IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2017.Doi Number

Dark Side of the Web: Dark Web Classification Based on TextCNN and Topic Modeling Weight

Gun-Yoon Shin[1], Younghoan Jang[1], Dong-Wook Kim[1], SungJin Park[2], A-ran Park[2], younghwan Kim[2], and Myung-Mook Han[1]

[1]Department of AI Software, Gachon University, Seongnam-si 13120, Republic of Korea
[2]Cyber Warfare, LIG Nex1, Seongnam-si 13488, Republic of Korea

Corresponding author: Myung-Mook Han (e-mail: mmhan@gachon.ac.kr).

This work was supported by the Korea Research Institute for Defense Technology Planning and Advancement (KRIT)—Grant funded by the Defense Acquisition Program Administration (DAPA) (KRIT-CT-21-037) and was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (RS-2023-00248132).

ABSTRACT The Dark Web is an internet domain that ensures user anonymity and has increasingly become a focal point for illegal activities and a repository for information on cyberattacks owing to the challenges in tracking its users. This study examined the classification of the Dark Web in relation to these cyber threats. We processed Dark Web texts to extract vector types suitable for machine learning classification. Traditional methods utilizing the entirety of Dark Web texts to generate features result in vectors including all words found on the Dark Web. However, this approach incorporates extraneous information in the vectors, diminishing learning effectiveness and extending processing duration. The research aimed to optimize the classification process by selectively focusing on keywords within each class, thereby curtailing word vector dimensions. This optimization was facilitated by leveraging the anonymity characteristic of the Dark Web and employing topic-modeling-based weight generation. These methods enabled the creation of word vectors with a constrained feature set, enhancing the distinction of Dark Web classes. To further improve classification performance, we integrated TextCNN with topic modeling weights. For validation, we employed two datasets and compared the performance of the model with other text classification algorithms, where the proposed model demonstrated superior effectiveness in Dark Web classification.

INDEX TERMS Dark Web, Dark Web analysis, text classification, topic modeling, model explanation

I. INTRODUCTION

The global adoption of internet services and the accessibility of affordable devices have caused a surge in data exchange across the globe. Although the visible portion of the internet, known as the surface web, appears to be an extensive repository of information, it represents a mere fraction of the entire web. This surface web comprises an insubstantial part of the total web, with over 90% being concealed within the Hidden Web, also termed the Deep Web, which harbors detailed and varied data. The Dark Web—a subset of the Deep Web—is notorious for its use in unlawful activities [1]. Accessing the Dark Web requires specialized browsers like Tor, I2P, and Freenet, which ensure user anonymity [2]. Although anonymity was once positively perceived, it has morphed into a veil for illicit activities, rendering it a platform where illegal actions occur unimpeded. Currently, 57% of online malevolent activities such as drug trafficking, pornography, theft of personal data, and hacking, originate from the Dark Web [3].

The illicit undertakings on the Dark Web have tangible negative consequences in the real world. Investigations have revealed that malware including ransomware and trojans is disseminated via phishing emails or used in cryptocurrency transactions on the Dark Web, where anonymity is assured [4]. This secrecy and anonymity emboldens users to actively engage in and promote malevolent activities, with confidence that their actions in this hidden sphere remain untraceable. For instance, unlike the surface web that employs coded language or abbreviations for information exchange, the Dark Web blatantly displays details such as drug names, transaction

methods, and prices or even advertises malware sales, transaction procedures, and source codes for public access. In previous studies, text mining has been utilized to classify and identify such behaviors, with recent advancements involving various machine learning algorithms for classification.

The application of machine learning to Dark Web classification involves the preprocessing of Dark Web text and its conversion into a vector format. This process, known as word embedding, transforms natural language into a form comprehensible to machines. Relevant technologies include the document-term matrix (DTM), term frequency-inverse document frequency (TF-IDF), word2vec, latent semantic analysis (LSA), and GloVe. These methods aim to effectively vectorize all words in the text; however, they do not exclude words irrelevant to learning nor identify those with significant learning impact. Additionally, this method of embedding increases the dimensional count in parallel with the increasing text word count. Generally, text learning involves processing all sentences to extract information from the entire text. However, the anonymity of the Dark Web facilitates the direct exposure of cyberthreat content without coded vocabulary or abbreviations. This characteristic enables learning by omitting non-essential components of the Dark Web and converting only the critical segments of each class into vectors.

In this regard, Ref. [5] examined and analyzed various dimensional reduction techniques to enhance text classification performance, whereas Ref. [6] sought to improve classification accuracy by applying dimensional reduction to word embedding. Furthermore, Ref. [7] aimed to improve performance via dimension reduction with autoencoders (AE), and Ref. [8] aimed to increase the accuracy of fake news detection models by implementing principal component analysis (PCA) and chi-square before convolutional neural network (CNN)-based learning. Thus, irrespective of the selected text classification algorithm, dimension reduction emerges as a viable strategy to improve performance.

Therefore, this research aimed to refine word dimensions by analyzing and filtering out irrelevant words and key terms for each category. Consequently, superfluous vector dimensions were eliminated, enhancing data quality, thus preventing the learning model from assimilating unnecessary word vectors and enhancing performance through the incorporation of highly influential word vectors for classification. Additionally, to facilitate user comprehension of the classified Dark Web results, we provided pertinent classes and specific web data. In this approach, principal words for each class were identified through topic modeling, and weights representing the relevance of the word to the class were computed. This enables a numerical assessment of the extent to which specific texts correlate with a class, based on word distribution and significance.

The methodology for classifying the website of the Dark Web in this study involved several steps. Initially, we preprocessed the acquired Dark Web data and utilized topic modeling to identify significant words and their impact on each malicious activity. Thereafter, we filtered out non-essential information from each dark website of the Dark Web and transformed it into a topic-embedded matrix. Following this, we trained the matrix using a CNN model to classify malicious activities, highlighting the classification through the influence and frequency of key words determined *via* topic modeling. The primary contributions of this study are outlined as follows:

- We introduced a Dark Web classification technique that integrates topic modeling weights with TextCNN. This method, by removing meaningless Dark Web content through topic modeling weights and constructing an embedded matrix, ensures precise classification.
- The learning process involved selecting only the keyword vectors pertinent to each class. This reduction in word vector dimensions allowed for improved performance, even with a smaller set of vectors learned.
- Through topic modeling, we calculated the primary words for each class and the respective weights of these words in relation to the topic. This calculation was applied to the text to provide users with a numerical indication of the class's relevance to the classified Dark Web, thereby simplifying the analysis results for user comprehension.
- Contrary to traditional Dark Web classification methods that focus on word frequency quantification, our approach utilized topic modeling to interpret the meanings of words present in the Dark Web. This enabled us to preserve classification performance without incorporating non-essential words found in the Dark Web.
- We evaluated the effectiveness of the model using two distinct datasets sourced from the Dark Web, which served as the ground truth.

