***Abstract*—Data-driven security has become essential as the cybersecurity scenario becomes more and more complex. Textual sources contain a lot of cybersecurity data, making it challenging for both security analysts and machines to analyze the data. We identify entities linked to cybersecurity and suggest a self- attention-based neural network model for the named entity recognition in cybersecurity in order to transform the textual information into structured data for additional automatic anal- ysis. We apply CNN to extract character features, which are subsequently concatenated into the word feature since we believe that a single word feature is insufficient to identify the item. Next, based on the current BiLSTM-CRF model, we incorporate the self-attention mechanism. Finally, using the labeled dataset to test the proposed model, we find that it performs better than the prior entity extraction approach.**

***Index Terms*—component, formatting, style, styling, insert**

1. INTRODUCTION

Current critical infrastructure vulnerability assessment pro- cedures rely heavily on the personnel of security operations centers (SOCs). (Feng et al., 2017). However, traditional manual and subjective audits fall short of the desired at- tack surface evaluation and dynamic CIs characteristics. For instance, certain commonly used tools for calculating the severity of vulnerabilities, such as the CVSS1 calculator, need user input and are based on qualitative assessments of vulnerability characteristics including exploitability, scope, and consequences. (Joh and Malaiya, 2011).However, textual sources including vendor announcements, blogs, whitepapers, and hacker forums contain enormous amounts of cybersecurity information [2]. The lengthy and ineffective process of manual labor is used in traditional threat analysis of information from textual sources. As a result, the security analysts are unable to fully utilize the verified cybersecurity information to promptly and accurately respond to cyber threats.

1. EXISTING WORKS
2. *Literature Review*

In the present age, cyberbullying is a severe issue, especially on social networking platforms like Facebook. The mental and emotional well-being of the target of technological harassment or intimidation may suffer. Researchers have investigated the use of natural language processing (NLP) techniques to

automatically identify instances of cyberbullying on social media platforms like Facebook in an effort to stop the practice. Recent studies on Facebook’s NLP for identifying cyberbul- lying are examined in this literature review.

Making use of machine learning as a base: A well-liked technique for identifying cyberbullying on social media sites like Facebook is the machine learning-based approach. A dataset of Facebook posts that have been manually classified as cyberbullying or not is used to train a machine learning model. The goal of this training is to teach the model the char- acteristics and patterns that set the two groups of posts apart. When the model has been trained, it may be used to determine whether recent Facebook posts qualify as cyberbullying. These classifications are provided by the model using the patterns and traits that were discovered during training. This technique is advantageous since it allows for the automatic detection of cyberbullying in sizable communication datasets.

Mishra et al developed a deep learning model which uses convolutional and recurrent neural networks to identify cy- berbullying. The dataset used to train the algorithm consisted of 10,000 Facebook posts that were either identified as cy- berbullying or not. 91% of the time, the software accurately recognised posts that constituted cyberbullying, demonstrating the effectiveness of this tactic. It is critical to keep in mind that the quality and representativeness of the training data affect how accurate the model is (Mishra, P., Maheshwari, P., Sharma, R., Kumaraguru, P. 2021).

Text analysis-based approach: In order to look for cases of cyberbullying, the text analysis-based technique evaluates Facebook postings’ content. Natural language processing tech- niques include sentiment analysis, keyword extraction, and language translation to identify patterns and characteristics of cyberbullying in Facebook posts. The approach also considers other factors that could contribute to cyberbullying, such the presence of photographs and the identity of the poster. One indication of cyberbullying may be the presence of a certain sort of image in a post, such as a graphic or violent image. The identity of the poster, including their age or gender, may also be taken into consideration while looking for signs of cyberbullying in the message (Kumar, A., Yadav, A. K., 2020).

Hybrid approach: The hybrid technique brings together the benefits of text analysis and machine learning. This approach trains a machine learning model to evaluate whether or not a communication is cyberbullying using a dataset of labeled Facebook postings. Following the algorithm’s discovery of a group of possibly bullying communications, text analysis techniques such as sentiment analysis, keyword extraction, and language translation may be used to further examine these postings and decide whether they constitute cyberbullying.

