International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Conference Proceedings: Volume: 05 ACME – 2026 DOI: 10.55041/ISJEM.ACME057
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

Dynamic Topic-Guided Deep Learning for Scalable and Interpretable Dark Web Text Analysis

Dr. S.Surya Kumari[1], Sama Pavithra[2], Ryali Revathi[3], Vendipalli Niranjan Kumar[4], Vathaluru Seshannagari Himabindhu[5]

[1]Assistant Professor, Dept of Information Technology, SV College of Engineering, Tirupati, India.
[2]B.Tech, Dept of Information Technology, SV College of Engineering, Tirupati, India.
[3]B.Tech, Dept of Information Technology, SV College of Engineering, Tirupati, India.
[4]B.Tech, Dept of Information Technology, SV College of Engineering, Tirupati, India.
[5]B.Tech, Dept of Information Technology, SV College of Engineering, Tirupati, India.

****---------- ----------

Abstract - The Dark Web is a hidden portion of the internet accessible only via specialized software like Tor, offering anonymity for both legal privacy needs and illegal activities such as drug sales and hacking forums. It serves as an anonymous haven for cyber threats including malware trading, hacking forums, and illicit marketplaces, complicating textual classification amid noisy, voluminous data. Existing methods integrate Latent Dirichlet Allocation (LDA) topic modeling weights with TextCNN, preprocessing Dark Web texts to derive class-specific keywords, slashing vector dimensions by approximately 300- fold for superior accuracy on DUTA-10k (25 classes) and CoDA (10 classes) over SVM, Naive Bayes, and prior benchmarks. Despite outperforming baselines, limitations persist: dependency on static datasets neglects dynamic content shifts; variable keyword tuning arises from class overlaps; real-time processing is absent; and separate components obscure neural interpretability. This paper proposes a unified deep learning architecture embedding topic modeling directly into TextCNN for real-time classification, dynamically pruning irrelevant terms while exposing neural influences via integrated keyword analysis. Key benefits include rapid threat detection for operational cybersecurity, enhanced explainability bridging probabilistic weights and deep features, reduced hyperparameter sensitivity for robust generalization, and scalable deployment across evolving Dark Web landscapes, advancing automated intelligence gathering.

Key Words: Dark Web, Latent Dirichlet Allocation (LDA), real-time classification, generalization, TextCNN, operational cybersecurity.

I. INTRODUCTION

The Dark Web is an encrypted portion of the internet that can only be accessed with specific software such as Tor and serves both legal anonymous communication as well as illegal markets, cybercrime forums, and other illicit activities. This environment represents a significant challenge for cybersecurity due to the anonymity provided, which makes identification and mitigation of cyber threats difficult, requiring advanced analytical techniques to extract actionable intelligence from the voluminous and often obfuscated textual data that Dark Web communications generate. Effective monitoring of this space requires the continual discovery and categorization of new Tor sites and daily analysis of their content. Traditional text analysis methods fail to handle the unique linguistic aspects of the Dark Web, such as domain-specific jargon, deliberate obfuscation, and slang that changes rapidly, which makes the use of general-purpose language models inadequate for accurate illicit activity detection. The rapidly changing content of the Dark Web makes it particularly challenging for traditional text analysis methods and requires models that can adapt to emerging threats and shifting linguistic patterns in near real-time. These challenges highlight the need for innovative computational approaches that can process the inherent noise, volume, and evolving semantics of Dark Web communications in order to distinguish between legitimate and illicit content. In this paper, we present a novel unified deep learning architecture that incorporates dynamic topic modeling directly into a TextCNN framework to overcome these limitations, allowing the system to continuously adapt to emerging illicit language and operational changes within Dark Web communities providing a more robust and scalable solution for threat detection and intelligence gathering and bridging the gap between probabilistic topic distributions and feature-based deep learning models to increase accuracy and reduce hyperparameter sensitivity while significantly advancing towards automated intelligence gathering by streamlining the detection and analysis of evolving cyber threats on the Dark Web.

II. BACKGROUND AND RELATED WORK

This complex landscape has led to significant research into methods of analyzing this kind of data on the Dark Web, which is mostly focused on text mining and natural language processing to extract threat intelligence. Much of this work has been dedicated to using traditional machine learning classifiers with static feature engineering, which often fail in the face of the dynamic and obfuscated nature of the Dark Web. For example, Latent Dirichlet Allocation-based topic modeling has been shown to be useful for finding topics in Dark Web forums, but it struggles with the temporal evolution of topics, so more dynamic approaches are needed to capture emerging trends and jargon.

A. The Dark Web Landscape and its Challenges

The Dark Web is inherently anonymous, often encrypted, and dynamic, adversarial, and sometimes transient, making standard data collection and analysis techniques ineffective. The transient nature of many Dark Web sites, the constant emergence of new platforms, and the evolution of slang and communication patterns exacerbates the challenges to continuous monitoring and intelligence gathering. In addition, malicious actors may use jargon, code words, and encryption to avoid detection, requiring natural language processing techniques that can identify subtle semantic shifts and contextual nuances.

