Model for forecasting the development of information threats in the cyberspace of Ukraine

Mariia Nazarkevych[1,2,†], Victoria Vysotska[1,3,*,†], Yurii Myshkovskyi[1,†], Nazar Nakonechnyi[1,†] and Andrii Nazarkevych[1,†]

[1] Lviv Polytechnic National University, 12 Stepana Bandera str., 79013 Lviv, Ukraine
[2] Ivan Franko National University of Lviv, 1 Universitetska str., 79000 Lviv, Ukraine
[3] Osnabrück University, 29 Neuer Graben, 49074 Osnabrück, Germany

Abstract

Approaches to the formation of models for forecasting the development of information threats in cyberspace have been developed, which is an urgent task when fake news and information manipulation can affect public sentiment, politics, and the economy. The program uses machine learning and Natural Language Processing (NLP) techniques to detect fakes in a dataset. In the developed method, we train the model on a data set where true and fake news or any other types of information are already marked. The model can then be used to classify new data. The dataset contains news that the average Ukrainian saw during the war in the Internet space on such social networks as Telegram, Facebook, and Twitter, on news sites. The language of the messages, which were in Ukrainian and Russian, was highlighted as a separate field. In a separate field, it was noted how many people liked and how many people shared this message. The data set contains some fake news and some real news. The F1 score is 0.98 for both classes (0-forgery, 1-not forgery). Such good results can be explained by the "laboratory" quality of the data set. In further experiments, we will test the model on real-time news.

Keywords

information threats, cyberspace, fake messages, machine learning

1. Introduction

Today, society is increasingly faced with various types of cyberattacks: failures in the provision of electronic services, blocking the work of state bodies, phishing attacks by e-mail, cybercrimes, violations of data integrity and confidentiality, information-psychological pressure on the population, cyberterrorism, cyberespionage, information expansion into the national information space of the country, blocking the work or destruction of strategically important enterprises for the economy and security of the state, life support systems and objects of increased danger [1, 2].

2. The main types of cyber-attacks

Malware is a type of program that can perform various malicious tasks. Some types of malware are designed to create persistent network access, some are designed to spy on a user to obtain credentials or other valuable information, and some are simply designed to disrupt operations. Some types of malware are designed to extort money from the victim. Probably, the most famous form of malicious software is a ransomware program—it is designed to encrypt the victim's files and then demand a ransom to obtain the decryption key [3].

Cyberspace, along with other territories, is recognized as one of the potential theaters of war, so the state's ability to protect its national interests is considered an important component of cyber security.

2.1. Distributed attacks

Criminals actively work on finding vulnerabilities in assets (management systems) and develop for this purpose unique in their characteristics: universal malicious software, encryption viruses, botnets that perform distributed attacks (DDoS) on operating networks, production systems that use cloud services, as well as supply chain attacks. Given the progress in artificial intelligence technologies over the next 5–10 years, the scope and consequences of such interventions will grow. The expansion of the use of cyberspace by terrorist organizations (cyberterrorism) is becoming a global trend [4].

The new resolution of the Government of Ukraine will allow timely response and planning of cyber protection measures. We are talking about the Resolution of the Cabinet of Ministers of Ukraine dated 04.04.23 No. 299

CPITS-II 2024: Workshop on Cybersecurity Providing in Information and Telecommunication Systems II, October 26, 2024, Kyiv, Ukraine
* Corresponding author.
† These authors contributed equally.
mariia.a.nazarkevych@ lpnu.ua (M. Nazarkevych);
victoria.a.vysotska@lpnu.ua (V. Vysotska);
yurii.myshkovskyi@lpnu.ua (Y. Myshkovskyi);
nazar.i.nakonechnyi@lpnu.ua (N. Nakonechnyi);
andrii.nazarkevych.ri.2023@lpnu.ua (A. Nazarkevych)
0000-0002-6528-9867 (M. Nazarkevych);
0000-0001-6417-3689 (V. Vysotska);
0009-0004-0051-026X (Y. Myshkovskyi);
0009-0000-2456-3498 (N. Nakonechnyi);
0009-0007-2078-8447 (A. Nazarkevych)

"Some issues of response of cyber security entities to various types of events in cyberspace" [5].

2.2. Ransomware or blackmailer

Highlight the following categories: Malware is an attack on a wide audience, in particular on the Internet. "Ransomware or Blackmailer", which is a partial case. Distributed "denial of service" DDoS attacks are attacks aimed at blocking the operation of a specific network resource. The attack can be implemented by the following three mechanisms: overflow of the communication channel, "denial of service")—a hacker attack on a comprehensive system to bring it to failure, that is, creating such conditions under which bona fide system users will not be able to access the provided system resources (servers), or this access will be closed.

Failure of the "enemy" system can also be a step towards mastering the system, if, in the next situation, the software releases some critical information—for example, the version, part of the software code, etc. DoS is a simplified variant of DDoS attacks. A distinctive feature is the clear manifestation of the moment of attack.

