



Article

Social Media Zero-Day Attack Detection Using TensorFlow

Ahmet Ercan Topcu ^{1,*}, Yehia Ibrahim Alzoubi ², Ersin Elbasi ¹ and Emre Camalan ³

¹ College of Engineering and Technology, American University of the Middle East, Egaila 54200, Kuwait; ersin.elbasi@aum.edu.kw

² College of Business Administration, American University of the Middle East, Egaila 54200, Kuwait; yehia.alzoubi@aum.edu.kw

³ Computer Engineering Department, Ankara Yildirim Beyazıt University, Ankara 06010, Turkey; emre@camalan.net

* Correspondence: ahmet.topcu@aum.edu.kw

Abstract: In the current information era, knowledge can pose risks in the online realm. It is imperative to proactively recognize potential threats, as unforeseen dangers cannot be eliminated entirely. Often, malware exploits and other emerging hazards are only identified after they have occurred. These types of risks are referred to as zero-day attacks since no pre-existing anti-malware measures are available to mitigate them. Consequently, significant damages occur when vulnerabilities in systems are exploited. The effectiveness of security systems, such as IPS and IDS, relies heavily on the prompt and efficient response to emerging threats. Failure to address these issues promptly hinders the effectiveness of security system developers. The purpose of this study is to analyze data from the Twitter platform and deploy machine learning techniques, such as word categorization, to identify vulnerabilities and counteract zero-day attacks swiftly. TensorFlow was utilized to handle the processing and conversion of raw Twitter data, resulting in significant efficiency improvements. Moreover, we integrated the Natural Language Toolkit (NLTK) tool to extract targeted words in various languages. Our results indicate that we have achieved an 80% success rate in detecting zero-day attacks by using our tool. By utilizing publicly available information shared by individuals, relevant security providers can be promptly informed. This approach enables companies to patch vulnerabilities more quickly.

Keywords: zero-day attack; Twitter; TensorFlow; machine learning; word classification



Citation: Topcu, A.E.; Alzoubi, Y.I.; Elbasi, E.; Camalan, E. Social Media Zero-Day Attack Detection Using TensorFlow. *Electronics* **2023**, *12*, 3554. <https://doi.org/10.3390/electronics12173554>

Academic Editor: Ashwin Ashok

Received: 20 July 2023

Revised: 10 August 2023

Accepted: 18 August 2023

Published: 23 August 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In the modern era, the Internet has emerged as the most efficient, accessible, and cost-effective medium for international communication. Social networking platforms, such as Twitter, YouTube, and Facebook, have become essential tools for billions of individuals to connect, exchange information, and enjoy various content. Moreover, governmental organizations and businesses have recognized the potential of these online networks to reach a broader audience and carry out effective advertising campaigns [1]. As of April 2023, approximately 5.18 billion people, accounting for 64.6% of the global population, were using the Internet. Out of these Internet users, around 4.8 billion individuals, making up 59.9% of the world's population, actively engaged with social networking websites [2]. These statistics highlight the widespread adoption and popularity of the Internet and social media platforms globally.

Indeed, the widespread use of the Internet and social media platforms brings significant privacy and security concerns for individuals, data, systems, businesses, and governments. The sheer volume of online activity makes it challenging to monitor every single action. Specific individuals' behaviors can have harmful consequences for themselves or others. It is crucial to be aware of these risks and take proactive measures to protect personal information, maintain secure systems, and promote responsible online

behavior [3]. On the other hand, businesses and Internet service providers frequently collect sensitive information for advertising purposes and other uses. This behavior has led to growing concerns among Internet users regarding the security and confidentiality of their personal data. Research conducted on security and privacy in cyberspace indicates that most individuals believe it is primarily the responsibility of the government to ensure security on the Internet [4]. These findings highlight the need for stronger regulations and measures to protect user privacy and enhance overall cybersecurity in online environments [5]. While it is important for governments to establish regulations and frameworks to ensure online security, it is also the responsibility of individuals to take proactive measures to protect themselves online.

One of the most severe cyber threats is the zero-day attack, which remains unrecognized by both the general Internet community and experts [6]. Zero-day attacks, by their very nature, lack pre-existing fixes or protections because they exploit unknown vulnerabilities. These attacks can be leveraged to breach security measures, install malicious software, extract sensitive information, or disrupt system performance [7]. Due to their novelty and unpredictability, zero-day attacks pose significant challenges for defenders, as no known defense mechanisms or countermeasures are available, making them difficult to detect and defend against. This emphasizes the importance of ongoing monitoring, strong cybersecurity measures, and collaborative efforts among experts, researchers, and organizations to identify and address these emerging threats [8] effectively.

This study aims to employ machine learning techniques, such as word categorization using TensorFlow, on Twitter data to identify and counter zero-day vulnerabilities promptly. Consequently, the purpose of this study utilizes tweets sourced from the Twitter platform in an attempt to predict and forecast the occurrence of zero-day attacks. By analyzing the content of these tweets, researchers aim to uncover patterns, indicators, or discussions that could indicate zero-day vulnerabilities or impending attacks. The goal is to leverage social media data as a proactive tool for early detection and mitigation of zero-day attacks, thereby enhancing cybersecurity measures and reducing the associated risks. This paper makes several noteworthy contributions.

- It addresses the importance of ongoing efforts in developing intelligence software capable of detecting dangerous language and preventing potential attacks globally. The study highlights the relevance of applications like CyberTwitter, which aid in data collection and threat identification. However, it emphasizes relying on Open-Source Threat Intelligence (OSINT) platforms to achieve these goals. The researchers have developed codes specifically for data collection from Twitter, focusing on vulnerability assessment and utilizing it as a valuable OSINT source for identifying potential threats. Twitter was selected for this study due to its widespread social media use. The researchers have developed a code that analyzes tweets specifically for relevant terms related to threats.
- Our primary focus is heightened performance in identifying zero-day attacks on social media. We gained access to tweets from individuals who had shared zero-day attacks in the past. We harnessed NLTK tools to capture target words in various languages, integrating these with Tensorflow modules. NLTK was chosen for text analysis due to its robust capabilities in NLP, text mining, and language tasks. NLTK's preprocessing complements TensorFlow's models, improving accuracy. The NLTK and TensorFlow combination proves effective in various scenarios. This approach allowed us to compile all zero-day-related terms across different languages on Twitter.
- The code utilizes machine learning techniques to identify additional similar words or phrases associated with potential threats. By examining the content of tweets, the code aims to determine whether any tweets contain information or discussions pertaining to security threats. Moreover, this code can be enhanced by incorporating the ability to learn new phrases, even beyond the realm of the security context. This flexibility allows the program to adapt and recognize emerging patterns or language trends that may be relevant to identifying potential threats.

- Another crucial aspect of the code implemented in this research is its capability to identify the user's identity and location. This feature provides an added layer of information that can assist in analyzing and contextualizing the detected threats.

The remaining sections of the paper are structured as follows. Section 2 offers a background of the study, covering the zero-day attack, cybersecurity on Twitter, and related literature. Section 3 outlines the research methodology employed in this study. Section 4 presents the findings and results derived from the research. Section 5 delves into a discussion and analysis of the aforementioned findings. Finally, Section 6 concludes the paper by summarizing the key insights and potential future directions.

2. Research Background

2.1. Zero-Day Attack

Cybersecurity threats encompass a wide range of challenges, including both technical and non-technical attacks [9]. Technical attacks involve exploiting vulnerabilities in software, networks, or systems, whereas non-technical attacks, like Social Engineering, manipulate human psychology to deceive individuals into revealing sensitive information or performing certain actions [10]. To minimize cyber threats, a comprehensive approach is essential. Conducting workshops, seminars, and simulations can raise awareness among employees, teaching them to recognize and avoid social engineering tactics. Game-based education fosters a proactive cybersecurity mindset [11]. On the technical front, protecting against zero-day attacks requires robust intrusion detection and prevention systems, as well as timely software patching and updates [7]. A multilayered defense strategy, incorporating encryption, access controls, and regular security assessments, can bolster overall cyber resilience and safeguard against evolving threats.

A zero-day attack is a type of cyberattack that exploits previously unknown vulnerabilities in software or systems. It uses security flaws that software makers have not found or fixed, leaving users open to possible attacks [4]. The term "zero-day" indicates that attackers discover and exploit the vulnerability before the software vendor becomes aware of it, leaving zero days to prepare a defense. Zero-day attacks can have severe consequences since they might result in unauthorized access, breaches of data, the execution of malicious software, or the interruption of system operation. Zero-day attacks provide substantial issues requiring proactive security measures, ongoing monitoring, and quick software upgrades to fix new vulnerabilities [12]. Systems remain susceptible until the patch is deployed since it might take a while for software providers to create a patch to cure the vulnerability [6].

