Journal of Trends in Computer Science and Smart Technology (ISSN: 2582-4104)
www.irojournals.com/tcsst/

Lyrebird Green Anaconda Optimization based Bayesian Hierarchical Neural Attention Harmonic Network for Illicit Dark Web Classification

Bollikonda Vinod Babu[1], **Kiran K V D.**[2]

[1]Research Scholar, [2]Professor, Department of Computer Science and Engineering,
Koneru Lakshmaiah Education Foundation, Vaddeswaram, Andhra Pradesh, India.

Email: [1]2002031023@kluniversity.in, [2]kiran_cse@kluniversity.in

Abstract

The quick development of internet technology has opened the door for a number of illegal activities targeted at users. The fact that these malevolent actions are usually carried out by anonymous people or organizations makes identification and tracking more difficult. In order to address these problems, a novel method for categorizing illegal dark content has been created called the Lyrebird Green Anaconda Optimization-based Bayesian Hierarchical Neural Attention Harmonic Network (LGAO_BHNAHN). To find and extract pertinent information, textural content extraction is done first. Following that, GPT-NEOX receives the extracted contents and uses them to process and produce text or passages. The Bayesian Hierarchical Neural Attention Harmonic Network (BHNAHN) is then used to classify illicit dark web content. However, the Bayesian Neural Network (BNN) and the Hierarchical Neural Attention classifier with Forward Harmonic analysis are combined to create BHNAHN. Additionally, Lyrebird Green Anaconda Optimization (LGAO), which combines the Lyrebird Optimization Algorithm (LOA) and Green Anaconda Optimization (GAO), is used to train BHNAHN. Lastly, GLOA-trained HNAHN is used to classify drug and arms types. The proposed framework makes a significant advancement in secure, real-time threat detection by combining GPT-NEOX with a novel Bayesian Hierarchical Neural Attention Harmonic Network optimized by the Lyrebird Green Anaconda Algorithm (LGAO). It achieves 93.30% accuracy, with a False Positive Rate (FPR) of 5.62% and a True Positive Rate (TPR) of 92.85% in classifying illicit dark web content.

DOI: https://doi.org/10.36548/jtcsst.2025.2.007
Received: 18.05.2025, received in revised form: 07.06.2025, accepted: 21.06.2025, published: 01.07.2025

Keywords: Dark Web, Illicit Dark Web Classification, Hierarchical Neural Attention Classifier, Lyrebird Optimization Algorithm, Green Anaconda Optimization.

1. Introduction

Natural Language Processing (NLP) is a branch of Artificial Intelligence (AI) that aims to understand and interpret human communication utilizing computational Machine Learning (ML) approaches [1]. The main objective of NLP is to capture the essence of human language, allowing algorithms to understand the meanings of whole sentences as expressed by individuals. An NLP model is used for analysing the expressiveness of phrases, discerning a person's emotions or desires on the basis of their word choices, and identifying similarities in intentions across diverse sentences [2] [3]. By utilizing NLP approaches, researchers can devise algorithms that evaluate data for detecting and categorizing various types of illicit content, such as illegal marketplaces, forums for cybercriminals, or discussions regarding drugs as well as weapons [4]. Currently, the Internet plays an essential role in both our personal and professional lives and can be categorized into three segments: the Deep Web, Surface Web and Dark Web. The Surface Web is the familiar portion of the Internet that most people access regularly [5]. Similarly, the Deep Web comprises content, which is not indexed by commercial search engines like Google, making it inaccessible for web crawlers. The dark web is accessible to the public but needs specialized encryption tools such as the Onion Router (TOR) for access. TOR allows the creation of hidden services that help websites to be hosted anonymously [6]. The dark web comprises hidden services accessible through anonymous communication systems like TOR, where service providers can conceal server locations while offering various network services. Unlike the surface web, accessing these services requires specialized browsers or proxy configurations [7]. Dark web crawling presents significant challenges due to TOR's unique characteristics, particularly the prevalence of unrelated websites creating weak inter-site connections that complicate efficient information tracking. The platform's anonymity has increasingly facilitated criminal activities, as individuals exploit law enforcement's difficulty in tracing actual IP addresses to distribute illicit content [8]. The primary objective of this work is to develop LGAO_BHNAHN, a novel framework for classifying illicit dark web content. Initially, textual content is extracted to identify relevant information, which is then processed by GPT-NEOX for advanced language modeling. The processed content is classified using BHNAHN, a model built by integrating a Bayesian Neural Network (BNN), a Hierarchical Neural Attention classifier, and Forward Harmonic Analysis. This classifier is trained using the Lyrebird Green Anaconda Optimization (LGAO) algorithm, which combines Genetic Algorithm Optimization (GAO) and the Lyrebird Optimization Algorithm (LOA). The system effectively classifies categories such as drugs, pornography, hacking, and arms under multilingual and dynamic dark web conditions, offering a real-time, accurate, and robust classification framework.

2. Literature Survey

Shin et al. [16] employed TextCNN for dark web classification, effectively capturing local n-gram structures to understand textual features, though scalability issues emerged when processing massive unstructured datasets. Jin et al. [13] developed Comprehensive Dark web Annotations (CoDA) for classification, enhancing cybersecurity through sophisticated AI monitoring tools, but potentially enabling malicious actors to refine evasion strategies. Sangher et al. [17] implemented classification methods that improved vulnerability exploitation response times, yet false negatives could result in missed threats and inadequate attack preparedness. Alaidi et al. [11] utilized BERT for dark web classification, though the approach lacked predictive capabilities for future illegal activities and remained vulnerable to poor data quality affecting mining accuracy. G. Cascavilla., *et al.* [5] introduced Bidirectional Encoder Representations from Transformers (Bert) for classifying illicit dark web content. This approach effectively classified content in multiple languages within the diverse linguistic environment of the dark web. However, it raised significant ethical concerns, including issues of privacy and the potential misuse of technology, which might result in surveillance challenges. A. Dalvi., *et al.* [18] devised Term Frequency-Inverse Document Frequency (TF-IDF) and Recurrent Neural Network (RNN) for classifying illicit dark web content. This technique could capture the order of words and their interdependencies, which was beneficial for understanding documents where meaning might change depending on context. However, the training durations expanded the processes of iterative testing as well as optimization. M. K. Alshammery and A. F. Aljuboori., [10] introduced TF-IDF for classifying illegal activities. This approach managed increasing data volumes and adapting to varying network conditions without the requirement for a total system. However, different components might not always integrate smoothly which leads to integration problems that complicate the overall functionality. C. A. S. Murty and H. R., [19] devised Support Vector Machines (SVM) algorithm for classifying illicit activities. This technique maximized the margin between classes, which led to better generalization on unseen data. Although, it might struggle with imbalanced databases, leading to poor classification performance for underrepresented classes.

3. Proposed Work

The dark web includes diverse illicit activities conducted by anonymous individuals or groups, making traceability demanding. The constantly evolving nature of illegal content makes its collection and classification complex, presenting significant challenges. This issue has recently gained urgent attention from both industry and academia because of its complex and time-sensitive nature. Therefore, the main goal of this research is to devise LGAO_BHNAHN for illicit dark web classification. First, dark web data obtained from a database is passed to textural content extraction. Then, textural content extraction is carried out to identify and extract relevant information.