Table 1
Types of attacks and software that detects them

Attack type	ESET Internet Security	Avast Premium Security	Bitdefender Total Security	Avira Internet Security
Malicious software	+	+	+	+
Distributed attacks, denial of service DDoS	+	+	+	+
denial of service DoS	+	+	+	+
Phishing (social engineering)	+	+	+	+
Using SQL injections	+	+	+	+
Cross-Site Scripting (XSS)	+	+	+	+
Botnets	+	+	+	+
Brute force attack	+	+	+	+
Drive-By Download	+	+	–	–
The Man in the Middle	+	+	–	–
Ransomware or blackmailer	+	+	–	–
Unsuccessful authorization attempts	+	+	+	+
Attempts to exploit vulnerabilities	+	+	+	+
Publication of fraudulent information	+	+	–	–
Network scanning	+	+	+	+

2.3. DoS attacks

DoS vulnerabilities are refusal of service stand separately in several security threats (Fig. 1). As a rule, this class of attacks includes events described in the news "Hackers attacked site X, disrupting its operation. The site was down for Y hours". Requests are made to the server that it cannot (does not have time to) process, as a result, it does not have time to process the requests of ordinary visitors and appears to them as not working. These attacks are not intended to steal data from the database but can help launch other types of attacks, i.e. clear the path. For example, some programs can cause exceptional situations due to errors in their code.

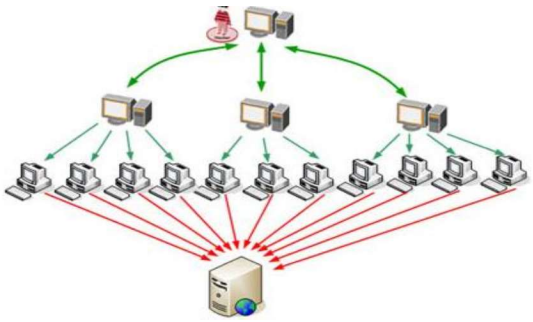

Figure 1: DOS attacks

It is impossible to protect against DOS attacks 100%, but it is possible to limit the number of login attempts from the same IP address in a certain amount of time. For example—no more than 5 in 10 minutes. When running out, show a "wait" message or offer to enter a CAPTCHA. Some systems ask to enter the CAPTCHA in general at each login attempt [6].

2.4. Phishing attacks

is the practice of sending emails that appear to be from trusted sources to obtain personal information or influence users to do something. It combines social engineering and technical techniques. It could be an email attachment that downloads malware to your computer. It can also be a link to an illegal website that can trick you into downloading malware and handing over your data. Spear phishing is a very targeted type of phishing. Attackers spend time researching targets and crafting messages that are personal and relevant. Therefore, spear phishing is very difficult to recognize and even more difficult to protect against it. One of the easiest ways a hacker can conduct a spear phishing attack is through email spoofing, where the information in the "From" section of an email is faked to make it look like the email is coming from someone you know, such as your management or company—partner. Another trick scammers use to give their story credibility is website cloning: they copy legitimate websites to trick you into entering personal information or login credentials [7].

2.5. Cros-site scripting attack

A cross-site scripting (XSS) attack occurs when a site has a vulnerability that allows the introduction of scripts (Fig. 2). Attackers use such vulnerabilities and introduce malicious JS scripts into the database site data. When the user subsequently requests this data, the user's web browser executes a malicious JS script. This would allow an attacker to steal browser cookies to hijack the session. Hackers can then use the session information to exploit additional vulnerabilities, possibly gain network information, and control the user's computer. This is especially important in an enterprise environment, as a single XSS attack (Fig. 2) can compromise an entire network [8].

In order not to become a victim of an XSS attack, the following security rules should be observed: all nested structures must be filtered. Encryption. When creating a filter, you must take into account the risk of encoding attacks. There are a lot of encoder programs that can be used to encrypt any attack so that more than one filter will not "see" it. Application of tags. There is one vulnerability related to tags url, bb, img, which have many parameters including lowsrc and dynsrc containing javacsript. These tags should be filtered [3].

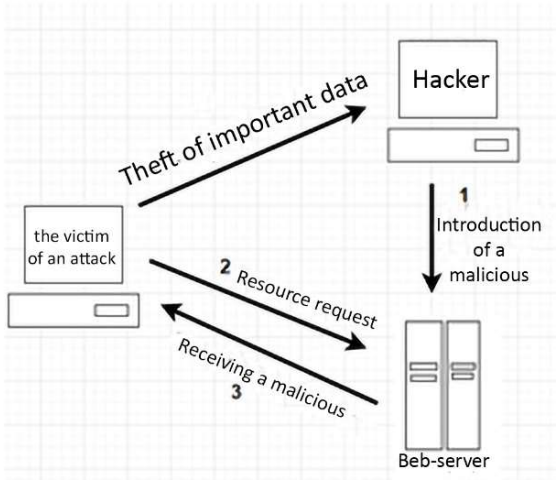

Figure 2: XSS attack

2.6. Brute force attack

A brute force attack, sometimes called a password attack, is one of the simplest forms of web attacks. The hacker simply tries different combinations of usernames and passwords over and over again until he gets into the user's account. Of course, one computer would need years to go through all the combinations. But when hackers gain control over several computers or develop a powerful software computing engine, things can become very simple. Bruteforce is one of the most popular methods of cracking passwords of online bank accounts, payment systems, or websites. But as the length of the password grows, this method becomes inconvenient due to the length of time it takes to go through all possible options [9].

