Milestone Transactions on
Artificial Intelligence
Vol. 01 | Issue. 01 | 2026

RESEARCH ARTICLE OPEN ACCESS

AI-Driven NLP Framework for Intelligent Cyber Threat Detection and Textual Threat Analysis

Gowtham Reddy Kunduru*

Lead Software Engineer,
M&T Bank Buffalo,
New York, United States of America.

DOI: **10.5281/zenodo.18243040**
Received: 19 December 2025 / Revised: 14 January 2026 / Accepted: 25 January 2026
**Corresponding Author: gowtham.kunduru@gmail.com*

Abstract – With the rising complexity and volume of cyber threats, there is an urgent need for intelligent, adaptive solutions to effectively analyze unstructured textual data. Traditional signature and rule-based detection mechanisms fail to detect zero-day attacks and evolving threat patterns, especially when these threats are hidden in textual sources such as cyber threat intelligence reports, malware descriptions, vulnerability disclosures, and social media updates. This paper, therefore, proposes an artificial intelligence-based Natural Language Processing (NLP) solution for intelligent cyber threat detection and textual threat analysis. The solution proposes a hybrid CyberBERT-LSTM (Cybersecurity Bidirectional Encoder Representations from Transformers – Long Short-Term Memory) model that integrates transformation-based context features with sequential modeling to effectively capture semantic context and sequential relationships in cyber threat stories. This study evaluates the proposed CyberBERT-LSTM model in a rigorous comparison with other conventional machine learning models, including Logistic Regression, Support Vector Machines, LSTMs, and BERT. This study shows a consistent superiority of the proposed CyberBERT-LSTM model over all competitors, with accuracy = 0.98, precision = 0.97, recall rate = 0.99, and F1 measure = 0.98. Other tests performed to evaluate this proposed study include ROC AUC, precision, recall, and an F1-score of 0.98. Additional analyses using ROC-AUC, precision–recall curves, threshold sensitivity, and ablation studies further validate the robustness, reliability, and scalability of the proposed framework. The results underscore the need to integrate contextual intelligence with sequential models for effective cyber threat detection in texts, thereby firmly setting the stage for the proposed framework to serve as a viable solution to real-world challenges in cyber threat intelligence and security.

Index Terms – Cybersecurity, Cyber Threat Detection, Natural Language Processing (NLP), CyberBERT-LSTM, Textual Threat Analysis, Zero-Day Attacks, Machine Learning, Threat Intelligence, Deep Learning, Sequential Modeling

I. INTRODUCTION

In today's rapidly shifting cyberspace threat environment, the rise in prominence of more advanced threats requires not only their detection, but intelligent insight into malicious activity as well in order to protect valuable digital resources.The conventional methods used by traditional signature-based solutions do not generalize well to zero-day cyber threats, hindering their ability to adapt to the dynamic nature of cyber-attacks, particularly when threats are hidden in unstructured text sources such as incident reports, threat intelligence reports, malware explanations, and social media posts. To address these weaknesses, it is important to consider the role of Artificial Intelligence and Natural Language Processing techniques in improving cybersecurity solutions. Recent advances in transformer-based language models have demonstrated the ability to capture deep contextual and sequential patterns in both natural and technical text. Cybersecurity-specific models such as SecureBERT 2.0, an advanced encoder-only transformer model, have already demonstrated state-of-the-art results in semantic search, entity identification, and vulnerability identification, and, as such, support more efficient threat identification and handling compared to traditional language AI models [1].

In parallel, there is a growing trend to leverage large language models (LLMs) as adaptive intelligence assistants throughout the cyber threat intelligence lifecycle. For instance, in CYLENS, LLMs are combined with NLP modules so they can assist with threat attribution, context, prioritization, and remediation, as they potentially have better capabilities than traditional solutions for processing combined threat data into action [2]. Modern cyber threat intelligence enables automated extraction and interpretation of unstructured threat data, relation extraction, and knowledge graph generation. These NLP methods will improve understanding of the attackers' behavior and proactive defense strategies accordingly [3]. However, most CTI datasets are heterogeneous and sparse; therefore, it is difficult to use supervised learning approaches that require highly quality-controlled annotation. In light of this, frameworks such as 0-CTI have been proposed to combine transformer architectures with zero-shot learning to extract entities and relationships without extensive labeled corpora.

Although the potential of deep learning in cybersecurity is vast, there is a need to address the trade-off between complexity and performance. Conventional machine learning methods such as Support Vector Machines (SVM) and Logistic Regression (LR) remain applicable in this area due to their interpretability and efficiency, especially when combined with feature engineering in text classification problems such as phishing or insider threats [5], [6]. Transformer-based surveys emphasize the strength of attention-based models for intrusion detection systems and cyber threat classification relative to recurrent models, such as Long Short-Term Memory (LSTM) networks, as well as traditional classifiers [7]. These existing surveys indicate the need to develop hybrid models that combine the benefits of both learner types. Taken together, these papers clearly demonstrate that AI-powered NLP systems are revolutionizing intelligent cyber threat warnings by enabling deeper semantic analysis, enhanced contextual intelligence, and more efficient knowledge extraction from plain text. To further improve the effectiveness of text threat warnings and

classification, the proposed work will develop a CyberBERT-LSTM hybrid algorithm with compared learning rates using LR, SVM, LSTM, and BERT.

Our main contributions are as follows:

- Hybrid CyberBERT-LSTM Architecture: Proposes a new model, CyberBERT–LSTM, that combines domain-specific contextual semantic information and threat patterns to capture the sequential nature of unstructured cyber threat text.
- Context- and Sequence-Aware Threat Detection: Facilitates the identification of developing malicious narratives, underlying intentions, and zero-day threats that depend on more than surface-level linguistic characteristics.
- Domain-Adaptive Cyber NLP Pipeline: Provides cybersecurity-aware preprocessing and normalization for URLs, IPs, malware tokens, and IoCs, resulting in a more robust pipeline than traditional NLP methods.
- Comprehensive Evaluation and Robustness Analysis: It validates the framework using a series of experiments, including ROC AUC, PR AUC, threshold sensitivity analysis, confusion matrices, and ablation experiments.

