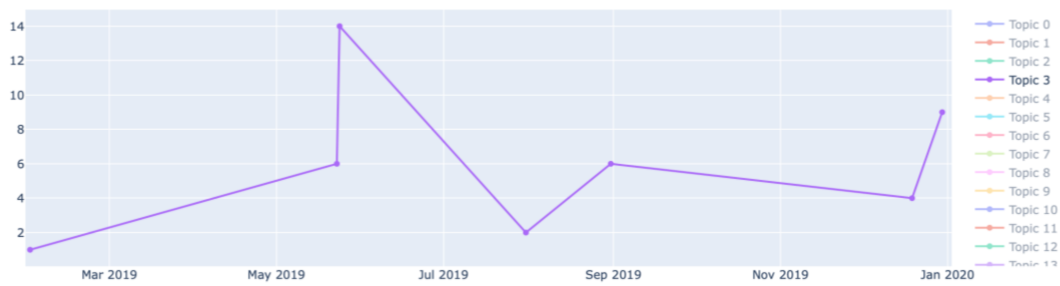

Figure 9. Time series of tweets posted in third topic

```
(4,
  '0.010*"chemical" + 0.008*"waste" + 0.008*"water" + 0.008*"anti" + '
  '0.008*"poc" + 0.008*"community" + 0.006*"know" + 0.006*"cut" + '
  '0.006*"think" + 0.006*"self"')
```

The fourth topic that was generated shows the post of the organization of young people that voting against climate change. This tweet had a lot of reposts and that's why it was generated in a single topic. The tweet was posted on 01.08.2019 and on time series for this topic Figure 10 the first spike of activity is especially on this day, the largest number of posts can be seen in January 2020, which relates to the increased number of reposts.

"Yesterday we marched with #FrontlineDetroit communities demanding a #GreenNewDeal

48217 is Michigan's most polluted zip code & has one of the highest % of POC

Due to historic redlining, POC are far more likely to live next to toxic waste & industrial facilities. We must change."

- @sunrisemvmt

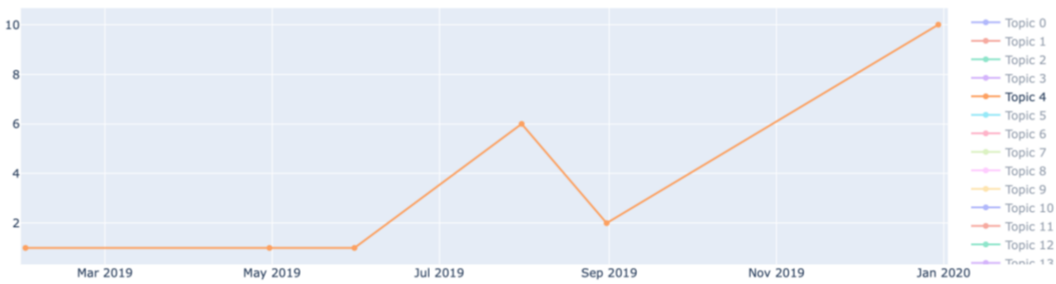

Figure 10. Time series of tweets posted in 4th topic

```
(5,
  '0.038*"fish" + 0.031*"stop" + 0.030*"eat" + 0.024*"food" +
0.022*"farmed" + '
  '0.019*"watch" + 0.018*"documentary" + 0.016*"foodie" + 0.016*"market" +
'
  '0.016*"friend"')
```

The next topic is about some documentary movie about farmed fish, which describes it as the most toxic one. This research was published on 05.07.2019, which relates to the second spike in timeline Figure 11 of usage this topic. The largest number of activities in this topic relates to the preview or promotion of this documentary in April 2019.

"Just watched a documentary on farmed fish.... Please twitter friends Stop eating it!!! It is the MOST toxic food in the market! #fish #toxic #foodies #HealthyLiving"

- reddawn1978

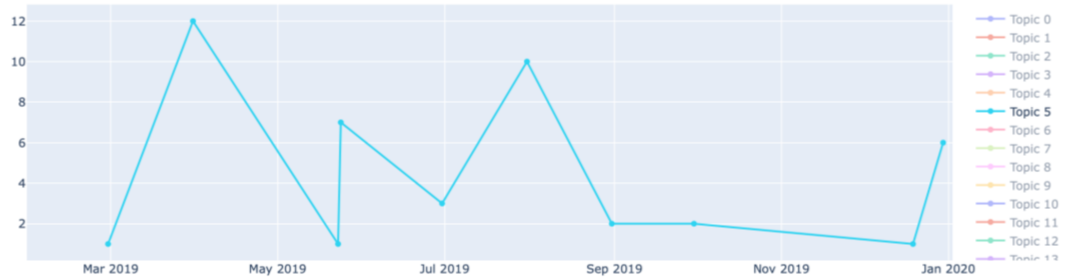

Figure 11. Time series of tweets posted in 5th topic

```
(6,
'0.012*"love" + 0.010*"ban" + 0.008*"food" + 0.008*"good" + '
'0.008*"pesticide" + 0.008*"bee" + 0.008*"thiacloprid" + 0.008*"tell" + '
'0.008*"@eu_commission" + 0.007*"little"')
```

The sixth topic that was generated tells about dangerous neonicotinoid pesticide that is toxic for people and bees. Most frequently these tweets related to a website, where you can vote to ban thiacloprid in the EU [32]. It is hard to find out when this voting was announced, but on 22 October was succeeded and this pesticide is now prohibited to use in European Union. Most probably the starting of collecting signatures for this voting relates to the largest number of tweets in the timeline Figure 12, in August 2019.

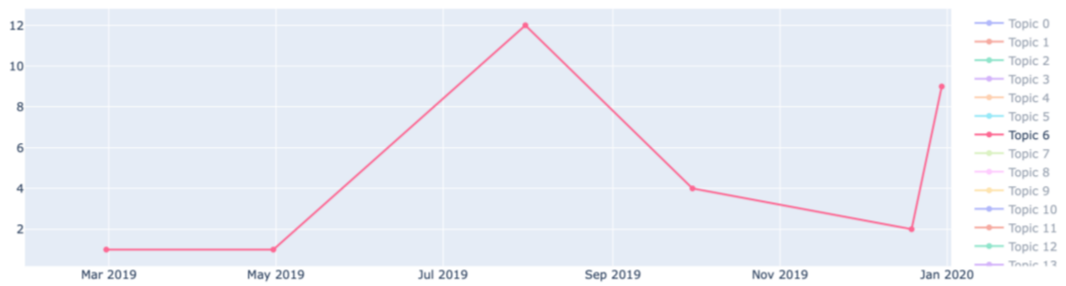

Figure 12. Time series of tweets posted in 6th topic

```
(7,
```

```
  '0.019*"toxicity" + 0.013*"safe" + 0.011*"waste" + 0.009*"silicone" + '
  '0.009*"cause" + 0.008*"health" + 0.006*"time" + 0.006*"•silicone" + '
  '0.005*"day" + 0.005*"mom"')
```

Here topic shows research about silicon toxicity [33]. This research was published on 27.08.2019, but in the timeline that shows the activity of tweets Figure 13 postings the highest spike is shown in June 2019. It can relate to the preview of this research, another group of users that is concerned about silicon usage in homeware or the generated topic was extended because of other similar for LDA modeling tweets.

Figure 13. Time series of tweets posted in 7th topic

```
 (8,
  '0.022*"chemical" + 0.015*"know" + 0.013*"lose" + 0.013*"mold" + '
  '0.013*"watch" + 0.012*"voice" + 0.012*"wanna" +
0.012*"@podspeechbubble" + '
  '0.012*"@jacediehl" + 0.012*"@realhartman"')
```

This tweet collection describes the advertisement of a podcast about a story about how James Arnold Taylor lost his voice because of mold. This podcast was launched on the YouTube platform on 31.03.2019, the first activity spike in this topic timeline Figure 14 relates exactly to this period. The second increase of usage of the topic of this tweet accounts for January 2020, it can relate to raising people's interest in this topic.

"Wanna know more about when I lost my voice from toxic mold? Watch @realhartman @JaceDiehl @podSpeechBubble"

- @JATactor

Figure 14. Time series of tweets posted in 8th topic

```
 (9,
  '0.041*"snake" + 0.030*"dead" + 0.017*"dose" + 0.016*"health" +
0.015*"low" '
  '+ 0.015*"trump" + 0.014*"kill" + 0.014*"tree" + 0.014*"guam" + '
  '0.014*"tylenol"')
```

In this topic, we can see reposts of scientific research about very toxic Guam snakes. Scientific workers want to kill these snakes by dropping nearby a dead mouse, that is poisoned by Tylenol, which causes the death of these tree snakes [34]. The research was published on 01.04.2019 and the same dating has a timeline that is related to this generated topic tweets distribution, as can be seen on timeline Figure 15.