This method allows for a more in-depth examination of Facebook postings since text analysis tools may provide extra context to the wording and tone of the posts. Furthermore, the machine learning model may discover postings that traditional text analysis approaches may overlook. The hybrid approach properly identified Facebook postings including cyberbullying 87% of the time. The hybrid technique to detect Facebook cyberbullying is often promising since it combines the benefits of text analysis with machine learning approaches to provide a more broad and accurate research of Facebook postings (Saini, A., Nigam, A., Jaiswal, A., 2021).

1. *Research Problem*

The practice of manually assessing vulnerabilities can lead to inaccurate information and require complicated analysis. Additionally, the variety, lack of completeness, and duplication of security data in modern repositories contribute to these security concerns. These problems are common in both public and manufacturer vulnerability reports, making it challenging to identify and address security weaknesses through direct analysis. [**?**]

Extracting pertinent information from the overwhelming influx of data is a difficult task. With the proliferation of social networks and widespread computing, the volume of digital text content is growing rapidly. One of the primary responsibilities of a security centers, allowing for correlation. [[2](#_bookmark1)]analyst is to gain situational awareness by identifying cyber threat- related information, enabling proactive monitoring and risk mitigation. To accomplish this, it is imperative to extract cyber threat-related information specific to an organization from the public internet, due to the growing number of information systems used in security operation

The passage discusses the challenges posed by cognitive cyber-physical systems of the Internet of Things (CPS-IoT) in the healthcare sector. These systems have increased cognitive complexity, making them vulnerable to traditional CPS-IoT threats and new threats related to their inherent cognitive functionalities. Moreover, the ubiquity of CPS-IoT increases the attack surface, making the public safety risks higher for critical infrastructure. Healthcare services and infrastructures are particularly vulnerable to major security risks, and the situation is exacerbated by the increasing interconnectedness of medical devices and services that are exposed to new cybersecurity vulnerabilities. Reports indicate a rise in cyber- attacks and medical identity theft globally. The current security solutions are unable to tackle the dynamicity, complexity, un- certainty, and high connectivity of CPS-IoT enabled healthcare services and critical infrastructures. The solution proposed is

to develop innovative techniques for building cognitive cy- bersecurity for CPS-IoT enabled healthcare ecosystems using cognitive architecture and artificial intelligence to enhance au- tomated intelligent cybersecurity decision-making mechanisms with expert-level ability. [[3](#_bookmark2)]

Detecting web attacks within URLs is indeed a challenging task, and there are several major issues that need to be addressed. Firstly, there is a need for an effective way to transform all kinds of URLs into representations, as different attacks can hide in various ways within their URLs. Secondly, different attacks exhibit different signatures in their URLs, which makes feature selection a challenging task. Thirdly, many deep learning models used for cyber security only have one model for detection, making it difficult to update the system when new attacks emerge. Finally, in the IoT cloud environment, the centralization feature can impact the application of distributed services, such as network security mechanisms for IoT applications. To address these challenges, novel security models, controls, and decisions must be dis- tributed at the edge of the cloud to support the new IoT paradigm, known as the EoT. [[4](#_bookmark3)]

That is a plausible explanation for the modest number of enti- ties of certain types in the annotated dataset. During an attack, it may not be immediately clear who or what is responsible, and it may take some time to gather the necessary information to identify the entities involved. Additionally, some attacks may be carried out by previously unknown entities or groups, making it difficult to label them in advance. As a result, the annotated dataset may not be representative of all types of attacks and entities involved, and further research and data collection may be necessary to improve the performance of automated detection systems. [[5](#_bookmark4)]

1. *Research Objectives*

Our proposed approach to cybersecurity involves enhanc- ing human cognitive capabilities in two ways. Firstly, we address conflicting vulnerability reports and prepare trust- worthy datasets by processing embedded security indicators. Subsequently, we utilize these datasets as a foundation for our ensemble meta-classifier methods, which combine various machine learning techniques to enhance predictive accuracy beyond that of individual algorithms. [[1](#_bookmark0)]