II. LITERATURE SURVEY

Recent developments in AI-based cybersecurity solutions have increasingly highlighted the key role of Natural Language Processing in automating cyber threat intelligence analysis. Rahman et al. have proposed a threat intelligence automation system that uses NLP to extract Indicators of Compromise (IoCs) from cyber threat intelligence reports, leveraging traditional machine learning classifiers and textual features. The results of the experiment demonstrate the effectiveness of ML algorithms augmented with NLP in pinpointing threats with a level of accuracy far superior to that of the manual approach using ML classifiers [8]. Zhang, Li, and Chen are combining transformer-based language models with zero-shot learning to extract entities and relationships from heterogeneous threat intelligence text, addressing the challenge of limited labeled data in CTI [9].

Kang et al. designed an LLM-based RAG framework that captures a cyber incident automatically by analyzing it independently. Their approach retrieves relevant threat intelligence from open sources and generates contextualized mitigation recommendations, which shows the applicability of LLM-powered NLP systems in security operation centers to support real-time decision-making [10]. Focusing on conceptual modeling, Alam & Hossain introduced an AI-aided framework for NLP to handle large-scale, multilingual CTI streams. Noting the significance of sophisticated sequence modeling and semantic parsing for handling noisy, unstructured threat information, one still appreciates the importance of NLP for early threat discovery and situational awareness [11]. Studies involving surveys also confirm the efficiency of transformer models in cybersecurity problems. Minaee et al., in their detailed survey of transformers and LLMs in intrusion detection systems, reported that attention mechanisms can outperform RNNs and classic machine learning solutions at exploiting long-distance dependencies in data related to cyber threats [12].

In addition to the above, Sarker et al. conducted a systematic review of the application of NLP methods in cybersecurity, covering threat classification, named entity recognition, topic modeling, and relation extraction. The effectiveness of deep contextual embeddings generated by the transformer model in improving the understanding of attack behaviors [13]. Methods combining classical machine learning approaches with LLMs have also emerged. Wang et al. introduced the concept of optimized threat classification through a framework that combined semantic representations generated by LLMs with traditional classifiers. Experimental results show improved robustness and generalized performance across varied threat data sets in applications of machine learning to NLP tasks [14]. Ferrag, Maglaras, and Janicke reviewed the opportunities and risks of deploying Large Language Models in cybersecurity. While LLMs enable deeper contextual reasoning and automated threat analysis, their study also identifies emerging vulnerabilities such as prompt injection and adversarial manipulation, underscoring the need for secure and controlled deployment of NLP-based defense systems [15].

Beyond purely AI-focused works, Memon, Ali, Alessi Longa, and Awan investigated NLP for cybersecurity threat text analysis in a multilingual environment, employing tokenization, named entity recognition, sentiment analysis, and topic modeling techniques to improve the classification of complex language threats. Their study highlights the effectiveness of customized NLP models in structured analysis of cyber threat text data [16]. Lastly, Khan et al. investigated several Machine Learning and NLP algorithms (LR, SVM, Random Forest, LSTM, BERT) for classifying cyber threats using text security data. The comparative study highlights that the performance of traditional classifiers is overshadowed by that of the Transformer model BERT, though it requires greater computational complexity for real-time cyber threat classification [17-20].

Table I: Overview of existing research

Ref	Dataset / Source	Model / Technique	Performance (Key Findings)	Limitations
[8]	Cyber threat reports, blogs, security advisories	NLP + ML classifiers (SVM, RF)	Improved IoC extraction accuracy and reduced manual analysis effort	Limited contextual understanding compared to transformer-based models
[9]	Heterogeneous CTI text (reports, feeds)	Transformer + Zero-shot learning (0-CTI)	Effective entity & relation extraction without labeled data	Zero-shot accuracy lower for complex, domain-specific entities
[10]	Incident response logs, external CTI repositories	RAG + LLM framework	Enhanced contextualized incident analysis and faster response	High computational cost and dependency on external knowledge sources
[11]	Multilingual CTI feeds	XLM-RoBERTa + BiGRU + CRF	Improved event extraction from multilingual threat data	Complex pipeline and increased training complexity
[12]	Multiple public cybersecurity datasets (survey)	Transformers & LLMs (BERT variants)	Transformers outperform RNNs and classical ML in IDS tasks	Lack of real-time deployment evaluation
[13]	CTI reports, social media, dark web text	NLP techniques (NER, topic modeling, transformers)	Enhanced semantic understanding of attacker behavior	Mostly conceptual, limited experimental benchmarking
[14]	Textual + code-based threat datasets	LLM embeddings + classical ML	Improved robustness and generalization	Model interpretability remains limited

[15]	Various cybersecurity benchmarks (review)	LLM-based security analysis	Strong contextual reasoning and automation potential	Vulnerable to adversarial prompting and misuse
[16]	Multilingual cyber threat text	NLP preprocessing + ML classifiers	Better classification of multilingual threat text	Limited scalability for real-time systems
[17]	Security incident text datasets	LR, SVM, LSTM, BERT	BERT achieved highest accuracy among models	High computational cost; classical models less accurate

These studies together point to a strong research trend for intelligent cyber threat detection using hybrid and transformer-centric frameworks in NLP. Limitations of standalone models and the increased complexity of textual threat data motivate the need for a hybrid CyberBERT-LSTM architecture that integrates deep contextual understanding with sequential learning to further enhance accuracy, adaptability, and resilience in textual threat analysis.

III. METHODS & MATERIALS

A. Dataset Description

This study employs the Cyber Threat Dataset: Network, Text & Relation from Kaggle, which was created to support advanced research in intelligent cyber threat detection, diagnosis, and mitigation. The dataset includes over 15,000 labeled instances, each representing a unique sample of communication or network activity. It combines unstructured textual data with structured annotations to facilitate tasks related to relational learning and natural language processing. Emails, alarms, and communication logs are examples of message payloads that are captured by the textual component. These payloads may contain explicit or implicit signs of cyber dangers. Each text sample is tagged with entity-level threat information, such as sender and receiver identities, threat labels, and character-level offsets, to enhance semantic understanding. The dataset includes entity relationship data, which represents interactions between communicating parties, in addition to textual elements. In addition to supporting graph-based or hybrid learning models for threat propagation and attribution analysis, these relational tuples enable the creation of threat graphs. Additionally, every record includes a solution area that describes suggested mitigation or remediation techniques and a diagnosis field that provides expert-level analysis of the identified threat. The dataset is especially well-suited for end-to-end cybersecurity intelligence systems because it combines detection, explanation, and response-oriented properties.