A number of tools, such as OSINT, and other methods for zero-day threat detection are available to examine abnormal activity. OSINT, a crucial component of efficient security intelligence processes, is a collection of signs pointing to heightened risk and highlighting certain threats [8]. These data are essential for assisting security officials in identifying possible risks and deciding what course of action to take next. The data obtained by OSINT can come from various publicly available sources, which is why it is described as open-source intelligence [13].

2.2. Cybersecurity on Twitter

Social media users often fall victim to security vulnerabilities by unknowingly exposing private data, images, or messages through their accounts and profiles. This is a significant contributing factor to the security flaws observed in these platforms. Both organizations and individuals increasingly rely on social media as a means of communication and find it convenient to target specific audiences [14]. Cybersecurity encompasses a set of principles, rules, techniques, recommendations, reliability, and technologies aimed at protecting the resources of companies and customers and the online environment [15]. These resources include personnel, infrastructure systems, facilities, network connections, and all the data transmitted and stored within the online environment. The primary objective

of cybersecurity is to establish and maintain effective safeguards against relevant security threats in the online realm [16].

The importance and global attention towards cybersecurity have significantly increased. Notably, an action document has been published, which features the official positions of over 50 countries on cyberspace, cybercrime, or cybersecurity. This demonstrates the growing recognition and efforts to address cybersecurity challenges internationally [17]. Multiple systems should be employed simultaneously to provide cybersecurity since the chances of an attack are growing as Internet usage grows. In the current climate, where attacks are growing exponentially more complicated, corporate cybersecurity is a notion that cannot be disregarded. There are several approaches to security intelligence, as described above. These can be terrifying, but tracking individuals is vital for the safety of governmental systems or communities when it comes to terrorism, threats, or cyberwarfare. Because of this, scientists and engineers are working to develop fresh approaches to intelligence technology [5].

Twitter is widely recognized as one of the most popular text-based social media platforms on the Internet. Its unique text-oriented format makes it particularly suitable for this study compared to other social platforms. Users can easily share personal information, news, meeting details, or relevant texts on Twitter. Therefore, it becomes a valuable platform to monitor and gather information about new zero-day threats. However, it is important to note that Twitter imposes certain restrictions on data collection. While users can access and retrieve information from Twitter using automated tools, like robots or crawlers, there are limitations in place to prevent the unrestricted collection of all users' information. These restrictions aim to safeguard user privacy and ensure responsible data usage [5]. By utilizing Twitter as a social media platform, this study can leverage its text-focused nature and the availability of publicly shared information to analyze and identify potential zero-day threats. The study acknowledges and adheres to Twitter's policies and guidelines concerning data collection and respects users' privacy while conducting the research.

Cyber-Twitter is a program designed for real-time social media research, specifically Twitter, to identify and assess risks and potentially hazardous communications. It functions as a system to discover and analyze cybersecurity intelligence on Twitter, acting as an OSINT source [8]. When the program is active, it collects data from Twitter and analyzes tweets to identify those that pertain to potential attacks or threats. This program allows for identifying, tagging, and extracting various real-world conceptual entities related to cybersecurity vulnerabilities. These entities may include the methods or means of an attack, the consequences of an attack, and the software, hardware, or vendors affected by the vulnerabilities. The program employs a security vulnerability concept extractor to perform these tasks [18].

2.3. Related Literature

Various research papers in the literature were identified, all of which centered on investigating zero-day attacks. Table 1 provides a summary of these studies, including their specific focus, and highlights the distinctions between this current study and the previously conducted ones. Altalhi and Gutub [19] examined Twitter data to identify and predict security threats. They compared various previously published papers that utilized Twitter streaming data to gather information on ongoing and potential cyberattacks. The investigation considered aspects such as detection scope, performance measurements, feature extraction methods, information summarization levels, algorithm complexity, and scalability over time. Several recommendations were proposed to improve the accuracy of forecasts. The results indicated that the SYNAPSE strategy achieved the highest summation value of 490 and an average score of 82, making it the preferred overall approach. On the other hand, the DataFreq scheme performed well, but could not match SYNAPSE in terms of average and total scores, positioning it as the second contender for enhancement [19].

Table 1. Summary of related work on zero-day attack detection.

Study	Focus	Method	Description
[19]	Predictions of cyberattacks	Survey	Comparing different proposed work against detection scope, performance measurements, feature extraction methods, information summarization levels, algorithm complexity, and scalability over time
[1]	Zero-day prediction	Autoencoder and deep anomaly detection	Test three datasets from the real world totaling 222,541 URLs
[20]	Zero-day detection	Deep learning technique	Developing an Intrusion IDS model with a high recall rate and minimal false negatives
[21]	Zero-day detection	tDCGAN	Generating synthetic malware and distinguishing it from real malware
[22]	Zero-day detection	Semi-supervised machine learning	Deploying Benford’s law that locates abnormal behavior based on the distribution of leading digits in numerical data
[23]	Zero-day detection	Neural Network classifier	Generating synthetic zero-day data and applying NN classifier to predict the zero-day attack
[24]	Zero-day detection	Zero-shot learning approach	Evaluating the effectiveness of machine learning-based IDSs in recognizing zero-day attack
[25]	Zero-day detection	Deep learning-based IDS	Using deep novelty-based classifiers and conventional clustering based on specialized layers of deep structures
[26]	Analogous zero-day detection	PlausMal-GAN	A malware training framework based on the generated analogous malware data using generative adversarial networks
[27]	Classify various ransomware variants	Multi-tier streaming analytics model	Numerically grouping ransomware variants into ancestor groups and statistically combining those from multiple-descendant families
This study	Zero-day detection	Tensorflow technique	Collecting and analyzing real data from the Twitter platform to detect potential zero-day attacks

Bu and Cho [1] presented a method for identifying zero-day attacks utilizing a convolutional autoencoder and deep character-level anomaly detection. They conducted rigorous tests using three datasets from the real world totaling 222,541 URLs, and their strategy outperformed other recent deep learning approaches. Evaluation of the receiver operating characteristic curve, tenfold cross-validation, and contrasts with deep learning techniques based on categorization was used to demonstrate the superiority of the suggested method. The findings showed that the suggested strategy improved sensitivity by 3.98% compared to classification-based deep learning approaches. The enhancement was credited to using an operation designed for the unique properties of URLs and installing a neural network structure optimized for URL modeling. The study highlighted the potential for boosting cybersecurity measures by demonstrating the efficacy of the deep character-level anomaly detection technique in detecting zero-day threats [1].

Hindy et al. [20] proposed using an autoencoder technique for detecting zero-day vulnerabilities. The objective was to develop an Intrusion Detection System (IDS) model with a high recall rate and minimal false negatives. To demonstrate the effectiveness of the model, its results were compared with those of a one-class Support Vector Machine (SVM). The study focused on assessing the one-class SVM’s performance when zero-day attacks deviated from expected behavior. The autoencoder’s encoding–decoding features proved to be highly beneficial for the proposed IDS model. The results of the study revealed that

autoencoders are capable of effectively identifying sophisticated zero-day attacks. The findings showed an accuracy range of 89–99% in zero-day detection, underscoring the efficacy of the approach [20].

Kim et al. [21] proposed a technique called transferred deep-convolutional generative adversarial network (tDCGAN), aiming to generate synthetic malware and effectively distinguish it from real malware. The method leverages actual data and data modified by tDCGAN through a deep autoencoder, which extracts essential features and enhances GAN training stability. The deep autoencoder learns malware characteristics, generates generic data, and passes this knowledge to facilitate reliable GAN training. Through transfer learning, the trained discriminator imparts its ability to recognize malware traits to the detection system. As a result, tDCGAN achieved an average classification accuracy of 95.74%, contributing to increased learning stability in malware detection [21].

Mbona and Eloff [22] presented a method for employing semi-supervised machine learning to detect zero-day attacks. The use of Benford's law, a statistical theory that locates abnormal behavior based on the distribution of leading digits in numerical data, is part of their approach. The study reveals that this method may successfully recognize key network characteristics that point to aberrant behavior, helping to identify zero-day attacks. Their findings emphasize using semi-supervised machine learning techniques for making the best choice of pertinent attributes. According to experimental findings, one-class SVMs performed best in identifying zero-day attacks, with a 74% Matthews correlation coefficient and an 85% F1 score. These results highlight the potential of semi-supervised machine learning approaches for identifying zero-day attacks, mainly when used in conjunction with suitable feature selection techniques [22].