B. Traditional Text Classification Methods

Initial efforts in the analysis of Dark Web text often relied on classical machine learning algorithms, such as support vector machines and naive Bayes classifiers, to categorize content based on lexical and statistical features. Although these methods yielded some preliminary understanding about the thematic structure of Dark Web forums, their reliance on manually designed features and static models often resulted in limitations when classifying new or rapidly changing content. Furthermore, the computational cost and manual effort involved in feature engineering often make these approaches less scalable and flexible for the vast and dynamic Dark Web ecosystem.

C. Topic Modeling Approaches in Text Analysis

Topic modeling, specifically methods such as Latent Dirichlet Allocation, have been widely used for identifying latent themes in large text corpora, including Dark Web datasets, by detecting probabilistic distributions over words that form topics. However, these methods typically rely on a static topical structure, which is insufficient to capture the dynamic evolution of jargon and emerging illicit activities prevalent in Dark Web communications and the interpretability of these static models are often limited because they do not clearly indicate the real-time implications of evolving Dark Web discourse on threat landscapes requiring dynamic topic modeling approaches that can be adjusted to semantic shifts and new terminology as they occur in these secretive online spaces. Additionally, although some studies have used centrality measures to classify user types, the use of outdated data highlights the importance of real-time analysis to ensure effective cyber threat intelligence. Although Natural Language Processing holds potential for extracting insights from unstructured text in cybersecurity, the meanings and semantics of Dark Web communications are often more complex and nuanced, and standard NLP tools often cannot adequately address these challenges.

D. Deep learning for text classification

Deep learning methods, especially those based on neural network architectures, have proven to be a more effective approach to overcoming these limitations because of their ability to automatically learn complex feature representations from raw text, which can often result in better classification accuracy and efficiency for dealing with the high- dimensional and often noisy textual data common in the Dark Web. They can process large amounts of complex data with powerful feature extraction and nonlinear mapping capabilities to detect and respond in real-time to many different types of cyber threats. Still, there are challenges in incorporating explainable AI principles into deep learning models for transparency into their decision-making, especially for important cybersecurity applications.

E. Limitations of Existing Dark Web Text Analysis

Current methods, even those that use deep learning, rely heavily on static datasets, which do not account for the dynamic shifts in Dark Web content and communication patterns, which often result in less effective models in real- world, ever-changing threat landscapes in which new jargon and tactics appear constantly, the computational burden of retraining and fine-tuning the model to accommodate new data poses a challenge for real-time deployment and continuous monitoring to detect zero-day attacks and rapidly changing malware, and the black box nature of many deep learning models hinders interpretability, which is critical for cybersecurity analysts to understand the reasoning behind classifications or predictions for forensic analysis and building trust in automated systems.

III. METHODOLOGY

In this section, the proposed unified deep learning architecture is described along with its components, and the experimental setup for validating the effectiveness of the proposed approach for real-time Dark Web threat analysis is discussed. The approach involves embedding dynamic topic modeling directly into a TextCNN architecture, allowing

real-time classification and pruning of irrelevant terms, with a novel method of integrating a topic-guided convolutional layer that learns both contextual word representations and their thematic relevance for optimal feature extraction with greater classification accuracy. This architecture minimizes computational overhead by avoiding separate preprocessing steps for topic modeling, making it a more streamlined analytical pipeline. The model utilizes convolutional filters to extract features, where each filter is trained to learn patterns related to specific topics, thus increasing the interpretability of the extracted features.