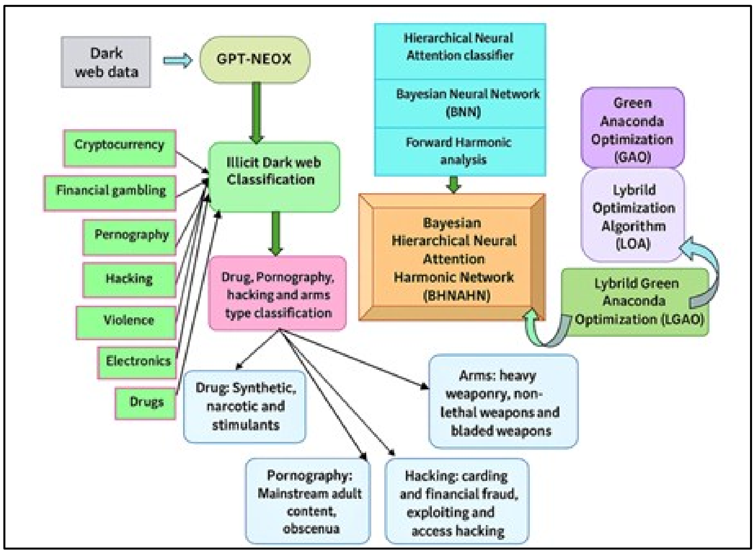

Dark web data
GPT-NEOX
Hierarchical Neural Attention classifier
Bayesian Neural Network (BNN)
Forward Harmonic analysis
Green Anaconda Optimization (GAO)
Lybrild Optimization Algorithm (LOA)
Cryptocurrency
Financial gambling
Pernography
Hacking
Violence
Electronics
Drugs
Illicit Dark web Classification
Drug, Pornography, hacking and arms type classification
Bayesian Hierarchical Neural Attention Harmonic Network (BHNAHN)
Lybrild Green Anaconda Optimization (LGAO)
Drug: Synthetic, narcotic and stimulants
Arms: heavy weaponry, non-lethal weapons and bladed weapons
Pornography: Mainstream adult content, obscenua
Hacking: carding and financial fraud, exploiting and access hacking

Figure 1. LGAO_BHNAHN for Illicit Dark Web Classification

After that, the extracted content is fed to GPT-NEOX for processing and generating text or passages. Moreover, illicit dark web classification into pornography, drugs, financial gambling, cryptocurrency, hacking, arms/weapons, violence, and electronics is accomplished using BHNAHN. Here, BHNAHN is an incorporation of BNN Hierarchical Neural Attention classifier, and Forward Harmonic analysis. Additionally, BHNAHN is trained by LGAO, which is designed by incorporating GAO and LOA. Finally, drug, pornography, hacking, and arms type classification is done utilizing HBFHNet and it is trained by LGAO. Figure 1 elucidates block diagram of LGAO_BHNAHN for illicit dark web classification, and similarly, Figure 2 details the simplified framework of the same.

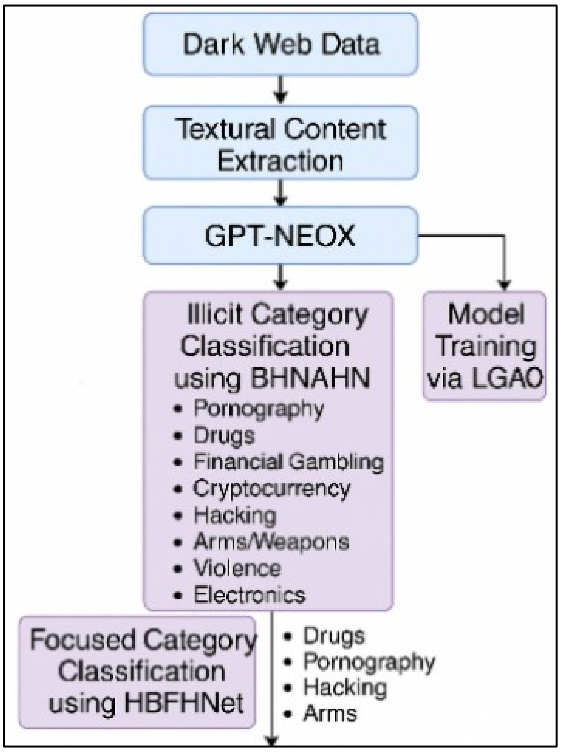


Figure 2. Framework of the Proposed Model

3.1 Web Data Collection

The TorBot [28] encompasses keywords such as pornography, financial gambling, drugs, hacking, cryptocurrency, arms/weapons, electronics, and violence. The collected web data from various websites can be represented based on the specified keywords.

$$C = \{C_1, C_2, \dots, C_x, \dots, C_m\} \tag{1}$$

Here, C implies TorBot database, C_x elucidates x^{th} data present in total $_m$ samples.

3.2 Textural Content Extraction

The initial technique for reading HTML files employs the urllib.request library in Python. Thereafter, it utilizes the Beautiful Soup library to carry out several key operations: starts with Remove all script and style elements, Extract the remaining text from the HTML, then Split the text into lines and eliminate any leading or trailing spaces on each line, and proceed with Break each of the multi-headlines into individual lines and finally Remove all blank and empty lines. The second technique is useful for extracting only the title and description from HTML pages, mainly targeting chosen marketplaces. This approach streamlines the content by eliminating duplicate words, concentrating mainly on the titles and descriptions. In addition, the excluded pages contained only product listings that did not have sufficient descriptions of items. This technique is not consistently applied to complete pages linked with a marketplace, as it requires thorough inspection and does not provide itself to generalization. It is employed mainly for marketplaces because of their reliable page templates. Moreover, the different pages from the Dark Web show diverse formats. Once this technique compiles the text from both approaches, proceed to the next phase of preparation.

3.2.1 Text Preprocessing

Textual data preprocessing is an essential feature of text classification tasks. Typically, the preprocessing phase consists of steps like tokenization, elimination of stop words, alteration to lowercase, and stemming. However, the tokenization technique varies from conventional NLP approaches, and the process implemented is broken down into six distinct steps that are explained as the any HTML tags using the Beautiful Soup Python library has to be eliminated, the proceed with eliminating URLs utilizing the re library, which offers consistent expression matching functions also changing contractions followed by eliminate all special characters such as currency signs as well as words that contain numbers. The next Preprocessing process involves eliminating articles, prepositions, and pronouns, as these components are frequently unnecessary for tasks like text classification. To acquire this, utilize the stop word collection from the NLTK Python library. Finally, apply the WordNetLemmatizer from the NLTK library for lemmatizing the processed text and effectively normalizing the database. Moreover, all the text is converted to lowercase. The outcome gained from this phase is elucidated as Z_x.

3.3 GPT-NEOX

The input given to this phase is Z_x. GPT-NeoX-20B [20] which refers to an autoregressive transformer decoder approach that primarily utilizes the architecture of GPT-3, with a few significant differences. The approach comprises of 20 billion parameters, with 19.9 billion categorized as non-embedding parameters, which is a suitable count for scaling laws evaluation.

3.3.1 Model Structure

This structure is the advanced version of GPT-3 with various differences. The model structure is nearly equal to that of GPT-J 2; however, it chose GPT-3 because there is no design of GPT-J.

- **Rotary Positional Embeddings**

The utilization of rotary embeddings instead of learned positional embeddings is found in GPT techniques because it provides good results in training large language models (LLMs). Rotary embeddings refer to a kind of fixed relative positional embedding, and they adjust the embedding space so that the attention token at position s gives to a token at location p depends linearly on the difference $s-p$. Moreover, this explains how the varied attention works based on the token's positions, and it is explained below,

$$Soft\max\left(\frac{1}{\sqrt{q}}\sum_{p,s} B_p^U Z_t^U Z_w B_s\right) \tag{2}$$

Here, B_p and B_s elucidates the token embeddings at location p and s. Z_t^U and Z_w indicates the query as well as key weights.

$$Soft\max\left(\frac{1}{\sqrt{q}}\sum_{p,s} B_p^U Z_t^U V_{\Theta(s-p)}^q Z_w B_s\right) \tag{3}$$

Here, $V_{\Theta(s-p)}^q$ implies a $q \times q$ block diagonal matrix, wherein the h^{th} block equivalent to a 2D rotation matrix described through the angle $\alpha\theta_h$ for hyperparameters that is expressed below,

$$QR = \left\{\theta_h = 10000^{\frac{-2h}{q}} \middle| h \in \{0,1,2,...,(q-1)/2\}\right\} \tag{4}$$

Instead of applying rotary embeddings to the complete embedding vector, which limits their application to only the first 25% of the sizes. Moreover, the experiments suggest that this technique offers an optimum balance between performance and computational effectiveness.