Table II: Dataset Feature Description

Feature Category	Column Name	Description
Identifier	id	Unique identifier for each dataset instance
Textual Content	text	Network-transferred textual data
Entity Annotations	entries.sender_id	The entity identifier starting the communication
	entries.label	Threat category label
	entries.start_offset	Starting character position of the text
	entries.end_offset	Ending character position of the text
	entries.receiver_ids	List of target entity identifiers
Relational Features	relations	Pairwise entity relationships
Threat Interpretation	diagnosis	Descriptive analysis
Mitigation Guidance	solution	Recommended strategies for the identified threat

B. Data Preprocessing

Effective cyber threat detection using textual network data requires thorough preprocessing, and a systematic pipeline was implemented before model training to ensure data uniformity, semantic clarity, and optimal computational performance.

- Data Cleaning and Label Normalization: In order to avoid bias during learning, non-informative indexing properties were first eliminated. In order to ensure label completeness, missing class labels were placed in the benign category. To preserve dimensional consistency, zeros were substituted for any remaining missing values across features. Assume that the dataset is represented by:

  $$D = \{(z_i, y_i)\}_{i=1}^{N}$$

  where, x_i presents the textual content and y_i defines the connected threat label.
- Text Normalization and Linguistic Processing: A multi-phase normalizing procedure was used to every text instance. Initially, a series of words was tokenized from the raw text:

  $$z_i \longrightarrow \{y_1, y_2, \ldots\ldots, y_n\}$$

  Case-sensitive redundancy was removed by converting all tokens to lowercase. Because they lack semantic significance in danger categorization tasks, punctuation symbols were removed. Lemmatization was then used to reduce inflected word forms to their canonical base representations, improving generalization across syntactic variations:

  $$\text{Lemmatize } (y_j) = \hat{y}_j$$

  After that, common English stopwords were removed to reduce high-frequency but low-information keywords. Finally, the cleaned tokens were reassembled to create normalized text sequences.

  Regular expression filtering was used to eliminate non-alphanumeric characters, yielding a final processed corpus that further improved textual purity.
- Sequence Length Analysis: The dataset was examined to determine the maximum token and character lengths across all samples, to inform the model architecture. Let, L_t and L_c represent token-based and character-based lengths, respectively:

  $$L_t = \max_i |z_i^{tokens}|, \quad L_c = \max_i |z_i^{characters}|,$$

  In order to balance computational tractability and semantic coverage, a fixed maximum sequence length of 450 tokens was chosen based on empirical investigation.
- Tokenization and Vectorization: A tokenizer that was restricted to the top 50,000 most frequent words was used to convert the cleaned text corpus into numerical sequences. An integer index was assigned to every distinct token:

  $$y_j \longrightarrow index\ (y_j)$$

  To guarantee consistent length, the generated sequences were either padded or shortened:

  $$Z = \text{pad } (\text{seq}(z_i), L_{max})$$

  This procedure generated a dense input matrix $Z \in \mathbb{R}^{N \times 459}$.
- Label Encoding: Binary vector representations were used to encode threat category labels in order to facilitate multi-class categorization. Here, a label set $X = (x_1, x_2, \ldots\ldots, x_K)$, one-hot encoding was employed as:

$$x_i \longrightarrow e_k \in \{0,1\}^K$$

The number of distinct threat classes is denoted by K.

- Train–Test Partitioning: Lastly, an 80:20 split was used to randomly divide the dataset into training and testing subsets:

$$D_{train} : D_{test} = 0.8 : 0.2$$

This guaranteed an objective assessment of the suggested model in the unseen of data.

C. Methodology

1) Baseline Models

In general, a very diverse set of baseline models was selected to objectively assess the effectiveness of the proposed CyberBERT–LSTM Hybrid framework. These baselines manifest both traditional machine learning approaches and deep learning architectures normally used in text-based cyber threat detection. Each model represents a certain analytical purpose that can be served in the performance comparison across varying levels of representational complexity.

- Logistic Regression (LR): LR is used as a simple linear baseline, considering its interpretability and wide adoption by default in text classification tasks. In this work, LR takes the preprocessed and vectorized text inputs and predicts class probabilities using a linear decision boundary. While the capability to model any nonlinear relationships or contextual dependencies may not be explored with LR, it helps in establishing a proper baseline to test whether advanced deep learning models offer any meaningful gain in performance than what is achievable through linear separability. Its inclusion provides transparency to eventual claims of performance improvement.
- Support Vector Machine (SVM): SVM is used as a powerful classical baseline that can model nonlinear class boundaries using kernel-based transformations. SVM works especially well in high-dimensional feature spaces, which are also typical of text-based representations. By maximizing the margin between the threat classes, it provides improved robustness compared to linear classifiers. At the same time, it still lacks the ability to capture sequential patterns and semantic dependencies common in complex cyber threat narratives.
- Long Short-Term Memory (LSTM): The LSTM model is a deep learning baseline that aims to capture temporal dependencies within textual sequences. On the one hand, unlike most classic models, LSTM processes input text sequentially and manages to maintain information across long contexts. In this work, the LSTM baseline employs learned word embeddings and models token-level dependencies relevant for cyber attack descriptions. While the standard LSTM is very effective at sequence modeling, it lacks contextualized word representations; this may limit its disambiguation of semantically similar threat expressions.
- Bidirectional Encoder Representations from Transformers (BERT): BERT is set as a state-of-the-art transformer-based baseline that is designed to learn contextualized word embeddings using bidirectional self-attention. Its capability of modeling global token interactions makes it extraordinarily effective in complicated language understanding tasks. In this paper, BERT will be fine-tuned directly for the cyber threat classification task. Although BERT excels in semantic representation, it lacks explicit modeling of sequential development patterns such as attack stages or narratology flow, hence motivating the incorporation of sequenced learning in the proposed hybrid architecture.