The remainder structure of this paper is described as follows. Section II provides a review of preceding studies pertinent to Dark Web classification. Section III describes the Dark Web classification method proposed in this research, along with an explanatory approach employing topic modeling weights. Section IV details various evaluations conducted using two datasets sourced from the Dark Web, complemented by descriptions of the experimental methodologies. Conclusively, the findings of

this study and the prospective avenues for future research are summarized in Section V.

II. RELATED WORKS

Section II offers a comprehensive review of existing research on Dark Web classification. This section is methodically segmented into three distinct areas: Dark Web analysis, topic modeling, and text classification. Within Dark Web analysis, we examine various methodologies that have been explored to date. The segment on topic modeling delves into techniques for topic analysis and the extraction of significant words. In the text classification portion, methods for feature extraction and the classification of texts are scrutinized.

A. Dark Web Analysis

The Dark Web—a segment of the Deep Web—is frequently utilized for malicious or illegal activities owing to its inherent anonymity [9]. The majority of information available for collection is in textual format, and analyses may encompass various elements such as page links [10], HTML tag data [11], community-specific information [12], or risk assessments derived from marketplace and vendor data [13]. Techniques commonly employed in this domain include machine learning, influence analysis, authorship identification, and social network analysis.

In machine learning methodologies, Dark Web or forum data is processed using machine or deep learning techniques to achieve detection, classification, and clustering outcomes. A crucial aspect of integrating these methodologies is the construction of data in a format conducive to the model, with feature extraction playing a key role. Studies have involved constructing a Dark Web dataset and assessing suspicious websites through ranking algorithms [14]. One investigation extracted diverse elements such as text, HTML, images, and graphs from the Dark Web and evaluated them using a learning-to-rank-based domain prioritization [15]. Another study combined five different term-weighting methods for feature generation and executed Dark Web classification [16].

Additional approaches include employing term extraction methods such as TF-IDF and bag of words (BOW), where classification was conducted using naive Bayes (NB), support vector machines (SVM), and linear regression (LR) [1, 17]. Another research focused on classifying marketplace attributes via group clustering and decision trees (DT) [18]. There have been endeavors to detect the Dark Web using edge computing and feature weighting [19], and to generate software quality metrics for classifying Dark Web characteristics [20]. Research involving the creation of a DOM tree from dark-web HTML and the application of a Graph CNN for phishing website detection is notable as well [11].

In a recent study, a web-to-graph conversion of the Dark Web was executed, and a classification methodology utilizing a graph neural network (GNN) was introduced [21]. Researchers have made various propositions for understanding the historical context of item sales and vendor activities in marketplaces [22], and for detecting malicious event scenarios therein [23]. A novel analysis was conducted on forum similarities based on the number of posts per class of malicious behavior [12]. Predominantly, these studies employed text data for Dark Web classification and processed texts through methods like BOW, DTM, n-gram, and term frequency-inverse document frequency (TF-IDF), converting words in the document into formats apt for base learning. Recently, deep learning algorithms such as neural networks (NN) and graph neural networks (GNN), along with traditional algorithms like NB, SVM, and DT, have been utilized for Dark Web classification. However, owing to the characteristic of Dark Web to overtly expose malicious behavior, these studies often end up using all textual information available, which may include non-essential content. The words associated with malicious activities are placed on the Dark Web both intuitively and explicitly. Thus, it is crucial to focus on learning only the essential keywords for Dark Web classification through dimensional reduction, rather than processing all words in the text.

B. Topic Modeling

Topic modeling, an unsupervised learning method, is designed to discern abstract topics within a document collection. This model processes documents and identifies key terms for user presentation. Employing this technology, one can ascertain the probabilities of keywords related to specific topics and the correlation probability between topics and documents. Researchers have introduced LSA to overcome the limitation of DTM or TF-IDF, which only calculates word frequency without considering the meaning of the words [24]. To address the issue where LSA does not account for topic distribution within a document, latent Dirichlet allocation (LDA) was proposed. LDA employs the Dirichlet distribution to estimate and present probabilities of topic distribution within a document and word distribution by topic [25]. Studies on supervised learning methods have been conducted to establish optimal solutions for the limitations of LDA, which cannot utilize labeled data [26, 27]. Furthermore, a study incorporated word embedding into topic modeling to learn both word frequency and relationships, thereby enhancing performance [28]. The approaches for topic extraction from brief texts, primarily produced by social networks [29, 30], or based on user preferences or intentions, have also been proposed [31].

Recent investigations have employed topic modeling to identify duplicate Dark Web forums, generating topics based on the presence or absence of interactions between forum users and the content of their posts [32]. In another study, vector spaces were developed to ascertain the themes

of Dark Web forums [33], while text mining and social network analysis (SNA) were utilized to discern themes within these forums [34]. The cohesion of topics predominantly discussed in each Dark Web forum was quantified using LDA [35], and similar forum patterns were detected using the hidden Markov model (HMM) and LDA [36]. Additionally, LDA has been applied to trend analysis on the Dark Web [37]. Most studies concerning the Dark Web leverage topic modeling to pinpoint the primary topics in specific forums, thereby confirming the trajectory of the forum and detecting illicit activities. Furthermore, these studies have examined the relationships between topics across various forums and analyzed trends within the Dark Web.

strategy combining mutual information and LDA [41]. Additional research classified Dark Web forums by integrating TF-IDF, random projection, and PCA [42], extracting features according to class, and applying machine learning [43]. This study also employed TF-IDF for feature extraction and executed classification using K-means and DT [18]. Numerous analyses of the Dark Web have been conducted using word2vec [44], and there is active research on transforming extracted text features into matrices *via* embedding and implementing deep learning [45, 46]. In the Dark Web context, text classification is predominantly utilized for preprocessing and feature extraction [16-18, 40-44].