- **Parallel Attention + FF Layers**

This technique calculates the Attention and Feed-Forward (FF) layers simultaneously and then sums their outcomes. This technique is mainly useful for efficiency because every residual addition with op-sharding requires an all-reduce operation during both the forward and backward passes. By running these layers in parallel, this technique decreases the outcomes locally before performing a single all-reduce. This technique results in a 15% enhancement in throughput in the Mesh Transformer while upholding similar loss curves during earlier training in comparison with running in series. However, due to an error, this approach applied two separate Layer Norms instead of employing a shared Layer Norm.

$$\alpha + C\left(WK_1\left(\alpha\right)\right) + \varpi\left(WK_1\left(\alpha\right)\right) \tag{5}$$

Here, C expounds attention layer and ϖ signifies FF layers. Moreover, the codebase decouples the layer normalization, which is explained beneath,

$$\alpha + C\left(WK_1\left(\alpha\right)\right) + \varpi\left(WK_2\left(\alpha\right)\right) \tag{6}$$

- **Initialization**

For the FF resultant layers preceding the residual connections, this technique employed a specific initialization strategy. This technique aids in preventing the activations from enhancing excessively because of deeper and wider structures. The factor of 2 is included to offset the effects of having both parallel and FF layers arranged in parallel. For all additional layers in the model, it applies the small initialization strategy.

- **All Dense Layers**

While GPT-3 utilizes an integration of alternating dense and sparse layers through this technique, it utilizes only dense layers for simplifying the implementation. The outcome gained from the GPT-NEOX is indicated as content G_x that relates to the above keywords.

3.4 Illicit Data Web Classification Using BHNAHN

It includes the procedure of detecting and categorizing online content that is illegal or harmful, including forums related to illicit activities. A main demand in this field is the fast evolution of the internet, where newer platforms and communication approaches frequently emerge that often affect the present classification systems. Here, BHNAHN was employed for illicit data web classification. BHNAHN is an integration of BNN, Hierarchical Neural Attention classifier, and forward harmonic concept. At first, content G_x is concatenated, with H_1, thereby the addition of weight $\sum H_1$ is obtained. Similarly, content G_x is given to BNN and its resultant is concatenated with $\sum H_1$ to obtain BNN resultant P_1. Moreover, content G_x is subjected to Hierarchical Neural Attention classifier and then, weight H_2 is concatenated to its resultant. The acquired outcome is again concatenated with BNN resultant to accomplishing the resultant P_2 output. At last, the forward harmonic concept is applied to both resultants for obtaining the final classified resultant as P_3 classifies dark web illicit activities such as drugs, hacking, financial gambling, and violence.

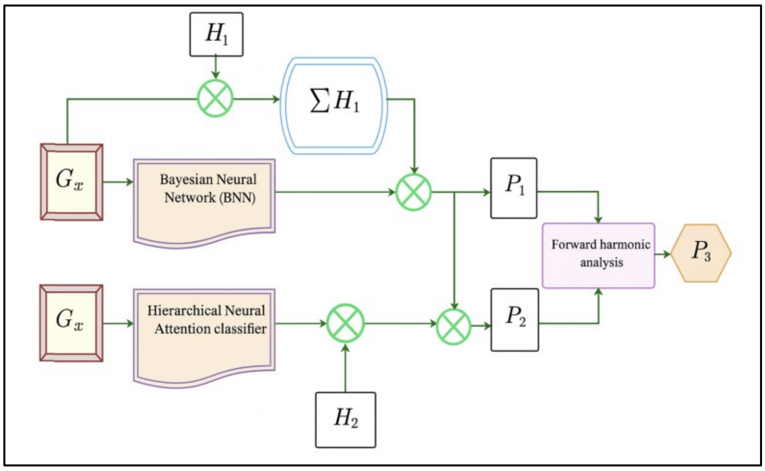

H_1
$\sum H_1$
G_x
Bayesian Neural Network (BNN)
P_1
Forward harmonic analysis
P_3
G_x
Hierarchical Neural Attention classifier
P_2
H_2

Figure 3. Structure of BHNAHN

3.4.1 BNN

BNN [21] [22]employs Bayesian inference to evaluate the uncertainty in its predictions. Unlike traditional neural networks that offer point estimates for their weights, the BNN approach designs these weights as distributions. This permits the BNNs to accommodate the uncertainty linked with both the data and the technique itself, resulting in credible intervals for predictions rather than mere single-point predictions. The input passed to the BNN is expounded as content, G and the updated expression for the BNN is explained below,

$$P_1 = \left[\sum_{f=1}^{\mu} \beta_f(r,y) N_f(r,y) + \rho(r,y)\right] * \left[\sum H_1 * G_x\right] \tag{7}$$

Here, G expounds the content and H_1 elucidates weight, r elucidates the data location and the time is expressed as y. Moreover, the weight is signified as $\beta_f\left(r,y\right)$ and $\rho\left(r,y\right)$ indicates the bias. The outcome gained from BNN is expounded as P_1.

3.4.2 Hierarchical Neural Attention Classifier

The Hierarchical Neural Attention classifier is devised for tackling the demands linked with explosion approaches. This classifier utilizes a single encoder-decoder structure that sequentially predicts the class label at every hierarchical level, employing a dynamic illustration of web data improved by modifications to the attention mechanism. In this framework, the content G is passed as the input, and the output expression is determined below,

$$P_2 = \left[\left[soft\max\left(\mathrm{A}_c\tau_c\right)\right] * H_2\right] * \left[\left[\sum_{f=1}^{\mu}\beta_f\left(r,y\right)N_f\left(r,y\right)+\rho\left(r,y\right)\right] * \left[\sum H_1 * G_x\right]\right] \quad (8)$$

Here, H_2 indicates weight and the resultant gained from this phase is expressed as P_2. Moreover, by applying the concept of harmonic analysis concept, the updated expression is computed as follow,

$$\begin{aligned} P_3 = &\left[\left[\sum_{f=1}^{\mu}\beta_f\left(r,y\right)N_f\left(r,y\right)+\rho\left(r,y\right)\right] * \left[\sum H_1 * G_x\right]\right] \\ &\left[\left[\frac{1+2\cos\pi\kappa}{2}\right] + \left[\left[soft\max\left(\mathrm{A}_c\tau_c\right)\right] * H_2\right] * \left[\frac{1-2\cos\pi\kappa}{2}\right]\right] \end{aligned} \quad (9)$$

The resultant accomplished by this phase is expressed as P_3, which classifies the dark web into drugs, cryptocurrency, hacking, financial gambling, violence, electronics, and arms/weapons.

3.4.3 Tuning of Weight Utilizing LGAO

The hyperparameters of BHNAHN are trained using LGAO, an integrated optimization algorithm combining the Lyrebird Optimization Algorithm (LOA) and Green Anaconda Optimization (GAO). LOA is a population-based algorithm inspired by the vocal mimicry and adaptive behavior of the lyrebird, a bird species known for imitating diverse natural and artificial sounds. GAO, on the other hand, is based on the hunting strategies of green anacondas, particularly their ambushing and constricting behavior. LGAO is well-suited for the evolving and dynamic nature of illicit content on the dark web, as it can effectively adapt to changing data patterns. It operates in two main phases: an exploration phase, where LOA's strategies broadly search the solution space to maintain diversity, and an exploitation phase, where GAO's local search techniques refine the solutions.

- **Solution Encoding**

It is utilized for attaining an efficient solution for classifying illicit dark web content within a search space ψ, thereby $\psi = [1 \times \Upsilon]$, implying Υ the learning factor of BHNAHN.

- **Fitness Computation**

It is a specific type of objective function that quantifies the nearness of a given solution to the optimal solution of a problem that is expounded below,

$$T = \frac{1}{m}\sum_{x=1}^{m}[E - P_3]^2 \tag{10}$$

where, E implies targeted resultant. The steps of LGAO are expressed beneath,

Step 1: Solution Initialization

Initialize Lyrebird's location, which is identified arbitrarily in the search space and it is computed in the expression below,

$$D = \{D_1, D_2, ..., D_u, ..., D_v\} \tag{11}$$

where, D_u implies u^{th} member, whereas D_v signifies total variables in D population.