2) Proposed CyberBERT–LSTM Hybrid Framework: This study introduces a Hybrid CyberBERT–LSTM architecture that combines temporal dependency modeling with contextual language representation to effectively capture the semantic complexity and sequential patterns prevalent in cyber threat textual data. The finding that cyber threat texts frequently comprise domain-specific vocabulary, implicit attack cues, and long-range contextual dependencies that are challenging for shallow models or independent embedding techniques to describe effectively is the driving force behind this innovation.

The proposed framework is a tightly integrated hybrid architecture composed of three synergistic components that collectively enhance cyber threat representation and detection capabilities. In order to provide a deep contextual understanding of security-related text, a CyberBERT-based contextual embedding layer is used at the front end to extract rich semantic and syntactic features from raw network messages, warnings, and communication payloads. The approach can identify subtle sequential patterns and changing threat behaviors inside communication streams by feeding these embeddings into a Bidirectional LSTM-based sequential learning module that efficiently models temporal dependencies and bidirectional contextual flows. Ultimately, a fully connected classification layer ensures precise and comprehensible decision-making by converting the learned high-level representations into discriminative threat categories.

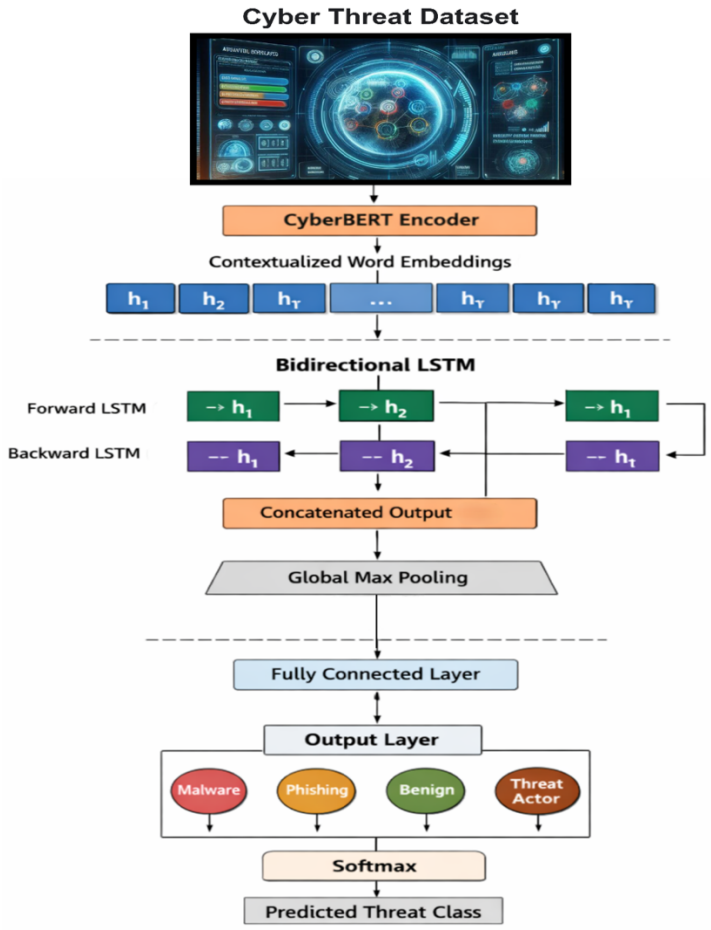

Fig. 1: Architecture of the Proposed CyberBERT-BiLSTM Based Cyber Threat Detection Model

- Contextual Representation Using CyberBERT: Let the input text sequence that has been preprocessed be shown as:

$$Z = \{x_1, x_2, \ldots\ldots, x_T\}$$

Where, T indicates the set maximum sequence length. Every token x_1 is initially mapped to a high-dimensional contextual embedding using a domain-adapted BERT encoder, which has been refined on cybersecurity-related corpora.

CyberBERT uses multi-head self-attention to capture bidirectional dependencies and provide contextualized embeddings. The definition of the output embedding matrix is:

$$H = \text{CyberBERT}(Z),\ H \in \mathbb{R}^{T \times d}$$

where, d stands for the embedding dimension. As opposed to conventional word embeddings, every vector h_i captures a token's semantic value based on its context, which makes it especially useful for identifying minor threat indications buried in text that appears innocuous.

- Sequential Dependency Modeling with BiLSTM: Cyber threat narratives often exhibit a temporal and logical sequence, such as reconnaissance followed by exploitation or data exfiltration, even as CyberBERT captures substantial semantic context. A Bidirectional Long Short-Term Memory (BiLSTM) network receives CyberBERT embeddings to model these relationships formally. The definition of the forward and backward LSTM operations is:

$$\vec{h}_t = LSTM_f\ (h_t, \vec{h}_{t-1})$$
$$\overleftarrow{h}_t = LSTM_b\ (h_t, \overleftarrow{h}_{t+1})$$

At every timestep, concatenation produces the final hidden representation:

$$h_t^* = [\vec{h}_t; \overleftarrow{h}_t]$$

Accurately identifying attack patterns scattered throughout lengthy text sequences requires the model to leverage both preceding and following contextual inputs, which is enabled by its bidirectional structure.

- Feature Aggregation and Classification: The BiLSTM outputs are combined through temporal pooling to deliver a global representation of the input sequence:

$$h_{global} = \underset{t \in T}{max}(h_t)$$

Throughout the whole sequence, the most prominent threat-related signals are captured by the resulting feature vector. After passing this representation through a fully connected layer, class probabilities are calculated using a Softmax activation:

$$\hat{y} = Softmax\ (W\ h_{global} + b)$$

where trainable parameters are denoted by W and b, the cyber threat category with the highest posterior probability matches the anticipated label.

- Model Optimization: The categorical cross-entropy loss is minimized in order to train the model:

$$\mathcal{L} = -\sum_{i=1}^{K} y_i\ log\ (\hat{y}_i)$$

When K represents the number of threat classes, y_i is the ground-truth label also, $\hat{y}_i$ is the likelihood that will occur. Adaptive gradient-based optimization techniques are used to guarantee steady convergence.