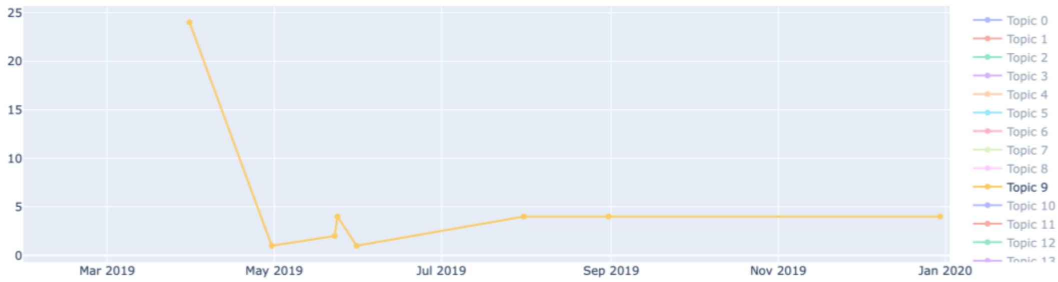

Figure 15. Time series of tweets posted in 9th topic

```
 (10,
  '0.041*"hair" + 0.024*"chemical" + 0.016*"handful" + 0.014*"try" + '
  '0.014*"strategy" + 0.014*"prevent" + 0.014*"loss" + '
  '0.013*"pic.twitter.com/xsfcwuycoz" + 0.013*"gentle" + 0.013*"blast"'),
```

The tenth topic that was generated, aware users about gentle products that can improve the quality of your hair [35]. This list of products that help your hair looks better was posted on 01.04. On the timeline of usage of this topic in Figure 16, the first spike is dedicated especially on this date.

"Before blasting your hair with toxic chemicals, why not try a handful of gentle strategies for preventing hair loss and strengthening the hair you have?"

- @Hair_So_Fine

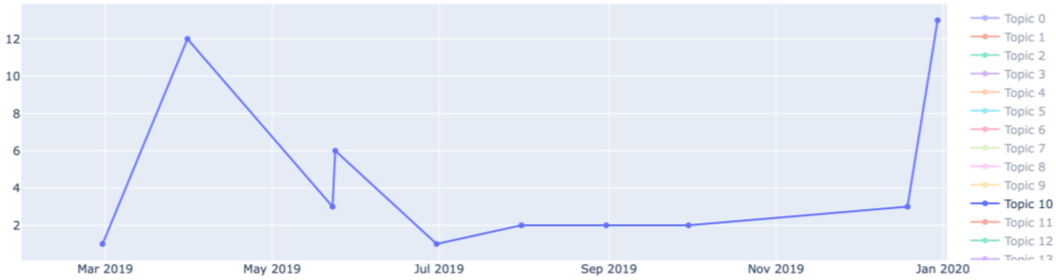

Figure 16. Time series of tweets posted in 10th topic

```
(11,
 '0.036*"toxicity" + 0.008*"cause" + 0.005*"industry" + 0.005*"fda" + '
 '0.005*"year" + 0.005*"dangerous" + 0.005*"arsenic" + 0.005*"find" + '
 '0.005*"depend" + 0.005*"abstract"')
```

This group of tweets tells us about the toxicity of arsenic in different branches of our lives. There is plenty of typical tweets that was reposted only a couple of time, that's why together they made one topic. One of these common tweets is a repost of an article about how heavy metals go away from your body through sweat [36]. This research was posted on February 22, 2012, so in this case, it can't be precisely determined which spike in timeline Figure 16 relates to this article because the splash of interest to this topic can be at any time, most probably it was in the end of May 2019, where the largest number of tweets in this topic is shown. Another popular tweet in this category is precisely about why arsenic is a danger was posted at the beginning of October 2019, in this case too, the peak of activity can't be exactly shown because it was mixed with other tweets related to this topic and timeline on Figure 17 can't show it exactly:

Rice and especially brown rice can be high in arsenic, a toxic heavy metal found naturally in the environment. Long-term consumption of arsenic can increase risk of diseases such as type 2 diabetes and cancer

- *@AHealthBlog*

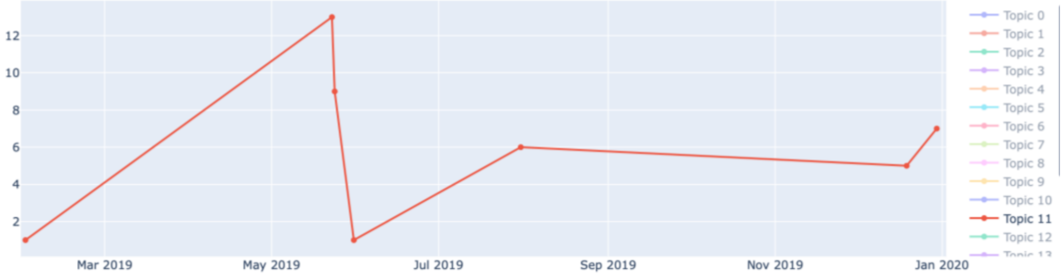

Figure 17. Time series of tweets posted in 11th topic

```
(12,
 '0.023*"toxicity" + 0.014*"mouse" + 0.014*"use" + 0.014*"new" + '
 '0.012*"water" + 0.012*"study" + 0.012*"liver" + 0.011*"cbd" + '
 '0.011*"chemical" + 0.011*"level"')
```

The twelfth topic group describes the popular research article about how CBD affects a mouse's liver [37]. The research was published on 24.05.2019 and on the timeline Figure 18, which shows how popular was this topic through time, the highest peak of activity is shown exactly at the end of May 2019.

Figure 18. Time series of tweets posted in 12th topic

```
(13,
 '0.051*"test" + 0.026*"free" + 0.026*"expose" + 0.024*"prep" + '
 '0.024*"succeed" + 0.024*"taker" + 0.024*"advice" + 0.024*"method" + '
 '0.019*"waste" + 0.016*"material"')
```

In these generated words the topic about preparation for California Basic Educational Skills Test (CBEST), TExES Exam, FTCE Test, NCLEX PN Test, and CSET Test is represented. In these tweets, users call the old system of preparation toxic and offer different articles and videos with advice on how to succeed on these exams. This test is held approximately every 3 months, in September, December, February, April, and July, this approximately corresponds to the peaks, that can be seen on a timeline of this topic in Figure 19.

Mostly tweets look like this and mostly the same, but with links to different test preparation:

Toxic CSET Test Prep Methods Exposed. Free Advice By Test Takers Who Succeeded

-@ csetprep

Toxic NCLEX PN Test Prep Methods Exposed. Free Advice By Test Takers Who Succeeded

-@ nclex_pn_review

Toxic TExES Test Prep Methods Exposed. Free Advice By Test Takers Who Succeeded

-@ TExES_Practice

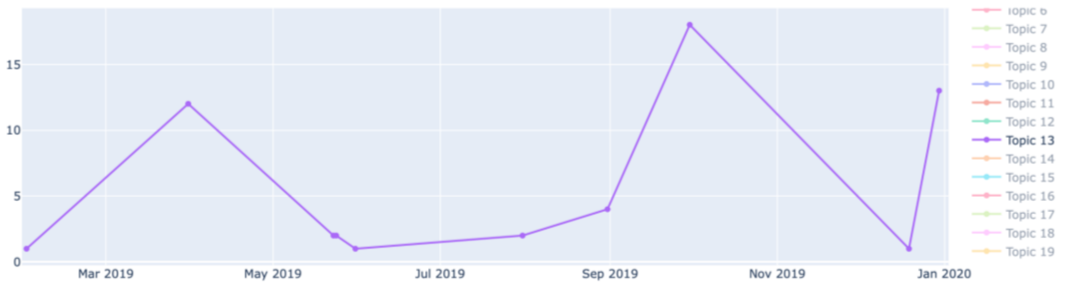

Figure 19. Time series of tweets posted in 13th topic

```
 (14,
  '0.019*"toxicity" + 0.012*"good" + 0.011*"4" + 0.010*"help" +
0.007*"time" + '
  '0.007*"drug" + 0.006*"waste" + 0.006*"ton" + 0.006*"million" + '
  '0.006*"skin"')
```

This topic was generated not very accurate, and it is hard to determine the main topic. This group describes two or more cases, first one is linked to a recommendation to some beauty blog, which was posted on 23.05.2019 and corresponds to the first and most probably to the second peak in the generated timeline in Figure 20. The next popular tweet topic in this group is the advertising of the documentary movie ‘Laugh Addict: Toxic Drug or Best Medicine?’ that was launched on 27 of August 2018. In this case, the announce date doesn’t correspond to the timeline and the spike of activity of this tweet can’t be predicted exactly, but most probably it is related to the end of 2019.