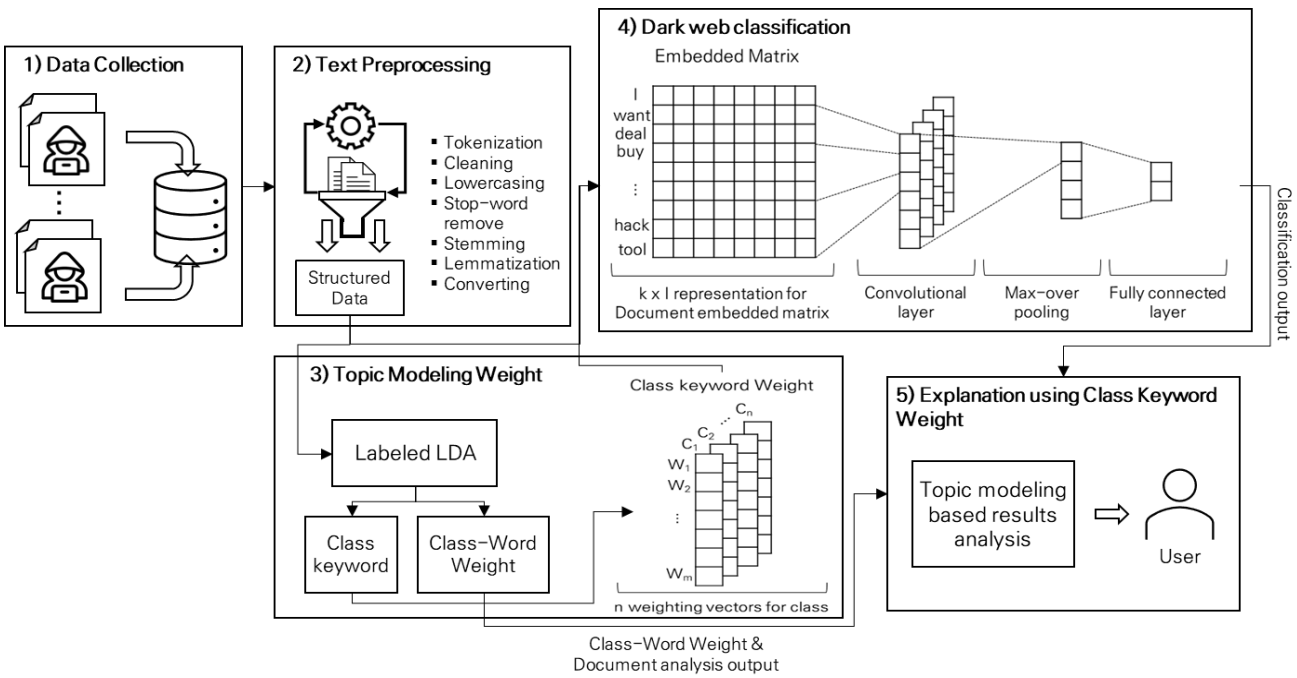

FIGURE 1. Illustration of dark web classification framework

C. Text Classification

Text classification is a pivotal aspect of natural language processing, and it is designed to assign labels or classes to textual units such as sentences, paragraphs, or documents [38]. It is applied in diverse fields including question answering, sentiment analysis, news categorization, and extends to various sources such as codes, web data, emails, chats, social media, and reviews. Its relevance is particularly notable in cybersecurity domains such as cyberbullying, abusive language detection, phishing websites, and malicious communities [39]. In the Dark Web where text predominates, the applicability of text classification is especially significant [40].

Researchers have introduced various methodologies for extracting and classifying keywords from the Dark Web. One approach utilized SVM for classification [40], whereas another adopted a two-step dimensionality reduction

III. PROPOSED METHOD

This research introduces a Dark Web classification framework that leverages TextCNN and topic modeling weights. The present methodology includes five key components, as illustrated in Fig. 1: 1) data collection, 2) text preprocessing, 3) topic modeling weights, 4) Dark Web classification, and 5) explanations utilizing topic keyword weights. Initially, we acquired relevant Dark Web data, and through preprocessing, convert this web data into a structured format suitable for model input. Subsequently, we calculate class-important word weights using label LDA (LLDA) in topic modeling. These weights are derived by identifying significant words within the class, assessing their impact, and prioritizing them accordingly. Thereafter, we apply these weights to the structured data to omit non-essential words and create an embedding matrix. This matrix is used as input for TextCNN to classify Dark Web content, and the impact of specific terms on the results is

evaluated by integrating the output, class-word weight, and Dark Web analysis outcome.

A. DATA Collection

For effective Dark Web classification, data with predefined classes are essential. In this study, two Dark Web datasets, previously validated through analyses, were utilized. DUTA-10k [14] includes 25 main classes and 10,367 samples, whereas CoDA [47] includes 10 classes and 8,855 samples. We restricted our data collection to English-language content from the Dark Web. In case of DUTA-10k, certain classes were notably smaller in size. Therefore, we eliminated classes that represented the bottom 1% of the entire Dark Web dataset, along with the "empty" class, which predominantly comprised non-textual or sparse content.

B. Text Preprocessing

Given that the collected data are authored by various users in distinct styles, it was imperative to process them into a uniform format. Text preprocessing was employed to optimize the unstructured data [48]. This preprocessing stage enabled us to acquire data suitable for model input. We first identified and eliminated empty data that lacked informational content. Following this, we removed white spaces, performed tokenization to break and create tokens [49], and eliminated stop words and special symbols that were irrelevant or minimally related to malicious behavior. Considering the computing language's distinction between uppercase and lowercase letters, we standardized all words to lowercase [50]. To reduce the vocabulary size, we combined words with similar meanings through lemmatization and stemming [51].

C. Topic Modeling Weight

In this study, LLDA [26] was utilized to determine the keywords for each class and their respective impact on each class. LLDA was developed to overcome the limitations of traditional LDA, which does not accommodate learning from data with predefined classes. Although LLDA employs Dirichlet distributions for probability calculations, similar to LDA, it distinguishes itself by learning the label set in advance. The comprehensive LLDA process for N words in a document, documents D, and topics K, is depicted in Fig. 2.

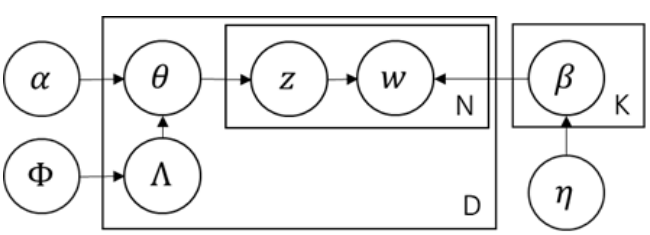

FIGURE 2. Graphical model of LLDA process

Here, α represents the Dirichlet distribution prior for the document and topic, η represents the Dirichlet distribution prior for the topic and word, and Φ represents the label distribution prior for the document. These three elements represent the hyperparameters of the LLDA. Furthermore, θ indicates the distribution of topics for each document combined with Λ, which maps the labels and topics, β indicates the multinomial distribution for the topic, and Λ represents the topic presence/absence indicator. The notation z represents the probability of assigned words for the topic, and w represents the observed word for the topic [28, 52, 53].

In this research, we employed the LLDA-based topic modeling weight process, inputting Dark Web-structured data to calculate the keywords for each class, the class-based impact of each word, and the frequency of each word within the document. Utilizing these outputs, we generated a class-specific keyword weight. The procedure of this algorithm is stated as follows.