Peppes and Alexakis [23] presented an approach involving the generation of authentic zero-day-type data in tabular format and evaluating a neural network trained with and without synthetic data to detect zero-day attacks. By employing Generative Adversarial Networks (GANs), they successfully created a larger dataset of synthetic information on zero-day exploits. This dataset was used to train a Neural Network classifier specifically designed for identifying zero-day attacks, and its performance was evaluated. The results showed that after approximately 5000 iterations, the synthetic zero-day attack data in tabular format reached a state of equilibrium, producing data that closely resembled the original data samples.

To evaluate the effectiveness of machine learning-based IDSs in recognizing zero-day attack scenarios, Sarhan et al. [24] developed a unique zero-shot learning approach. The learning models translate data characteristics to semantic attributes that distinguish between known attacks and benign activity during the attribute learning step. To identify zero-day attacks as malicious, the models create connections between known and zero-day attacks during the deductive phase. The efficiency of the learning model in identifying unknown attacks is measured by a new assessment metric called Zero-day Detection Rate. Two machine learning models, as well as two current IDS datasets, were used to assess the proposed system. The findings show that several of the study's identified zero-day attack groups provide problems for ML-based IDSs since they are difficult to identify as malicious. Further investigation revealed that these assaults with low-zero-day detection percentages have distinctive characteristic distributions and a wider Wasserstein distance than assaults in other assault classes. These results show the need to consider certain feature distributions to overcome such obstacles and illustrate the shortcomings of ML-based IDSs in successfully identifying specific zero-day attack situations [24].

Soltani et al. [25] proposed a thorough IDS system that uses deep learning methods to counter new assaults efficiently. This system stands out as the first of its type to use both deep novelty-based classifiers and conventional clustering based on specialized layers of deep structures. The study presented DOC++, an improved version of DOC that acts as a deep novelty-based classifier. In the preprocessing stage, the Deep Intrusion Detection framework was used to improve the capacity of deep learning algorithms to recognize content-based assaults. Four distinct algorithms—DOC, DOC++, OpenMax,

and AutoSVM—were contrasted as the framework’s novel detectors. According to the results, the open set identification module was most successfully implemented by DOC++. Furthermore, the clustering and post-training stages validated the model’s applicability for supervised labeling and updating procedures, which showed good levels of thoroughness and uniformity. The results highlight the proposed framework’s robustness and effectiveness in handling novel attack situations, confirming its value as a deep learning-based IDS strategy [25].

Won et al. [26] developed a malware training methodology called PlausMal-GAN, which utilizes generative adversarial networks to create similar malware data. They used a combination of authentic and artificially generated malware images to train the discriminator, which acts as a detector, to recognize various virus characteristics. The proposed approach outperformed other methods, especially for equivalent zero-day malware images, which are considered analogous zero-day malware data. Moreover, the architecture offers significant advantages for antivirus systems, as it does not require time-consuming malware signature evaluation. This highlights the potential of PlausMal-GAN as an effective and efficient tool for enhancing malware detection capabilities [26].

Zuhair et al. [27] suggested a multi-tiered streaming analytics technique known as a hybrid machine learner, which utilizes 24 static and dynamic features to classify various ransomware variants belonging to 14 different families. The suggested methodology involves numerically grouping ransomware variants into ancestor groups and statistically combining those from multiple-descendant families. To evaluate the effectiveness of the approach, the methodology was applied to categorize ransomware variants among a dataset consisting of 40,000 samples, including malicious software, goodware, and different variations of ransomware, in both semi-realistic and realistic scenarios. In a realistic comparison test, the ransomware streaming analytics model demonstrated an average classification accuracy of 97%, an error rate of 2.4%, and a miss rate of 0.34%. These results indicate that the proposed model outperforms competing anti-ransomware technologies in terms of performance and accuracy [27].

This study stands out from previous research in the field due to its unique approach and data source. Unlike many previous studies that might have relied on traditional datasets or network traffic logs, this study utilizes data from the Twitter platform for its analysis. By tapping into this rich source of user-generated content, the study gains valuable insights into real-time discussions, trends, and potential zero-day attack indicators that might not have been captured through conventional means. Furthermore, the study employs the TensorFlow technique to process the vast amount of Twitter data effectively. This sets this study apart from others that might have utilized more manual or rule-based approaches, making it potentially more efficient and scalable in detecting and responding to emerging threats.

3. Research Method

Figure 1 illustrates the research methodology and showcases the inclusion of a library and default systems. The fundamental premise of this study is that individuals who discover vulnerabilities and zero-day risks often attempt to share them with others online, particularly through social media platforms. The research aims to identify these vulnerabilities early on and raise awareness about them by promptly addressing concerns. Given the widespread usage of Twitter as a popular social media platform, it is chosen as the open data environment for this study. The primary objective of this research is to expand the collection of data from social media, particularly Twitter, and analyze them using techniques such as neural language processing and word categorization learning. By employing these methods, the study aims to gain insights and understanding from the social media data, specifically focusing on vulnerabilities, risks, and related discussions.

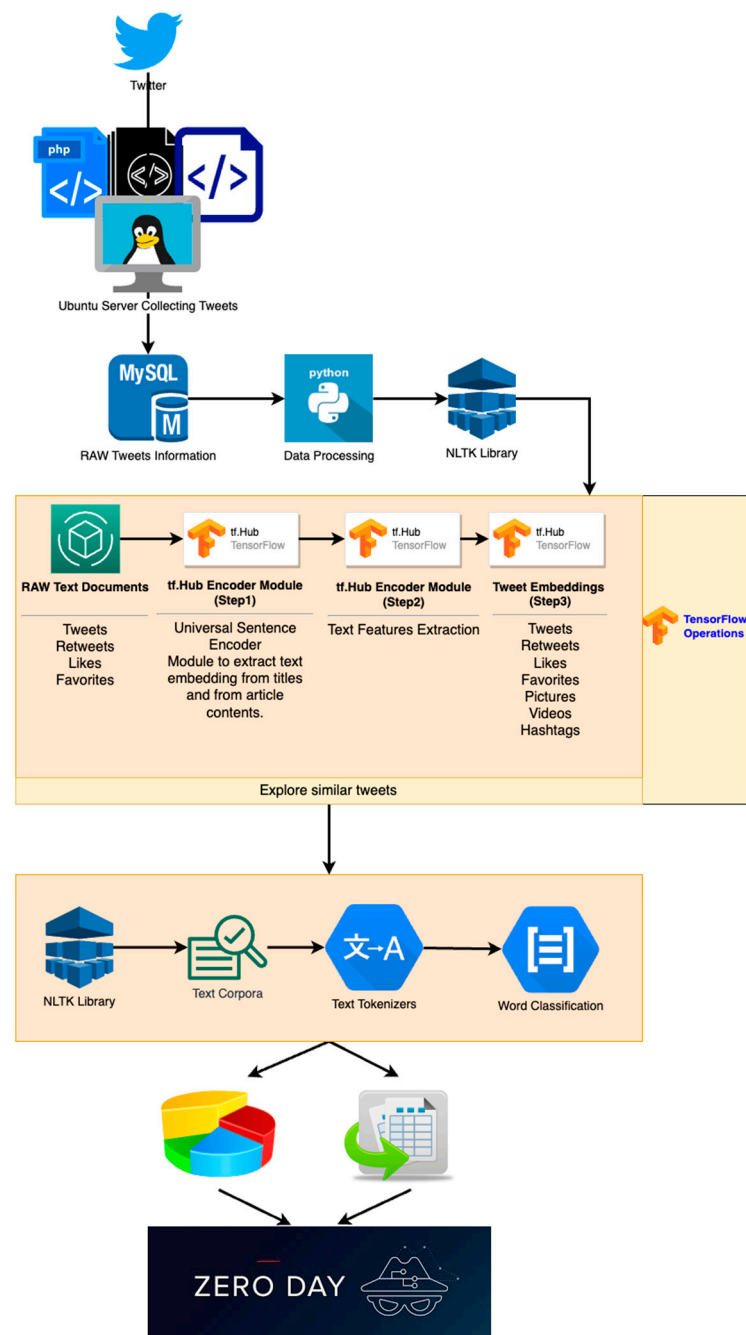


Figure 1. Research methodology ([21]).

The system encompasses a range of codes that require coordination, and as a result, the study incorporates a user-friendly Web UI for easy application. The Web UI allows users to input a parameter through an input box, initiating the serialization of the parameter, including figures and codes. When a user enters a term, such as “zero-day”, and clicks the “search” button, the PHP and Python programs in the Web UI invoke system processes. These scripts operate on the Ubuntu operating system and are capable of fetching specific tweets and storing them in MariaDB. The TensorFlow libraries are employed to process the raw text tweets obtained from Twitter, enabling the reading and parsing of the tweets. The text analysis procedures make use of powerful text analytics, Natural Language Processing (NLP), and the text mining toolkit Natural Language Toolkit (NLTK), which provides a diverse range of natural language algorithms.