The primary goal of this study was to determine whether a doc2vec-based language model could effectively serve as a natural language filter within our proposed system. The spe- cific contributions of our research are outlined below: First, we have proposed an architecture for an independent system that can identify content related to cyber threats from a vast amount of publicly available textual documents. Second, by utilizing cybersecurity-specific training data and custom preprocessing techniques, we have trained the doc2vec model to function as a domain-specific language filter for our autonomous system. Third, we have evaluated the performance of the language filter by utilizing the holdout method and optimizing the hyperparameters of the doc2vec model. [[2](#_bookmark1)]

The proposed cognitive cybersecurity methodology and the- ory aim to provide a comprehensive approach to defend

against dynamic and adaptive attacks to the CPS-IoT-enabled healthcare ecosystem. The methodology includes a cognitive architecture that models human cognitive behavior to antici- pate and respond to new security and privacy threats. It also considers trade-offs and other contributing factors to get ahead of attackers’ cognitive decision cycle, accounts for uncertain- ties, and optimizes temporal feedback loops. Additionally, the methodology integrates innovative mechanisms for security, privacy, metrics, and dynamic security knowledge base to enhance threat prevention, detection, incident response, and mitigation of impacts.Privacy-aware collaboration, computa- tional techniques, adaptive data collection, and actuation are also included in the methodology. Furthermore, the approach integrates cross-cutting techniques such as AI predictive an- alytics, run-time verification, evidence collection, and tracing for evidence-based risk management and dynamic forensics. The proposed methodology and theory provide a compre- hensive and systematic approach to defend against emerging attacks to the CPS-IoT-enabled healthcare ecosystem. [[3](#_bookmark2)]

The paper proposes a distributed system for web attack de- tection from URLs using deep learning techniques, including CNNs and NLP models. The system can represent all kinds of URLs, distinguish anomalous requests from normal ones, and apply multiple concurrent models to enhance system stability. Additionally, the paper proposes a generic distributed web attack detection system on edge devices of the cloud. The paper is organized into different sections, including a brief review of related works, the architecture and methodology of the proposed system, the datasets and experiment settings, experimental results and discussion, and the conclusion and future work. [[4](#_bookmark3)]

BERT transformers have been proven to be effective in natural language processing tasks, including named entity recognition (NER). In this paper, the authors propose using BERT trans- formers to improve NER performance in scenarios where the annotated dataset is small or lacks diversity. Additionally, they suggest an automatic dataset augmentation method to extend the training dataset with sentences containing automatically labeled named entities. The paper explores the performance of different BERT models, including a multilingual model, a model fine-tuned on Russian data, and a model fine-tuned on cybersecurity texts. The authors also introduce a new method of dataset augmentation for NER tasks and investigate the impact of different parameters on the quality of the NER system. Overall, the goal of the paper is to improve the performance of NER systems in challenging conditions where the annotated dataset is limited, and diverse named entities may be present. [[5](#_bookmark4)]

1. METHODOLOGY

In the modern digital era, cyberbullying is becoming a bigger issue, especially on social media platforms like Facebook. It is a form of abuse, pressure, or harassment aimed towards those who use digital technology. It could be challenging to identify cyberbullying on Facebook due to the enormous volume of text data that users produce every day. The use of natural language processing (NLP) methods

to automatically identify cyberbullying on social media sites has been studied. In this project, we create a Facebook cyberbullying detection system using NLP approaches.

1. *Data Collection*

A dataset of Facebook posts and comments is compiled as part of the initial phase of this investigation. We will com- pile the information using the Facebook API, which enables us to extract publicly accessible posts and comments about cyberbullying. We will filter the posts and comments based on specified hashtags and terms related to cyberbullying. The dataset will then undergo preprocessing to ensure that all of the data is text-based and to remove any irrelevant information.

1. *Preprocessing*

The dataset must then be prepared for analysis. Emojis, images, links, and any other non-text data from the dataset will be removed. After the text has been tokenized, or broken down into individual words, the data will be normalized using stemming and lemmatization techniques. Stopwords, or common words such as ”the” and ”and,” that do not advance the topic of the text, will be removed. To keep everything consistent, the remaining text will be changed to lowercase.