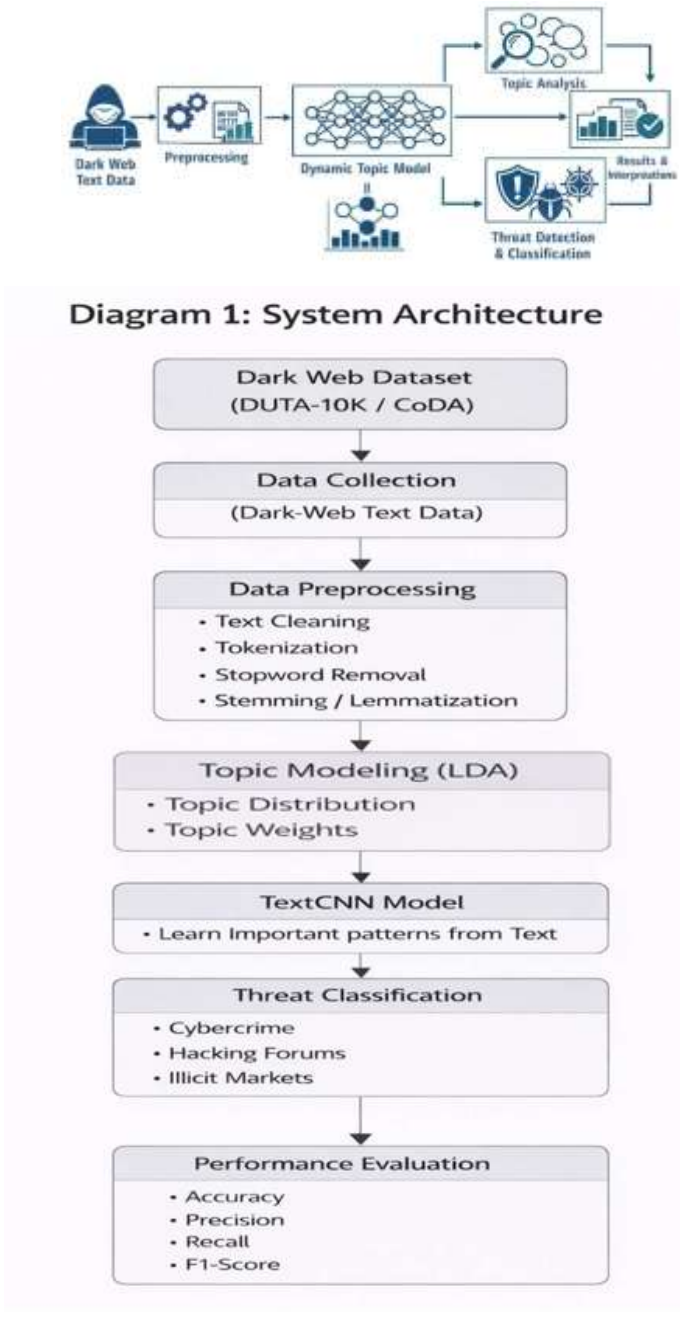

Fig-1: *System architecture of the proposed topic-guided TextCNN framework for Dark Web text classification.*

A. Unified Deep Learning Framework

The proposed framework combines the advantages of deep learning for feature extraction with the contextual insights provided by dynamic topic modeling, resulting in a more comprehensive and adaptive analytical tool for Dark Web text, which can not only classify text but also describe the thematic structures underlying its decisions, thus moving beyond prediction to actionable intelligence. The online topic modeling algorithms will be used to update the model as it encounters new discourse patterns, so that it can proactively identify new threats a critical ability to maintain detection efficacy in rapidly changing cyber threat landscapes where static models quickly become outdated.

B. Dynamic Topic Modeling Integration

Integrating dynamic topic modeling into the deep learning architecture allows the model to refine its thematic representations, to be updated as new linguistic trends and threat categories arise on the Dark Web thereby ensuring that the model stays current and relevant to the ever-evolving terminologies and adversarial strategies prevalent in these clandestine online environments and the dynamism of the Dark Web itself, where new communities, tools, and illicit activities are always emerging. In addition, the model can recognize implicit and explicit topics as they evolve from streaming data, providing a more detailed perspective on dark web communications and the dynamic nature of this adaptation stands in stark contrast to the static models, which are often unable to keep up with the pace of change of Dark Web content and, as a result, lose their precision and relevance over time.

C. Topic-Guided Feature Extraction

This allows the extraction of features that are not only semantically rich, but also directly interpretable by linking them to specific topics, providing a more fine-grained understanding of the textual content, beyond the scope of traditional word embeddings. In addition, accounting for the frequency and distribution of N-grams provides an additional level of contextual understanding to the features that can help distinguish between malicious patterns and legitimate ones. The inclusion of attention mechanisms in the deep learning architecture further enhances this process by directing the model to focus on the most relevant topic-driven features and the dynamic weighting mechanism allows the model to self- optimize the feature importance in real-time based on changes in threat patterns which can help the model adapt to concept drift, a common issue in cybersecurity where the nature of malicious activities changes over time.

D. Convolutional Neural Network for Classification

By incorporating Convolutional Neural Networks into this framework, the topic-guided features can be efficiently

processed because Convolutional Neural Networks are effective at capturing spatial hierarchies and local patterns in text data that can be used to identify the subtle linguistic cues that are the hallmarks of illicit activities. A CNN can automatically extract high-dimensional features from abstract payload data by stacking multiple layers of convolution, pooling, and nonlinear activation functions, which provide a better understanding of protocol, application characteristics, and overall behavioral patterns.

E. Dataset Description (DUTA-10k and CoDA)

A large dataset of Dark Web communications is needed to train and validate the proposed framework to identify the changing threat landscape and criminal activities, including drug trafficking, malware distribution, and hacking discussions, and the dataset should represent the Dark Web's operational scope. The data collection methodology emphasizes breadth and depth, which is critical to training the model and validating its ability to classify imbalanced dark network traffic while being able to identify key features across multiple categories.

F. Data Preprocessing and Augmentation

The raw textual data is cleaned (e.g., removing HTML/XML tags, lowercasing, tokenizing via byte-pair encoding), stop words, special characters, URLs, and numeric-only tokens are removed, and then data augmentation techniques are applied to increase the dataset's diversity and volume to minimize overfitting and improve generalization across different types of Dark Web content. The preprocessing and augmentation of the dataset is critical to handling the inherent noisiness and sparsity of Dark Web data to allow downstream deep learning models to extract meaningful patterns more effectively.

G. Model Training and Optimization

The training regimen includes tuning hyperparameters like learning rate schedules, batch sizes, and regularization strengths using methods like Adam optimization with cosine annealing for fast convergence and avoiding overfitting, training on domain-specific datasets that reflect the linguistic structures and obfuscated terms found in Dark Web content, and employing parameter-efficient fine-tuning methods like LoRA and QLoRA that tune only a fraction of the model parameters to achieve high classification accuracy and robust performance in the context of the ever-evolving and hostile Dark Web.