3.1. Python Programming Language

Python is renowned for its clarity and flexibility. Python is simple to learn and use for both novice and seasoned developers due to its emphasis on clear and concise code [28]. Building websites, analyzing data, computation in science, AI, and automation are just a few of the fields where Python is extensively used. Its enormous features are a result of its extensive core library and support for a variety of third-party libraries and frameworks. Python's significant advantages are flexible typing, automated handling of memory, and support for various programming models. Python is a great option for machine learning applications because of its simplicity, plenty of libraries, robust community support, and adaptability [29]. It makes the development process more straightforward, makes data manipulation easier, and allows developers to use cutting-edge techniques and instruments, enabling them to create robust and scalable machine learning models [29].

3.2. TensorFlow Machine Learning Library

TensorFlow is a widely used open-source machine learning platform that can be described as a library utilized in both research and production settings. It offers a range of APIs suitable for desktop, mobile, online, and cloud development, catering to users with varying levels of expertise, from beginners to advanced practitioners. One of the key strengths of TensorFlow is its ability to encapsulate various deep learning (or neural networking) models and techniques, making them easily accessible and useful. It provides a high-performance execution environment in Python, allowing developers to create applications using the framework through a user-friendly frontend API, C++ [30].

This study uses the TensorFlow library to read and parse unprocessed tweets collected from Twitter. To accomplish this, the study employs the text embedding package called `tf.Hub`, which offers convenient and flexible methods for handling raw Twitter data. Additionally, the `tf.Transform` package, specifically the `preprocess_fn` function, is employed, which operates within a hidden TensorFlow context. This context allows for the execution of various TensorFlow operations, including invoking `tf.Hub` modules, without requiring intricate procedures or steps [30]. By leveraging these functionalities, the study achieves simplicity and efficiency in processing and transforming the raw Twitter data.

TensorFlow utilizes vector representations to learn from and compare tweets, transforming each analyzed tweet into a vector. In the case of raw data from Twitter, such as tweets used for tasks like object recognition or speech recognition, all the necessary information is stored within the data themselves. However, in NLP systems, words are typically represented by distinct atomic symbols. For example, "cat" might be represented as `Id537`, while "dog" could be represented as `Id143`. These encodings are arbitrary and do not provide the system with any meaningful knowledge about potential connections between different symbols.

Representing words as distinct, discrete entities can lead to a lack of data and pose challenges in effectively training mathematical models. To overcome these challenges, vector representations are often employed. In TensorFlow, neural probabilistic language models are commonly trained using the maximum likelihood method, as shown in Equation 1 [30], to predict the next word based on the highest degree of similarity.

In this context, "H" represents historic words, while " W_t " represents the target word. By leveraging vector representations, TensorFlow language models can learn from the contextual information encoded in the historic words and make predictions for the target word. This approach allows for more effective language modeling and helps address some of the limitations of discrete word representations.

$$P(w_t|h) = \text{softmax}(\text{score}(w_t, h)) = \frac{\exp\{\text{score}(w_t, h)\}}{\sum_{\text{Word } w' \text{ in Vocab}} \exp\{\text{score}(w', h)\}} \quad (1)$$

A skip-gram model can be employed to address the computational cost associated with calculating and normalizing probabilities for language modeling. Compared to the SoftMax

classifier, the skip-gram model offers faster processing. Here is an example illustrating how the skip-gram model works:

Example Tweet: “Security researchers have found a zero-day vulnerability in a popular building on KDE Desktop”

Normalization Context: ([Security, have], researchers), ([researchers, found], have), ([have, found], zero-day), ...

In the skip-gram model, a new dataset is created by forming pairs from the normalization context:

New Dataset: (researchers, security), (researchers, have), (have, researchers), (have, found), (zero-day, have), (zero-day, found), ...

By generating such pairs, the skip-gram model can efficiently capture the relationships between words in a given context. This approach offers a more computationally efficient alternative to the SoftMax classifier, allowing for faster processing and training of language models.

In TensorFlow, vectors can have various dimensions, representing different aspects or categories. For example, dimensions can be used to capture relationships, such as male–female, verb–tense, or country–capital. In the context of this study, vector dimensionality is exemplified by terms related to zero-day vulnerability, including variations like ZeroDay-Vulnerability, oday-vulnerability, 0 day-vulnerability, zero-day-vulnerability, and zeroday-vulnerability. These vectors, generated through TensorFlow, capture the essential information of the tweet in a numerical format, facilitating further analysis and processing within the framework of the study. Figure 2 displays the TensorFlow vectors associated with the tweet, representing the numerical representations derived from the underlying text.

```
ecamalan@ubuntu:/mnt/hgfs/Shared/twitter2$ python3 tensorflow.py
[["the", "popular", "#steam", "game", "client", "for", "#windows", "has", "a", "#zeroday", "privilege", "escalation", "vulnerability", "that", "can", "allow", "an", "attacker", "with", "limited", "permissions", "to", "run", "a", "program", "as", "an", "administrator"], ["", "#cybersecurity", "#infosec", "#cyberthreats", "#cyberdefense", "#cybersecurity", "https://www", "bleepingcomp"]]
[["the", "popular"], ["the", "#steam"], ["popular", "the"], ["popular", "#steam"], ["popular", "game"], ["#steam", "the"], ["#steam", "popular"], ["#steam", "game"], ["#steam", "client"], ["game", "popular"], ["game", "#steam"], ["game", "client"], ["game", "for"], ["client", "#steam"], ["client", "game"], ["client", "for"], ["client", "#windows"], ["for", "game"], ["for", "client"], ["for", "#windows"], ["for", "has"], ["#windows", "client"], ["#windows", "for"], ["#windows", "has"], ["#windows", "a"], ["has", "for"], ["has", "#windows"], ["has", "a"], ["has", "#zeroday"], ["a", "#windows"], ["a", "has"], ["a", "#zeroday"], ["a", "privilege"], ["#zeroday", "has"], ["#zeroday", "a"], ["#zeroday", "privilege"], ["#zeroday", "escalation"], ["privilege", "a"], ["privilege", "#zeroday"], ["privilege", "escalation"], ["privilege", "vulnerability"], ["escalation", "#zeroday"], ["escalation", "privilege"], ["escalation", "vulnerability"], ["escalation", "that"], ["vulnerability", "privilege"], ["vulnerability", "escalation"], ["vulnerability", "that"], ["vulnerability", "can"], ["that", "escalation"], ["that", "vulnerability"], ["that", "can"], ["that", "allow"], ["can", "vulnerability"], ["can", "that"], ["can", "allow"], ["can", "an"], ["allow", "that"], ["allow", "can"], ["allow", "an"], ["allow", "attacker"], ["an", "can"], ["an", "allow"], ["an", "attacker"], ["an", "with"], ["attacker", "allow"], ["attacker", "an"], ["attacker", "with"], ["attacker", "limited"], ["with", "an"], ["with", "attacker"], ["with", "limited"], ["with", "permissions"], ["limited", "attacker"], ["limited", "with"], ["limited", "permissions"], ["limited", "to"], ["permissions", "with"], ["permissions", "limited"], ["permissions", "to"], ["permissions", "run"], ["to", "limited"], ["to", "permissions"], ["to", "run"], ["to", "a"], ["run", "permissions"], ["run", "to"], ["run", "a"], ["run", "program"], ["a", "to"], ["a", "run"], ["a", "program"], ["a", "as"], ["program", "run"], ["program", "a"], ["program", "as"], ["program", "an"], ["as", "a"], ["as", "program"], ["as", "an"], ["as", "administrator"], ["an", "program"], ["an", "as"], ["an", "administrator"], ["administrator", "as"], ["administrator", "an"], ["", "#cybersecurity"], ["", "#infosec"], ["#cybersecurity", "#infosec"], ["#cybersecurity", "#cyberthreats"], ["#infosec", "#cyberdefense"], ["#cyberthreats", "#cybersecurity"], ["#cyberthreats", "#infosec"], ["#cyberthreats", "#cyberdefense"], ["#cyberthreats", "#cybersecurity"], ["#cyberdefense", "#infosec"], ["#cyberdefense", "#cyberthreats"], ["#cyberdefense", "#cybersecurity"], ["#cyberdefense", "https://www"], ["#cybersecurity", "#cyberthreats"], ["#cybersecurity", "#cyberdefense"], ["#cybersecurity", "https://www"], ["#cyberdefense", "https://www"], ["https://www", "#cybersecurity"]]
[["the", "popular", "#steam", "game", "client", "for", "#windows", "has", "a", "#zeroday", "privilege", "escalation", "vulnerability", "that", "can", "allow", "an", "attacker", "with", "limited", "permissions", "to", "run", "a", "program", "as", "an", "administrator"], ["", "#cybersecurity", "#infosec", "#cyberthreats", "#cyberdefense", "#cybersecurity", "https://www"], ["bleepingcomp"]]
```

Figure 2. The TensorFlow vector results.