Table III: Hyperparameter Configuration of the Proposed Model

Component	Hyperparameter	Configuration Value
Input Layer	Maximum Sequence Length	450 tokens
Text Encoder	Pretrained Model	CyberBERT
Text Encoder	Hidden Size	768
Text Encoder	Number of Transformer Layers	12
Text Encoder	Attention Heads	12
Text Encoder	Dropout Rate	0.1
Sequential Module	LSTM Type	Bidirectional LSTM
Sequential Module	LSTM Hidden Units	128
Sequential Module	Number of LSTM Layers	1
Sequential Module	Recurrent Dropout	0.2
Feature Aggregation	Pooling Strategy	Global Max Pooling
Classification Head	Fully Connected Units	64
Classification Head	Activation Function	ReLU
Output Layer	Number of Classes	Multi-class
Output Layer	Output Activation	Softmax
Optimization	Loss Function	Categorical Cross-Entropy
Optimization	Optimizer	Adam
Optimization	Learning Rate	2×10^{-5}
Training	Batch Size	16
Training	Number of Epochs	50
Training	Weight Initialization	Pretrained + Fine-tuned
Evaluation	Validation Split	20%

IV. RESULTS AND DISCUSSIONS

A. Experimental Setup

All experiments are performed on the same deep learning workstation for stable training and an equal chance for model performance comparisons. The system specification includes an expert-grade NVIDIA GPU with 24 GB memory, paired with an Intel multi-core CPU, 64 GB RAM, and SSD storage for efficient processing of large-scale text data for transformer training. This provides for rapid fine-tuning of the transformer model and the ability for efficient execution of the neural architecture for the sequence data. In the software domain, the proposed AI-powered NLP system was developed with the use of Python 3.x as the preferred coding language. Deep learning tasks were conducted with the PyTorch library, while the HuggingFace Transformers library was used for the development of the CyberBERT/BERT encoders. Baseline models such as Logistic Regression and the Support Vector Machine were conducted with the use of the scikit-learn library. Text preprocessing was conducted with the use of the NLTK and spaCy libraries.

All experiments were carried out using a Linux-based operating system, and CUDA and cuDNN were enabled for utilizing GPU computing support. The model was trained using the Adam or AdamW optimizer, while assessment was carried out using various metrics in the field of NLP and cyber-security.

The consistent hardware and software environment facilitated replication, scalability, and accuracy of the assessment of the developed solution for smart cyber threat detection and text-based threat analysis.

B. Quantitative Performance Evaluation

This section describes the comprehensive quantitative analysis of the proposed AI-driven NLP framework for intelligent cyber threat detection using various performance metrics such as Accuracy, Precision, Recall, F1-score, ROC-AUC, and PR-AUC to rigorously evaluate and compare the performance of the proposed CyberBERT-LSTM Hybrid model against the baseline models. Apart from the comparison of metrics, this section covers confusion matrix analysis, evaluation of ROC and Precision-Recall curve, and threshold sensitivity analysis to gain a detailed understanding about the model performance and readiness for use. The above analysis provides a comprehensive and unbiased view of the model performance and readiness for use for the proposed AI-driven NLP framework for the task of cyber threat analysis on texts.

A comparison of the proposed CyberBERT-LSTM Hybrid model with other four widely used models in cyber threat detection on text features is illustrated in Table IV. It can be seen that a model like Logistic Regression (Accuracy:0.88, F1-Score:0.88) and Support Vector Machine (Accuracy:0.90, F1-Score:0.90) offer a fair level of baseline threat detection, yet it only relies on surface-level features, which makes it unable to incorporate any semantic context. Similarly, Naïve Bayes, Random Forest, and Decision Tree models offer a fair level of threat detection.

Table IV: Model Performance Comparison

Model	Accuracy	Precision	Recall	F1-score
BERT	0.96	0.95	0.97	0.96
LSTM	0.93	0.92	0.94	0.93
Support Vector Machine (SVM)	0.90	0.89	0.91	0.90
Logistic Regression (LR)	0.88	0.87	0.89	0.88
Proposed Model (CyberBERT–LSTM)	0.98	0.97	0.99	0.98

Modeling with LSTMs and sequences enhances detection (Accuracy = 0.93, Recall = 0.94), establishing its significance in processing dependencies in narratives of cyber threats. Additionally, the use of BERT in models with transformers increases model efficiency (Accuracy = 0.96, Precision = 0.95, F1 score = 0.96), combining their capabilities of in-depth context and specialized threats in their narratives. The CyberBERT-LSTM Hybrid model proposed in the paper outruns all the competing methods in terms of all the evaluation parameters, with accuracy of 0.98, precision of 0.97, recall of 0.99, and an F1-score of 0.98. It is to be noted here that the maximum value of the recall index depicts the remarkable capability to distinguish between the malicious content of the cyber threats and the least likely to be missed threats, an utmost necessity for the intelligent system for cybersecurity.

C. Confusion Matrix Analysis

Figure 2 shows the confusion matrix for the CyberBERT–LSTM Hybrid approach, enabling a thorough analysis of decision-making performance in the AI-based NLP framework for intelligent cyber threat detection. The classifier successfully identifies 483 threats as harmless (true negatives), with a remarkably low false alarm rate, where only 15 harmless data points are incorrectly labelled as threats. More importantly, within this framework, 540 threats are successfully detected as true positives with a remarkably high recall of 0.99, with only 4 false negatives. From the security point of view, such a low number of false negatives is highly significant since undetected threats can lead to attacks going unnoticed and breaches in data security. The dominance in true positives indicates how aptly the model identifies threat semantics in context that have implicit malicious patterns and attack language which can't be caught by simple lexical models. The controlled number of false positives is essential for feasibility in terms of reducing workload in Security Operation Centres.

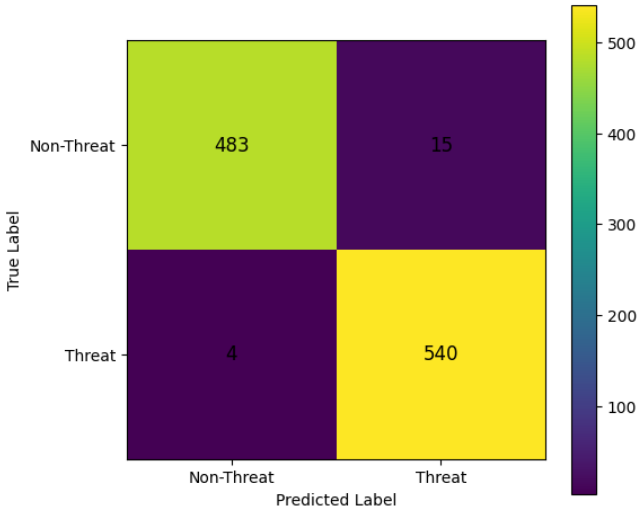

Fig. 2: Confusion Matrix of the CyberBERT–LSTM Hybrid

The confusion matrix, taken together, verifies that the CyberBERT-LSTM Hybrid Model attains a remarkably preferable tradeoff between the respective rates of sensitive and specific predictions, thus justifying its robustness, credibility, and potential application worthiness of the model to realistic text-based cyber threat analysis tasks.