Cosmetic vs Drug Do Firming Creams Firm? Dark Eye Circles Dose vs Toxicity Visit #FullyEposed blog by #BoardCertDerm

*-@ FryFace_*

,"THE FILM HELPING PTSD, ADDCITION HELPS FAMILIES RECOVER WATCH 4 limited time FREE !!!! on http://LaughAddictFILM.com ""Laugh Addict: Toxic Drug or Best Medicine?"" Robin Williams, depression, trauma

-@ LaughOutNOW

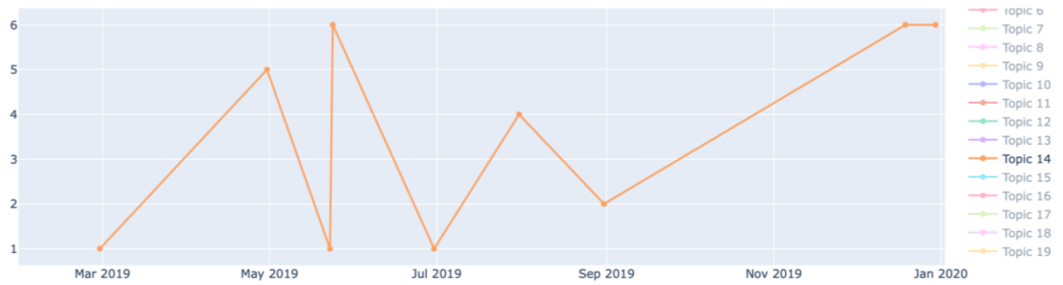

Figure 20. Time series of tweets posted in 14th topic

```
(15,
 '0.023*"food" + 0.017*"toxicity" + 0.010*"@youtube" + 0.010*"u.s" + '
 '0.009*"world" + 0.008*"new" + 0.008*"pesticide" + 0.008*"line" + '
 '0.008*"emperor" + 0.008*"robert"')
```

This topic is about a podcast in which a specific metabolite that gives sugar its dangerously damaging effects is described. The podcast was announced on 22.05.2019, the first peak on the timeline on Figure 21 that shows how popular was this topic corresponds especially on this date. Other peaks are most probably connected with other variety of tweets that contain words ‘food’ and ‘toxicity’ in it or just the late popularity of this podcast.

Robert Lustig and Fat Emperor - The Bottom Line on Processed Food Toxicity via @YouTube

#IvorCummins So worth the listen!

- @Swamper60

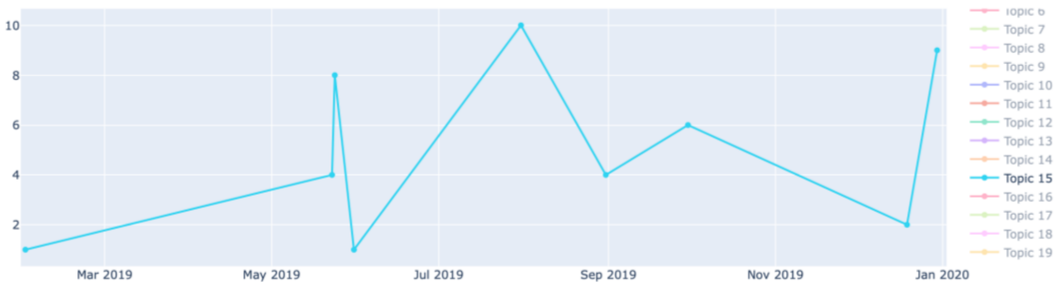

Figure 21. Time series of tweets posted in 15th topic

```
(16,
  '0.021*"toxicity" + 0.014*"protect" + 0.010*"help" + 0.010*"water" + '
  '0.009*"shellfish" + 0.009*"chemical" + 0.009*"metabolic" + 
0.009*"activity" '
  '+ 0.009*"coupling" + 0.009*"fatty"')
```

The 16th theme combines two different topics, both of them connected with scientific research. The most popular tweet that can be seen from the percentage of words popularity is research that describes how deep learning algorithms can be used to forecast shellfish toxicity to protect people from the dangerous effects of harmful algal blooms [38] According to the publishing date of this paper 19.12.2019 the highest popularity of this topic usage relates to the same date on Figure 22. Less popular, is the biological research about metabolic coordination of neutrons and astrocytes [39], it was released on 23.05.2019, and as the generated timeline of topic activity displays, that exactly in this time the first peak of usage this topic was recognized. In this case, the topic was generated very precisely.

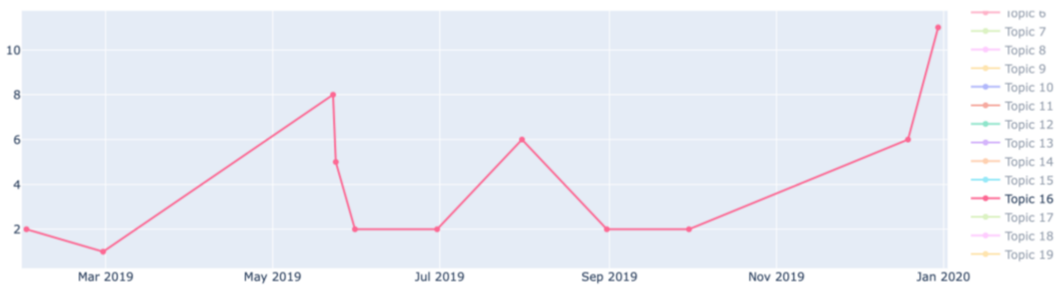

Figure 22. Time series of tweets posted in 16th topic

```
(17,
  '0.029*"air" + 0.026*"city" + 0.019*"pledge" + 0.018*"test" + '
  '0.018*"toxicity" + 0.018*"gas" + 0.017*"georgia" + 0.007*"lead" + '
  '0.007*"utensil" + 0.007*"water"')
```

This topic clearly describes very important news that in Georgia in the air the cancer-causing toxin is found. About this incident, a lot of sources were writing, but the main one, as can be seen from word popularity percentage is an article that some cities in Georgia pledge to test the air for toxic gas [40]. The article was posted on 31.07.2019, but the accident occurs a month earlier, on the timeline of the popularity of this theme in Figure 23, the peak relates to August 2019, which means exactly that users find out about the importance of this news right after it was published. Then the popularity of this topic decreases a then, at the end of 2019 another peak can be seen, which is most probably relate to some other events that contain words 'air' and 'city', which are accidentally counted by LDA modeling as the same topic.

Figure 23. Time series of tweets posted in 17th topic

```
(18,
  '0.018*"clean" + 0.018*"radiation" + 0.015*"nuclear" + 0.013*"cannabis"
+ '
  '0.011*"soil" + 0.010*"fire" + 0.009*"toxicity" + 0.008*"eu" + '
  '0.008*"health" + 0.007*"county"')
```

The 18th generated theme tall about an article that became very popular because it contains two very popular problems: its cannabis and nuclear radiation. This research is about how cannabis can help our planet to clean up the toxic radiated soil [41]. This article was launched on 16.03.2017, so the timeline of this topic in Figure 24 can show only that the interest in this topic increased by 15 times from February to December 2019.

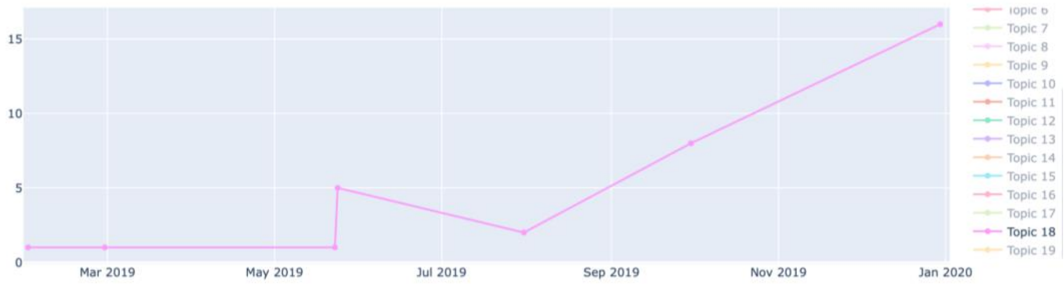

Figure 24. Time series of tweets posted in 18th topic

```
(19,
 '0.011*"like" + 0.010*"right" + 0.008*"study" + 0.008*"waste" + '
 '0.008*"health" + 0.008*"link" + 0.007*"cancer" + 0.007*"food" + '
 '0.007*"find" + 0.007*"water"')
```

The last topic counts many different topics that can't be combined personally, because of very low popularity, so LDA modeling finds their rather similar. In Figure 25 the timeline of the popularity of this topic is shown and it can be mentioned that mostly such tweets were posted in October. Here I put some examples of this topic group tweets:

'I remember one time, someone tried to convince me flouride is poisoning us. And cited a Harvard Study that shows toxicity but if you read it, it says the water also contained large traces of lead. The moral is read the study."

-@ AccessibleDan

RMPatPM study assesses impact of radiation dose to bladder neck on acute urinary toxicity & health-related quality of life in #ProstateCancer patients treated w/ MRI-guided high-dose-rate #brachytherapy combined w/ external beam #radiotherapy

-@ RadMedPM

A Turkish food engineer was sentenced to 15 months in jail after publishing the results of a study that linked toxic pollution to a high incidence of cancer in western Turkey"

- @PAPUfromPH

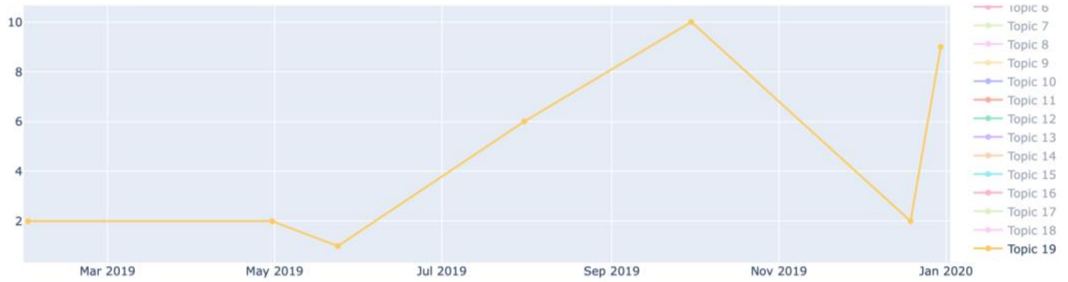

Figure 25. Time series of tweets posted in 19th topic

In figure 26 the overview of all topics' time distribution is shown. From this timeline, it can be seen that the 10th topic about the article that describes military bases that is toxic for people have the highest peak.