3. Model for forecasting the development of information threats

One of the most common cyber threats is the penetration of false information into the information space of Ukraine. Among them, false news occupies an important place. This news is also called fake news. The information space of Ukraine needs the development of new protection systems, as an uncontrolled process leads to the penetration of false information, which users spread in every way. The development of methods and tools for monitoring and detecting misinformation on the Internet is an urgent task in the conditions of the modern digital age when fake news and manipulation of information can affect public sentiment, politics, and the economy. NLP is a rapidly developing technology that helps businesses get the most out of artificial intelligence. Analytical research predicts an increase in the global NLP market from USD 20.98 billion in 2021 to USD 127.26 billion in 2027, with a compound annual growth rate (CAGR) [10] of 29.4%. Today, texts are analyzed using artificial intelligence using methods of NLP to analyze text messages and search for signs of manipulation or fake information. For example, artificial intelligence can detect suspicious speech patterns that are typical of disinformation. Also practiced is such an approach as fact-checking based on automated systems, which consists of an automated fact-checking system that can quickly compare information with reliable sources and determine whether it is reliable. For this, databases with verified information and algorithms for its analysis are used. For social networks, users' behavior is monitored, identifying the disseminators of disinformation and detecting networks engaged in the manipulation of mass consciousness. Blockchain technology is widely used to ensure transparency of information, which will reduce the number of fake news, as all information will be transparently tracked. It is necessary to develop crowdsourced platforms for fact-checking, where users can verify information themselves and provide their results, also effectively contribute to the detection of disinformation.

Effective development of disinformation detection methods requires a combination of technological innovations with international cooperation, regulatory measures, and increasing the level of digital literacy of users. Several methods and technologies based on artificial intelligence (AI) [11–13], machine learning [14, 15], and NLP are used to classify information as true and false [16, 17]. These methods allow you to automate the process of verifying the authenticity of information and quickly identify disinformation. NLP is becoming an important part of modern systems. It is intensively used in search engines, language interfaces, document processors, etc. Computers are very good at dealing with structured data. If the texts are in free form, computers face difficult tasks. The goal of NLP is to develop algorithms that would allow computers to recognize free text and understand live speech. The amount of variation possible is one of the biggest challenges in NLP. Context is of great importance for understanding the meaning of individual sentences. People are very good at this because they learn to understand the content over many

years. We apply our knowledge to understand the context and know exactly what the other person is talking about. To overcome this problem, researchers in the field of NLP have begun to develop various applications using machine learning-based approaches. To develop such applications, we have to collect huge arrays of text and then train an algorithm to perform various tasks, such as text categorization, sentiment analysis, or topic modeling. At the same time, the algorithms learn to detect patterns that repeat in the input text and get the content embedded in it.

Natural language has syntactic ambiguity, which is shown in the proverb "Time is not a horse, you can't drive it and you can't stop it". For NLP, it is unclear whether the sentence is about a horse or time. The Ukrainian language has a case ambiguity: in the phrases "Everyone was excited before the concert" and "It's not necessary to give before!" the word before means time or place, which completely changes the meaning of the phrase. There is also a referential ambiguity: in the phrase "Open the shelf and take out the wet umbrella, I want to dry it", the pronoun she will refer to the wet umbrella by its semantic meaning, but for the machine, which has a complete lack of understanding of reality, this pronoun will refer to both the shelf and to the umbrella. One of the challenges that arises in the process of NLP can be considered the problem of the presence of synonyms, as a result of which one concept can be expressed by several different words. As a result, documents that use synonyms may not be identified by the system. The influence of the above phenomena is especially noticeable when creating machine translation systems. The problem lies in the difficulty of establishing a concrete mapping of the valid semantic-syntactic structure of a sentence into its internal logical representation, which is automatically generated by the system.

3.1. General norms for the formation of messages

Postulates are not explicitly stated in the editing literature, although they are always used when processing messages. We think that fixing them will allow you to better understand the features of editing. Let's list the postulates that, in our opinion, should be adopted in the editing. The message must necessarily contain new information for the recipient. The message must have a defined modality. The message must be adapted to the time, place, and situation in which it will be perceived by the recipient. The author must use language and word meanings known to the recipients. The message must be adapted to the recipient's thesaurus. In the message, mechanisms should be implemented only for the perception of information by the recipient. In the message, means must be implemented that force the recipient to perceive it. The message must be protected from noise. The message must comply with the norms adopted at a specific time in a specific society. In addition to these postulates, which directly follow from the editing axiom, one more should be added to their number.

Any general (postulate) or specific norm can be violated if it leads to the set goal. Solving these types of ambiguities is possible by introducing additional values that will increase the program's knowledge of a particular industry. Today, there are no programs that "understand" all types of ambiguities in a wide range of industries, but there are programs that can correctly respond to ambiguities in very narrow areas.