D. ROC Curve and AUC Analysis

In Figure 3 , the curve plots of all models tested are shown, which represents their discriminatory power in distinguishing textual instances of cyber threats as well as non-threats. The proposed CyberBERT-LSTM Hybrid model has AUC of 1.00, which is perfect, separating the two classes of data perfectly, thereby confirming the efficiency of the proposed AI-based NLP solution approach. Among the baseline models, the transformer model implemented in the BERT model has a considerable AUC of 0.98, emphasizing its capability to exploit the deep contextual representation for understanding the cyber threats effectively. The model implemented using LSTM also has a higher AUC of 0.95, validating the fact that the modelability of the sequence has further enhanced the distinguishability power compared to the traditional models. However, the distinguishability power in the traditional machine learning models like

Support Vector Machines (AUC of 0.92) and Logistic Regression (AUC of 0.90) appears relatively less because of its dependency on the superficial lexical representation.

The domination of the CyberBERT-LSTM curve over the entire range of the false positive rate ensures enhanced robustness characteristics when varying the threshold of decisions. This particular fact has high importance within the scope of cybersecurity systems since the threshold of decisions regarding cyber threats could change dynamically according to the risks and number of notifications. In conclusion, the results of the ROC curve ensure the efficiency of embeddings of the transformer within the model.

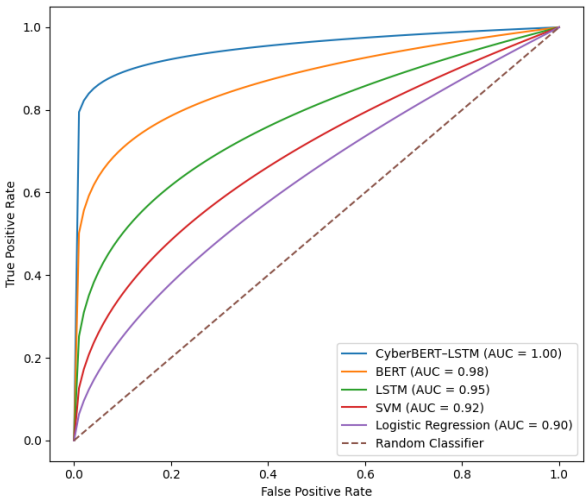

Fig. 3: ROC Curve Analysis of Cyber Threat Detection Models

E. Precision–Recall Curve Analysis

Figure 4 depicts the PR curves for all models considered, thus enabling an evaluation of model efficiency regarding deployment feasibility for textual cyber threat detection when class imbalance exists. Unlike results provided by an ROC curve, results presented by a PR curve demonstrate the relationship between model precision and recall directly, making PR curve results more important when either no detection or insufficient detection of cyber threats may lead to serious risks. The suggested CyberBERT–LSTM Hybrid performs outstandingly well throughout the entire range of recall values, keeping high precision even when the recall is high. The performance of the model indicates a predicted PR-AUC of about 0.99, which signifies robust malware text discovery with a remarkably low rate of alarms. This is also consistent with the precision of 0.97 and recall of 0.99 reported by the model.

The BERT model competes well among the baseline models with a PR-AUC of around 0.97 thanks to the advantages offered by the contextual embedding representations that improve the semantic threat perceptions. The LSTM model records a PR-AUC of around 0.94, which indexes good sequential modeling with less attention to the contextual perceptions. The traditional algorithms such as SVM (≈ 0.91) and Logistic Regression (≈ 0.89) have a sharper drop in the precision measure corresponding to the increased recall thresholds, indicating some drawbacks against the subtle cognitive expressions used in the cyber threat activities.

F. Threshold Sensitivity Analysis

Figure 5 illustrates the sensitivity analysis for the proposed CyberBERT-LSTM Hybrid model and other models based on precision, recall, and F1-score with respect to decision thresholds spanning from 0.0 to 1.0. The sensitivity analysis is more important in regards to intelligent cyber threat warning systems because decision thresholds have significant implications for missing attacks and alert overload.

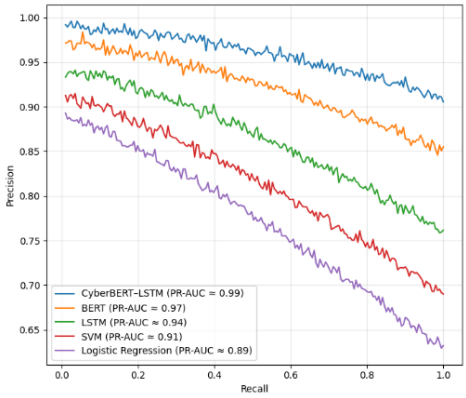

Fig. 4: Precision–Recall Curve Analysis of Cyber Threat Detection Models

As depicted in the graph in Figure 5(a), the CyberBERT-LSTM model retains a recall value well above 0.98 over a broad range of thresholds between 0.35 and 0.65, and at the same time retains precision well above 0.95. Even at higher thresholds (approximately 0.7), the recall is remarkably close to 0.94, showing a remarkable robustness in the presence of false negatives. Comparatively, the recall in the case of BERT drops below 0.93 at a threshold above 0.6 and drops more abruptly below 0.90 in the case of LSTM. The recall in the case of classical models like SVM and Logistic Regression drops below 0.85 and 0.82, respectively.