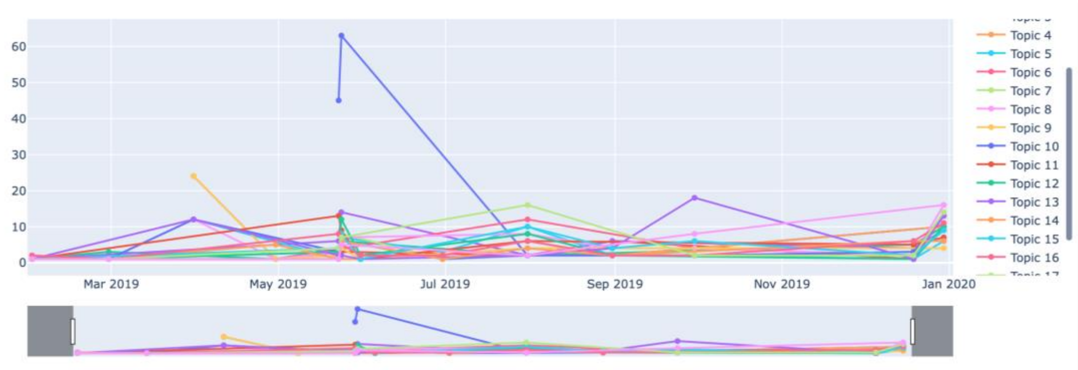

Figure 26. Time series of tweets in all topics in STEM group

Figure 27 describes how many tweets overall were posted in each group. Here it can be seen that the 0th topic group has the highest number of tweets; it proves that the highest interest of the public was about the article that investigates some US military bases that are toxic for humans. In the second place, we see the 8th topic group, which is about a famous person, who lost his voice because of mold.

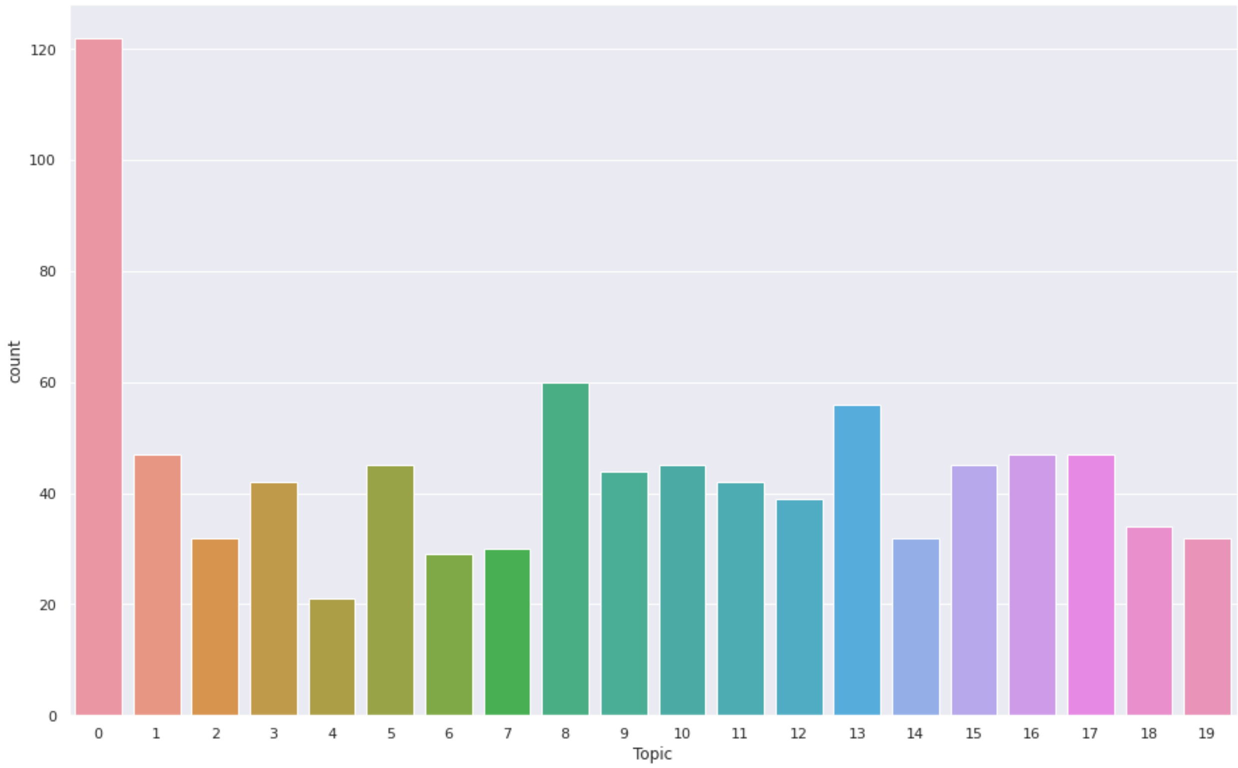

Figure 27. Number of tweets in each topic in STEM usage of word toxic.

4.2 Topic modelling with Latent Dirichlet Allocation (LDA) in case of social usage of word toxicity

In my next step, the LDA topic modeling of tweets in social usage of the word 'toxicity' was done. Here I decide to generate 14 topics because the coherence value of this number of topics, the according to graph in Figure 29 was the most suitable one. In this pack of generated topics, it is rather complicated to understand the real meaning, because, in the social context of using of word toxic, the definitions can vary.

Topic description

```
(0,
  '0.041*"love" + 0.021*"ass" + 0.019*"life" + 0.018*"rid" + 0.017*"fuck"
+ '
  '0.013*"thing" + 0.011*"trait" + 0.009*"time" + 0.008*"baby" + '
  '0.008*"worth"')
```

This topic most probably describes some unlucky love experience and advises to get rid of everything that is abusive in our own lives.

```
(1,
  '0.015*"know" + 0.014*"tell" + 0.010*"think" + 0.009*"toxicity" + '
  '0.008*"relationship" + 0.008*"like" + 0.007*"wait" + 0.007*"love" + '
  '0.006*"2" + 0.006*"trait"')
```

Consider the most frequently used words in this topic it can be said that posts that were assigned to this topic explain some ‘toxic’ thoughts and talks. As in most other topics, here we can see the most common words such as relationship, love, and trait.

```
 (2,
  '0.019*"relationship" + 0.015*"work" + 0.010*"toxicity" + 0.008*"read" +
'
  '0.007*"like" + 0.007*"world" + 0.006*"habit" + 0.006*"sign" +
0.006*"claim" '
  '+ 0.006*"song"')
```

The second topic most popular words are ‘relationship’ and ‘work’, then it is most likely to assume that this topic tells the problem of abusive relationship and work, that didn’t meet user’s requirements.

```
 (3,
  '0.021*"like" + 0.021*"leave" + 0.015*"relationship" + 0.013*"2019" + '
  '0.012*"energy" + 0.010*"want" + 0.010*"friend" + 0.009*"friendship" + '
  '0.008*"ass" + 0.008*"grow"')
```

The third topic can be described as a new year greeting. Here, users write that they want to leave all negative things in 2019 and enter the new year without any toxic traits. There are some examples of such tweets:

2019 has been a year where I've really begun to understand that toxicity, can be present in ALL forms of relationships and with some relationships it just takes a longer time to notice it.

-@ katahnaa

I'm tryna stay away from toxicity, 2019 was my cleanse lol.

-@ TriaTria

```
(4,
 '0.026*"twitter" + 0.014*"stan" + 0.012*"need" + 0.011*"right" + '
 '0.011*"toxicity" + 0.010*"fandom" + 0.008*"want" + 0.008*"support" + '
 '0.007*"woman" + 0.007*"know"')
```

The 4th topic describes the problem of toxicity in twitter and harassment against the woman. Another branch that can be seen in this post is that almost all of them encourage to support women's rights.

```
(5,
  '0.019*"toxicity" + 0.011*"day" + 0.010*"friend" + 0.009*"relationship"
+ '
  '0.009*"leave" + 0.007*"block" + 0.007*"hope" + 0.006*"boy" +
0.006*"trait" '
  '+ 0.006*"year"')
```

Considering the most popular words of this topic it can describe some bad days or unreliable friendship situations.

```
(6,
 '0.031*"good" + 0.020*"person" + 0.016*"toxicity" + 0.015*"think" + '
 '0.014*"life" + 0.011*"time" + 0.010*"relationship" + 0.008*"nigga" + '
 '0.008*"thing" + 0.007*"maybe"')
```

The 6th topic can cover a couple of social concerns: it can be aware of toxic relationships, pointing to noxious people, or just a storytelling reminder that in the life of a good person something toxic can interrupt too.

```
(7,
  '0.039*"year" + 0.029*"life" + 0.025*"2020" + 0.019*"family" +
0.018*"cut" + '
  '0.018*"new" + 0.015*"let" + 0.012*"friend" + 0.011*"feel" +
0.010*"know"')
```

This topic, one of the most understandable ones, mostly similar to the third topic, one of the main differences, that here people write about the new year 2020, but in previous, there were more mentions of the previous year, 2019.