3.3. NLTK Toolkit

NLTK is widely recognized as a leading Python programming environment for handling human language data. It provides a comprehensive set of text processing libraries encompassing functionalities such as categorization, tokenization, stemming, tagging, parsing, and semantic reasoning. Additionally, NLTK offers convenient wrappers for powerful NLP libraries widely used in the industry. One of the key strengths of NLTK is its extensive collection of resources, including over 50 corpora and lexical resources, such as

WordNet. These resources enable researchers and developers to access a vast array of linguistic data and lexical information for analysis and processing tasks. Furthermore, NLTK provides a platform for active discussion and community support through its discussion forum [31]. The choice of NLTK as the library for text analysis operations in this study is justified by its robust capabilities in text analytics, NLP, and text mining. Its broad range of natural language techniques and user-friendly interfaces make it a powerful tool for various language-related tasks. NLTK offers a wide range of popular algorithms, including tokenization, sentiment analysis, part-of-speech tagging, stemming, topic division, and named entity identification. These algorithms play a crucial role in analyzing tweets and enabling the evaluation, preparation, and comprehension of written language from raw Twitter data. When processing raw Twitter data using NLTK, it is important to consider that tweets often contain multiple languages and symbols. To illustrate a practical NLTK example for Twitter processing, let us consider the process depicted in Figure 3.

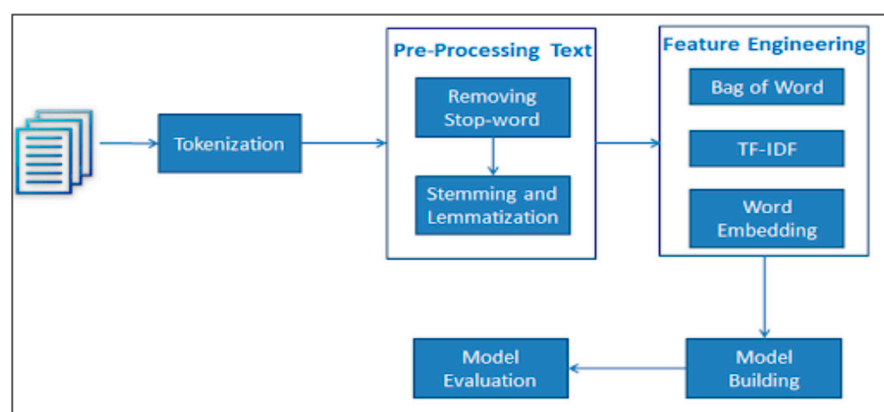


Figure 3. Word classification using NLTK.

The NLTK Tokenizer module is applied to the raw data, allowing it to be broken down into individual tokens. These tokenized data are then translated into English to facilitate further analysis and processing, as depicted in Figure 4. Lastly, word classification is performed as a final step in the NLTK pipeline. Given that there are numerous tweets and that the majority of them are similar, tweet categorization is one of the most crucial markers for zero-day vulnerability research tools. This classification step involves categorizing words based on their semantic or grammatical characteristics, allowing for a deeper understanding and interpretation of the text [32].

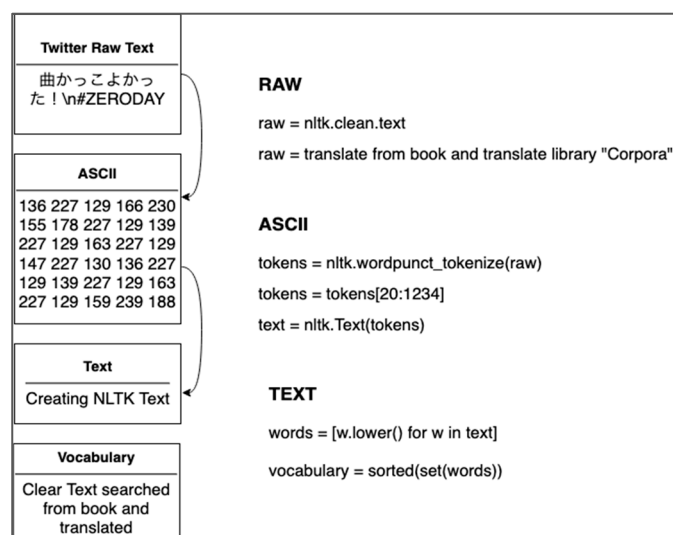


Figure 4. Text cleaning using NLTK.

4. Experimental Results

Figure 5 depicts a scenario where Jonathan tweets about discovering a tweet mentioning the phrase “zero-day” and stating that the message had been present for over 90 days. However, the actual tweet referred to is not provided in the given information. Nonetheless, with the help of the described tool, it becomes possible to identify both the specific word mentioned (“zero-day”) and the Twitter account that shared the tweet. This example highlights the notion that if someone claims to have discovered a vulnerability in an application like Zoom and that vulnerability still poses a potential threat, then, from the perspective of Twitter, it can be considered a zero-day vulnerability. The tool described in the study enables the identification of such instances, aiding in the detection and analysis of relevant information on social media platforms like Twitter.



Figure 5. An example of a zero-day tweet.

4.1. Code Developed

The research tool described in this study is designed to be accessible from any browser, ensuring compatibility across a wide range of browser types. The tool’s frontend web UI is implemented using HTML, PHP, jQuery, and Bootstrap technologies. In Figure 6, it is demonstrated that specific code components were developed to ensure compatibility specifically with the Microsoft Edge browser. The second essential component of the system, depicted in Figure 7, consists of PHP, Bash, and Python scripts. These scripts play a crucial role in the tool’s functionality, as they are responsible for tasks such as searching, gathering, computing, and locating zero-day data on Twitter. By leveraging these programs, the tool can effectively extract and process relevant information related to zero-day vulnerabilities. Collectively, the frontend web UI and the backend scripts form a comprehensive system that enables users to utilize the research tool across various browsers, facilitating the exploration and analysis of zero-day data on Twitter.

To retrieve a more significant number of tweets for analysis, the crawler in this study follows a specific process. Firstly, it utilizes the search bar to look up new terms, initiating the search for relevant tweets. However, Twitter’s search results typically display only the top 10 tweets, along with the best hashtag, images, and videos.

To overcome this limitation and gather a more extensive dataset, the crawler simulates human-like behavior by navigating down the page, mimicking human scrolling. This process allows the robots to access additional pages of tweets beyond the initially displayed ones. In this study, the robots collect and store all the relevant tweets found on five pages, ensuring a larger sample size for analysis.

The NLTK Library and TensorFlow Library play essential roles in this process. These libraries are utilized to extract and process the gathered tweets, specifically focusing on identifying and learning phrases associated with zero-day vulnerabilities. Figure 1 illustrates the integration of NLTK and TensorFlow libraries within the research framework, highlighting their significance in tweet collection and phrase learning.

```

<!DOCTYPE html>
<html lang="en">
<?php header('Content-Type: charset=utf-8'); ?>
<?php require_once("../inc/connect.php"); ?>
<?php
$id = $_GET['id'];

if($id == 1) {
    $id = 0;

    mysqli_query($conn, "TRUNCATE TABLE `twits`");
    header("Location: index.php");
}
?>
<head>

<meta charset="utf-8">
<meta http-equiv="X-UA-Compatible" content="IE=edge">
<meta name="viewport" content="width=device-width, initial-scale=1">
<meta name="description" content="">
<meta name="author" content="ubuntu" >
<meta http-equiv="Content-Type" content="text/html; charset=utf-8" />

<title>INVESTIGATING CYBER THREATS ON SOCIAL MEDIA USING MACHINE LEARNING</title>

<!-- Bootstrap Core CSS -->
<link href="../vendor/bootstrap/css/bootstrap.min.css" rel="stylesheet">

<!-- MetisMenu CSS -->
<link href="../vendor/metisMenu/metisMenu.min.css" rel="stylesheet">