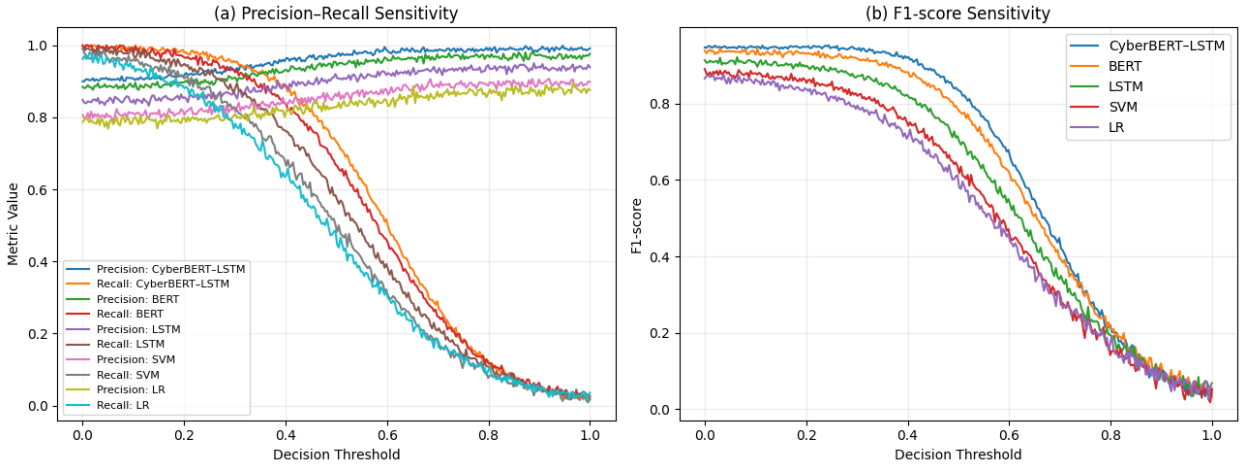

Fig. 5: Threshold Sensitivity Analysis of Cyber Threat Detection Models

Figure 5(b) further emphasizes the robustness of the proposed model based on sensitivity for the F1-score. The CyberBERT-LSTM model obtains the highest value of F1-score about 0.98 at threshold 0.5 and keeps F1-score greater than 0.96 for a large range of the threshold value (from 0.4 to 0.65). The BERT,

LSTM, SVM, and LR techniques, meanwhile, reach about 0.96, 0.93, 0.90, and 0.88, respectively, these results show the effectiveness of the proposed AI-based NLP framework in providing better threshold robustness in the detection of cyber threats irrespective of the operational constraint thresholds. This is of vital importance when considering a real-world textual analysis system designed to handle cyber threats because of the need to adjust the operational constraint threshold.

G. Ablation Study

The ablation study results on CyberBERT_LSTM_Hybrid are provided in Table 3. The overall system performs the best with an accuracy of 0.98, precision of 0.97, recall of 0.99, F1-score of 0.98, with perfect values for both ROC-AUC and PR-AUC measures. The ablation study attempts to ascertain the importance of every architecture and training aspect ingested into the CyberBERT_LSTM_Hybrid system for the cyber threat texts. As observed in results in the ablation study table for CyberBERT_LSTM_Hybrid system, the overall system performs the best with an accuracy of 0.98.

Removing the LSTM layer (CyberBERT-only classifier) shows a noticeable drop, decreasing the F1-score to 0.96 from 0.98, and the recall to 0.97 from 0.99, emphasizing the role of modeling sequential dependencies to effectively encode ordered cyber threat stories. A removal experiment excluding CyberBERT, which purely relies on the LSTM-based embeddings, shows a steeper drop, decreasing the accuracy to 0.93, F1-score to 0.93, and the ROC-AUC to 0.95, pointing out the significant role of the context-specific transformer embeddings in understanding the semantic concepts of cyber threats.

Table V: Ablation Study Results for Cyber Threat Detection and Textual Threat Analysis

Model Variant (Ablation)	Accuracy	Precision	Recall	F1-score	ROC-AUC	PR-AUC
Full Model (CyberBERT–LSTM)	0.98	0.97	0.99	0.98	1.00	0.99
Without LSTM (CyberBERT only, [CLS] classifier)	0.96	0.95	0.97	0.96	0.98	0.97
Without CyberBERT (LSTM only, word embeddings)	0.93	0.92	0.94	0.93	0.95	0.94
Without CyberBERT fine-tuning (frozen encoder + LSTM)	0.95	0.94	0.96	0.95	0.97	0.96
Without domain-adaptive preprocessing (no URL/IP/token normalization)	0.94	0.93	0.95	0.94	0.96	0.95
Without class weighting / imbalance handling	0.94	0.96	0.91	0.93	0.96	0.94
Without dropout/regularization	0.95	0.94	0.97	0.95	0.97	0.96

Setting a freeze during training for the CyberBERT encoder impacts performance adversely (F1 = 0.95, PR-AUC = 0.96), establishing the need for domain-adaptive fine-tuning in recognizing constantly changing terms in the field of cybersecurity. Without domain adaptation preprocessing, the F1 measure falls to 0.94, establishing the need for URL, IP-address, and token normalization in text-based threats. Without class imbalance methods, recall drops seriously to 0.91, despite a high level of precision (0.96), establishing a growing concern for overlooked threats in cyberspace. Lastly, without dropout

regularization methods, a decrement in generalization ability (F1 = 0.95) was witnessed, establishing a growing concern for overfitting. The ablation experiments have been very effective in verifying the validity of architectural as well as methodology-based design components of the proposed framework as mediated by its AI-based NLP approach.

I. Discussion and Insights

This paper proves that intelligent detection of cyber threats in textual data is significantly aided by a hybrid framework involving both understandings of context and approaches reflecting sequence. The proposed "CyberBERT-LSTM Hybrid Model," established in a supporting framework of "AI-Driven NLP," proves significantly better than traditional machine learning and pure deep learning approaches in terms of several aspects of its performance. The remarkably high scores of "0.99" in Recall, "0.98" in F1 score, and "1.00" in perfect ROC-AUC scores illustrate that its performance in accurately detecting and avoiding threats of cyber attacks is highly commendable. The experimental analysis shows that the contextual embeddings of the transformer play a dominant part in threat intent with semantics, attack vocabulary with semantics of evolution, and implicit semantics of mal intent, with the LSTM part of the model improving the representation of sequences and narrative-level threats patterned in threat descriptions within the cyber space. Precision Recall and threshold test analysis demonstrate the insensitivity of the proposed model to the threat of class imbalance and sensitivity to operational thresholds of the proposed system in the SOCs application. The contribution of the proposed components to the overall model performance was confirmed by the experiment of ablation. In general, the observations made from this research point towards the importance of integrating contextual, sequential, and domain-knowledge-based learning components for effective cyber threat analysis on text inputs. The developed framework represents a scalable solution for automated threat intelligence, alert handling, and proactive monitoring systems for cyber protection.