Classification of fakes

Fake (forgery) is false [18], often sensational information, distributed under the guise of news, that is, it is fake news. Fakes are created to gradually, step by step, form relationships, that is, to create reactions to a certain social group. The biggest danger from fakes is their cumulative effect. Fakes distort reality and undermine trust in the media. Scientists from the University of Western Ontario distinguish five types of fakes:

- intentionally created fakes
- jokes perceived as truth
- large-scale hoaxes
- intentionally one-sided coverage of events
- stories in which the "truth" is contradictory (for example, a terrorist for some is a freedom fighter for others).

Fakes were first mentioned in 1981 when journalist Janet Cook won a Pulitzer Prize for her story "Jimmy's World" for The Washington Post. Stephen Glass worked for the Washington magazine The New Republic from 1995 to 1998 and did not care about sensationalism, he simply invented them—half of his articles in TNR were fabricated. According to the observations of David Peterson [19], the editor of the Viralgranskaren project (Sweden), fakes are created:

- Viral sites create an instant response in the audience.
- Pranksters, to weigh in on the audience, are set on intellectuals.
- Scammers hook and lead to the goal.
- Ideological and political views, so that it is almost impossible to convince.
- Foreign players.
- And finally, ordinary people who do not create, but distribute.

4. The method of detecting fake messages

We will use the Python Natural Language Toolkit (NLTK) [20] package to build the corresponding applications. Be sure to install this package before reading further. Enter the following command in a terminal window: $ pip3 install nltk The use of neural networks and machine learning is based on labeled data: Neural networks are trained on a large number of examples of true and false information. During training, the model analyzes various characteristics of the text—vocabulary, syntax, presentation style, as well as sources of information. The model then learns to distinguish between true and false information based on these features. Classification algorithms based on the method of support vectors, decision trees, and deep neural networks are used to build models that classify texts as true or false based on statistical features.

NLP is carried out by analyzing linguistic features. The system analyzes the text for emotional color, level of bias, and degree of confidence or uncertainty in the presentation

of facts. For example, fake information often contains sensational or emotionally charged headlines and phrases. Search by keywords and phrases is also used. NLP technologies help find patterns or keywords often used in fake news, including elements of conspiracy theories or exaggerated claims. Fact-checking is used by checking literary sources. Machine learning can automatically find links to information sources and verify their credibility using databases of trusted news organizations or official sources. You can also compare it with other sources. Algorithms can compare information with other available facts and detect inconsistencies. This is especially useful for checking news that is shared on social media. Metadata analysis can be performed by establishing the publication time and examining the change history. Models can use metadata (time of creation, geographic location) to detect suspicious material. For example, fast-spreading news from new or unknown accounts can be filtered out as potentially fake. Some systems use AI to analyze text writing style and identify possible signs of automated content generation or bot use. Basic techniques used in NLP Tokenization Also called word segmentation, tokenization is one of the simplest and most important techniques (see Fig. 3). This is an important preprocessing step in which a long string of text is broken into smaller units called tokens. Tokens include words, symbols, and sub-words. They are the building blocks of NLP, and most NLP models process raw text at the token level. The most common tokenization process is the space/unigram. In this process, the entire text is broken into words by separating them with spaces.

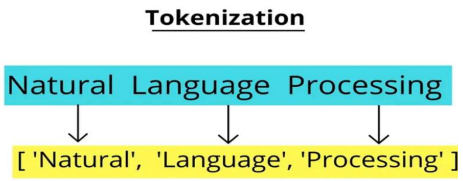

Figure 3: Tokenization

German Verarbeitung natürlicher Sprache is capital letters for nouns are mandatory and there are noun declensions according to 4 cases—the change occurs only in the ending of the adjective natürlicher de er ending in the genitive case of the feminine gender for Sprache, so the literal equivalent in German for Ukrainian and English yes

(processing natural language)

And for Polish Przetwarzanie języka naturalnego

(natural language processing)

In French Traitement du langage naturel

For Russian Processing of natural language

In Ukrainian Processing of natural language in the nominative and there are 5 more cases, so there are possible options for a stable keyword combination

Processing of natural language

Natural language processing

Etc

But for Tokanization will be

Natural language processing

Change of endings. Without conducting a preliminary morphological analysis based on the modified Potter algorithm, it is not possible for Ukrainian-language texts to correctly tokenize and lemmatize, as well as to determine the set of keywords in messages and news (see Fig. 4).

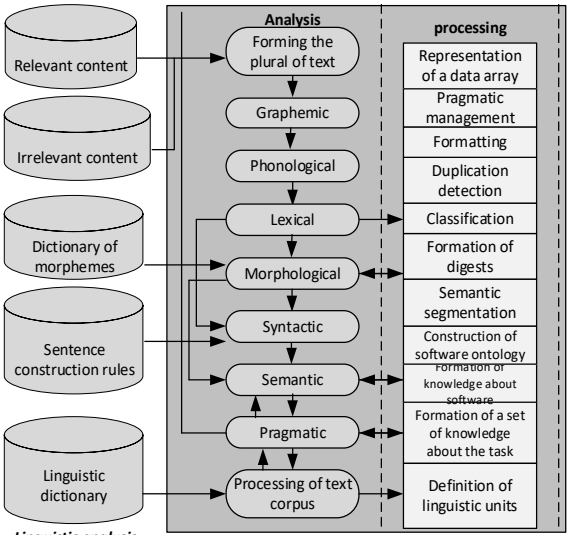

Figure 4: Classification of the main methods of natural language processing

A token is an atomic meaningful object from a sequence within [1, N] characters. Identifies tokens based on regular expressions and by location in character set/sentence and context. This is not grapheme analysis as separating a group of characters between punctuation marks. Tokens are identified by the rules of the lexer, taking into account already grammatical features from the previous step of MA, according to the natural language of the input text, in particular:

- Marking a set of incoming text characters into a set of tokens.
- Identification of a separate token as a logical linguistic unit of the text (word, mathematical sign, number, punctuation mark, etc.).
- Establishing a relationship between a token and a token—the specific text of the token (“for”, “1979”, “+”, “variable”, “.”, “p.”, “;”, etc.).
- Identification of additional token attributes (for example, a period as a sentence boundary or part of a contraction).
- Forming a tuple of tokens as input information for CA.

The lexical analyzer does not check the correctness of the links in the tuple of tokens. The parser recognizes parentheses, punctuation marks, and math symbols as characters, but does not check that each character “(” is matched by another “)”, and that each math character is between two specific numbers.

4.1. Stemming and lemmatization

After tokenization, the next preprocessing step is stemming, or lemmatization (Fig. 5). These methods generate a root word from the various existing variants of the word. Stemming and lemmatization [21, 22] are two different ways of trying to identify a root word. Creating roots works by removing the end of a word. This NLP technique may or may not work depending on the word. For example, this will work on “sticks” but not on “sticking” or “stuck”. Lemmatization is a more sophisticated technique that uses

morphological analysis to find the base form of a word, also called a lemma.

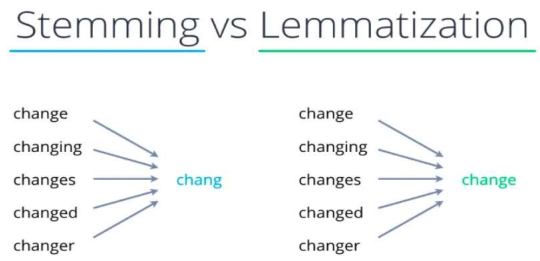

Figure 5: Stemming and lemmatization

4.2. Morphological segmentation

Morphological segmentation is the process of dividing words into morphemes that make them up. A morpheme is the smallest unit of language that carries meaning. Some words, such as "table" and "lamp", contain only one morpheme. But other words can contain several morphemes. For example: the word "energy saving" contains two morphemes: energy and conservation. Similar to stemming and lemmatization, morphological segmentation can help preprocess the input text [23, 24].

4.3. Morphological analysis

There are two types of POS tags in this case. Based on the rules of Stochastic POS Taggers Rule-based POS Tagger: For words with ambiguous meaning, a rule-based approach based on context information is applied [25]. This is done by checking or analyzing the meaning of the previous or next word. Information is analyzed from the word environment. Therefore, words are marked with the grammatical rules of a particular language, such as the use of capital letters and punctuation marks. If a word is most often marked with a certain tag in the training set, then the test sentence is assigned this specific tag. This method is not always accurate. Another way is to calculate the probability of a certain tag appearing in a sentence. Thus, the final tag is calculated by checking the maximum probability of a word with a given tag.

4.4. Sentiment analysis

Sentiment analysis, also known as emotion intelligence or opinion research, is the process of analyzing text to determine whether it is generally positive, negative, or neutral. As one of the most important NLP techniques for text classification, sentiment analysis is commonly used for applications such as user-generated content analysis. It can be used for a variety of text types, including reviews, comments, tweets, and articles [26, 27].

For example, the analysis and identification of psychological effects laid down by the author of the textual content depends on the availability of a personalized dictionary of the author and a sentiment dictionary of this region (not all words have the same emotional colors and in different languages and different regions, even different people of specific people—a simple translation will not help to get a real description of a person's psychological state). Statistical methods are used in content analysis to identify the state of social consciousness or emotional coloring to promote relevant political and/or commercial advertising in social networks.

In linguistic monitoring, in addition to the listed set of methods, regular expressions and a bag of words are used to study the functioning of language in a specific scientific, political, or mass media discourse. The purpose of monitoring is also recognition of fakes/propaganda and disinformation in the case of information threats, identification of foreign language borrowings, plagiarism/rewriting, grammatical/stylistic errors, vocabulary of emotions/feelings, thematic /spatial/ temporal vocabulary, etc.

5. Processes of machine learning

Machine learning methods have set new accuracy records in fields such as NLP [28, 29]. The success was facilitated by a large amount of training data and the availability of huge capacities for parallel calculations using modern graphics processors. Each search query in Google triggers several AI models at once, such as text recognition and personalization of the output of results. The spam detection system in Gmail works in the same way, identifying fraudulent messages (Fig. 6) [30, 31]. The method of detecting fake news is show in Fig. 7.