```

Figure 6. Extracting data from Web UI.

```

for _ in range(1):
    driver.refresh();
    time.sleep(3)
    for x in range(NUM):
        driver.execute_script("window.scrollTo(0, document.body.scrollHeight);")
        time.sleep(3)
        tweets = driver.find_elements_by_class_name('original-tweet')

        for tweet in tweets:
            usernames = tweet.find_elements_by_class_name('username')
            user = usernames[0].text.split('@')[1].encode('utf-8')
            user = user.decode('utf-8')

            dates = tweet.find_elements_by_class_name('js-short-timestamp')
            date = str(dates[0].text.encode('utf-8'))
            date = date.decode('utf-8')

            twits = tweet.find_elements_by_class_name('TweetTextSize')
            twit = twits[0].text.encode('utf-8')
            twit = twit.decode('utf-8')

```

(a) Scrolling page code

Figure 7. Cont.


```

<div class="navbar-default sidebar" role="navigation">
  <div class="sidebar-nav navbar-collapse">

    <ul class="nav" id="side-menu">

      <li class="sidebar-search">
        <form class="form-horizontal" role="form" method="post" action="searchword.php">
          <div class="input-group custom-search-form">
            <input type="text" class="form-control" name="formcontrol" placeholder="Search...">
            <span class="input-group-btn">
              <input id="submit" name="submit" type="submit" value="Send" class="btn btn-primary">
              <!--
              <button class="btn btn-default" type="button" id="submit" value="Send">
                <i class="fa fa-search"></i>
              </button>
              <!--
            </span>
          </div>
        </form>
        <!-- /input-group -->
      </li>
      <li>
        <a href="index.php"><i class="fa fa-dashboard fa-fw"></i> Dashboard</a>
      </li>
    </ul>
  </div>

```

(b) Code for the input parameter in HTML and PHP

```

<?php
$formcontrol = $_POST["formcontrol"];

$old_path = getcwd();
chdir('/mnt/hgfs/Shared/twitter2/');
// $command = escapeshellcmd('python tSearchAndInsert.py $formcontrol 2 > log.txt 2>&1');
// $output = shell_exec($command);

$cmd = "python tSearchAndInsert.py ".$formcontrol." 50 > log.txt 2>&1";
shell_exec($cmd);

chdir($old_path);

header("Location: index.php?id=$formcontrol");
?>

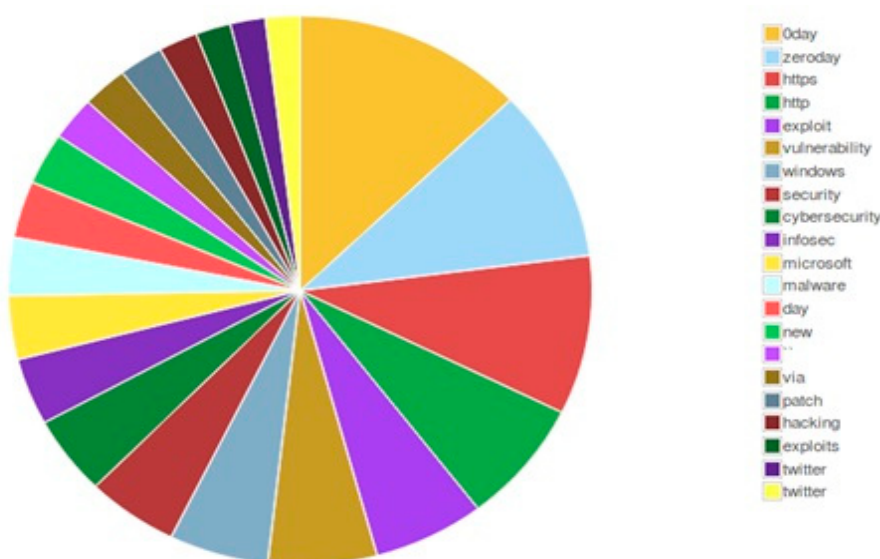
```

(c) Code for word parameter for search HTML & PHP

Figure 7. Samples of the codes developed in this study.

4.2. Searched Words

The terms “0-day”, “0 day”, “zero-day”, and “zeroday” are specific search terms used to investigate zero-day vulnerabilities. These terms hold significance both in the context of examining zero-day vulnerabilities and within the TensorFlow machine learning framework. Upon conducting the search using these terms, a substantial number of tweets, specifically, over 4500, were retrieved as results. Figure 8 presents the classification of these tweets, showcasing how they are categorized or grouped based on specific criteria or characteristics. This classification provides insights into the nature and distribution of the tweets related to zero-day vulnerabilities within the collected dataset. By organizing and categorizing the tweets, the study gains a better understanding of the content, trends, and patterns associated with discussions surrounding zero-day vulnerabilities. This information can contribute to further analysis and research on the subject matter.



(a) Categorized searched words

id	Search Word	User Name	Tweet	Date	Reply	ReTweet	Favori	wc	assurance	Tensorflow
1	vulnerability	Bitdefender	#Bitdefender researchers found a new security vulnerability affecting all modern Intel CPUs. Here's everything you need to know about #SWAPGS and how to stay protected.	Aug 7	5	31	94	25	130	1
2	ZeroDay	mirko_ross	#Blackhat market will raise when #Apple put the #bounty for #ZeroDay #Exploits to + USD 1 Mio. https://buff.ly/2Z1Mwbqu cc @DrJDrooghaag @KaiGrunwitz @HermesenKai @avrohomg	Aug 23	0	3	5	24	8	2
3	Oday	patrickwardle	As a developer, I struggle with Apple's new "security" enhancements as they often break security tools As a hacker, they often don't make any difference ex: 's secure "User-Approved Kext Loading" Bypassed (10.14.3) for the 3rd time: https://vimeo.com/325752606 #Oday	Mar 21	2	123	313	40	438	3
4	zero-day	ImNotTheWolf	Critical Security Update: Coinbase security team discovers Zero-Day exploit on Firefox (https://lunardigitalassets.com/news/critical-security-update-coinbase-security-team-discovers-zero-day-exploit-on-firefox/?fbclid=IwAR31veubeQQNZA-Mfd6lqVSw57ZakPuX7-XqIDy-kOsT9vgStlUd)	Jun 19	2	10	15	46	27	3

(b) Frontend of Web UI

Figure 8. Categorization of searched words (accessed on 19 July 2023).

4.3. Performance Evaluation

To collect data from Twitter, this study implemented a crawling procedure that emulates human behavior. Robots were programmed to open web browsers in a shadow mode, navigate through Twitter, extract relevant data, and store it in a database. The procedure considers the scrolling behavior commonly observed in humans while browsing Twitter. Unlike the scrolling limitation imposed by the scroll-down count on Twitter's interface, the robots in this study can scroll as much as needed. During the data gathering process, the response time of Twitter plays a critical role. Figure 9 visually represents the dependency of the model on Twitter's response time. The graph illustrates that the model's performance is influenced by the time it takes for Twitter to respond to requests and provide the necessary data. Understanding and accounting for the response time factor is crucial for optimizing

the crawling procedure and ensuring efficient data collection from Twitter. It enables researchers to manage and adapt their approach accordingly to mitigate any delays or limitations imposed by the platform's response time.

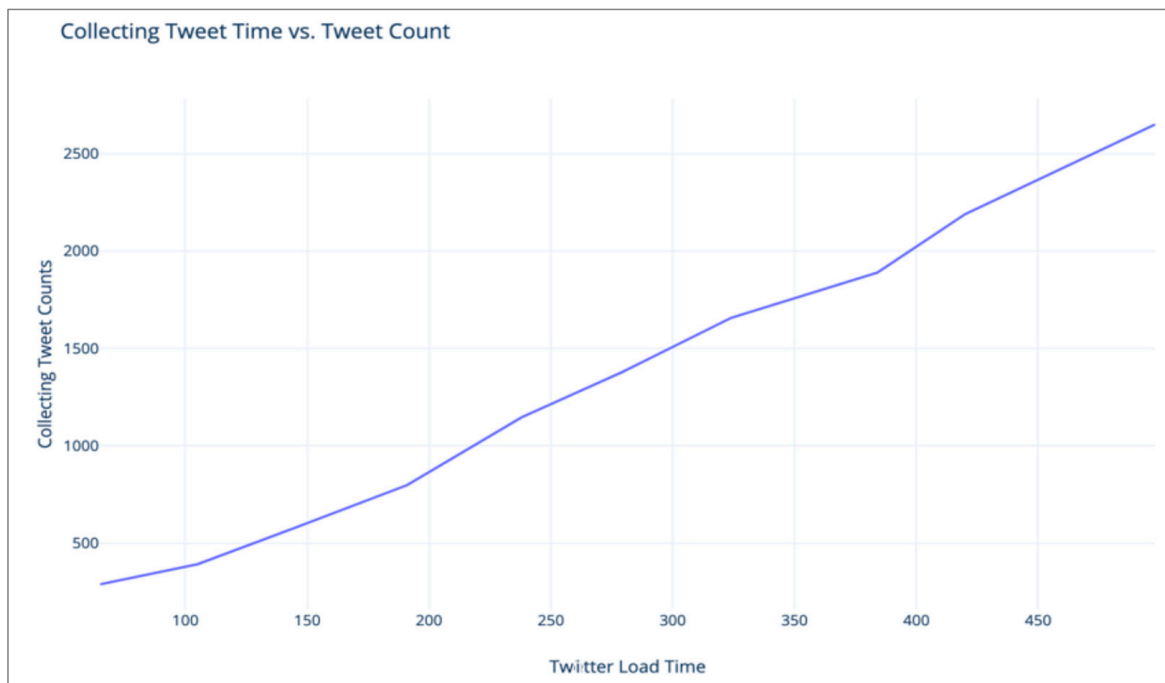


Figure 9. Tweet time vs. number of tweets.