1. For each topic k:
2. Generate $\beta \sim Dir(\eta)$
3. For each document d:
4. Generate word frequency for d
5. For each topic k:
6. Generate $\Lambda \sim \text{Bernoulli}(\Phi)$
7. Generate $\theta \sim Dir(\alpha \times \Lambda)$
8. For each i in $\{1, \dots, N\}$:
9. Generate $z_i \sim Mult(\theta)$
10. Generate $w_i \sim Mult(\beta)$
11. For each topic k:
12. Generate Class Keyword List
13. Generate Class-word Weight Matrix
14. For each Class
15. Generate Class keyword Weight Matrix
$= Class \times Class\ Keyword\ List$

The word frequency refers to the number of words in a document and can be calculated using (1). For all documents, the number of words was counted to create a DTM using the following formula:

$$DTM = \begin{pmatrix} d_1 t_1 & \cdots & d_1 t_N \\ \vdots & \ddots & \vdots \\ d_m t_1 & \cdots & d_m t_N \end{pmatrix}, \quad (1)$$

where d represents a set of documents composed of m in total and t represents a vocabulary set comprising N in total.

In this study, the class keyword list, encompassing words significantly influencing class classification from the learned vocabulary, was derived using LLDA. This list was curated by selecting words with substantial influence, based on the predetermined number of keywords for each class.

The corresponding equation, Equation (2), is articulated as follows:

$$Class\ Keyword\ List = Influence(t, l) \quad (2)$$

where t represents a specific class, and l indicates the number of keywords extracted. Using $Influence()$, we returned the desired number of keywords as a list.

The class-word weight matrix quantifies the impact of each word on a class in probabilistic terms. Utilizing this matrix, we can discern which words in a document are indicative of each class and the extent of their impact. Equation (3) facilitates this calculation, where P represents the degree of influence of a specific word on a class, c represents a class set consisting of k in total, and t represents a vocabulary set consisting of N in total.

$$CW\ Matrix = \begin{pmatrix} P(c_1 t_1) & \cdots & P(c_1 t_N) \\ \vdots & \ddots & \vdots \\ P(c_k t_1) & \cdots & P(c_k t_N) \end{pmatrix} \quad (3)$$

The class keyword weight matrix enumerates the high-impact keywords for each class and is formulated based on the class keyword list and the class-word weight matrix. This matrix ranks keywords according to their impact, specific to the specific class and the number of keywords extracted. Formula (4) is employed for this computation, where c represents a class set consisting of k in total. r represents the priority of the influence of the keyword and l represents the number of keywords extracted. Table I summarizes the matrix generated according to the Dark Web classes, with the keyword count set to five.

$$CKW\ Matrix = \begin{pmatrix} c_1 r_1 & \cdots & c_1 r_l \\ \vdots & \ddots & \vdots \\ c_k r_1 & \cdots & c_k r_l \end{pmatrix} \quad (4)$$

TABLE I
EXAMPLE OF CKW MATRIX EXTRACTION

Class	r_1	r_2	r_3	r_4	r_5
Gambling	Casino	game	time	slot	online
Drugs	gener	weight	money	buy	weed
Hacking	Hack	Facebook	time	account	hacker

This class keyword weight matrix, derived from the aforementioned methodology, was incorporated as an input into the Dark Web classification model alongside structured data. Both the class-word weight and word frequency were leveraged to provide comprehensive explanations to users.

D. Dark Web Classification

In the field of Dark Web classification, we developed a classifier utilizing both a basic CNN model and the TextCNN model [54]. This model demonstrated exceptional performance through hyperparameter tuning and the use of static vectors. In this approach, pre-trained word vectors are processed using a CNN for classification; the performance is dependent on the word vector employed for document embedding. Consequently, we generated pre-trained word vectors based on the class keyword weight matrix devised in Section III.C. Thereafter, each document was converted into an embedded matrix using these vectors. The pre-training of word vectors is exclusively conducted with words specified in the class keyword weight (CKW) matrix, which diminishes the size of the word vector created from thousands of existing words. This reduction accelerates processing speed as well as enhances performance by focusing on key terms. Newly trained word vectors are then utilized to transform each document into an embedded matrix, which serves as input for the classification model.

E. Explanation using Topic Keyword Weight

To elucidate the results obtained in Section III.D, we employed class-word weights and word frequencies for providing user-centric explanations. We ascertained the predicted class of the classified documents and assessed the keywords from the predicted class within the document, along with their proportion relative to the total word count in the document.

$$Influence\ rate = \frac{\sum Keyword_p}{\sum w_d} \quad (5)$$

$$Influence\ keyword = \{[w] | w \in d\ \cap keyword_p\} \quad (6)$$

The influence rate, calculated using Equation (5), signifies the ratio of keywords pertaining to the predicted class within the document's total word content. Employing these factors, we assessed the impact level of the predicted classes. The influenced keywords are determined using Equation (6), denoting the set of words present in the document that are also keywords of the predicted class. This enabled us to pinpoint keywords with direct influence.

IV. EXPERIMENT AND RESULT

In this section, the effectiveness of the proposed model is appraised utilizing two distinct datasets. To assess the performance of the model, a classification matrix was implemented, and further validation was conducted through comparative analysis with algorithms and prior studies previously employed in this domain.

A. Data set

The present experiments were performed using the DUTA-10k [14] and CoDA [47] datasets. The DUTA-10k dataset, an expansion of the initial labeled Dark Web dataset proposed in 2016 [1], encompasses a broader array of sources, thereby aggregating a more diverse set of onion sites. The data collection channels included the surface web, the Tor network, and various hyperlinks, leading to the accumulation of numerous standard and malicious webpages in a variety of languages. This dataset contains

25 main classes, 41 subclasses, and 10,367 entries. However, for our study, we focused solely on the main classes, excluding those in the bottom 1% owing to their small size and the "empty" classes predominantly filled with non-textual or sparse content. The experiments were then conducted using the classes and data sizes delineated in Table II.

TABLE II
DUTA-10K DATASET STATISTICS USED FOR DARK WEB

Class	count
Hosting & Software	1,949
Cryptocurrency	798
Down	714
Locked	682
Personal	419
C. Credit Cards	392
Social-Network	293
Drugs	290
Services	284
Pornography	226
Marketplace	189
Hacking	182
Forum	128
Total	6,524

TABLE III
CODA DATASET STATISTICS USED FOR DARK WEB

Class	count
Pornography	1,195
Drugs	1,172
Financial	1,003
Gambling	787
Cryptocurrency	763
Hacking	649
Arms/Weapons	599
Violence	485
Electronics	426
Others	2,921
Total	10,000

The CoDA dataset was developed to mitigate the class-data imbalance observed in DUTA-10k. Owing to the intrinsic complexities of the Dark Web, which blur the distinctions between illegal, legal, and other subclasses, these subclassifications were omitted. This adjustment reduced the original count of 25 classes to 10. Additionally, 18 mask identifiers were implemented to safeguard against personal data exposure. The statistical composition of the CoDA dataset is detailed in Table III.