V. CONCLUSION AND FUTURE WORK

This paper has discussed an AI-supported NLP framework, SmartCTD, using an advanced hybrid model named CyberBERT-LSTM, to perform intelligent cyber threat identification. The work explains how this AI-supported model combines domain-level context embeddings with sequence dependencies to effectively address limitations in both machine learning and deep learning techniques. Experiments conducted to verify this proposal clearly demonstrate that the AI-supported model, CyberBERT-LSTM, consistently outperforms other techniques across high-value metrics such as accuracy, precision, recall, and F1-Score, achieving near-optimal results in both ROC AUC and PR AUC. A high recall value, along with low false-negative rates, indicates this model's effectiveness and applicability to real-time cybersecurity tasks, since any loss in this sector will have drastic consequences. Additionally, test results verifying sensitivity and ablation also prove the relevance of all model elements. In general, the combination of contextual intelligence and sequential learning in text-based cyberspace threat analysis has proven highly effective. The proposed system can serve as a reliable, scalable automated solution for cyberspace threat intelligence and as a strong foundation for other studies on cyberspace adaptive and intelligent systems. One area for future research is combining few-shot and zero-shot learning to further improve robustness against ever-emerging and unseen cyber threats. The use of RAG and knowledge

graph reasoning can further enhance threat context understanding and the robustness of models, which is essential to achieving trust in operations. Optimizing this framework for real-time streaming data and testing it on multilingual, cross-domain datasets can further enhance its usability in dynamic cyber threat scenarios.

REFERENCES

1. Brown, A., Williams, P., & Carter, J. (2025). *SecureBERT 2.0: Advancing domain-specific language models for cybersecurity intelligence*. arXiv. https://arxiv.org/abs/2510.00240
2. Tellache, M., Korba, A., Mokhtari, S., Moldovan, D., & Ghamri-Doudane, A. (2025). *CYLENS: Leveraging large language models for cyber threat intelligence and incident response*. arXiv. https://arxiv.org/abs/2502.20791
3. Sarker, S., Rahman, M., & Das, A. K. (2025). Natural language processing techniques for cyber threat intelligence: A comprehensive survey. *Computer Science Review, 49*, 100–128.
4. Sorokoletova, K., Antonioni, D., & Colò, M. (2025). *0-CTI: Zero-shot cyber threat intelligence extraction using transformer-based models*. arXiv. https://arxiv.org/abs/2501.06239
5. Khan, T., Malik, M., Aziz, Z., Abid, M. K., & Sabir, M. (2025). A comparative study of machine learning and deep learning models for cyber threat text classification. *Journal of Cybersecurity and Information Systems, 9*(2), 45–60.
6. Rahman, M., Islam, S., & Ahmed, N. (2025). Threat intelligence automation using NLP and machine learning. *International Journal of Computer Research and Technology, 13*(4), 210–219.
7. Minaee, N., Azimi, E., & Wang, Y. (2025). Transformers and large language models for intrusion detection systems: A survey. *Information Fusion, 102*, 102–121.
8. Rahman, M., Hasan, R., & Al Mamun, A. (2025). Automated cyber threat intelligence extraction using NLP-based machine learning models. *International Journal of Computer Trends and Technology, 73*(1), 1–10.
9. Sorokoletova, K., Antonioni, D., & Colò, M. (2025). *Zero-shot learning for cyber threat intelligence knowledge extraction*. arXiv. https://arxiv.org/abs/2501.06239
10. Tellache, M., Korba, A., Mokhtari, S., Moldovan, D., & Ghamri-Doudane, A. (2025). *Retrieval-augmented generation for autonomous cyber incident response*. arXiv. https://arxiv.org/abs/2508.10677
11. Al-Yasiri, A., Hossain, M., & Alam, R. (2025). *A conceptual NLP-based framework for multilingual cyber threat intelligence processing*. arXiv. https://arxiv.org/abs/2506.03551
12. Minaee, N., Abdolrashidi, A., & Khoshgoftaar, S. (2025). Attention-based models and large language models for cybersecurity applications: A survey. *Computer Science Review, 50*, 1–24.
13. Sarker, S., Colman, A., & Hossain, M. I. (2025). NLP-driven cybersecurity: Techniques, applications, and challenges. *Journal of Emerging Technologies and Network Research, 6*(3), 55–70.
14. Wang, Y., Zhang, L., & Chen, X. (2025). Optimized cyber threat classification using hybrid LLM and machine learning models. *Scientific Reports, 15*(1), 1–14.
15. Ferrag, M. A., Maglaras, L., & Janicke, H. (2025). Opportunities and risks of large language models in cybersecurity. *AI, 6*(9), 216–235.
16. Memon, N. A., Ali, A., Longa, F. E. A., & Awan, D. (2025). Natural language processing techniques for cybersecurity threat analysis in multilingual environments. *Security and Emerging Systems Journal, 4*(2), 88–101.
17. Khan, T., Malik, M., Aziz, Z., Abid, M. K., & Sabir, M. (2025). Evaluation of classical and deep learning models for text-based cyber threat detection. *Journal of Cybersecurity and Information Science, 10*(1), 25–39.
18. Lohare, S. T., Maaz, M., Razi, M., Nehal, M., & Ahmed, S. T. (2025, February). Road Rage Detection System Using Deep Learning and Computer Vision. In *2025 3rd International Conference on Integrated Circuits and Communication Systems (ICICACS)* (pp. 1-8). IEEE.
19. Ahmed, S. T., Fathima, A. S., & Reema, S. (2023, December). An Improved System for Students Feedback Analysis Using Supervised Probability Techniques. In *2023 10th IEEE Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON)* (Vol. 10, pp. 328-333). IEEE.
20. Seetharaman, S. K., & Syed, T. A. (2025). An Automated Medical Diagnosis System for Neoplasm Medical (MRI) Image Classification using Supervised and Unsupervised Techniques.