I made my peace with everyone, I apologized where it was needed, forgave where it was wanted, and moved on. The very little people I have in my life right now is all I need. Don't need any toxicity in 2020, no thank you.

-@ ft_bbym

yesterday trump got impeached and i broke up with my boyfriend. we are not bringing toxicity into 2020

-@ priceyspicy

this year was filled with internal toxicity. wasn't necessarily poisonous to anyone but myself, and that slipped out to other people unintentionally. it is also a year of healing with suppressed traumas. lived and learned. 2020 will be much different. i feel it in my bones.

-@ _wolfesque

```
 (8,
  '0.037*"like" + 0.029*"shit" + 0.018*"talk" + 0.011*"wanna" +
0.010*"life" + '
  '0.009*"point" + 0.009*"trait" + 0.008*"bad" + 0.008*"time" +
0.007*"want"')
```

The 8th topic is very hard to define, but it seems to be connected with some bad conversations with people that cause bad lifetime spending for some people.

```
 (9,
  '0.017*"word" + 0.016*"need" + 0.016*"bitch" + 0.012*"damn" +
0.010*"start" '
  '+ 0.010*"toxicity" + 0.008*"watch" + 0.007*"hate" + 0.007*"wanna" + '
  '0.007*"mfs"')
```

In this generated topic it is hard to find the main idea. Because most tweets, in which people just impress feelings about everyday life in social networks can contain these words.

```
(10,
 '0.019*"masculinity" + 0.009*"drop" + 0.009*"know" + 0.009*"2020" + '
 '0.009*"laura" + 0.008*"ingraham" + 0.008*"feel" + 0.007*"company" + '
 '0.007*"end" + 0.006*"gillette"')
```

On this topic, two rather popular types of posts are described. The first one is complaining about some show of Laura Ingraham, that user society considers as 'toxic'.

@TheZebraCo Why are you advertising on Laura Ingraham's show? Advertisers have been dropping this show in droves due to host Laura Ingraham's highly-toxic and inflammatory rhetoric. #DropFox

-@ DebbieTrumbo

Another important topic that can be recognized from generated words is the article about the advertisement of Gillette company, which seems very controversial to the public. Most of the posts contain links to articles of Washington Times Journal [44] or some other magazines. Another part of tweets in this generated topic is about consequences of this occasion, for example:

"It's only by challenging ourselves to do more that we can get closer to our best" @ProcterGamble 's @Gillette is taking action to end toxic masculinity with their new Super Bowl ad #TheBestMenCanBe

-@ Sidney_W

The main article [44] of this topic was posted on 31 of July 2019, especially on this date we can see a peak on the timeline in Figure 30 that was generated for this topic frequency tweets distribution.

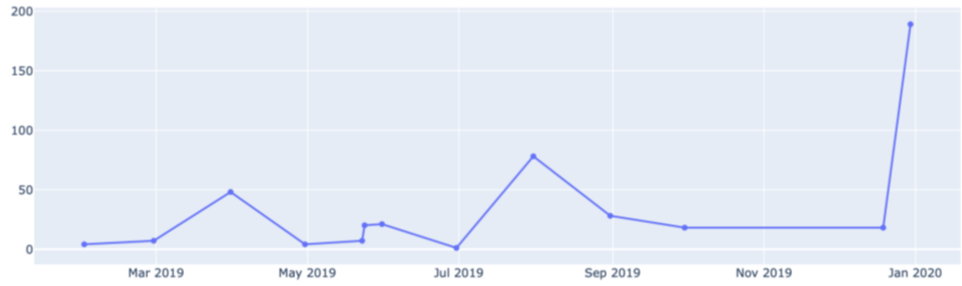

Figure 30. Time series of tweets posted in 10th social topic

```
(11,
  '0.026*"fuck" + 0.024*"fucking" + 0.017*"trait" + 0.013*">" +
0.011*"know" + '
  '0.009*"swear" + 0.008*"hurt" + 0.007*"life" + 0.007*"weekend" + '
  '0.006*"twitter"')
```

In this topic, most probably some bad habits are discussed. The word train is very popular through all these generated topics, but in this case, the weight of this word is the highest one. So, it can be assumed that most of the posts with a description of user's bad habits was assigned to this topic.

```
 (12,
  '0.025*"hate" + 0.015*"girl" + 0.014*"try" + 0.013*"play" + 0.013*"want"
+ '
  '0.011*"help" + 0.011*"change" + 0.009*"today" + 0.008*"friend" + '
  '0.007*"good"')
```

In the 10th topic, the common reflections of a whole group of socially toxic tweets can be seen. Here, the most frequent word is 'hate', so it can be guessed that users in this group express their feelings about things, that they don't like.

```
(13,
  '0.033*"like" + 0.018*"man" + 0.015*"shit" + 0.014*"want" + '
  '0.013*"relationship" + 0.011*"let" + 0.010*"social" + 0.010*"toxicity"
+ '
  '0.009*"love" + 0.008*"masculinity"')
```

The 13th topic clearly tells us about abusive masculinity that is more frequently can be seen in the relations of males and females. 'Toxic masculinity' term usually means the aggressive desire to dominate. Originally the domination, patriarchate position of man in the family was usual, but nowadays, in the time of woman concerning about their rights and abilities this way of living sometimes described as harassment against feminism.

As it can be seen from the topic description, only in a couple of generated groups the main ideas can be recognized completely. This happened because the social meaning of the word toxic usually uses such words as ‘relationship’, ‘trait’, ‘masculinity’, ‘work’, ‘person’, ‘live’. So, it is rather hard for machine learning to define the real idea of every post because most people use the same words, even if they are telling about different things. Another problem why it can be like that is that there is not enough data for more efficient results, or the number of topics is generated not correctly, even if calculations of coherence score on Figure 29 clearly show this result.

Figure 31, which shows the timeline of all topic’s frequency in social usage of word ‘toxic’ group, proves that machine learning text classification wasn’t successful in this case. All topics, that were generated have a rather similar usage timeline as can be seen from Figure 31. But, with manual analysis of parsed content, the results are not the same, different news and events that happen in the world cause a surge of public concern at different times. They are more popular at the beginning and completely subside after a time.

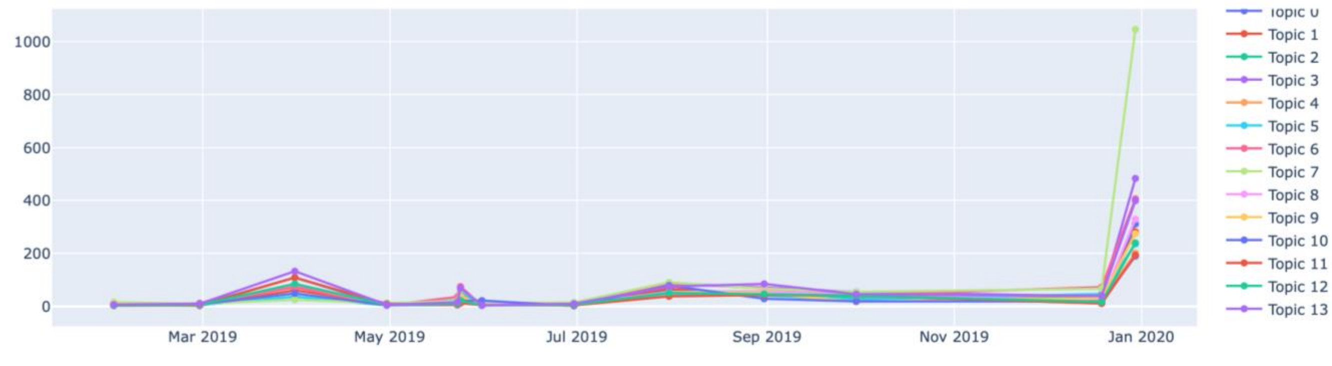

Figure 31. Time series of tweets in all topics in social group of tweets

Only one visible high peak can be recognized from the 7th topic at the end of the 2019 year. This topic is about greetings to the new year and finding a solution to improve your life by leaving everything ‘toxic’ and meet the new year in happiness. As it can be seen from Figure 32, where all topics are compared by the number of posts that they contain, topic number 7 is the most popular one. That’s why on a timeline such a big peak can be recognized. The next topic by the number of tweets in it is the third one. This topic contains the same information as the previously discussed one, but the algorithm defines it as different because here people mention 2019 year instead of 2020 as in the 7th topic. 6th and 13th generated topic is on the third place of usage, and these two topics talk about masculinity, that people don’t like and the abusive relationship that considers as one of the biggest concerns in the word cloud on Figure 4.