B. Performance evaluation according to number of keywords

The effectiveness of the proposed model is contingent on the quantity and quality of the vocabulary assimilated into word vectors, which is related to the keyword criteria employed. Therefore, determining and applying an appropriate number of keywords for each dataset is crucial. To establish these criteria, we evaluated performance indicators relative to the keyword count. Table IV details the vocabulary size of the dataset in relation to the number of keywords. Notably, a reduction of over 300 keywords was observed compared to training without the use of topic modeling weights.

Table V showcases the experimental outcomes for the DUTA-10k dataset, indicating improved performance with an incremental increase in keyword count. Optimal performance was recorded when utilizing 90 keywords. Conversely, Table VI, detailing the CoDA dataset results, demonstrates that peak performance was attained with 20 keywords. These findings suggest that the requisite number of keywords varies depending on the extent of word overlap across different classes within the dataset. In scenarios with substantial word overlap, a higher keyword count is essential. Notably, the CoDA dataset exhibited proficient classification capabilities even with a reduced keyword count, attributed to the distinct differentiation of words in each class.

TABLE IV
VOCABULARY SIZE OF DATASET BASED ON NUMBER OF KEYWORDS

Number of keywords	Vocabulary size	
	DUTA-10k	CoDA
None	432,894	146,071
10	78	70
20	169	124
30	240	173
40	305	233
50	372	279
60	429	326
70	483	367
80	548	416
90	616	454
100	673	494

TABLE V
PERFORMANCE COMPARISON OF DUTA-10K BASED ON NUMBER OF KEYWORDS

Number of keywords	Precision	Recall	Accuracy	F1-score
None	0.7607	0.6220	0.6608	0.6220
10	0.6981	0.8832	0.7944	0.7526
20	0.8381	0.9061	0.8755	0.8645
30	0.8538	0.9240	0.8849	0.8817
40	0.8597	0.9136	0.9066	0.8798
50	0.8509	0.8885	0.9015	0.8595
60	0.9027	0.8957	0.8957	0.8962
70	0.9214	0.9175	0.9175	0.9173
80	0.9173	0.9153	0.9153	0.9144

90	**0.9224**	**0.9203**	**0.9203**	**0.9202**
100	0.9165	0.9124	0.9124	0.9126

TABLE VI
PERFORMANCE COMPARISON OF CODA BASED ON THE NUMBER OF KEYWORDS

Number of keywords	Precision	Recall	Accuracy	F1-score
None	0.8480	0.9109	0.8708	0.8725
10	0.9217	0.9554	0.9354	0.9361
20	**0.9656**	**0.9835**	**0.9695**	**0.9742**
30	0.9510	0.9771	0.9544	0.9633
40	0.9552	0.9817	0.9619	0.9676
50	0.9461	0.9730	0.9514	0.9586
60	0.9376	0.9664	0.9469	0.9507
70	0.9352	0.9716	0.9459	0.9521
80	0.9489	0.9762	0.9559	0.9616
90	0.9499	0.9464	0.9464	0.9463
100	0.9298	0.9686	0.9449	0.9471

C. Keyword Selection Results

Further analysis was conducted to ascertain the relevance of the extracted keywords to their respective classes. This examination verified that word vectors learn terms effectively representative of classes, facilitating the creation of an accurate embedding matrix. The top five keywords for each class in the DUTA-10k dataset are listed in Table VII, where user IDs and repetitive words were prevalent among Dark Web users. Based on the experimental findings from Section IV.B, a larger keyword size is necessary to discern between classes as the keyword count increases. In contrast, Table VIII, representing the CoDA dataset, predominantly features unique keywords with minimal overlap. This aligns with the observations from Section IV.B, where satisfactory results were achieved even with a smaller keyword set.

TABLE VII
KEYWORD EXTRACTION RESULTS BY CLASS IN DUTA-10K

Class	K1	k2	k3	k4	k5
Hosting & Software	com	file	php	hostinform	home
Cryptocurrency	btc	blockchainbdgpzk	adderss	blockchain	info
Down	hy2qtbfmejpdovq2	tag	site	ahmia	txt
Locked	file	index	home	universidad	login
Personal	file	home	universidad	arhivachovtj2jrp	localhost
C. Credit Cards	file	home	universidad	card	index
Social-Network	file	afg	sort	home	www
Drugs	file	home	universidad	torpharmzxholobn	product
Services	file	home	universidad	index	domain
Pornography	file	home	universidad	localhost	escort
Marketplace	file	product	type	bestshop3	home
Hacking	ccc	hack	turkish	armi	file
Forum	rrcc5uuudhh4oz3c	cmd	topic	profil	file

TABLE VIII
KEYWORD EXTRACTION RESULTS BY CLASS IN CODA

Class	K1	k2	k3	k4	k5
Pornography	video	free	sex	girl	teen
Drugs	gener	money	weight	buy	weed
Financial	money	gener	card	time	buy
Gambling	casino	game	time	online	slot
Cryptocurrency	money	crypto	bitcoin	btc	address
Hacking	hack	facebook	time	account	money
Arms/Weapons	gun	time	weapon	use	money
Violence	time	kill	murder	hitman	anonym
Electronics	money	gener	iphone	files	appl
Others	url	time	normal	onion	tor

D. Class Influence Analysis According to Weights

In this research, we performed a keyword influence analysis on the model results, utilizing weights derived from topic modeling. This analysis is instrumental in identifying the keywords that contribute to the prediction of classes and in assessing the extent of their impact within the document. The impacts and keywords corresponding to the predicted classes are illustrated in Fig. 3.

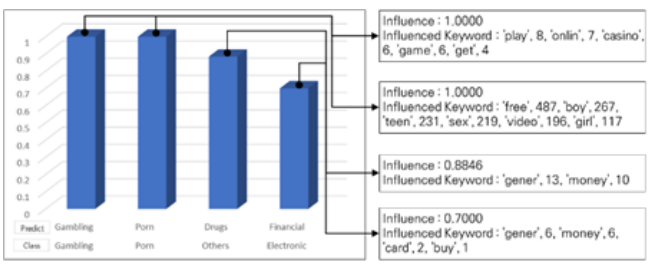

FIGURE 3. Examples of output explanation based on topic modeling weight

We randomly selected documents for examination to assess the impact results based on their classifications. For documents categorized under Gambling and Pornography, the frequency of keywords designated as influential was observed, affirming the precise detection of the class. Conversely, in the categories of Other and Electronic documents, the inclusion of irrelevant keywords led to less accurate classifications.

These insights enable us to furnish explanations that aid users in comprehending the model's rationale behind the classified outcomes. Additionally, by evaluating the influence of keywords, we can pinpoint those contributing to false positives and verify their impact.

E. Comparison of Classification Performance According to Model

To assess the effectiveness of our proposed method, we undertook a comparative analysis using various algorithms typically employed in text classification. For this experiment, we selected algorithms commonly used in text classification and maintained all parameters of each algorithm at their default settings. Table IX displays the performance comparison results based on the DUTA-10k dataset. These results demonstrate that our method outperforms the others. Notably, in the case of DUTA-10k, despite the challenge posed by overlapping words between classes, which could impede the text classification process, our method still achieved commendable classification performance.