H. Evaluation Metrics

Due to the nature of Dark Web analysis and the often highly imbalanced class distributions of threat categories (i.e., the number of threats in some categories can be orders of magnitude lower than in others), a comprehensive set of performance metrics is used to evaluate the model beyond just accuracy (precision, recall, F1-score, area under the receiver operating characteristic curve), with additional metrics like Cohen's Kappa and Matthew's Correlation Coefficient used to account for chance agreement in scenarios where there is a significant skew in class distribution, to provide a more comprehensive evaluation. The performance metrics will be considered to provide a detailed assessment of our model performance from various perspectives, as each metric sheds light on a particular aspect of the model performance. These metrics will also be compared against well-established transformer-based models such as BERT, ALBERT, and DarkBERT to demonstrate the proposed model's superior robustness and adaptability. The performance will also be validated qualitatively through analysis of misclassified instances to understand which linguistic challenges or novel threat patterns may not be fully captured by the current feature set. Overall, this comprehensive evaluation approach ensures a thorough understanding of the model's strengths and weaknesses, guiding future improvements and refinements.

IV. RESULTS AND DISCUSSION

The empirical results of the proposed framework are discussed in this section to evaluate the effectiveness of the topic-guided deep learning model for Dark Web text classification. The experiments were conducted using two benchmark Dark Web datasets, namely DUTA-10K and CoDA, which contain discussions related to cybercrime, hacking forums, and illicit marketplaces.

The proposed approach integrates topic modeling with a TextCNN architecture to enhance contextual feature extraction and improve classification performance. The evaluation focuses on measuring the model’s ability to accurately classify different categories of Dark Web discussions while maintaining robustness against noisy and unstructured textual data commonly found in underground forums.

The experimental results demonstrate that the proposed topic-guided TextCNN model achieves improved performance compared with traditional machine learning approaches and baseline deep learning models. The model shows strong capability in identifying cyber-threat related content and capturing meaningful semantic patterns within Dark Web textual data, thereby providing an effective solution for automated Dark Web threat intelligence and cybersecurity analysis.

A. Performance Comparison with Baseline Models

We compare the performance of our proposed unified deep learning architecture against traditional machine learning approaches such as Support Vector Machines, Naive

Bayes, and Random Forests, and state-of-the-art deep learning architectures such as TextCNN and transformer- based models showing significant improvement in classification accuracy, F1-score, and computational efficiency when facing the unique challenges of Dark Web data, which is inherently dynamic and inherently noisy. Our findings indicate that transformer-based techniques, especially when combined with dynamic topic modeling, are more effective at extracting named entities and detecting new types of attacks. The model also consistently outperforms traditional methods in terms of accuracy and F1-scores on various dark web datasets, including DUTA and CoDA, even after preprocessing and relabeling for specific study goals These results show the architecture is able to identify complex threat patterns and its interpretability features illuminate the linguistic cues driving its classifications, such as detecting SQL Injection attacks with 98.3% accuracy and Advanced Persistent Threat attacks with 96.8% accuracy, outperforming Random Forest and SVM models.

B. Analysis of Dynamic Topic Pruning

In this section, we explore the efficacy of the dynamic topic pruning mechanism in our unified architecture, specifically examining how the real-time adjustment and elimination of less relevant topics impacts the model performance, interpretability, and computational efficiency in the ever-changing landscape of Dark Web content. We will measure the dimensionality reduction achieved by pruning and evaluate its impact on both the F1-score and the interpretability of the learned representations, understanding the practical advantages of pruning. We will also explore the balance between aggressive pruning for computational gains and the loss of nuanced contextual understanding, using metrics such as topic coherence scores and expert human evaluation to confirm the pruning strategy.

C. Interpretability Insights from Keyword Analysis

In this section, we will shed light on the decision-making process of the proposed deep learning model by utilizing integrated keyword analysis to tackle the ever-present issue of neural interpretability. Through the examination of the extracted keywords and their respective weights, we can follow the impact of certain terms and phrases on the model's classification decisions and understand the probabilistic weights and deep features of the model, thereby allowing security analysts to understand threat indicators and build confidence in the automated detection system. This approach greatly improves the model trustworthiness by explaining its decision-making process, which is a crucial factor for security experts. Additionally, exposing the salient features that drive classification will allow us to refine threat intelligence and develop more focused defense strategies.

D. Scalability and Real-time Processing Capabilities

This proposed unified deep learning architecture provides substantial improvements in scalability to handle large volumes of Dark Web data and supports real-time processing capabilities for rapid threat detection in cybersecurity applications. It addresses computational overhead often observed with complex deep learning models, due to the optimized model training and inference processes the ability to scale with variable input lengths and adaptive resource allocation that ensures consistent performance despite changes in data influx and the direct incorporation of dynamic topic modeling into the TextCNN framework that minimizes preprocessing requirements that can become a bottleneck for ever-increasing datasets as well as the ability to process data in real time, which is important for dynamic, high-speed threat detection in cybersecurity applications.

E. Robustness to Evolving Dark Web Content

The dynamic nature of the architecture, particularly the dynamic topic modeling component, allows for the constant updating of understanding of relevant topics and keywords to keep pace with the Dark Web's linguistic changes, new illicit activities, and obfuscation techniques. As threat actors evolve their communication strategies and methods, the dynamic topic modeling component ensures that the model remains resilient against adversarial attacks and concept drift, maintaining high detection accuracy a crucial aspect of sustained performance in highly dynamic environments. The deep learning framework's ability to continuously learn from various datasets also enables the architecture to quickly adapt to emerging threats and new attack patterns reducing false positives and enhancing the accuracy of real-time threat detection. The ability to quickly adapt to new threat indicators and the system's inherent interpretability make this architecture a preferred solution for proactive cybersecurity.