1. *Extraction*

In the third stage, relevant attributes are extracted from the preprocessed dataset. The text data will be converted into numerical features using NLP techniques such as bag- of-words, which represents the text as a collection of words, term frequency-inverse document frequency, which gauges the significance of each word in a document, and word embed- dings, which represents words as vectors. The cyberbullying detection model will be trained using these features.

1. *Model Development*

The creation of a model for identifying cyberbullying is the fourth phase. The model will be developed using a range of machine learning approaches, including logistic regression, decision trees, and support vector machines. We will also look into the use of deep learning techniques like transformers and recurrent neural networks. The features of the preprocessed dataset will be used to train the models.

1. *Model Evaluation*

The effectiveness of the developed cyberbullying detection model is assessed in the final stage. Accuracy, precision, recall, and F1-score are a few of the measures that will be used to gauge the model’s efficacy. To ensure that the model generalizes effectively to fresh data, it will be assessed on a separate dataset. The evaluation will assist us in determining how effectively the model detects instances of cyberbullying on Facebook.

1. MODEL TRAINING
2. *Data Processing*

We first obtained the necessary data from a variety of sources, including kaggle, github, and other internet sites, due to a lack of data.The dataset was then mixed with some personally gathered data and modified before being ready for pre-processing. Next, we looked to see if any fields were null as well as the overall count. Later, we eliminated redundant information and combined all the words into a single string. We receive the finalized data for the primary work. Then we go through some Natural Language Processing techniques and focused on training the data.

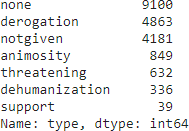


Fig. 1: Data frequency

1. *Equation and Calculation*

True Positive: Situation in which you accurately foresee a favorable outcome.

False Positive: When you expect a great outcome yet it turns out to be negative.

False Negative: Suppose you forecast a negative outcome, but it turns out to be a positive one.

True Negative: Scenario where you accurately foresee a disastrous outcome

a)Accuracy = (TP+TN)/(TP+TN+FP+FN)

b)Recall = TP/(TP+FN) c)Precision = TP/(TP+FP)

d)F1 Score = (2 \* Precision \* Recall) / (Precision + Recall)

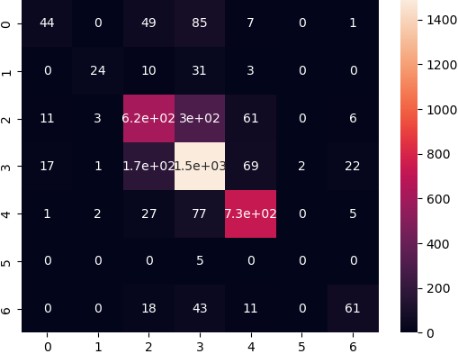


Fig. 2: Confusion Matrix

1. *Classification Result*

Here we train our dataset with the Logistic regression model.The performance of this model is a little satisfactory but we want to train some other model in our future work to improve the accuracy.The anayzing report of Logistic regres- sion model is given below:

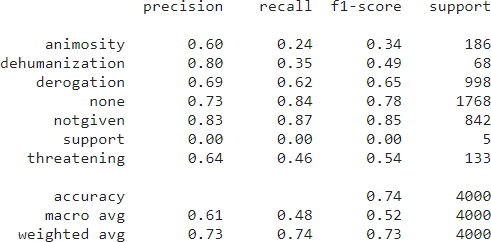


Fig. 3: Classification Report

CONCLUSION

In this work, we suggested a system design that might automatically extract data about cyber threats to help human operators. Using the neural embedding technique doc2vec, we conducted tests to see if the Natural Language Filter portion of the system could be implemented in practice.Our analysis shows that a doc2vec-based natural language model trained with cybersecurity-specific text data and unique preprocessing might be employed as a Natural Language Filter for the proposed autonomous system with 74% accuracy.We will further work on it with different techniques so that we can get maximum accuracy among the existing system.