TABLE IX
PERFORMANCE COMPARISON ACCORDING TO TEXT CLASSIFICATION ALGORITHM IN DUTA-10K

Model	Precision	Recall	Accuracy	F1-score
LLDA	0.6454	0.5431	0.5495	0.5431
ETM	0.4346	0.3130	0.3360	0.3130
NB	0.4800	0.2759	0.5793	0.2881
KNN	0.6245	0.5470	0.7212	0.5590
SVM	0.7788	0.7388	0.8269	0.7539
DT	0.6792	0.6765	0.7654	0.6759
RF	0.8162	0.6889	0.8009	0.7297
TextCNN	0.7607	0.6220	0.6608	0.6220
LSTM	0.7438	0.6180	0.7364	0.7377
BiLSTM	0.7622	0.6830	0.7661	0.7625
GRU	0.7637	0.6311	0.7545	0.7574
Proposed	**0.9224**	**0.9203**	**0.9202**	**0.9203**

The classification outcomes utilizing the CoDA dataset are summarized in Table X, affirming that our proposed method achieved the most effective performance. Additionally, owing to the more distinct differentiation of keywords by class in CoDA compared to that in DUTA-10k, other algorithms exhibited high performance.

TABLE X
PERFORMANCE COMPARISON ACCORDING TO TEXT CLASSIFICATION ALGORITHM IN CODA

Model	Precision	Recall	Accuracy	F1-score
LLDA	0.8559	0.8509	0.8509	0.8518
ETM	1.0000	0.8216	0.8216	0.9021
NB	0.9223	0.5910	0.7179	0.6524
KNN	0.8486	0.8443	0.8383	0.8449
SVM	0.9421	0.9118	0.9164	0.9258
DT	0.8407	0.8265	0.8247	0.8330
RF	0.9222	0.8269	0.8593	0.8651
TextCNN	0.8480	0.9109	0.8725	0.8708
LSTM	0.8347	0.8304	0.8338	0.8305
BiLSTM	0.8394	0.8324	0.8327	0.8336
GRU	0.8488	0.8570	0.8548	0.8507
proposed	**0.9552**	**0.9817**	**0.9676**	**0.9619**

Furthermore, we executed comparative evaluations of performance with other studies that employed the same datasets. For the DUTA-10k dataset, the inherent complexity of the Dark Web sometimes blurs the distinction between normal and malicious classes, rendering clear differentiation challenging. Consequently, certain studies have amalgamated these classes into a single category. Therefore, during comparative validation with studies using DUTA-10k, we scrutinized the number of classes utilized. Table XI showcases a comparison of performance evaluations with previous studies on DUTA-10k. The proposed method was confirmed to exhibit superior performance in classifying the same number of classes. In cases where the number of classes varied, models with fewer classified classes occasionally exhibited superior performance, but the proposed model consistently outperformed other models. This reinforces the capability of our proposed method to maintain its efficacy across various class classifications.

TABLE XI
PERFORMANCE COMPARISON WITH PREVIOUS STUDIES IN DUTA-10K

Model	Precision	Recall	Accuracy	F1-score	class
[1]	**0.9750**	**0.9740**	**0.9660**	**0.9740**	9
[15]	–	–	–	0.9414	7
[19]	0.9000	0.8900	0.8900	0.8900	6
[55]	0.8011	0.7994	–	0.8001	**13**
[56]	0.8200	0.7800	0.9608	0.8000	12
Proposed	**0.9224**	**0.9203**	**0.9202**	**0.9203**	**13**

Table XII offers a comparative performance analysis for the CoDA dataset. The proposed method outstripped previous studies in performance, underscoring its aptitude for effective Dark Web classification.

TABLE XII
PERFORMANCE COMPARISON WITH PREVIOUS STUDIES IN CODA

Model	Precision	Recall	Accuracy	F1-score
[47]	0.9251	0.9250	–	0.9249
[55]	0.9446	0.9445	–	0.9446
Proposed	**0.9552**	**0.9817**	**0.9676**	**0.9619**

V. CONCLUSIONS AND FUTURE WORK

This research introduces a Dark Web classification methodology integrating topic modeling weights with TextCNN. The Dark Web data collected included an extensive vocabulary, not all of which was conducive to effective classification. Thus, a preprocessing step to eliminate extraneous words was essential. We tackled this challenge using topic modeling weights. Initially, structured data were created via text preprocessing, followed by the identification of principal keywords for each class using LLDA. These keywords were then applied to the Dark Web data to exclude irrelevant words, resulting in a significant reduction of the word vector size by approximately 300 times. The word vectors were instrumental in forming an embedded matrix, which served as the input for the Dark Web classification model. We utilized topic modeling weights in our classification process to discern the influential keywords and their respective impact. The efficacy of our proposed approach was assessed using two labeled Dark Web datasets, validated through comparative analysis with various text classification algorithms and prior studies. The experimental outcomes verified the efficiency of the proposed method in reducing the vocabulary size required for learning and its superior performance.

In future research, we intend to develop real-time classification techniques for the Dark Web, considering its dynamic nature. Both accuracy and processing speed are paramount in real-time scenarios. Therefore, the removal of unnecessary words through topic modeling weights is anticipated to yield complementary benefits. Additionally, we aim to devise a strategy that consolidates topic modeling weights and deep learning within a unified model, to elucidate the aspects of the neural network influencing the overall results.

REFERENCES

[1] M. W. Al Nabki *et al.*, "Classifying illegal activities on tor network based on web textual contents" in *Proc. 15th Conference of the European Chapter of the Association for Computational Linguistics, Long [Papers]*, vol. 1, no. 2017, Apr., pp. 35-43. doi:10.18653/v1/E17-1004.

[2] G. Cascavilla *et al.*, "Cybercrime threat intelligence: A systematic multi-vocal literature review," *Comput. Sec.*, vol. 105, p. 102258, 2021. doi:10.1016/j.cose.2021.102258.

[3] J. Saleem *et al.*, "The anonymity of the dark web: A survey," *IEEE Access*, vol. 10, pp. 33628-33660, 2022. doi:10.1109/ACCESS.2022.3161547.

[4] M. Balduzzi and V. Ciancaglini, 2015, "Cybercrime in the deep web," Black Hat. Amsterdam: EU.

[5] S. R. Rahamat Basha and J. K. Rani, "A comparative approach of dimensionality reduction techniques in text classification," *Eng. Technol. Appl. Sci. Res.*, vol. 9, no. 6, pp. 4974-4979, 2019. doi:10.48084/etasr.3146.