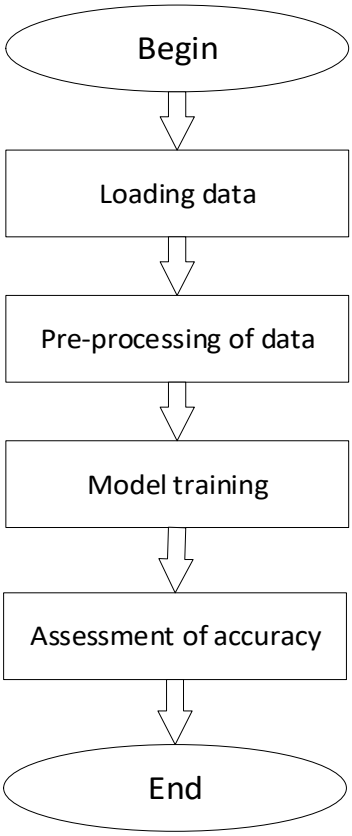

Figure 6: Processes of machine learning

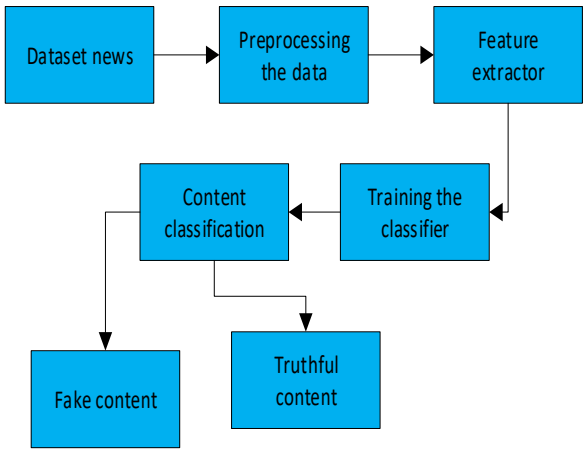

Figure 7: The method of detecting fake news

6. Experiments

For this study, a dataset was formed, which includes more than a thousand fake and real news. The dataset format is shown in Fig. 8. In this dataset, the news that the average Ukrainian saw during the war in the Internet space in such social networks as Telegram, Facebook, Twitter, and on news sites was formed.

A separate field was allocated to the language of the messages, which were in Ukrainian and Russian. In a separate field, it was noted how many people liked and how many people shared this message. The dataset contains part of fake news and part of true news. Well, for clarity, in the next field, we enter the author of the message and the web address of the site from where this news was read [32–34].