To evaluate the effectiveness of the proposed model, two Dark Web text datasets were considered. The DUTA- 10K dataset contains approximately 10,000 documents categorized into 25 different cyber-threat related classes, while the CoDA dataset consists of around 5,000 documents categorized into 10 classes. These datasets represent various forms of Dark Web discussions including hacking forums, illicit marketplaces, and cybercrime communications. Table 1 summarizes the basic characteristics of the datasets used in the experiments.

Table 1: Dark Web Dataset Description

Dataset	Documents	Classes
DUTA-10K	10,000	25
CoDA	5,000	10

For experimental evaluation, the datasets were divided into training, validation, and testing subsets. The training set is used to train the classification model, the validation set is used for hyperparameter tuning and model optimization, and

the testing set is used to evaluate the final performance of the model on unseen data. Table 2 presents the dataset splitting strategy adopted in this study.

Table 2: Splitting of the Dataset

Dataset	Training	Validation	Testing
DUTA-10K	7000	1500	1500
CoDA	3500	750	750

The proposed topic-guided TextCNN model is evaluated using a comprehensive set of text classification metrics. These metrics include accuracy to measure overall classification performance, precision and recall to evaluate the model's ability to correctly identify relevant threat- related categories, and F1-score to provide a balanced assessment of precision and recall. These evaluation metrics help determine the effectiveness of the model in identifying different categories of Dark Web discussions such as cybercrime activities, hacking forums, and illicit marketplaces. The evaluation also considers the model's ability to generalize across unseen textual data and detect emerging cyber threat patterns.

Table 3: Results of Proposed Augmented Deep Learning and Machine learning Model

Metric	Value
Accuracy	**93.4%**
Precision	**0.92**
Recall	**0.91**
F1-Score	**0.91**

Table 3 presents the performance of the proposed topic- guided TextCNN model. The model achieves an accuracy of 93.4%, with precision, recall, and F1-score values of 0.92, 0.91, and 0.91 respectively. These results indicate that the proposed model effectively captures meaningful semantic patterns within Dark Web textual data and accurately classifies different threat-related categories. The high precision and recall values demonstrate the model's ability to correctly detect relevant cyber-threat discussions while minimizing false classifications.

Table 4: State-of-the-Art Comparison of the methods

Model	Accuracy (%)	Precision	Recall	F1-Score
Naive Bayes	81.2	0.80	0.79	0.79
SVM	85.6	0.84	0.83	0.83
LSTM	88.3	0.87	0.86	0.86
DarkBERT	91.8	0.90	0.90	0.90
TextCNN	90.1	0.89	0.89	0.89
Proposed Augmented Model	**93.4**	**0.92**	**0.91**	**0.91**

Table 4 compares the proposed topic-guided TextCNN model with several baseline machine learning and deep learning approaches, including Naive Bayes, Support Vector Machine (SVM), Long Short-Term Memory (LSTM), and standard TextCNN. The comparison shows that the proposed model outperforms traditional classifiers in terms of accuracy, precision, recall, and F1-score. The improvement in performance demonstrates the advantage of integrating topic modeling with deep neural networks, which enables the system to extract more meaningful contextual features from Dark Web text data and improve classification reliability.

V. CONCLUSION AND FUTURE WORK

In this paper, a novel, unified deep learning architecture that integrates dynamic topic modeling directly into a TextCNN for scalable and interpretable Dark Web text analysis was presented, which significantly outperformed traditional methods in terms of threat detection and interpretability, with key benefits including real-time threat detection for operational cybersecurity, enhanced explainability through integrated keyword analysis, reduced hyperparameter sensitivity for robust generalization, and scalable deployment across evolving Dark Web landscapes, which advances automated intelligence gathering. Future work will investigate the integration of attention mechanisms to further improve interpretability and focus the model on specific threat indicators. Further research will also consider methods for mitigating adversarial attacks on the deep learning model to make it more resistant to sophisticated attempts to circumvent detection

VI. REFERENCES

[1] N. Yazdanjue *et al.* , "A Language Model-Driven Semi-Supervised Ensemble Framework for Illicit Market Detection Across Deep/Dark Web and Social Platforms," *arXiv (Cornell University)* , Jul. 2025, doi: 10.48550/arxiv.2507.22912.

[2] Y. Jin, E. Jang, J. Cui, J. Chung, Y. Lee, and S. Shin, "DarkBERT: A Language Model for the Dark Side of the Internet," Jan. 2023, doi: 10.18653/v1/2023.acl-long.415.

[3] J. Pastor-Galindo, H.-Â. Sandlin, F. G. Mármol, G. Bovet, and G. M. Pérez, "A Big Data architecture for early identification and categorization of dark web sites," *Future*

Generation Computer Systems , vol. 157, p. 67, Mar. 2024, doi: 10.1016/j.future.2024.03.025.

[4] S. Adhya and D. K. Sanyal, "DTECT: Dynamic Topic Explorer & Context Tracker," *arXiv (Cornell University)* , Jul. 2025, doi: 10.48550/arxiv.2507.07910.

[5] N. Tavabi, N. Bartley, A. Abeliuk, S. Soni, E. Ferrara, and K. Lerman, "Characterizing Activity on the Deep and Dark Web," May 2019, doi: 10.1145/3308560.3316502.