[6] K. N. Singh et al., "A novel approach for dimension reduction using word embedding: An enhanced text classification approach," *International Journal of Information Management Data Insights*, vol. 2, no. 1, p. 100061, 2022. doi:10.1016/j.jjimei.2022.100061.

[7] W. Alkhatib et al., "Multi-label text classification using semantic features and dimensionality reduction with autoencoders" in *Language, Data, and Knowledge: First International Conference, LDK 2017*, 2017, pp. 380-394. doi:10.1007/978-3-319-59888-8_32.

[8] M. Umer et al., "Fake news stance detection using deep learning architecture (CNN-LSTM)," *IEEE Access*, vol. 8, pp. 156695-156706, 2020. doi:10.1109/ACCESS.2020.3019735.

[9] R. Basheer and B. Alkhatib, "Threats from the dark: A review over dark web investigation research for cyber threat intelligence," *J. Comput. Netw. Commun.*, vol. 2021, pp. 1-21, 2021. doi:10.1155/2021/1302999.

[10] A. Alharbi *et al.*, "Exploring the topological properties of the Tor Dark Web," *IEEE Access*, vol. 9, pp. 21746-21758, 2021. doi:10.1109/ACCESS.2021.3055532.

[11] L. Ouyang and Y. Zhang, "Phishing web page detection with html-level graph neural network" in 20th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom), vol. 2021. IEEE. IEEE, 2021, Oct., pp. 952-958. doi:10.1109/TrustCom53373.2021.00133.

[12] H. Alnabulsi and R. Islam "Identification of illegal forum activities inside the dark net international conference on machine learning and data engineering (iCMLDE)," IEEE, 2018, Dec., pp. 22-29.

[13] D. R. Hayes *et al.*, "A framework for more effective dark web marketplace investigations," *Information*, vol. 9, no. 8, p. 186, 2018. doi:10.3390/info9080186.

[14] M. W. Al-Nabki *et al.*, "Torank: Identifying the most influential suspicious domains in the tor network," *Expert Syst. Appl.*, vol. 123, pp. 212-226, 2019. doi:10.1016/j.eswa.2019.01.029.

[15] M. W. Al Nabki *et al.*, "Supervised ranking approach to identify infLuential websites in the darknet," *Appl. Intell.*, vol. 53, no. 19, pp. 22952–22968, 2023. doi:10.1007/s10489-023-04671-9.

[16] T. Sabbah *et al.*, "Hybridized term-weighting method for dark web classification," *Neurocomputing*, vol. 173, pp. 1908-1926, 2016. doi:10.1016/j.neucom.2015.09.063.

[17] S. He *et al.*, "Classification of illegal activities on the dark web" in *Proc. 2nd International Conference on Information Science and Systems*, 2019, Mar., pp. 73-78. doi:10.1145/3322645.3322691.

[18] S. Nazah *et al.*, "An unsupervised model for identifying and characterizing dark web forums," *IEEE Access*, vol. 9, pp. 112871-112892, 2021. doi:10.1109/ACCESS.2021.3103319.

[19] R. Li *et al.*, "Edge-based detection and classification of malicious contents in tor darknet using machine learning," *Mob. Inf. Syst.*, vol. 2021, pp. 1-13, 2021. doi:10.1155/2021/8072779.

[20] G. Cascavilla *et al.*, "'When the Code becomes a Crime Scene' Towards Dark Web Threat Intelligence with Software Quality Metrics" in IEEE International Conference on Software Maintenance and Evolution (ICSME), vol. 2022. IEEE, 2022, Oct., pp. 439-443. doi: 10.1109/ICSME55016.2022.00055

[21] T. Guo and B. Cui, "Web page classification based on graph neural network" in *Innovative Mobile and Internet Services in Ubiquitous Computing, Proceedings 15th International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing*. Springer International Publishing, 2022, pp. 188-198. doi:10.1007/978-3-030-79728-7_19.

[22] M. Ball et al., "Data capture and analysis of darknet markets". Available at SSRN 3344936, *SSRN Journal*, 2021. doi:10.2139/ssrn.3344936.

[23] Z. Ursani et al., "The impact of adverse events in darknet markets: An anomaly detection approach" in IEEE European Symposium on Security and Privacy Workshops (EuroS&PW), vol. 2021. IEEE, 2021, Sept., pp. 227-238. doi:10.1109/EuroSPW54576.2021.00030.

[24] P. W. Foltz, "Latent semantic analysis for text-based research," *Behav. Res. Methods Instrum. Comput.*, vol. 28, no. 2, pp. 197-202, 1996. doi:10.3758/BF03204765.

[25] D. M. Blei et al., "Latent dirichlet allocation," *J. Mach. Learn. Res.*, vol. 3, no. Jan., pp. 993-1022, 2003.

[26] D. Ramage *et al.*, "Labeled LDA: A supervised topic model for credit attribution in multi-labeled corpora" in *Proc. 2009 Conference on Empirical Methods in Natural Language Processing*, 2009, Aug., pp. 248-256. doi:10.3115/1699510.1699543.

[27] J. Mcauliffe and D. Blei, "Supervised topic models," *Adv. Neural Inf. Process. Syst.*, vol. 20, 2007.

[28] A. B. Dieng *et al.*, "Topic modeling in embedding spaces," *Trans. Assoc. Comp. Linguist.*, vol. 8, pp. 439-453, 2020. doi:10.1162/tacl_a_00325.