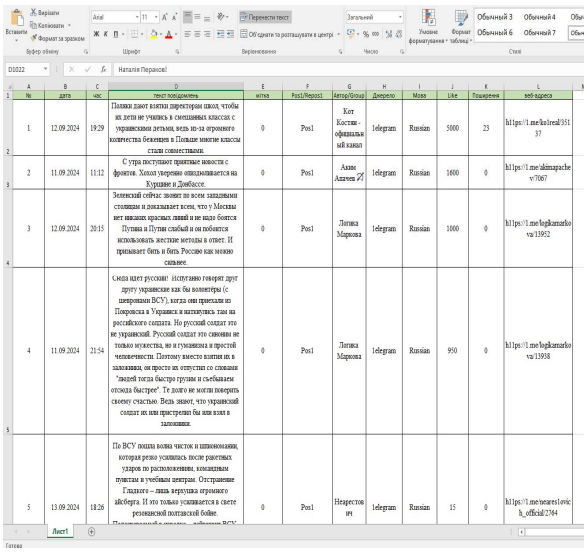

Figure 8: Format dataset

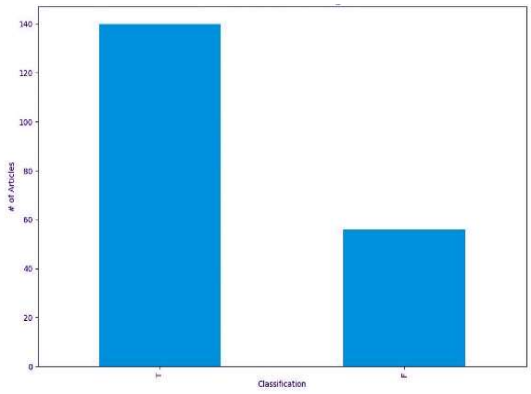

Figure 9: Analysis of fake and real news

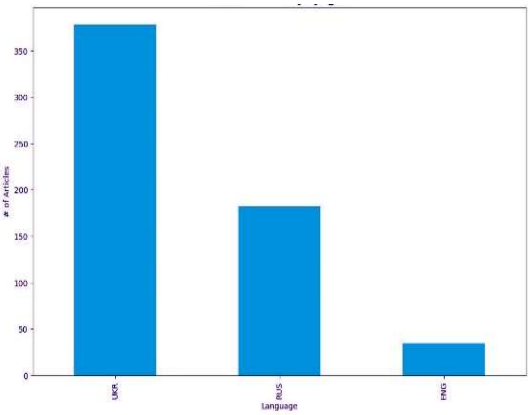

Figure 10: Analysis of Ukrainian, Russian, and English news

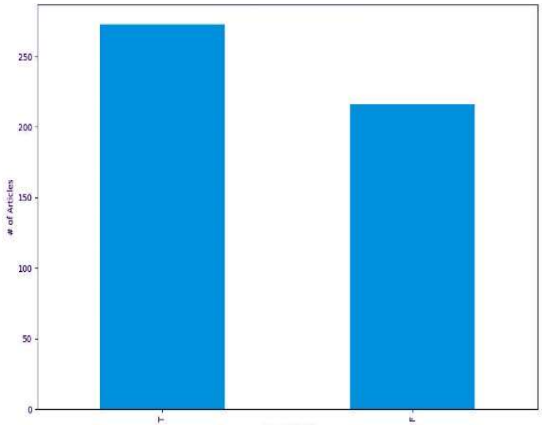

Figure 11: Classification Fake, True in the Telegram

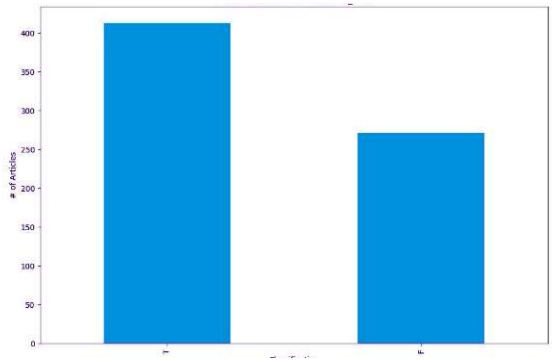

Figure 12: Classification Fake, True in the WWW

BOW and Logistic Regression functions were used for the forecast model. The results of the model are shown in Fig. 13.

```
print(classification_report(y_test, y_pred_lr))
```

	precision	recall	f1-score	support
0	0.98	0.99	0.98	2971
1	0.99	0.98	0.98	3029
accuracy			0.98	6000
macro avg	0.98	0.98	0.98	6000
weighted avg	0.98	0.98	0.98	6000

Figure 13: Model results

The F1 score is 0.98 for both classes (0-forgery, 1-not forgery). Such good results can be explained by the "laboratory" quality of the data set. In further experiments, we want to focus on validating the model on real-time news.

7. Conclusions

An analysis of attacks in the cyberspace of Ukraine was carried out. It is noted that for each attack it is necessary to form a countermeasure, which is expressed in the development of new software, new hardware, etc.

One of the most common threats is the penetration of false information in social networks and chatbots, and it is necessary to detect fakes and delete this type of news in every possible way.

A dataset of fake on real news has been created.

A program with machine learning was organized that would allow us to evaluate the current news as real or fake.

Acknowledgments

The research was carried out with the grant support of the National Research Fund of Ukraine "Information system development for automatic detection of misinformation sources and inauthentic behaviour of chat users", project registration number 187/0012 from 1/8/2024 (2023.04/0012). Also, we would like to thank the reviewers for their precise and concise recommendations that improved the presentation of the results obtained.

References

[1] O. Trofymenko, Monitoring the State of Cyber Security in Ukraine, Legal Life of Modern Ukraine: Mater. International Science and Practice Conference, 1 (2019) 642–646.

[2] O. Trofymenko, et al., Cybersecurity of Ukraine: Analysis of the Current State, Ukrainian Inf. Secur. Res. J. 21(3) (2019) 150–157.

[3] V. I. Yashchuk, The Role and Place of the Cyber Security Strategy of Ukraine in Ensuring the Information Security of yhe State (2024).

[4] Some Issues of Response by Cyber Security Entities to Various Types of Events in Cyberspace: Resolution of the Cabinet of Ministers of Ukraine dated (04.04.2023 No. 299).

[5] E. Altulaihan, M. A. Almaiah, A. Aljughaiman, Anomaly Detection IDS for Detecting DoS Attacks in IoT Networks Based on Machine Learning Algorithms, Sensors, 24(2) (2024) 713.

[6] M. A. Tamal, et al., Unveiling Suspicious Phishing Attacks: Enhancing Detection with an Optimal Feature Vectorization Algorithm and Supervised Machine Learning, Frontiers in Computer Science, 6 (2024) 1428013.

[7] A. Hannousse, S. Yahiouche, M. C. Nait-Hamoud, Twenty-Two Years Since Revealing Cross-Site Scripting Attacks: A Systematic Mapping and a Comprehensive Survey. Computer Science Review, 52 (2024) 100634.

[8] R. Alhamyani, M. Alshammari, Machine Learning-Driven Detection of Cross-Site Scripting Attacks, Information, 15(7) (2024) 420.

[9] R. A. Febrian, Y. Muhyidin, D. Singasatia, Analisis Penyerangan Bruteforce Terhadap Secure Shell (Ssh) Menggunakan Metode Penetration Testing, Scientica: Jurnal Ilmiah Sains dan Teknologi, 2(11) (2024) 151–162.

[10] M. A. Paranjape, S. Sathe, M. A. A. Abkari, Study On Awareness and Perceptions of Individual Investors Towards Cagr On Equity Shares, J. Econom. 17 (2024).

[11] O. Mykhaylova, et al., Person-of-Interest Detection on Mobile Forensics Data—AI-Driven Roadmap, in: Workshop on Cybersecurity Providing in Information and Telecommunication Systems, CPITS, vol. 3654 (2024) 239–251.