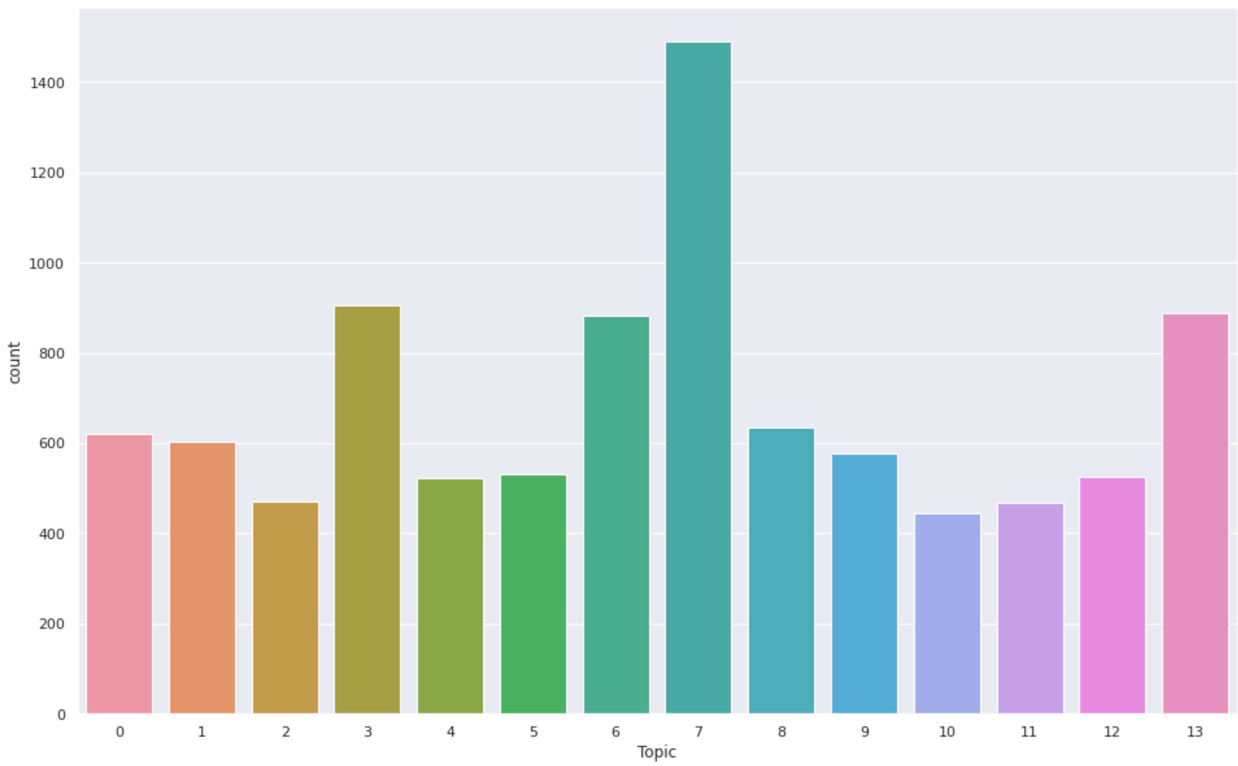

Figure 32. Number of tweets in each topic in a social usage of word toxic.

4.3 Text classification modelling

The next step of my research was creating of Scikit-Learn Multi-Class Text Classification model. This algorithm will predict for which class, for STEM or social usage of word ‘toxicity’, does this tweet rely on future data analyses. Without required manual interruption it can save a lot of time and can be applied to other research configurations.

The description of my code and code itself can be found in Chapter 2.3. In this part, the main results and their analysis will be presented. After all, data preprocessing, vectorizing, and classifier training I make a comparison of the most suitable for this case model.

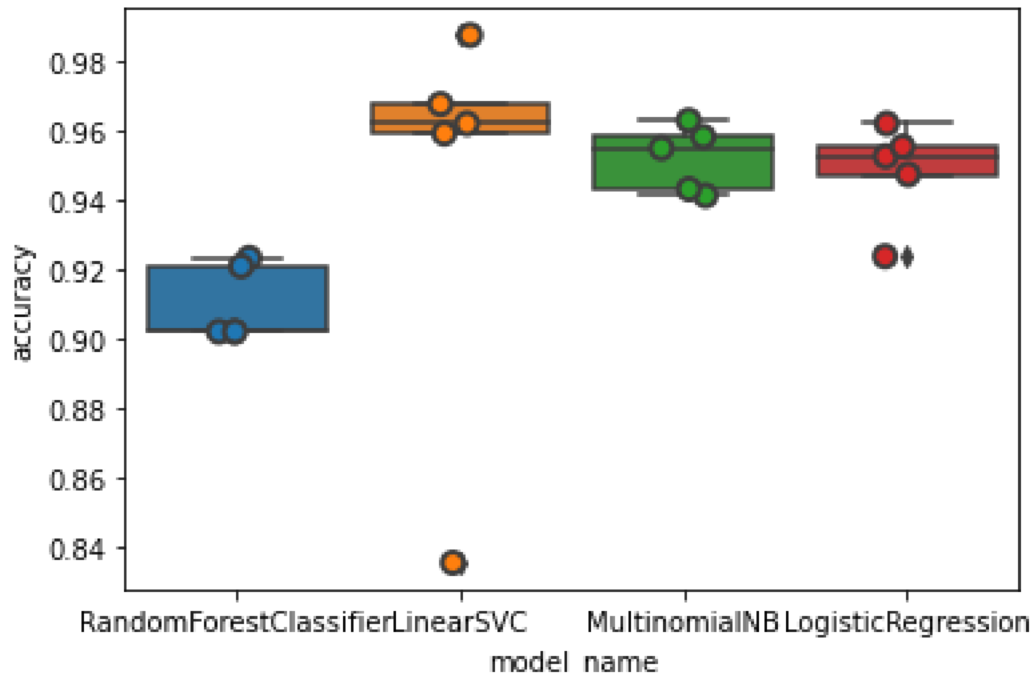

Figure 33. Accuracy comparison of different machine learning models.

In Figure 33 the graph representation of comparison is shown. On this plot, all accuracies of these 4 models are shown by intervals, but for choosing the most precise model it’s better to calculate the median accuracy.

Model name	Accuracy
LinearSVC	0.942170
LogisticRegression	0.948019
MultinomialNB	0.951887
RandomForestClassifier	0.909906

Table 1. Median accuracy of classifying models

In table 1 mostly all models have the same, a rather high mean accuracy, but for better results, the most precise, Multinomial Naive Bayes model was chosen.

At the next step, after teaching the model, the efficiency of it should be discussed. For these purposes usually, the confusion matrix is generated to show the discrepancies between predicted and actual labels. The confusion matrix is a performance measure for a machine learning classification task where the output can be two or more classes. This is a table with 4 different combinations of predicted and actual values, the general case is shown in Figure 34.

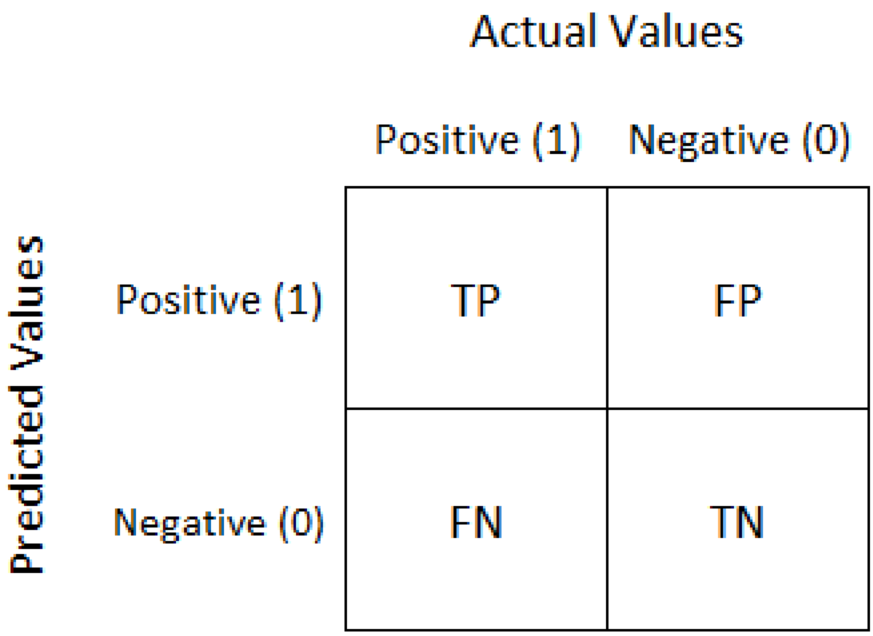

Figure 34. The principal description.

Interpretation of the of the values:

- TP - True Positive: positive is predicted and it is true.
- TN is a true-negative decision: a negative value is predicted, and it is true.
- FP - False-positive decision (Error Type 1): predicted as a positive value, and it is incorrect.
- FN- False Negative Solution (Error Type 2): Interpretation: You predicted a negative value, and it is incorrect.

Here I describe predicted values as positive and negative and actual values as true and false, as it shown on Figure 35.