Step 2: Evaluation of Objective Function

It is beneficial for assessing the discrepancy between the observed outcome and the target outcome, which is calculated using Eq. (9).

Step 3: Phase 1-Escaping Strategy (Exploration Phase)

In this phase, the location of each population member is updated within the search space by simulating the escaping behavior of the lyrebird from danger to find secure regions. In LOA, every individual in the population recognizes the locations of additional members that exhibit superior objective function values as safe zones. Moreover, the collection of safe regions for every member of the LOA population is computed below,

$$R_u = \left\{ I_e, T_e < T_u \;\; and \;\; e \in \left\{1,2,...,v\right\}\right\}, \quad where\, u = 1,2,...,v \tag{12}$$

Here, R_u expounds on the group of secure regions for u^{th} lyrebird, I_e implies e^{th} row of I^{th} matrix that obtains a superior fitness value than u^{th} LOA member.

In LOA, it is considered that the lyrebird escapes arbitrarily to one of the chosen secure regions. By using the lyrebird displacement modeling, every member of the LOA attains a new location that is computed below,

$$D^{p1}_{u,g} = D_{u,g} + s_{u,g}.\left(SW_{u,g} - O_{u,g}.D_{u,g}\right) \tag{13}$$

$$D^{k+1}_{u,g} = D_{u,g} + s\ .\left(SW_{u,g} - O_{u,g}.D_{u,g}\right) \tag{14}$$

$$D^{k+1}_{u,g} = D_{u,g} + \left(1 - O_{u,g}\right) + s.SW_{u,g} \tag{15}$$

The standard expression of GAO for illicit data web classification is explained as follows,

$$A^{P}_{o,n} = A_{o,n} + \left(1 - 2j_{o,n}\right)\frac{WE_n - LE_n}{z} \tag{16}$$

Let us consider,

$$A^{P2}_{o,n} = D^{k+1}_{u,g} \tag{17}$$

$$A_{o,n} = D^{k}_{u,g} \tag{18}$$

$$j_{o,n} = s_{u,g} \tag{19}$$

Then, the expression becomes,

$$D^{k+1}_{u,g} = D^{k}_{u,g} + \left(1 - 2s_{u,g}\right)\frac{WE_n - LE_n}{z} \tag{20}$$

$$D^{k}_{u,g} = D^{k+1}_{u,g} - \left(1 - 2s_{u,g}\right)\frac{WE_n - LE_n}{z} \tag{21}$$

Substitute Eq. (21) in (15), the expression becomes

$$D^{k+1}_{u,g} = \left(D^{k+1}_{u,g} - \left(1 - 2s_{u,g}\right)\frac{WE_n - LE_n}{z}\right)\left(1 - O_{u,g}\right) + s.SW_{u,g} \tag{22}$$

$$D^{k+1}_{u,g} - D^{k+1}_{u,g}\left(1 - 2s_{u,g}\right) = \left(2s_{u,g} - 1\right)\left(\frac{WE_n - LE_n}{z}\right)\left(1 - O_{u,g}\right) + \frac{s.SW_{u,g}.k}{k} \tag{23}$$

$$\left(1-1+O_{u,g}\right)D_{u,g}^{k+1}=\left(2s_{u,g}-1\right)\left(\frac{WE_n-LE_n}{z}\right)\left(1-O_{u,g}\right)+\frac{s.SW_{u,g}.k}{k} \tag{24}$$

The updated expression of LGAO is explained in the beneath expression,

$$D_{u,g}^{(k+1)}=\frac{\left(2s_{u,g}-1\right)\left(WE_n-LE_n\right)\left(1-O_{u,g}\right)+s.SW_{u,g}.k}{t.O_{u,g}} \tag{25}$$

where, SW_u implies the chosen secure region for u^{th} lyrebird, $SW_{u,g}$ elucidates its g^{th} dimension, $s_{u,g}$ implies the arbitrary numbers from $[0,1]$ interval and $O_{u,g}$ indicates the numbers, which are arbitrarily chosen as 1 or 2. Moreover, k indicates the iteration counter and WE_n, LWE_n implies the upper as well as lower bounds, respectively.

If the fitness value is improved, the new location takes the place of the preceding location for the corresponding member and it is expressed in the expression below,

$$I_u=\begin{cases} I_u^{P1}, & T_u^{P1}\le T_{u,} \\ I_u, & else \end{cases} \tag{26}$$

Here, I_u^{P1} implies the new location computed for u^{th} lyrebird on basis of escaping strategy and T_u^{P1} elucidates its fitness value.

Step 4: Phase 2-Hiding Strategy (Exploitation Phase)

During this phase, the locations of population members are updated within the search space in accordance with a design strategy that mimics the lyrebird's capacity to conceal itself in its environment. In LOA, new positions for every member of the population are determined using Eq. (8), which is based on the lyrebird's movement toward nearby regions that are suitable for hiding.

$$D_{u,g}^{P2}=D_{u,g}+\left(1-2s_{u,g}\right).\frac{\varepsilon_g-\sigma_g}{k} \tag{27}$$

This new location replaces the prior one for the corresponding member if it improves the value of the fitness explained utilizing the expression below,

$$I_u=\begin{cases} I_u^{P2}, & T_u^{P2}\le T_{u,} \\ I_u, & else \end{cases} \tag{28}$$

Here, I_u^{P2} illustrates the new location computed for u^{th} lyrebird on basis of the hiding scheme, $I_{u,g}^{P2}$ indicates its g^{th} dimension.

Step 5: Reevaluation of Fitness

The maximal fitness is evaluated until it reaches an optimum solution.

Step 6: Termination

LGAO is terminated after obtaining an optimum solution by continuously performing the above steps. Algorithm 1 presents the pseudocode of LGAO.

Algorithm 1. Pseudo Code of LGAO

```
Input: Fitness Function
Begin
Set LOA population size and maximum number of iterations, Create the initial population matrix randomly using Eq. (11), Evaluate the objective function for everyone using Eq. (10) and identify the current best solution
For each iteration do
    For each LOA member do
            Determine the type of lyrebird defense strategy against predator using Eq. (12)
            If strategy is Phase 1 then
                    Identify candidate secure regions using Eq. (13)
                    Compute new position of LOA member using Eq. (26)
                    Update LOA member using Eq. (26)
            Else (strategy is Phase 2)
                    Compute new position of LOA member using Eq. (27)
                    Update LOA member using Eq. (28)
            End If
    End For
End For - Save the optimal solution
End
```

3.5 Types Classification of Drug, Pornography, Hacking and Arms Using LGAO_BHNAHN

The input passed to this phase is drug AM_x, pornography AS_x, hacking AK_x and arms AH_x. The dark web is classified for various illicit activities, including the trade of drugs, pornography, hacking services, and arms. Drug-related activities on the dark web often involve synthetics, narcotics, and stimulants. Pornography varies widely, encompassing everything from legally ambiguous adult content to illegal materials, including mainstream adult content and coercive content. Hacking services are another prominent category, where individuals can buy or sell personal data, including carding and financial fraud, cryptojacking, and access brokering. Lastly, the trade of arms includes heavy weaponry, non-lethal weapons, and bladed weapons. Here, the type classification is performed using BHNAHN, which is trained using LGAO. BHNAHN is the integration of BNN and the Hierarchical Neural Attention classifier.

3.5.1 BNN

BNN employs Bayesian inference to measure uncertainty in its predictions. Unlike regular neural networks that offer single values for their weights, BNNs represent weights as distributions. This permits BNNs to account for uncertainty in both the data and the model itself, producing credible intervals for predictions instead of just single-point estimates.

3.5.2 Hierarchical Neural Attention Classifier

The Hierarchical Neural Attention classifier is modeled to handle the challenges associated with explosion approaches. It employs a single encoder-decoder structure that predicts class labels one level at a time. It improves this process by modifying the attention mechanism, utilizing a dynamic representation of web data. The classified outcomes gained for drug is UY_x, pornography UK_x, hacking UL_x and arms UN_x.