[12] V. Buhas, et al., Cybersecurity Role in AI-Powered Digital Marketing, in: Workshop on Digital Economy Concepts and Technologies Workshop, DECaT, vol. 3665 (2024) 1–11.

[13] V. Buhas, et al., AI-Driven Sentiment Analysis in Social Media Content, in: Workshop on Digital Economy Concepts and Technologies Workshop, DECaT, vol. 3665 (2024) 12–21.

[14] V. Zhebka, et al., Optimization of Machine Learning Method to Improve the Management Efficiency of Heterogeneous Telecommunication Network, in: Workshop on Cybersecurity Providing in Information and Telecommunication Systems, vol. 3288 (2022) 149–155.

[15] V. Zhebka, et al., Methodology for Predicting Failures in a Smart Home based on Machine Learning Methods, in: Workshop on Cybersecurity Providing in Information and Telecommunication Systems, CPITS, vol. 3654 (2024) 322–332.

[16] O. Romanovskyi, et al., Prototyping Methodology of End-to-End Speech Analytics Software, in: 4th International Workshop on Modern Machine Learning Technologies and Data Science, vol. 3312 (2022) 76–86.

[17] I. Iosifov, O. Iosifova, V. Sokolov, Sentence Segmentation from Unformatted Text using Language Modeling and Sequence Labeling Approaches, in: IEEE 7th International Scientific and Practical Conference Problems of Infocommunications. Science and Technology (2020) 335–337. doi: 10.1109/PICST51311.2020.9468084.

[18] A. Ghai, P. Kumar, S. Gupta, A Deep-Learning-based Image Forgery Detection Framework for Controlling

the Spread of Misinformation, Information Technology & People, 37(2) (2024) 966–997.
[19] E. R. Peterson, et al., The Impact from Galaxy Groups on Cosmological Measurements with Type Ia Supernovae, arXiv preprint arXiv:2408.14560 (2024).
[20] J. Shen, et al., Citekit: A Modular Toolkit for Large Language Model Citation Generation, arXiv (2024). doi: 10.48550/arXiv.2408.04662
[21] O. Toporkov, R. Agerri, Evaluating Shortest Edit Script Methods for Contextual Lemmatization, arXiv (2024). doi: 10.48550/arXiv.2403.16968.
[22] M. Medykovskyy, Methods of Protection Document Formed from Latent Element Located by Fractals, in: 10th Int. In Scient. and Techn. Conf. Comp. Sci. and Infor. Techn. (CSIT) (2015) 70–72. doi: 10.1109/STC-CSIT.2015.7325434.
[23] R. Groenendijk, L. Dorst, T. Gevers, HaarNet: Large-Scale Linear-Morphological Hybrid Network for RGB-D Semantic Segmentation, International Conference on Discrete Geometry and Mathematical Morphology (2024) 242–254.
[24] M. Nazarkevych, et al., Evaluation of the Effectiveness of Different Image Skeletonization Methods in Biometric Security Systems, Int. J. Sensors Wireless Commun. Control, 11(5) (2021) 542–552.
[25] V. Vysotska, et al., NLP Tool for Extracting Relevant Information from Criminal Reports or Fakes/Propaganda Content, in: IEEE 17th International Conference on Computer Sciences and Information Technologies (CSIT) (2022) 93–98.
[26] J. O. Krugmann, J. Hartmann, Sentiment Analysis in the Age of Generative AI, Customer Needs and Solutions, 11(1) (2024) 3.
[27] K. Alieksieieva, A. Berko, V. Vysotska, Technology of Commercial Web-Resource Processing, in: 13th International Conference: The Experience of Designing and Application of CAD Systems in Microelectronics, CADSM (2015).
[28] V. Hrytsyk, M. Nazarkevych, Real-Time Sensing, Reasoning and Adaptation for Computer Vision Systems, International Scientific Conference Intellectual Systems of Decision-making and Problems of Computational Intelligence, Proceedings (2022) 573–585.
[29] I. Tsmots, et al., Basic Components of Neuronetworks with Parallel Vertical Group Data Real-Time Processing, Advances in Intelligent Systems and Computing II: Selected Papers from the International Conference on Computer Science and Information Technologies, CSIT (2018) 558–576.
[30] I. Khomytska, V. Teslyuk, The Multifactor Method Applied for Authorship Attribution on the Phonological Level, In COLINS (2020) 189–198.
[31] I. Tsmots, et al., The Method and Simulation Model of Element Base Selection for Protection System Synthesis and Data Transmission, Int. J. Sensors Wireless Commun. Control, 11(5) (2021) 518–530.
[32] N. Pasieka, et al., Harmful Effects of Fake Social Media Accounts and Learning Platforms, in: Cybersecurity Providing in Information and Telecommunication Systems, vol. 2923 (2021) 258–271.
[33] N. Pasieka, et al., Lego Technology as a Means of Enhancing the Learning Activities of Junior High School Students in the Conditions of the New Ukrainian School, International Conference on Interactive Collaborative Learning (2022) 530–541.
[34] P. Skladannyi, et al., Improving the Security Policy of the Distance Learning System based on the Zero Trust Concept, in: Cybersecurity Providing in Information and Telecommunication Systems, vol. 3421 (2023) 97–106.