The Twitter API has a limitation where the maximum number of tweets that can be retrieved is 150 per hour. However, the tool used in this study demonstrates a significantly higher capability, gathering 294–1000 tweets within just five minutes. This implies that individuals involved in Twitter-related initiatives would find it challenging to monitor and extract real-time data at such a rapid pace. The application developed in this study operates in real time, enabling swift and efficient data extraction from Twitter.

Table 2 presents a comparison of tweet rates for various metrics. These metrics include reply and retweet counts, favorite counts, word classification counts from the NLTK libraries, and vector counts obtained from the TensorFlow libraries. This table offers insights into the volume and frequency of different metrics associated with the collected tweets. It demonstrates the robustness and capabilities of the tool in capturing and analyzing various aspects of tweet data, providing valuable information for further analysis and understanding.

Table 2. Finding zero-day tweet success rate comparison.

Technique	Success State	Success Rate	Reply Count	Favorite Count	Retweet	NLTK	TensorFlow
Reply	X	10%	381	121	54	24	1
Favorite	X	20%	3	988	636	34	0
Retweet	✓	60%	44	920	899	25	2
NLTK	✓	60%	0	17	9	61	2
TensorFlow	✓	80%	0	2	0	52	5

Upon concluding the application procedure, it was observed that the tool developed in this study can effectively uncover new information from tweets by utilizing word

classifications through the TensorFlow and NLTK libraries. By analyzing real data, the tool has the capability to identify and categorize tweets based on specific word classifications. The ability of the tool to analyze real-time data and discover new insights from tweets, along with the need for manual intervention to handle certain cases, underscores its potential for monitoring and capturing up-to-date information about vulnerabilities and flaws discussed on Twitter.

The experiments conducted during the study also revealed that manual troubleshooting of the system may be required. This is primarily because approximately 50% of the tweets discussed newly discovered vulnerabilities or flaws. While the tool can identify and classify such terms, it is important to differentiate between terms that are newly connected and those that may have been mentioned in the past. This distinction allows users to discern between old and newly learned terms, enhancing the accuracy and relevance of the extracted information.

5. Discussion and Research Limitations

The objective of this study is to introduce a novel tool that utilizes a machine learning approach to predict potential zero-day attacks by collecting and analyzing tweets from the Twitter platform. Various combinations of zero-day attack terms were employed to search for relevant tweets. The tool was developed using Python codes executed in a TensorFlow environment. Data preparation was carried out using the NLTK toolkit, which allowed for effective preprocessing before the data were analyzed using Python codes.

In contrast to conventional studies using traditional datasets or network logs, this research employs Twitter data. This offers real-time insights into discussions, trends, and zero-day indicators unavailable through usual methods. The study distinguishes itself by utilizing Tensor-Flow to efficiently process vast Twitter data. Unlike manual or rule-based methods, this approach enhances efficiency and scalability in detecting and addressing emerging threats.

The evaluation of the developed tool demonstrated its remarkable capability to gather a substantial number of tweets, ranging from 294 to 1000, within a span of just five minutes. In comparison, other methods were limited to retrieving only 150 tweets in the same timeframe. Furthermore, the success rate of the TensorFlow environment in predicting potential zero-day attacks was found to be 80%, while the NLTK toolkit achieved a success rate of 60%. This highlights the superior performance and effectiveness of the TensorFlow-based approach in analyzing and predicting zero-day attacks compared to the NLTK toolkit.

The pie chart depicted in Figure 8 of the study's application section reveals more significant keywords in the tweets than just the specific "zero-day" word combinations. This indicates that the developed tool possesses the capability to learn and identify other relevant terms related to the topic at hand. For instance, in order to aid in preventing actual acts of terror, the program can search for phrases, such as "terror, threat, terrorism, attack", along with the word "bomb". Simultaneously, when searching for terms like "zeroday, 0 day, and zero-day", the program can also uncover the concept of a zero-day vulnerability, which may be associated with terms like "windows, exploit, steam". By incorporating these term categories into the tweet search process, the program has the ability to identify and highlight new instances of zero-day vulnerabilities. This demonstrates the versatility of the tool in discovering and monitoring various aspects of cybersecurity, including both specific terms related to zero-day vulnerabilities and broader topics, such as terrorism and threats.

Furthermore, this study effectively addresses the notion that if there is a vulnerability or flaw, the corresponding tweet discussing it would likely receive significant engagement in terms of likes, ratings, favorites, and retweets. It is expected that individuals who discover and recognize the existence of a vulnerability would appreciate and acknowledge it by engaging with the tweet. However, the study also reveals an interesting observation that people tend to retweet and engage with content that they find aesthetically pleasing or positive. In the case of a newly discovered flaw, it may not necessarily be perceived as aesthetically pleasing or appealing to users. As a result, a tweet discussing a vulnerability

might not receive significant engagement in terms of endorsements or retweets. This finding highlights the complex nature of user behavior on social media platforms like Twitter. While significant engagement is typically associated with positive or visually appealing content, the perception and response to content related to vulnerabilities and flaws may differ. Understanding these dynamics is crucial for effectively monitoring and addressing cybersecurity issues on social media platforms.

Not only can the tool focus on identifying and collecting tweets related to zero-day threats, but it also has the capacity to discover additional keywords or tweets pertaining to unusual activities. This is made possible by the program's ability to learn and identify new terms that are frequently used in conjunction with the selected keywords. For example, when the keywords "Nike" and "Adidas" are entered, the application may also retrieve tweets mentioning "New Balance". This observation indicates that people often discuss "New Balance" about Adidas and Nike. In other words, the tool has the capability to learn, match, and list more relevant phrases beyond the specified list of keywords. This feature allows the tool to provide a more comprehensive and dynamic approach to collecting and analyzing social media data.

The method proposed in this study demonstrates the feasibility of quickly identifying new zero-day exploits. However, it is acknowledged that the method still has certain limitations and may require manual intervention or encounter occasional obstacles during the process. Additionally, it is noted in the study that the laptop computer utilized for the virtualization system and hard disk were not adequately equipped to support the extensive scope of the research. Consequently, the crawlers used in this study were limited in their capabilities. They could not refresh at a higher frequency than once per second, perform real-time searches, or directly target specific individuals or groups on Twitter. Due to these limitations, the crawlers relied on searching Twitter users by utilizing the built-in search functionality provided by Twitter. They retrieved search results and scrolled through them to collect relevant data. Although the crawlers could gather valuable information, the laptop's constrained performance and the crawlers' limitations restricted their ability to operate at higher speeds or target specific entities on the platform. These limitations highlight the need for robust computational resources and more advanced crawling techniques to enhance the efficiency and effectiveness of data collection in future research endeavors.

6. Conclusions and Future Research Directions

Cyberattacks present a significant risk in virtual environments and real-life scenarios, as online activities can lead to tangible consequences. With the increasing prevalence of social media platforms, monitoring them for potential criminal activities has become crucial, as individuals may openly discuss and plan such activities online. This study aimed to develop a method to monitor social media platforms specifically for mentions of zero-day security vulnerabilities. These vulnerabilities, if exploited by cybercriminals or cyberterrorists, can lead to severe online crimes. The tool created in this study utilizes the TensorFlow machine learning framework, incorporating concepts like word categorization and maximum probability to identify potential dangers in the cyber environment. This utilization provides a robust and efficient means of monitoring social media platforms for security risks.

The findings of this study demonstrated the effectiveness of the tool developed in detecting zero-day vulnerabilities. The program's ability to identify relevant text by regularly searching for different keywords was showcased. However, it is essential to acknowledge that these keywords may evolve and change over time. Developing a self-learning, self-adapting, and self-evolving control mechanism with a word dictionary is crucial to address this challenge. Implementing such a control tool makes it possible to stay updated with the current discussions and topics of interest among users. This tool enables adaptation to changes in language use, emerging keywords, and evolving threat landscapes. As a result, the tool's effectiveness in identifying and responding to zero-day vulnerabilities would be enhanced.