3.5.3 Tuning of Weight Utilizing LGAO

The hyperparameters of the BHNAHN are trained using LGAO, which integrates LOA and GAO. LOA is motivated by the lyrebird's vocal mimicry and social behaviors, serving as a novel population-based algorithm. On the other hand, GOA draws from the green anaconda's exceptional hunting skill, which focuses on constricting the prey. LGAO is well-suited for diverse problems and datasets, making it particularly effective for the dynamic nature of illicit content on the dark web, where new types of content continually emerge.

- **Solution Encoding**

It is employed for accomplishing an effectual solution for classifying illicit dark web within a search space ψ, thereby $\psi = [1 \times \lambda]$, such that λ signifies the learning factor of BHNAHN.

- **Fitness computation**

It is used to assess the efficiency of an obtained solution for classifying illicit dark web that is computed below,

$$T = \frac{1}{m} \sum_{x=1}^{m} \left[E - UY_x \right]^2 \tag{29}$$

Here, UY_x implies the classified output for drug.

4. Result and Discussion

The outcome of the LGAO_BHNAHN for illicit dark web classification is clarified in the following part.

4.1 Experimental Setup

The Python software is utilized to test the LGAO_BHNAHN for dark web illegal classification. The diagram depicts the individual contributions of every one of the four primary components the GPT-NEOX, Bayesian Neural Network (BNN), Hierarchical Attention Classifier, and LGAO Optimization Algorithm to the system in general. This split allows readers to clearly view the setup of every module, including the critical hyperparameters such as the number of layers, activation functions, learning rates, and optimization methods. The significance of this visualization lies in its ability to increase comprehension and replicability of the proposed methodology. It enhances the transparency of the experimental design by providing practitioners and researchers with the ability to replicate or augment each component without ambiguity. The hybrid character of the system is once more emphasized through this modular architecture, where biological inspiration-based optimization and state-of-the-art deep learning are integrated to enhance performance in illegal dark web classification tasks.

4.2 Dataset and Preprocessing Description

Pornography, financial gambling, drugs, hacking, cryptocurrency, weapons, electronics, and violence are among the keywords included in the TorBot dataset. About 12,000 labelled dark web entries were collected using keyword filters and focused crawling to create the TorBot dataset used in this study. Numerous illegal domains are represented in these entries, such as financial gambling (1,500), weapons (1,800), hacking (2,200), drugs (3,000 entries), and pornography (2,500 entries). Relevant web content is gathered by using these keywords. Ten carefully chosen features are included in this database in addition to nine main features.

4.3 Experimental Results

Table 1 elucidates the experimental results of LGAO_BHNAHN. It comprises input text, GPT outcome, illicit dark web classification, and classified output.

Table 1. Experimental Outcomes of LGAO_BHNAHN

Input	GPT Output	Classification	Output
Initial Access Brokers (IABs) infiltrate networks and sell unauthorized access to other cybercriminals, including ransomware operators.	Initial Access Brokers (IABs) are specialized cyber threat actors who exploit vulnerabilities, conduct credential harvesting, or use malware and abuse RDP to gain unauthorized access to computer networks. They then monetize this access by selling it on dark web marketplaces or directly to other cybercriminal groups, such as ransomware operators, Advanced Persistent Threats (APTs), and cyber espionage actors.	Hacking	Access brokering

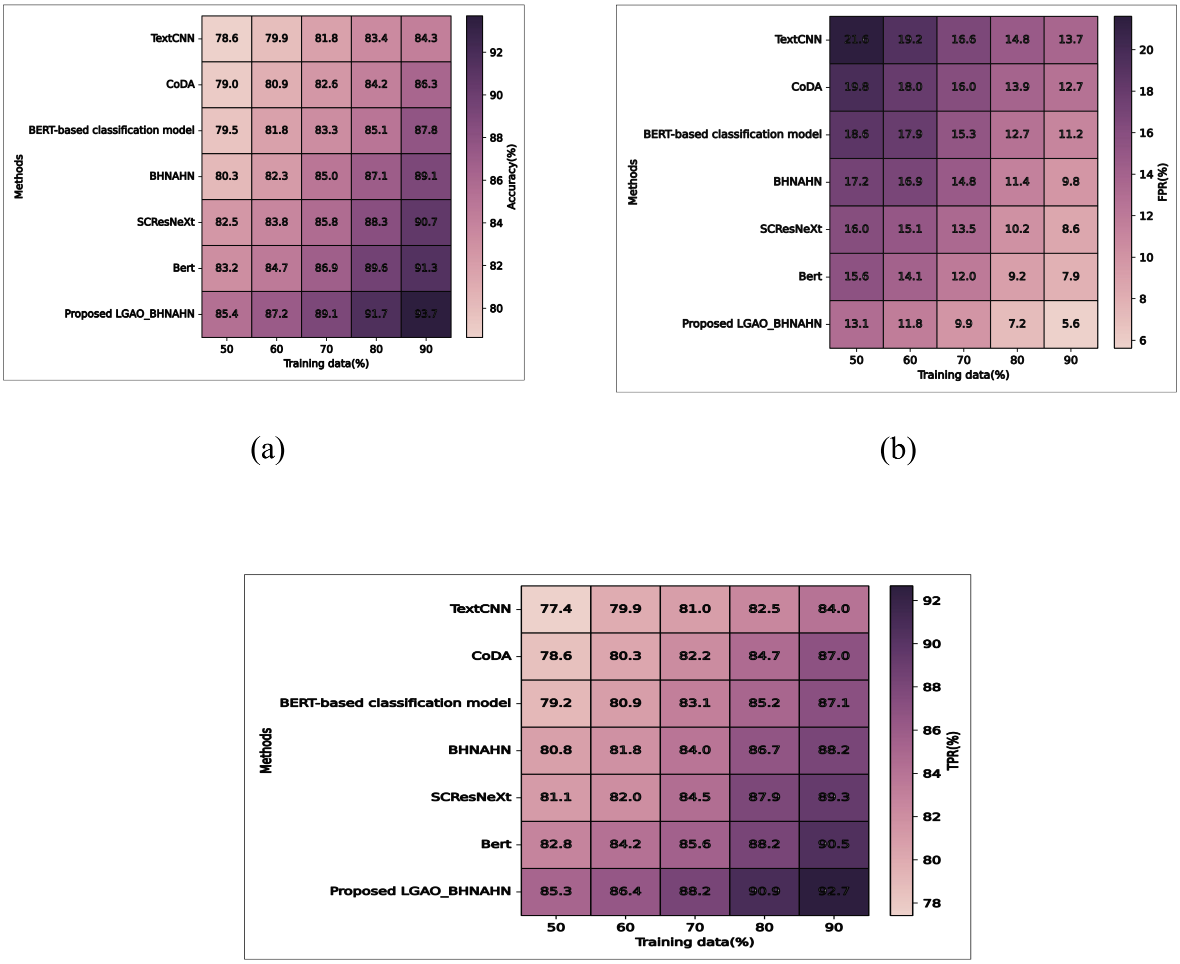
A content analysis found the most common categories of sexual violence in mainstream online pornography were incest, aggression and attacks, image-based sexual abuse	A content analysis of mainstream online pornography identified incest, aggression, attacks, and image-based sexual abuse (sexual violence). These findings highlight the prevalence of harmful and exploitative themes in such content, raising concerns about their impact on viewers and societal perceptions of sexual violence.	Pornography	violent or coercive content
Synthetic drugs are chemically engineered substances that mimic the effects of natural drugs. How are synthetic drugs made? Unlike natural drugs derived from plants or other organic	Synthetic drugs are lab-created dangerous and unpredictable artificial chemical substances designed to mimic natural drugs. Unlike natural drugs derived from plants or organic sources, synthetic drugs are chemically engineered to produce similar effects. Their composition is often modified to enhance potency, alter effects, or evade legal restrictions.	Drugs	synthetic drugs
An edged weapon or bladed weapon is a hand-to-hand combat weapon with a cutting edge. Bladed weapons include swords, daggers, knives, and bayonets	An edged weapon, also known as a bladed weapon, used for stabbing, is a hand-to-hand combat tool with a sharp cutting edge. Examples include swords, daggers, knives, and bayonets.	Arms/Weapons	bladed weapons

4.4 Comparative Assessment

Four configurations are compared for the illicit dark web classification of LGAO_BHNAHN: setup 1 involves weapons, setup 2 involves pornography, setup 3 involves hacking, and setup 4 involves drugs.