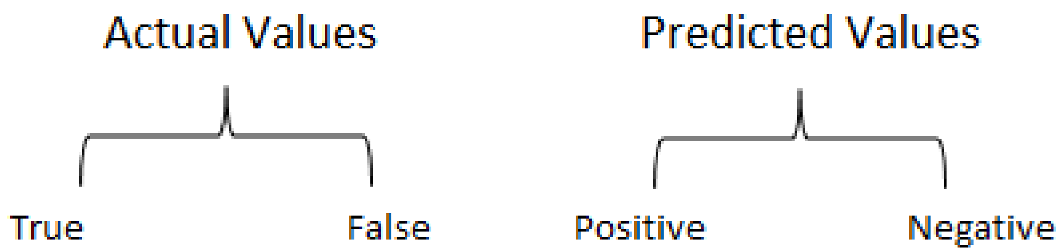

Figure 35. Description of Actual and Predicted values

Further calculations to find out the real efficiency of the model are based on all discussed previously values. Firstly, I calculated how much of all the positive classes, did the model predict correctly. It should be as high as possible. This number calls Recall and can be calculated by this equation:

$$\textbf{\textit{Recall}} = \frac{TP}{TP + FN}$$

Then the Precision value should be calculated, which means how many of all the classes, did the model predicts correctly. It should be as high as possible too. Usually, it calculates by this formula:

$$\textbf{\textit{Precision}} = \frac{TP}{TP + FP}$$

It is difficult to compare two models with low accuracy and high responsiveness or vice versa. Therefore, to make them comparable, usually, the F-measure is used. The F-measure helps to measure Recall and Precision at the same time. It uses the harmonic mean instead of the arithmetic mean, penalizing extreme values more. The equation for this value is:

$$\textbf{\textit{F - measure}} = \frac{2 * Recall * Precision}{Recall + Precision}$$

In Figure 37 the confusion matrix for the MultinominalNB model for this research is shown. As it can be seen, most of the predictions end up on the diagonal, this means that the predicted label is equal to the actual label. In general, especially this distribution is needed for high-quality predictions [42]. However, there are a number of misclassifications, that can be caused by imperfection of machine learning modeling or lack of input data.

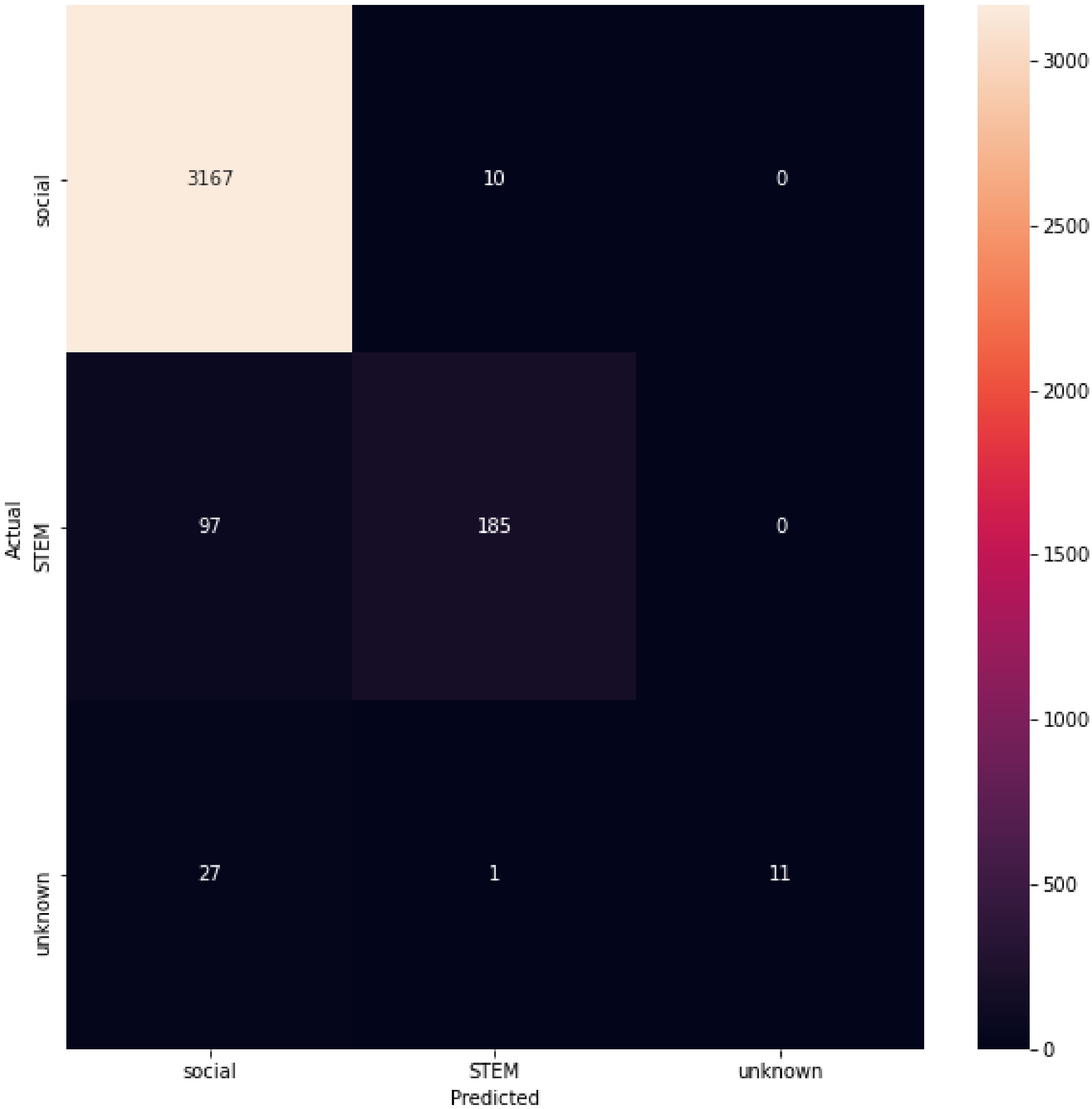

Figure 36. Confusion matrix of MultinominalNB model

	Precision	Recall	F-measure
Social	0.96	1.00	0.98
STEM	0.94	0.66	0.77
Unknown	1.00	0.28	0.44

Table 2. Values of efficiency of prediction model

In Table 2 the calculations of Recall, Precision, and F-measure values for this research data are shown. After the analyzes of this model, it can be said that the group of social usage of word toxic is the most successfully predicted one because it has the biggest number of tweets and input data in it. The dependence of input sufficient to the model efficiency can be recognized from the F-measure value on the ‘unknown’ group of tweets. It has the lowest precision, because of the lack of ‘teacher’ data.

Conclusion

This paper analyses the use of the word 'toxicity' in the English-speaking social network Twitter. Data from the social network was collected and transcribed, and a comparative analysis was made to determine how often the word 'toxicity' is used in different scenarios of using this word.

In this research, the topic modeling of all analyzed data was done. The code for this task was performed in the python programming language with Latent Dirichlet Allocation modeling library. For each of two classified manually groups depending on usage of word ‘toxicity’: STEM - science usage or social, each generated topic was described according to most frequent word. For STEM usage of the word ‘toxicity’, 20 topics were generated by the machine learning approach. They can be described rather comprehensibly, the patterns and dependencies in the generated data can be found rather transparent. According to this, it could be said that Latent Dirichlet Allocation modeling works correctly with this group of tweets.

For the group with social usage of the term ‘toxicity,’ 14 topics were generated. In this case, there were a lot of inaccuracies in dependencies and the patterns could be barely found. Additionally, the interpretation of these generated topics couldn’t be done very precisely, because mostly the most frequent words in it were the same. It is possible to conclude, that machine learning in the case of social studies works worse than with the science tweets concept. Failure in this topic modeling can be due to the lack of parsed data too, but the number of tweets in the STEM category was about 10 times smaller than in the social content group. So most probably that the input data should be enough for a proper topic generation. Another case why in this case machine learning wasn’t successful is that Latent Dirichlet Allocation modeling ability wasn’t enough, and, in that case, a more suitable library can be founded.

For future research, related to this topic the text classification modeling was done to modernize the data sorting process. The code was performed on python programming language and the Multinomial Naive Bayes model was used. The classification algorithm then was tested on the same data set in order to compare the manual and machine learning sorted values. After analysis of comparison output, it can be said that the taught classifying mechanism is rather precise. The accuracy of classifying is proportional to the size of the input dataset for each class. The best classifying of tweets was performed in the case of social usage of wort ‘toxicity’, especially since this group contains the largest amount of data.