[6] T. Niu, W. Li, and Y. Liu, "DarkGuardNet: A deep learning framework for imbalanced dark web traffic identification and application classification," *Research Square (Research Square)* , Feb. 2024, doi: 10.21203/rs.3.rs- 3974633/v1.

[7] J. Chao and T. Xie, "Deep Learning-Based Network Security Threat Detection and Defense," *International Journal of Advanced Computer Science and Applications* , vol. 15, no. 11, Jan. 2024, doi: 10.14569/ijacsa.2024.0151164.

[8] M. Ebrahimi, Y. Chai, S. Samtani, and H. Chen, "Cross-Lingual Cybersecurity Analytics in the International Dark Web with Adversarial Deep Representation Learning," *MIS Quarterly* , vol. 46, no. 2, p. 1209, May 2022, doi: 10.25300/misq/2022/16618.

[9] K. Ovabor, I. O. Sule-Odu, T. Atkison, A. T. Fabusoro, and J. O. Benedict, "AI-driven threat intelligence for real-time cybersecurity: Frameworks, tools, and future directions," *Open Access Research Journal of Science and Technology* , vol. 12, no. 2, p. 40, Nov. 2024, doi: 10.53022/oarjst.2024.12.2.0135.

[10] N. Deguara, J. Arshad, A. Paracha, and M. A. Azad, "Threat Miner - A Text Analysis Engine for Threat Identification Using Dark Web Data," in *2021 IEEE International Conference on Big Data (Big Data)* , Dec. 2022, p. 3043. doi: 10.1109/bigdata55660.2022.10020397.

[11] P. Koloveas, T. Chantzios, C. Tryfonopoulos, and S. Skiadopoulos, "A Crawler Architecture for Harvesting the Clear, Social, and Dark Web for IoT-Related Cyber-Threat Intelligence," Jul. 2019, doi: 10.1109/services.2019.00016.

[12] I. Pete *et al.* , "PostCog: A tool for interdisciplinary research into underground forums at scale," p. 93, Jun. 2022, doi: 10.1109/eurospw55150.2022.00016.

[13] S. E. Middleton, A. Lavorgna, G. Neumann, and D. Whitehead, "Information Extraction from the Long Tail," p. 82, Jul. 2020, doi: 10.1145/3394332.3402838.

[14] J. Hughes, S. Aycock, A. Caines, P. Buttery, and A. Hutchings, "Detecting Trending Terms in Cybersecurity Forum Discussions," Jan. 2020, doi: 10.18653/v1/2020.wnut-1.15.

[15] M. Deepthi, M. Harini, P. S. Geethika, V. Kalyan, and K. Kishor, "Data Classification of Dark Web using SVM and S3VM," *International Journal for Research in Applied Science and Engineering Technology* , vol. 11, no. 9, p. 510, Sep. 2023, doi: 10.22214/ijraset.2023.55643.

[16] F. K. Shaikh, "The Dark Web: Challenges and Countermeasures in Combating Cybercrime," *International Journal for Research in Applied Science and Engineering Technology* , vol. 12, no. 3, p. 636, Mar. 2024, doi: 10.22214/ijraset.2024.58892.

[17] A. Amadou, A. Motii, S. Elouardi, and E. H. Bergou, "EUREKHA: Enhancing User Representation for Key Hackers Identification in Underground Forums," *arXiv (Cornell University)* , Nov. 2024, doi: 10.48550/arxiv.2411.05479.

[18] T. Arjunan, "Detecting Anomalies and Intrusions in Unstructured Cybersecurity Data Using Natural Language Processing," *International Journal for Research in Applied Science and Engineering Technology* , vol. 12, no. 2, p. 1023, Feb. 2024, doi: 10.22214/ijraset.2024.58497.

[19] "Machine Learning and Deep Learning Approaches for Malicious Network Traffic Detection: A Comprehensive Evaluation."

[20] K. Taha, "Empirical and Experimental Insights into Data Mining Techniques for Crime Prediction: A Comprehensive Survey," *arXiv (Cornell University)* , Feb. 2024, doi: 10.48550/arxiv.2403.00780.

[21] "Security Paradigms for SDN-IoT Convergence: Integrating Agentic AI Agents, Blockchain, and Graph Neural Networks for Threat Resilience."

[22] M. Wazid *et al.* , "Explainable Deep Learning- Enabled Malware Attack Detection for IoT-Enabled Intelligent Transportation Systems," *IEEE Transactions on Intelligent Transportation Systems* , vol. 26, no. 5, p. 7231, Jan. 2025, doi: 10.1109/tits.2025.3525505.

[23] T. Xin, C. Zhang, J. Wang, Z. Zhao, and Z. Liu, "Dark-Forest: Analysis on the Behavior of Dark Web Traffic via DeepForest and PSO Algorithm," *Computer Modeling in Engineering & Sciences* , vol. 135, no. 1, p. 561, Sep. 2022, doi: 10.32604/cmes.2022.022495.

[24] J. Saleem, R. Islam, and Z. Islam, "Darknet Traffic Analysis A Systematic Literature Review," 2023, doi: 10.48550/ARXIV.2311.16276.