4.4.1 Assessment Concerning Setup-1

An evaluation of LGAO_BHNAHN for training data = 90% is shown in Figure 4 with reference to setup 1. The accuracy achieved by LGAO_BHNAHN is 93.70% in Figure 4a), while the accuracy obtained by the classical approaches was 84.28%, 86.29%, 87.81%, 89.08%, 90.67%, and 91.35%. LGAO_BHNAHN's performance is 9.312%, 6.136%, 6.019%, 4.842%, 3.621%, and 2.334% better than that of the current approaches.



(a)

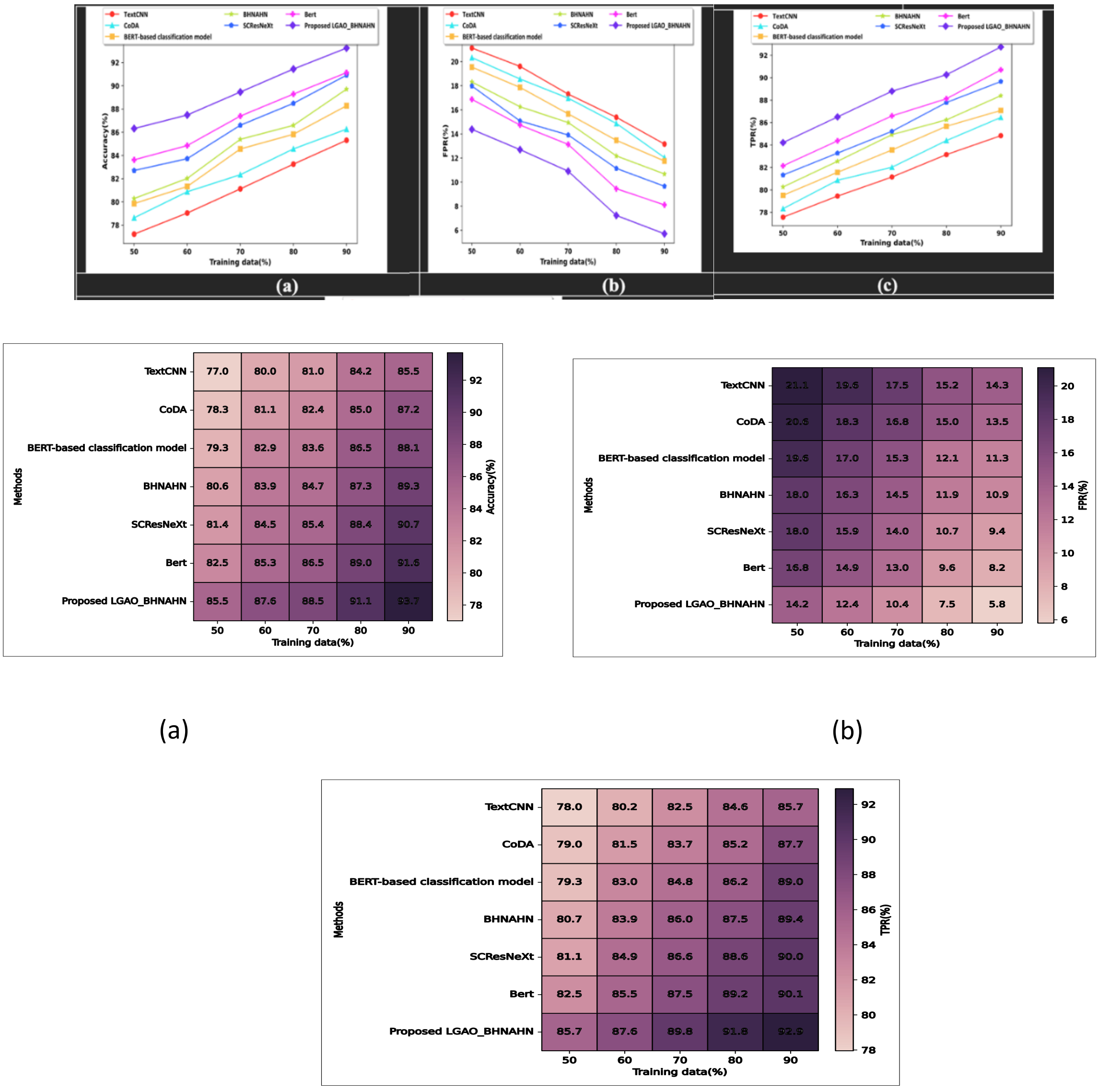
(b)

(c)

Figure 4. Analysis of LGAO_BHNAHN Concerning Setup-1, a) Accuracy, b) FPR, c) TPR

4.4.2 Assessment Concerning Setup-2

Figure 5 uses 90% training data under Setup 2 to assess LGAO_BHNAHN performance. In every metric, the suggested model performs noticeably better than traditional methods. With an accuracy of 93.71%, LGAO_BHNAHN outperforms baseline techniques (85.49%-91.62%) by 2.23%-8.77%.



(a)

Methods	50	60	70	80	90
TextCNN	77.0	80.0	81.0	84.2	85.5
CoDA	78.3	81.1	82.4	85.0	87.2
BERT-based classification model	79.3	82.9	83.6	86.5	88.1
BHNAHN	80.6	83.9	84.7	87.3	89.3
SCResNeXt	81.4	84.5	85.4	88.4	90.7
Bert	82.5	85.3	86.5	89.0	91.6
Proposed LGAO_BHNAHN	85.5	87.6	88.5	91.1	93.7

(b)

Methods	50	60	70	80	90
TextCNN	21.1	19.6	17.5	15.2	14.3
CoDA	20.6	18.3	16.8	15.0	13.5
BERT-based classification model	19.6	17.0	15.3	12.1	11.3
BHNAHN	18.0	16.3	14.5	11.9	10.9
SCResNeXt	18.0	15.9	14.0	10.7	9.4
Bert	16.8	14.9	13.0	9.6	8.2
Proposed LGAO_BHNAHN	14.2	12.4	10.4	7.5	5.8

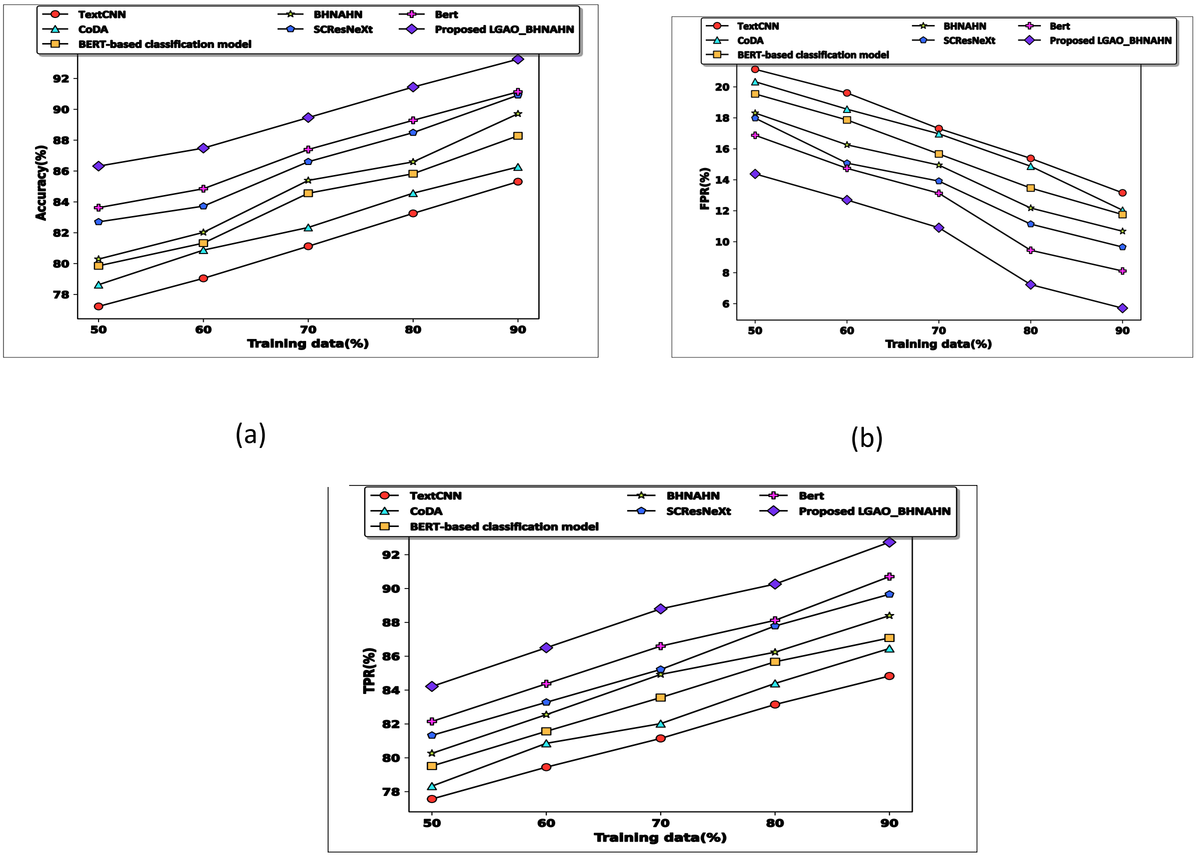
(c)