[29] L. Shi et al., "Dynamic topic modeling via self-aggregation for short text streams," *Peer-to-Peer Netw. Appl.*, vol. 12, no. 5, pp. 1403-1417, 2019. doi:10.1007/s12083-018-0692-7.
[30] F. Kou et al., "A semantic modeling method for social network short text based on spatial and temporal characteristics," *J. Comp. Sci.*, vol. 28, pp. 281-293, 2018. doi:10.1016/j.jocs.2017.10.012.
[31] L. Shi et al., "A user-based aggregation topic model for understanding user's preference and intention in social network," *Neurocomputing*, vol. 413, pp. 1-13, 2020. doi:10.1016/j.neucom.2020.06.099.
[32] S. A. Ríos and R. Muñoz, "Dark web portal overlapping community detection based on topic models" in *Proc. ACM SIGKDD Workshop on Intelligence and Security Informatics*, 2012, Aug., pp. 1-7. doi:10.1145/2331791.2331793.
[33] H. M. Alghamdi and A. Selamat, "Topic detections in Arabic dark websites using improved vector space model" in 4th Conference on Data Mining and Optimization (DMO), vol. 2012. IEEE, 2012, Sept., pp. 6-12. doi:10.1109/DMO.2012.6329790.
[34] G. L. Huillier *et al.*, "Topic-based social network analysis for virtual communities of interests in the dark web," *ACM SIGKDD Explor. Newsl.*, vol. 12, no. 2, pp. 66-73, 2011. doi:10.1145/1964897.1964917.
[35] M. Rashed *et al.* in *Proc. 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*. Evaluation of extremist cohesion in a darknet forum using ERGM and LDA, 2019, Aug., pp. 899-902.
[36] N. Tavabi *et al.*, "Characterizing activity on the deep and dark web" in *Companion Proceedings of the 2019 World Wide Web Conference*, 2019, May, pp. 206-213. doi:10.1145/3308560.3316502.
[37] K. Porter, "Analyzing the DarkNetMarkets subreddit for evolutions of tools and trends using LDA topic modeling," *Digit. Investig.*, vol. 26, pp. S87-S97, 2018. doi:10.1016/j.diin.2018.04.023.
[38] S. Minaee *et al.*, "Deep learning--Based text classification: A comprehensive review," *ACM Comput. Surv.*, vol. 54, no. 3, pp. 1-40, 2022. doi:10.1145/3439726.
[39] H. Liu *et al.*, "A fuzzy approach to text classification with two-stage training for ambiguous instances," *IEEE Trans. Comput. Soc. Syst.*, vol. 6, no. 2, pp. 227-240, 2019. doi:10.1109/TCSS.2019.2892037.
[40] C. A. S. Murty and P. H. Rughani, "Dark web text classification by learning through SVM optimization," *J. Adv. Inf. Technol.*, vol. 13, no. 6, 2022. doi:10.12720/jait.13.6.624-631.
[41] M. Faizan and R. A. Khan, "A two-step dimensionality reduction scheme for dark web text classification" in *Ambient Communications and Computer Systems: RACCCS*. Singapore: Springer, 2020, pp. 303-312. doi:10.1007/978-981-15-1518-7_25.
[42] T. Sabbah and A. Selamat, "'Hybridized feature set for accurate Arabic dark web pages classification' in *Intelligent Software Methodologies, Tools and Techniques*," *Proc. 14*: 14th International Conference, SoMet 2015, Naples, Italy, September 15-17, 2015. Springer International Publishing, 2015, pp. 175-189. doi:10.1007/978-3-319-22689-7_13.
[43] A. H. M. Alaidi *et al.*, "'Dark web illegal activities crawling and classifying using data mining techniques" iJIM, 2022, vol. 16, no. 10, pp. 123. DOI: 10.3991/ijim.v16i10.30209.
[44] N. Deguara *et al.*, "Threat miner-A text analysis engine for threat identification using dark web data" in IEEE International Conference on Big Data (Big Data), 2022, pp. 3043-3052. doi:10.1109/BigData55660.2022.10020397.
[45] H. Ma *et al.*, "Dark web traffic detection method based on deep learning" in 10th Data Driven Control. and Learning Systems Conference (DDCLS), vol. 2021. IEEE. IEEE, 2021, May, pp. 842-847. doi:10.1109/DDCLS52934.2021.9455619.
[46] M. Kadoguchi *et al.*, "Deep self-supervised clustering of the dark web for cyber threat intelligence" in IEEE International Conference on Intelligence and Security Informatics (ISI), vol. 2020. IEEE, 2020, Nov., pp. 1-6. doi:10.1109/ISI49825.2020.9280485.
[47] Y. Jin *et al.*, 2022, Shedding new light on the language of the dark web. arXiv preprint arXiv:2204.06885. doi:10.18653/v1/2022.naacl-main.412.
[48] B. Thapa, 2022, Sentiment Analysis of Cyber Security content on Twitter and Reddit. arXiv preprint arXiv:2204.12267. doi:10.5121/csit.2022.120708.
[49] L. Hickman *et al.*, "Text preprocessing for text mining in organizational research: Review and recommendations," *Organ. Res. Methods*, vol. 25, no. 1, pp. 114-146, 2022. doi:10.1177/1094428120971683.
[50] G. George *et al.*, "Big data and data science methods for management research," *Acad. Manag. J.*, vol. 59, no. 5, pp. 1493-1507, 2016. doi:10.5465/amj.2016.4005.
[51] S. Bird *et al.*, *Natural Language Processing with Python: Analyzing Text with the Natural Language Toolkit*. O'Reilly Media," Inc., 2009.
[52] B. Jing *et al.*, "Cross-domain labeled LDA for cross-domain text classification" in *IEEE International Conference on Data Mining (ICDM)*, vol. 2018. IEEE, 2018, Nov., pp. 187-196. doi:10.1109/ICDM.2018.00034.
[53] Y. Bai and J. Wang, "News classifications with labeled LDA" in *7th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K)*, vol. 1, no. v. IEEE, 2015. doi:10.5220/0005610600750083.
[54] Y. Kim, 2014, Convolutional neural networks for sentence classification. arXiv preprint arXiv:1408.5882.
[55] Y. Jin *et al.*, 2023, DarkBERT: A language model for the dark side of the Internet. arXiv preprint arXiv:2305.08596. doi:10.18653/v1/2023.acl-long.415.
[56] G. Cascavilla et al., *"Illicit Darkweb Classification via Natural-Language Processing: Classifying Illicit Content of Webpages Based on Textual Information" SECRYPT2022*, 2022. doi:10.5220/0011298600003283.

Gun-Yoon Shin received the M.S. and Ph.D. degrees in computer engineering from Gachon University, Republic of Korea, in 2018 and 2023, respectively. Currently, he is a postdoctoral researcher at Gachon University, Korea. His research interests include authorship attribution, unknown attack detection, network anomaly detection, information security, machine learning, and artificial intelligence.

Younghoan Jang earned his Master's degree in Computer Science from Gachon University in 2017 and obtained his Ph.D. in Computer Science from the same university in 2021. His research interests include network traffic, anomaly detection, IoT, and EMS.

Dong-Wook Kim received the M.S. and Ph.D. degrees in computer engineering from Gachon University, Republic of Korea, in 2017 and 2022, respectively. He is currently a postdoctoral researcher at Gachon University. His research interests include insider threats, information security, data mining, and machine learning.

Sungjin Park received his B.S degree in Cyber Security from Ajou University in 2020. he is currently a research engineer in LIG Nex1. He is interested in Cyber Threat Intelligence, AI Security and Offensive.

A-ran Park received her M.S. degree from the Graduate School of Information Security, Korea University, in 2017. She has been with LIG Nex1, since 2017. Her current research interests include digital forensics, information security and cyber security.

YOUNGHWAN KIM received the M.S. degree in computer science from HANKUK University of Foreign Studies in 1993. He has been working with LIG Nex1, since 2010. His research interests include combat management system of ship and network security.

Myung-Mook Han received his M.S. degree in computer science from the New York Institute of Technology in 1987 and his Ph.D. in information engineering from Osaka City University in 1997. From 2004 to 2005, he was a visiting professor at the Georgia Tech Information Security Center (GTISC), Georgia Institute of Technology. Currently, he is a professor in the Department of Software, Gachon University, Korea. His research interests include information security, intelligent systems, data mining, and big data.