The tool presented in this study is currently in its early stages of development. However, with further improvements and enhancements, it has the potential to expand its capabilities and collect more data from various social networking sites that contain textual content. By recognizing the existing constraints and identifying areas for improvement, researchers can build upon the methodology discussed in this study and contribute to advancing intelligence research, leading to more effective identification and mitigation of zero-day exploits in the digital landscape.

The future expansion of this study may involve incorporating additional machines and browsers to enhance data collection capabilities. Moreover, the study will encompass a wider range of social media platforms, allowing for deeper analysis and more precise insights. While Twitter served as the primary source of data for this study, incorporating other social media platforms will enable a more comprehensive understanding of the subject matter. To enhance the validity and accuracy of the findings, future research will involve verifying the zero-day vulnerabilities identified through various Open-Source Intelligence (OSINT) sources. This verification process will help distinguish genuine zero-day vulnerabilities from previously known vulnerabilities, thereby ensuring the study's results are reliable and reflective of the current cybersecurity landscape.

Author Contributions: Conceptualization, E.C.; methodology, A.E.T. and Y.I.A.; validation, A.E.T. and E.E.; analysis, Y.I.A.; writing—original draft preparation, E.C. and E.E.; writing—review and editing, Y.I.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Data available on request due to restrictions on privacy or ethical.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bu, S.-J.; Cho, S.-B. Deep character-level anomaly detection based on a convolutional autoencoder for zero-day phishing URL detection. *Electronics* **2021**, *10*, 1492. [CrossRef]
2. Statista. Number of Internet and Social Media Users Worldwide as of April 2023. Available online: <https://www.statista.com/statistics/617136/digital-population-worldwide/> (accessed on 26 June 2023).
3. Marinho, R.; Holanda, R. Automated emerging cyber threat identification and profiling based on natural language processing. *IEEE Access* **2023**, *11*, 58915–58936. [CrossRef]
4. Cheng, X.; Zhang, J.; Tu, Y.; Chen, B. Cyber situation perception for Internet of things systems based on zero-day attack activities recognition within advanced persistent threat. *Concurr. Comput. Pract. Exp.* **2022**, *34*, e6001. [CrossRef]
5. Pattnaik, N.; Li, S.; Nurse, J.R. Perspectives of non-expert users on cyber security and privacy: An analysis of online discussions on twitter. *Comput. Secur.* **2023**, *125*, 103008. [CrossRef]
6. Zahoor, U.; Rajarajan, M.; Pan, Z.; Khan, A. Zero-day ransomware attack detection using deep contractive autoencoder and voting based ensemble classifier. *Appl. Intell.* **2022**, *52*, 13941–13960. [CrossRef]
7. Ahmad, R.; Alsmadi, I.; Alhamdani, W.; Tawalbeh, L.A. Zero-day attack detection: A systematic literature review. *Artif. Intell. Rev.* **2023**, *5*, 1–79. [CrossRef]
8. Yadav, A.; Kumar, A.; Singh, V. Open-source intelligence: A comprehensive review of the current state, applications and future perspectives in cyber security. *Artif. Intell. Rev.* **2023**, *15*, 1–32. [CrossRef]
9. Malatji, M.; Marnewick, A.; von Solms, S. Validation of a socio-technical management process for optimising cybersecurity practices. *Comput. Secur.* **2020**, *95*, 101846. [CrossRef]
10. Fatima, R.; Yasin, A.; Liu, L.; Wang, J. How persuasive is a phishing email? A phishing game for phishing awareness. *J. Comput. Secur.* **2019**, *27*, 581–612. [CrossRef]
11. Fatima, R.; Yasin, A.; Liu, L.; Jianmin, W. Strategies for counteracting social engineering attacks. *Comput. Fraud Secur.* **2022**, *2022*, S1361–S3723. [CrossRef]
12. Ali, S.; Rehman, S.U.; Imran, A.; Adeem, G.; Iqbal, Z.; Kim, K.-I. Comparative evaluation of AI-based techniques for zero-day attacks detection. *Electronics* **2022**, *11*, 3934. [CrossRef]
13. Fjelland, R. Why general artificial intelligence will not be realized. *Humanit. Soc. Sci. Commun.* **2020**, *7*, 10. [CrossRef]
14. Mishra, A.; Alzoubi, Y.I.; Anwar, M.J.; Gill, A.Q. Attributes impacting cybersecurity policy development: An evidence from seven nations. *Comput. Secur.* **2022**, *120*, 102820. [CrossRef]
15. Fourati, M.; Jedidi, A.; Gargouri, F. A deep learning-based classification for topic detection of audiovisual documents. *Appl. Intell.* **2022**, *53*, 8776–8798. [CrossRef]

16. Mishra, A.; Alzoubi, Y.I.; Gill, A.Q.; Anwar, M.J. Cybersecurity enterprises policies: A comparative study. *Sensors* **2022**, *22*, 538. [CrossRef]
17. Mishra, A.; Jabar, T.S.; Alzoubi, Y.I.; NathMishra, K. Enhancing privacy-preserving mechanisms in cloud storage: A novel conceptual framework. *Concurr. Comput. Pract. Exp.* **2023**, e7831. [CrossRef]
18. Mittal, S.; Das, P.K.; Mulwad, V.; Joshi, A.; Finin, T. Cybertwitter: Using twitter to generate alerts for cybersecurity threats and vulnerabilities. In Proceedings of the IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM '16), San Francisco, CA, USA, 18–21 August 2016; IEEE: New York, NY, USA; pp. 860–867.
19. Altalhi, S.; Gutub, A. A survey on predictions of cyber-attacks utilizing real-time twitter tracing recognition. *J. Ambient Intell. Humaniz. Comput.* **2021**, *12*, 10209–10221. [CrossRef]
20. Hindy, H.; Atkinson, R.; Tachtatzis, C.; Colin, J.-N.; Bayne, E.; Bellekens, X. Utilising deep learning techniques for effective zero-day attack detection. *Electronics* **2020**, *9*, 1684. [CrossRef]
21. Kim, J.-Y.; Bu, S.-J.; Cho, S.-B. Zero-day malware detection using transferred generative adversarial networks based on deep autoencoders. *Inf. Sci.* **2018**, *460*, 83–102. [CrossRef]
22. Mbona, I.; Eloff, J.H. Detecting zero-day intrusion attacks using semi-supervised machine learning approaches. *IEEE Access* **2022**, *10*, 69822–69838. [CrossRef]
23. Peppes, N.; Alexakis, T.; Adamopoulou, E.; Demestichas, K. The effectiveness of zero-day attacks data samples generated via GANs on deep learning classifiers. *Sensors* **2023**, *23*, 900. [CrossRef]
24. Sarhan, M.; Layeghy, S.; Gallagher, M.; Portmann, M. From zero-shot machine learning to zero-day attack detection. *Int. J. Inf. Secur.* **2023**, *22*, 947–959. [CrossRef]
25. Soltani, M.; Ousat, B.; Siavoshani, M.J.; Jahangir, A.H. An adaptable deep learning-based Intrusion Detection System to zero-day attacks. *J. Inf. Secur. Appl.* **2023**, *76*, 103516. [CrossRef]
26. Won, D.-O.; Jang, Y.-N.; Lee, S.-W. PlausMal-GAN: Plausible malware training based on generative adversarial networks for analogous zero-day malware detection. *IEEE Trans. Emerg. Top. Comput.* **2022**, *11*, 82–94. [CrossRef]
27. Zuhair, H.; Selamat, A.; Krejcar, O. A multi-tier streaming analytics model of 0-day ransomware detection using machine learning. *Appl. Sci.* **2020**, *10*, 3210. [CrossRef]
28. Matplotlib. Matplotlib: Visualization with Python. 2022. Available online: <https://matplotlib.org/> (accessed on 26 June 2023).
29. Python. Python 3.11.1 documentation. Available online: <https://docs.python.org/3/> (accessed on 25 June 2023).
30. TensorFlow. TensorFlow core. Available online: <https://www.tensorflow.org/tutorials/text/word2vec> (accessed on 28 June 2023).
31. NLTK. Natural Language Toolkit. Available online: <https://www.nltk.org/> (accessed on 28 June 2023).
32. Mohammed, A.; Kora, R. An effective ensemble deep learning framework for text classification. *J. King Saud Univ.-Comput. Inf. Sci.* **2022**, *34*, 8825–8837. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.