Methods	50	60	70	80	90
TextCNN	78.0	80.2	82.5	84.6	85.7
CoDA	79.0	81.5	83.7	85.2	87.7
BERT-based classification model	79.3	83.0	84.8	86.2	89.0
BHNAHN	80.7	83.9	86.0	87.5	89.4
SCResNeXt	81.1	84.9	86.6	88.6	90.0
Bert	82.5	85.5	87.5	89.2	90.1
Proposed LGAO_BHNAHN	85.7	87.6	89.8	91.8	92.9

Figure 5. Analysis of LGAO_BHNAHN, a) Accuracy, b) FPR, c) TPR

4.4.3 Assessment Concerning Setup-3

Figure 6 expounds on an estimation of LGAO_BHNAHN for training data =90% is calculated regarding setup 3



(a)

(b)

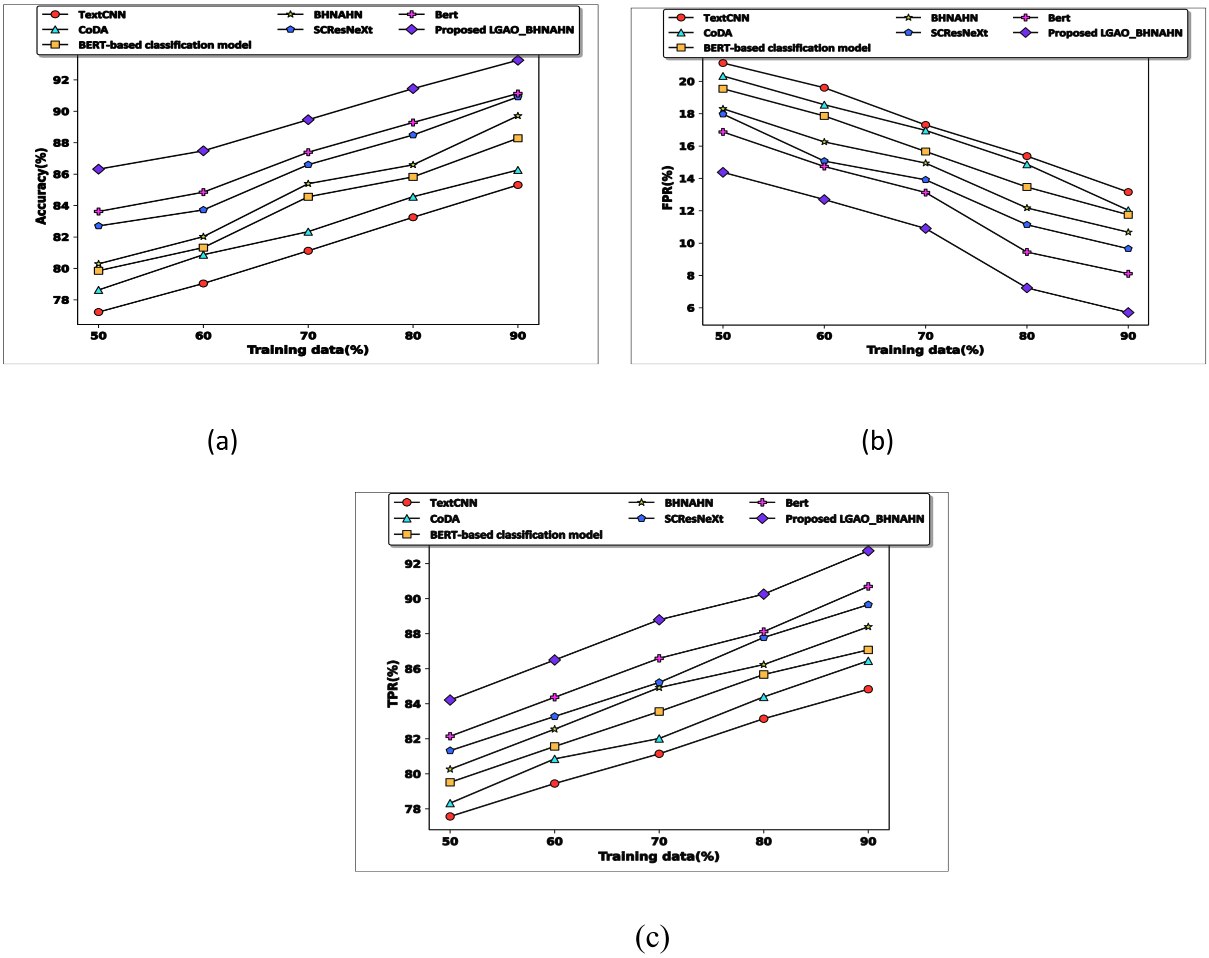
(c)

Figure 6. Analysis of LGAO_BHNAHN Concerning Setup-3, a) Accuracy, b) FPR, c) TPR

Figure 6 b) shows the assessment concerning FPR, with LGAO_BHNAHN attaining an FPR of 5.71%. The enhancement in performance of LGAO_BHNAHN to that of existing methodologies are 8.524%, 6.772%, 6.099%, 4.678%, 3.318% and 2.192%.

4.4.4 Assessment Regarding Setup-4

Figure 7 shows an estimation of LGAO_BHNAHN with a training data percentage of 90% evaluated under setup-4. This expounds a clear advancement in TPR performance with LGAO_BHNAHN over classical techniques, which are 8.890%, 7.175%, 5.820%, 4.696%, 2.752% and 2.362%.



(a)

(b)

(c)

Figure 7. Analysis of LGAO_BHNAHN Concerning Setup-4, a) Accuracy, b) FPR, c) TPR

Figure 7 b) explains an estimation of LGAO_BHNAHN in regard to FPR, where LGAO_BHNAHN recorded an FPR of 5.62%, in comparison with prior models that displayed FPR values of 13.96%, 12.76%, 11.90%, 10.76%, 9.74% and 8.08%. The performance gains of LGAO_BHNAHN compared to traditional strategies are 7.891%, 6.767%, 5.604%, 4.136%, 2.964% and 2.175%.

4.5 Comparative Discussion

Table 2 describes the comparative discussion. LGAO_BHNAHN attains superior results in comparison with TextCNN, CoDA, the BERT-based classification model, BHNAHN, SC ResNeXt and Bert for setup 4. The accuracy value attained by LGAO_BHNAHN is 93.30%, FPR of 5.62%, and the TPR accomplished is 92.85%. It is noted that LGAO_BHNAHN has achieved an accuracy of 93.30%, FPR of 5.62%, TPR of 92.85%.