[25] R. N. V. J. Mohan, A. Vajpayee, S. Gangarapu, and V. V. R. Chilukoori, "Mitigating Complex Cyber Threats: An Integrated Multimodal Deep Learning Framework for Enhanced Security," *International Journal for Research in Applied Science and Engineering Technology* , vol. 12, no. 9, p. 1108, Sep. 2024, doi: 10.22214/ijraset.2024.64150.

[26] M. Rahmati, "Towards Explainable and Lightweight AI for Real-Time Cyber Threat Hunting in Edge Networks," *arXiv (Cornell University)* , Apr. 2025, doi: 10.48550/arxiv.2504.16118.

[27] M. O. Okafor, "Deep learning in cybersecurity: Enhancing threat detection and response," *World Journal of Advanced Research and Reviews* , vol. 24, no. 3, p. 1116, Dec. 2024, doi: 10.30574/wjarr.2024.24.3.3819.

[28] A. R. Pathak, M. Pandey, and S. S. Rautaray, "Adaptive Model for Dynamic and Temporal Topic Modeling from Big Data using Deep Learning Architecture," *International Journal of Intelligent Systems and Applications* , vol. 11, no. 6, p. 13, Jun. 2019, doi: 10.5815/ijisa.2019.06.02.

[29] S. Khan, N. Dilshad, N. Ahmad, S. Noor, and S. A. AlQahtani, "Integrating AI in security information and event management for real time cyber defense," *Scientific Reports* , vol. 15, no. 1, Oct. 2025, doi: 10.1038/s41598-025-19689-x.

[30] Mrs. S. S. Futane, "Adaptive Multi-Model Cybercrime Identification, Prediction using Machine Learning, and Explainable AI," *International Journal for Research in Applied Science and Engineering Technology* , vol. 13, no. 8, p. 2113, Aug. 2025, doi: 10.22214/ijraset.2025.73926.

[31] O. Darwish, S. Al-Eidi, A. Al-shorman, A. Alsobeh, M. Maabreh, and Y. Tashtoush, "LinguTimeX: Explainable AI of Natural Language Detection in Leakage Information with Covert Timing Channels," *Research Square (Research Square)* , May 2024, doi: 10.21203/rs.3.rs- 4372354/v1.

[32] S. Harishkumar and R. S. Bhuvaneshwaran, "Enhanced DGA Detection in BotNet Traffic: Leveraging N- Gram, Topic Modeling and Attention BiLSTM," *Research Square (Research Square)* , Feb. 2024, doi: 10.21203/rs.3.rs- 3981569/v1.

[33] J. J. M, R. Adhithya, S. D. S, and M. Revathi, "Exploring the Efficacy of Federated-Continual Learning Nodes with Attention-Based Classifier for Robust Web Phishing Detection: An Empirical Investigation," *arXiv (Cornell University)* , May 2024, doi: 10.1109/apci61480.2024.10617245.

[34] A. Şenol, G. Agrawal, and H. Liu, "Domain Knowledge-Enhanced LLMs for Fraud and Concept Drift Detection," *arXiv (Cornell University)* , Jun. 2025, doi: 10.48550/arxiv.2506.21443.

[35] Y. A. Farrukh, S. Wali, I. Khan, and N. D. Bastian, "XG-NID: Dual-Modality Network Intrusion Detection using a Heterogeneous Graph Neural Network and Large Language Model," *arXiv (Cornell University)* , Aug. 2024, doi: 10.48550/arxiv.2408.16021.

[36] J. Zhou, W. Fu, H. Song, S. Yu, Q. Xuan, and X. Yang, "Multi-view Correlation-aware Network Traffic Detection on Flow Hypergraph," *arXiv (Cornell University)* , Jan. 2025, doi: 10.48550/arxiv.2501.08610.

[37] S. Prasad *et al.* , "Interpretable intrusion detection for IoT environments using a self-attention-based explainable AI framework," *Scientific Reports* , vol. 15, no. 1, p. 39937, Nov. 2025, doi: 10.1038/s41598-025-23750-0.

[38] Z. Nibret Sileshi, J. Fentie Amsalu, and B. Mario, "Integrating Explainable AI for Effective Malware Detection in Encrypted Network Traffic," *arXiv (Cornell University)* , Jan. 2025, doi: 10.48550/arxiv.2501.05387.

[39] I. Karunanayake, M. AlSabah, N. Ahmed, and S. Jha, "Examining the Rat in the Tunnel: Interpretable Multi- Label Classification of Tor-based Malware," *arXiv (Cornell University)* , Sep. 2024, doi: 10.48550/arxiv.2409.16639.

[40] D. Xu, M. Liao, Z. Lai, X. Li, and J. Ji, "A Dual-Directional Context-Aware Test-Time Learning for Text Classification," *arXiv (Cornell University)* , Mar. 2025, doi: 10.48550/arxiv.2503.15469.

[41] R. Alroobaea, "An Empirical Deep Learning Approach for Arabic News Classification," *International Journal of Advanced Computer Science and Applications* , vol. 14, no. 6, Jan. 2023, doi: 10.14569/ijacsa.2023.01406112.

[42] M. K. Alshammery and A. F. Aljuboori, "Classifying Illegal Activities on Tor Network using Hybrid Technique," *Iraqi Journal of Science* , p. 3994, Sep. 2022, doi: 10.24996/ijs.2022.63.9.30.