Literature

1. Ashley A. Anderson, Dominique Brossard, Dietram A. Scheufele, Michael A. Xenos, Peter Ladwig, The “Nasty Effect:” Online Incivility and Risk Perceptions of Emerging Technologies, Journal of Computer-Mediated Communication, Volume 19, Issue 3, 1 April 2014, Pages 373–387.
2. Georgakopoulos S.V., Tasoulis S.K., Vrahatis A.G., and Plagianakos V.P. 2018. Convolutional Neural Networks for Toxic Comment Classification. In Proceedings of the 10th Hellenic Conference on Artificial Intelligence (SETN ’18). Association for Computing Machinery, New York, NY, USA, Article 35, 1–6.
3. User’s manual R-Studio // R-tools Technology Inc 2020. P. 155-160.
4. Why Twitter is a toxic place for women [Electronic resource]. URL: https://www.amnesty.org/en/latest/research/2018/03/online-violence-against-women-chapter-1/ (accessed: 15.03.2021).
5. Jack Dorsey On Deleting Tweets, Banning Trump, And Whether An Unbiased Twitter Can Exist [Electronic resource] // BuzzFeed News. URL: https://www.buzzfeednews.com/article/charliewarzel/11-things-twitters-ceo-told-us-about-sean-hannity-deleting (accessed: 28.04.2021).
6. Bendix A. Jack Dorsey says Twitter makes it “super easy” to harass and abuse others, and addressing the problem is his biggest worry [Electronic resource] // Business Insider. URL: https://www.businessinsider.com/jack-dorsey-twitter-abuse-ted-2019-4 (accessed: 11.04.2021).
7. Twitter Is Trying To Fix The Dunk And Ratio [Electronic resource] // BuzzFeed News. URL: https://www.buzzfeednews.com/article/alexkantrowitz/twitter-is-trying-to-fix-the-dunk-and-ratio (accessed: 11.04.2021).
8. Grothaus M. A woman on Twitter is abused every 30 seconds [Electronic resource] // Fast Company. 2018. URL:

https://www.fastcompany.com/90283504/a-woman-on-twitter-is-abused-every-30-seconds (accessed: 11.04.2021).

9. Why Twitter is a toxic place for women [Electronic resource]. URL: https://www.amnesty.org/en/latest/research/2018/03/online-violence-against-women-chapter-1/ (accessed: 15.04.2021).
10. Amnesty reveals alarming impact of online abuse against women [Electronic resource]. URL: https://www.amnesty.org/en/latest/news/2017/11/amnesty-reveals-alarming-impact-of-online-abuse-against-women/ (accessed: 15.04.2021).
11. Chandrasekharan E, Samory M, Srinivasan A, Gilbert E. The Bag of Communities: Identifying Abusive Behavior Online with Preexisting Internet Data. Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. New York, NY, USA: ACM; 2017. pp. 3175–3187.
12. Zafarani, R., Abbasi, M., & Liu, H. (2014). Introduction. In Social Media Mining: An Introduction (pp. 1-10). Cambridge: Cambridge University Press. doi:10.1017/CBO9781139088510.002
13. Irfan, R., King, C., Grages, D., Ewen, S., Khan, S., Madani, S., . . . Li, H. (2015). A survey on text mining in social networks. The Knowledge Engineering Review, 30(2), 157-170. doi:10.1017/S0269888914000277
14. Aggarwal C.C. (2011) An Introduction to Social Network Data Analytics. In: Aggarwal C. (eds) Social Network Data Analytics. Springer, Boston, MA. https://doi.org/10.1007/978-1-4419-8462-3_1
15. Morgun O.M. Critical Discourse Analysis in the Methodology of Political Science // Political Linguistics. no. 3. C. 122-129.
16. Krzyżanowski M., Forchetner B. Theories and concepts in critical discourse studies: Facing challenges, moving beyond foundations // Discourse and society. Vol. 27. No. 3. P. 253–262. DOI: 10.1177/0957926516630900.
17. Harris Z. Discourse Analysis // Language. Vol. 28. No. 1. P. 1–30.
18. Ezhova T.V. On the differences in the understanding of the term 'discourse' // Vestnik of the Orenburg State Pedagogical University. № 1. C. 61-65.

19. Jørgensen, M., & Phillips, L. J. (2002). Discourse analysis as theory and method. SAGE Publications Ltd, 100-104 https://www.doi.org/10.4135/9781849208871
20. Irekponor V.E. marquisvictor/Optimized-Modified-GetOldTweets3-OMGOT: Python. 2021. https://github.com/marquisvictor/Optimized-Modified-GetOldTweets3-OMGOT
21. University of Miami Miller School of Medicine. "Cyberbullying Linked to Increased Depression and PTSD." ScienceDaily. ScienceDaily, 22 January 2020. <www.sciencedaily.com/releases/2020/01/200122080526.htm>.
22. BMJ. "Cyberbullying linked to post traumatic stress for victims and perpetrators: Children should be screened for symptoms, suggest researchers." ScienceDaily. ScienceDaily, 23 June 2020. <www.sciencedaily.com/releases/2020/06/200623185242.htm>.
23. ИСУ ИТМО [Electronic resource]. URL: https://isu.ifmo.ru/pls/apex/f?p=2143:LOGIN:111393547097680 (accessed: 01.05.2021).
24. Topic Modeling in Python with Gensim [Electronic resource] // ML+. 2018. URL: https://www.machinelearningplus.com/nlp/topic-modeling-gensim-python/ (accessed: 10.05.2021).
25. Tong, Zhou & Zhang, Haiyi. (2016). A Document Exploring System on LDA Topic Model for Wikipedia Articles. The International journal of Multimedia & Its Applications. 8. 01-13. 10.5121/ijma.2016.8401.
26. David M. Blei, Andrew Y. Ng, and Michael I. Jordan. 2003. Latent dirichlet allocation. J. Mach. Learn. Res. 3, null (3/1/2003), 993–1022.
27. Fairclough, N. (1995). Critical discourse analysis: The critical study of language. London: Longman.
28. UK T.D. Guardian. A Trail of Toxicity: The US Military Bases Making People Sick [Electronic resource] // Reader Supported News. URL: https://readersupportednews.org/news-section2/318-66/56715-a-trail-of-toxicity-the-us-military-bases-making-people-sick (accessed: 10.05.2021).

29. HOME | súki candles | Australia [Electronic resource] // Suki Candles. URL: https://www.sukicandles.com.au (accessed: 01.05.2021).
30. Dog Deaths Linked To Toxic Algae Showcase Rising Global Toxicity [Electronic resource] // URL: https://www.bitchute.com/video/7cbPeFr8jlVF/
31. Blue-Green Algae, Toxic Water. A Major Health Issue. [Electronic resource] // Florida Sportsman. 2019. URL: http://www.floridasportsman.com/2019/04/30/blue-green-algae-toxic-water-major-health-issue/ (accessed: 05.05.2021).
32. Victory! The EU has banned a pesticide that's toxic to bees and humans [Electronic resource] // SumOfUs. URL: https://actions.sumofus.org/a/the-pesticide-thiacloprid-is-toxic-to-bees-and-humans-ban-it-now-8 (accessed: 05.05.2021).
33. Silicone Safety: Risks, Exposure Sources, Is Silicone Toxic & More [Electronic resource] // Healthline. 2019. URL: https://www.healthline.com/health/body-modification/is-silicone-toxic (accessed: 05.05.2021).
34. https://coursesite.lehigh.edu/mod/page/view.php?id=1274898
35. https://hairsofine.com/collections/thinning-hair-treatment-collection
36. Sears M.E., Kerr K.J., Bray R.I. Arsenic, cadmium, lead, and mercury in sweat: a systematic review // J Environ Public Health. 2012. Vol. 2012. P. 184745.
37. nutraingredients-usa.com. New mouse study finds level of liver toxicity for CBD [Electronic resource] // nutraingredients-usa.com. URL: https://www.nutraingredients-usa.com/Article/2019/05/24/New-mouse-study-finds-level-of-liver-toxicity-for-CBD (accessed: 05.05.2021).
38. Forecast to help shellfish growers weather toxicity [Electronic resource]. URL: https://phys.org/news/2019-12-shellfish-growers-weather-toxicity.html (accessed: 18.05.2021).

39. Ioannou M.S. et al. Neuron-Astrocyte Metabolic Coupling Protects against Activity-Induced Fatty Acid Toxicity // Cell. 2019. Vol. 177, № 6. P. 1522-1535.e14.
40. Goodman B., MA, Miller y. Georgia City Pledges To Test Air for Toxic Gas [Electronic resource] // WebMD. URL: https://www.webmd.com/cancer/news/20190731/georgia-city-pledges-to-test-air-for-toxic-gas (accessed: 18.05.2021).
41. How Cannabis Cleans Up Nuclear Radiation And Toxic Soil [Electronic resource]. URL: https://hightimes.com/news/how-cannabis-cleans-up-nuclear-radiation-and-toxic-soil/ (accessed: 18.05.2021).
42. Li S. Multi-Class Text Classification with Scikit-Learn [Electronic resource] // Medium. 2018. URL: https://towardsdatascience.com/multi-class-text-classification-with-scikit-learn-12f1e60e0a9f (accessed: 18.05.2021).
43. Google Colaboratory [Electronic resource]. URL: https://colab.research.google.com/drive/1YO9aRW3LkcHkyUZVVTeIJCZWcraAikPr#scrollTo=d6AL9BXfA0hM (accessed: 28.05.2021).
44. https://www.washingtontimes.com T.W.T. Gillette’s ‘toxic masculinity’ ad haunts P&G as shaving giant takes $8B writedown [Electronic resource] // The Washington Times. URL: https://www.washingtontimes.com/news/2019/jul/31/gillettes-toxic-masculinity-ad-haunts-pg-as-shavin/ (accessed: 28.05.2021).