Table 2. Comparative Discussion of LGAO_BHNAHN

Analysis	**Metrics**	**Text CNN**	**CoDA**	**BERT**	**BHNA HN**	**SCRes NeXt**	**BERT**	**LGAO_B HNAHN**
Setup-1	Acc (%)	84.28	86.29	87.81	89.08	90.67	91.35	93.70
	FPR (%)	13.72	12.66	11.16	9.85	8.58	7.90	5.62
	TPR (%)	84.04	86.98	87.09	88.18	89.31	90.51	92.67
Setup-2	Acc (%)	85.49	87.24	88.10	89.32	90.68	91.62	93.71
	FPR (%)	14.28	13.46	11.29	10.94	9.43	8.20	5.81
	TPR (%)	85.69	87.75	89.00	89.36	90.00	90.13	92.89
Setup-3	Acc (%)	85.31	86.26	88.28	89.71	90.91	91.13	93.25
	FPR (%)	13.16	12.04	11.75	10.68	9.66	8.11	5.71
	TPR (%)	84.83	86.46	87.08	88.40	89.66	90.70	92.74
Setup-4	Acc ((%)	85.00	86.60	87.87	88.91	90.73	91.09	93.30
	FPR (%)	13.96	12.76	11.90	10.76	9.74	8.08	5.62
	TPR (%)	85.52	86.56	87.64	89.01	90.09	90.83	92.85

5. Conclusion

BHNAHN with LGAO optimization and GPT-NEOX improves classification accuracy and adaptability while offering a novel, coherent structure that can manage the evolving security needs and complexity of dark web real-time surveillance systems. It is challenging to track different illegal activities performed by anonymous individuals or institutions on the dark web. The nature of dark web illegal content is continually evolving and being revised. The complexity and length of the process make it very challenging to gather and categorize such criminal activities. LGAO_BHNAHN was created to categorize illegal dark web content to ease this.

To identify and extract relevant information, textual content extraction is performed initially. GPT-NEOX subsequently processes the information. BHNAHN, which is developed by integrating a Bayesian Neural Network (BNN), a Hierarchical Neural Attention classifier, and Forward Harmonic Analysis, performs the classification task. LGAO, a combination of Genetic Algorithm Optimization (GAO) and Lyrebird Optimization Algorithm (LOA), is utilized to train BHNAHN. The precision, False Positive Rate (FPR), and True Positive Rate (TPR) of LGAO_BHNAHN were 93.30%, 5.62%, and 92.85%, respectively. In the future, research will focus on developing anomaly detection algorithms, which are useful for identifying emerging or evolving threats in particular, to detect unusual patterns that could indicate illicit activity.

References

[1] Makridakis, Spyros. "The forthcoming Artificial Intelligence (AI) revolution: Its impact on society and firms." Futures 90 (2017): 46-60.

[2] Rodríguez, John Ibañez, Santiago Rocha Durán, Daniel Díaz-López, Javier Pastor-Galindo, and Félix Gómez Mármol. "C 3-Sex: A conversational agent to detect online sex offenders." Electronics 9, no. 11 (2020): 1779.

[3] Pastor-Galindo, Javier, Mattia Zago, Pantaleone Nespoli, Sergio López Bernal, Alberto Huertas Celdrán, Manuel Gil Pérez, José A. Ruipérez-Valiente, Gregorio Martínez Pérez, and Félix Gómez Mármol. "Spotting political social bots in Twitter: A use case of the 2019 Spanish general election." IEEE Transactions on Network and Service Management 17, no. 4 (2020): 2156-2170.

[4] Ramírez Sánchez, Julián, Alejandra Campo-Archbold, Andrés Zapata Rozo, Daniel Díaz-López, Javier Pastor-Galindo, Félix Gómez Mármol, and Julián Aponte Díaz. "Uncovering cybercrimes in social media through natural language processing." Complexity 2021, no. 1 (2021): 7955637.

[5] Cascavilla, Giuseppe, Gemma Catolino, and Mirella Sangiovanni. "Illicit darkweb classification via natural-language processing: Classifying illicit content of webpages based on textual information." arXiv preprint arXiv:2312.04944 (2023).

[6] Yegneswaran, Shalini Ghosh Phillip Porras Vinod, and Ken Nitz Ariyam Das. "ATOL: A Framework for Automated Analysis and Categorization of the Darkweb Ecosystem." (2017).

[7] Hayes, Darren R., Francesco Cappa, and James Cardon. "A framework for more effective dark web marketplace investigations." Information 9, no. 8 (2018): 186.

[8] Chertoff, Michael. "A public policy perspective of the Dark Web." Journal of Cyber Policy 2, no. 1 (2017): 26-38.

[9] Alshammery, Mohammed Khalafallah, and Abbas Fadhil Aljuboori. "Classifying illegal activities on tor network using hybrid technique." Iraqi Journal of Science (2022): 3994-4004.

[10] Alaidi, Abdul Hadi M., M. Roa'a, H. T. H. S. ALRikabi, Ibtisam A. Aljazaery, and Saif Hameed Abbood. "Dark web illegal activities crawling and classifying using data mining techniques." iJIM 16, no. 10 (2022): 123.

[11] Wang, Gang, Hsinchun Chen, and Homa Atabakhsh. "Automatically detecting deceptive criminal identities." Communications of the ACM 47, no. 3 (2004): 70-76.

[12] Jin, Youngjin, Eugene Jang, Yongjae Lee, Seungwon Shin, and Jin-Woo Chung. "Shedding new light on the language of the dark web." arXiv preprint arXiv:2204.06885 (2022).

[13] Zhang, Ning, Mohammadreza Ebrahimi, Weifeng Li, and Hsinchun Chen. "Counteracting dark Web text-based CAPTCHA with generative adversarial learning for proactive cyber threat intelligence." ACM Transactions on Management Information Systems (TMIS) 13, no. 2 (2022): 1-21.

[14] Iqbal, Farkhund, Benjamin CM Fung, Mourad Debbabi, Rabia Batool, and Andrew Marrington. "Wordnet-based criminal networks mining for cybercrime investigation." IEEE access 7 (2019): 22740-22755.

[15] Shin, Gun-Yoon, Younghoan Jang, Dong-Wook Kim, Sungjin Park, A-Ran Park, Younghwan Kim, and Myung-Mook Han. "Dark side of the web: Dark web classification based on TextCNN and topic modeling weight." IEEE Access 12 (2023): 36361-36371.

[16] Sangher, Kanti Singh, Archana Singh, Hari Mohan Pandey, and Vivek Kumar. "Towards safe cyber practices: Developing a proactive cyber-threat intelligence system for dark web forum content by identifying cybercrimes." Information 14, no. 6 (2023): 349.

[17] Dalvi, Ashwini, Soham Bhoir, Nishavak Naik, Atharva Kitkaru, Irfan Siddavatam, and Sunil Bhirud. "A hybrid TF-IDF and RNN model for multi-label classification of the deep and dark web." International Journal of Advanced Computer Science and Applications 14, no. 7 (2023).

[18] Murty, C. A., and Parag H. Rughani. "Dark web text classification by learning through SVM optimization." J Adv Inf Technol 13, no. 6 (2022).

[19] Black, Sid, Stella Biderman, Eric Hallahan, Quentin Anthony, Leo Gao, Laurence Golding, Horace He et al. "Gpt-neox-20b: An open-source autoregressive language model." arXiv preprint arXiv:2204.06745 (2022).

[20] Xie, Xurong, Xunying Liu, Tan Lee, and Lan Wang. "Bayesian learning for deep neural network adaptation." IEEE/ACM Transactions on Audio, Speech, and Language Processing 29 (2021): 2096-2110.

[21] Mullachery, Vikram, Aniruddh Khera, and Amir Husain. "Bayesian neural networks." arXiv preprint arXiv:1801.07710 (2018).

[22] Kowsari, Kamran, Donald E. Brown, Mojtaba Heidarysafa, Kiana Jafari Meimandi, Matthew S. Gerber, and Laura E. Barnes. "Hdltex: Hierarchical deep learning for text classification." In 2017 16th IEEE international conference on machine learning and applications (ICMLA), IEEE, 2017, 364-371.