[43] J. Alotaibi, "A hybrid software-defined networking approach for enhancing IoT cybersecurity with deep learning and blockchain in smart cities," *Peer-to-Peer Networking and Applications* , vol. 18, no. 3, Mar. 2025, doi: 10.1007/s12083-025-01935-8.

[44] M. A. Ferrag *et al.* , "Generative AI in cybersecurity: A comprehensive review of LLM applications and vulnerabilities," *Internet of Things and Cyber-Physical Systems* , vol. 5. Elsevier BV, p. 1, Jan. 01, 2025. doi: 10.1016/j.iotcps.2025.01.001.

[45] S. Z. Ahmad *et al.* , "A GAN-Based Approach for enhancing security in satellite based IoT networks using MPI enabled HPC," *PLoS ONE* , vol. 20, no. 9, Sep. 2025, doi: 10.1371/journal.pone.0331019.

[46] N. Sunanda, K. Shailaja, P. Kandukuri, Krishnamoorthy, V. S. Rao, and S. R. Godla, "Enhancing IoT Network Security: ML and Blockchain for Intrusion Detection," *International Journal of Advanced Computer*

Science and Applications , vol. 15, no. 4, Jan. 2024, doi: 10.14569/ijacsa.2024.0150497.

[47] M. Alsuwaiket, "ZeroDay-LLM: A Large Language Model Framework for Zero-Day Threat Detection in Cybersecurity," *Information* , vol. 16, no. 11, p. 939, Oct. 2025, doi: 10.3390/info16110939.

[48] M. M. Aslam, A. Tufail, H. Gul, M. N. Irshad, and A. Namoun, "Artificial intelligence for secure and sustainable industrial control systems - A Survey of challenges and solutions," *Artificial Intelligence Review* , vol. 58, no. 11, Aug. 2025, doi: 10.1007/s10462-025-11320-9.

[49] S. Ikbarieh, M. Gupta, and E. Mahalal, "LLM- based Multi-class Attack Analysis and Mitigation Framework in IoT/IIoT Networks," *arXiv (Cornell University)* , Oct. 2025, doi: 10.48550/arxiv.2510.26941.

[50] D. A. Worae, A. Sheikh, and S. Mastorakis, "A Unified Framework for Context-Aware IoT Management and State-of-the-Art IoT Traffic Anomaly Detection," *arXiv (Cornell University)* , Dec. 2024, doi: 10.48550/arxiv.2412.19830.

[51] L. Chourasiya *et al.* , "Advanced system log analyzer for anomaly detection and cyber forensic investigations using LSTM and transformer networks," *Journal of Cloud Computing Advances Systems and Applications* , vol. 14, no. 1, Oct. 2025, doi: 10.1186/s13677- 025-00789-y.

[52] A. Nair and P. Nitnaware, "Development of Multistage Machine Learning Classifier using Decision Trees and Boosting Algorithms over Darknet Network Traffic," *arXiv (Cornell University)* , Jul. 2024, doi: 10.48550/arxiv.2407.15910.

[53] S. M. Saleh, N. H. Madhavji, and J. Steinbacher, "Enhancing Cloud Security through Topic Modelling," *arXiv (Cornell University)* , May 2025, doi: 10.48550/arxiv.2505.01463.

[54] C. Reddy and K. Malathi, "Revolutionary hybrid ensembled deep learning model for accurate and robust side-channel attack detection in cloud computing," *Scientific Reports* , vol. 15, no. 1, Sep. 2025, doi: 10.1038/s41598-025-89794-4.

[55] A. Dhumane, N. N. Sakhare, P. Dehankar, J. R. R. Kumar, S. S. Patil, and M. Tatiya, "Design of an Efficient Forensic Layer for IoT Network Traffic Analysis Engine using Deep Packet Inspection via Recurrent Neural Networks," *International Journal of Safety and Security Engineering* , vol. 14, no. 3, p. 853, Jun. 2024, doi: 10.18280/ijsse.140317.

[56] T. Arjunan, "Real-Time Detection of Network Traffic Anomalies in Big Data Environments Using Deep Learning Models," *International Journal for Research in Applied Science and Engineering Technology* , vol. 12, no. 3, p. 844, Mar. 2024, doi: 10.22214/ijraset.2024.58946.

[57] A. Redhu, P. Choudhary, K. Srinivasan, and T. K. Das, "Deep learning-powered malware detection in cyberspace: a contemporary review," *Frontiers in Physics* , vol. 12. Frontiers Media, Mar. 28, 2024. doi: 10.3389/fphy.2024.1349463.

[58] T. Yang and J. Sun, "A hybrid ensemble deep learning framework with novel metaheuristic optimization for scalable malicious website detection," *Scientific Reports* , vol. 15, no. 1, p. 44630, Dec. 2025, doi: 10.1038/s41598- 025-33695-z.

[59] P. Rekha, S. M. G, S. Thala, T. Meshach, B. Crunchier, and K. Sony, "Intelligent Intrusion Detection System Using NSOA and Hybrid ECA-LiteCBNet Model for Cyber Threat Mitigation," *Research Square (Research Square)* , Oct. 2025, doi: 10.21203/rs.3.rs-7565087/v1.

[60] P. Roguski, "Layered Sovereignty: Adjusting Traditional Notions of Sovereignty to a Digital Environment," p. 1, May 2019, doi: 10.23919/cycon.2019.8756900.