Jana Publication & Research

HOMOMORPHIC ENCRYPTION- ENABLED DEEP LEARNING MODEL FOR INTELLIGENT CYBER THREAT DETECTION

📋 VRC38

Document Details

Submission ID

trn:oid:::2945:312979460

Submission Date

Sep 18, 2025, 10:59 AM GMT+5:30

Download Date

Sep 18, 2025, 11:00 AM GMT+5:30

File Name

IJAR-53884.pdf

File Size

878.5 KB

9 Pages

3,256 Words

18,734 Characters

18% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

Filtered from the Report

▸ Bibliography

▸ Quoted Text

Match Groups

37 Not Cited or Quoted   15%
Matches with neither in-text citation nor quotation marks

8   Missing Quotations   3%
Matches that are still very similar to source material

0   Missing Citation   0%
Matches that have quotation marks, but no in-text citation

0   Cited and Quoted   0%
Matches with in-text citation present, but no quotation marks

Top Sources

10%   🌐 Internet sources

11%   📖 Publications

13%   👤 Submitted works (Student Papers)

Match Groups

37 Not Cited or Quoted 15%
Matches with neither in-text citation nor quotation marks

8 Missing Quotations 3%
Matches that are still very similar to source material

0 Missing Citation 0%
Matches that have quotation marks, but no in-text citation

0 Cited and Quoted 0%
Matches with in-text citation present, but no quotation marks

Top Sources

10%	🌐	Internet sources
11%	📖	Publications
13%	👤	Submitted works (Student Papers)

Top Sources

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1	Student papers	
Anna University on 2025-06-15		**1%**

2	Student papers	
Heriot-Watt University on 2023-10-31		**1%**

3	Publication	
Thangaprakash Sengodan, Sanjay Misra, M Murugappan. "Advances in Electrical ...		**<1%**

4	Internet	
www.coursehero.com		**<1%**

5	Internet	
pdfs.semanticscholar.org		**<1%**

6	Student papers	
Imperial College of Science, Technology and Medicine on 2021-01-21		**<1%**

7	Publication	
Pushpa Choudhary, Sambit Satpathy, Arvind Dagur, Dhirendra Kumar Shukla. "Re...		**<1%**

8	Publication	
Vanlalruata Hnamte, Jamal Hussain. "Dependable Intrusion Detection System usi...		**<1%**

9	Student papers	
University Politehnica of Bucharest on 2023-06-28		**<1%**

10	Publication	
Preetesh Purohit, Anuradha Purohit, Vrinda Tokekar. "Localization of try block an...		**<1%**

11	Student papers	
City of Glasgow College on 2023-02-28		<1%

12	Publication	
Dhanalekshmi Prasad Yedurkar, Ganesh R. Pathak, Manisha Galphade, Thompso...		<1%

13	Student papers	
Georgia Institute of Technology Main Campus on 2025-04-06		<1%

14	Student papers	
Johns Hopkins Unversity on 2025-02-24		<1%

15	Publication	
Mohamed H. Behiry, Mohammed Aly. "Cyberattack detection in wireless sensor n...		<1%

16	Student papers	
Queen Mary and Westfield College on 2010-07-26		<1%

17	Student papers	
Universiti Teknologi Malaysia on 2024-09-27		<1%

18	Internet	
arxiv.org		<1%

19	Internet	
faculty.ecnu.edu.cn		<1%

20	Internet	
vsip.info		<1%

21	Student papers	
Higher Education Commission Pakistan on 2025-02-07		<1%

22	Student papers	
National College of Ireland on 2023-07-31		<1%

23	Student papers	
Kaplan College on 2023-09-14		<1%

24	Student papers	
University of Hertfordshire on 2022-12-02		<1%

25	Student papers	
University of Hertfordshire on 2023-08-28		<1%

26	Internet	
dokumen.pub		<1%

27	Internet	
ir.kabarak.ac.ke		<1%

28	Internet	
link.springer.com		<1%

29	Student papers	
Bowie State University on 2022-02-23		<1%

30	Student papers	
Universiti Malaysia Pahang on 2025-03-04		<1%

31	Internet	
dr.lib.iastate.edu		<1%

32	Internet	
journalofbigdata.springeropen.com		<1%

33	Internet	
www.igi-global.com		<1%

34	Internet	
www.researchsquare.com		<1%

35	Internet	
www.scpe.org		<1%

36	Publication	
"Proceedings of 17th International Conference on Machine Learning and Comput...		<1%

37	Student papers	
University of Huddersfield on 2024-09-06		<1%

38	Publication	
Al-Nsour, Lubna H.. "Optimizing Hospital Pharmacy Inventory: A Robust Model Pr...		<1%

HOMOMORPHIC ENCRYPTION-ENABLED DEEP LEARNING MODEL FOR INTELLIGENT CYBER THREAT DETECTION

Abstract:

Cybersecurity is one of the emerging fields which address information security, protecting computer systems and networks from intrusions.Cyber-attack detection is crucial for safeguarding computer networks against unauthorized access and malicious activities. Traditional machine learning approaches often suffer from overfitting, leading to reduced accuracy. To overcome this, an optimized deep learning approach with a homomorphic encryption-based authentication protocol is proposed. Cyber threat data is collected from open sources and pre-processed using data cleaning and binning method. Feature extraction is carried out using TF-IDF, followed by dimensionality reduction through Principal Component Analysis. The processed data is then classified using an Extended Physics-Informed Neural Network (EPINN) for cyber-attack detection. A Signature-based Authentication Protocol ensures secure user authentication, while BFV encryption secures data storage in the cloud. Experimental results show high efficiency with 97.36% accuracy, 89.78% precision, and a low false positive rate of 2.17%. This approach enables automatic and robust cyber threat detection, improving proactive defense mechanisms for organizations.

*Keywords:*Cyber security; Feature extraction; Principal component analysis; Extended Physics-Informed Neural Network; Classification.

1.   Introduction

The term "cyber security" describes the procedures and policies intended to shield computers, networks, software, and data from harm, illegal access, and cyberattacks. In order to protect the confidentiality, integrity, and availability of information, it entails utilizing technologies, procedures, and policies to stop, identify, and address threats [1].Cybersecurity is "the art of protecting networks, devices, and data from unauthorized access or criminal use and the practice of ensuring confidentiality, integrity, and availability of information [2]. The proliferation of online and mobile banking has led to a steady increase in cybercrime, which includes a variety of frauds like identity theft, credit card scams, spam, and ATM theft. The financial industry's data is particularly sensitive, despite its substantial monetary value [3]. Hackers can use stolen financial information and banking credentials to generate income in a variety of ways.Due to the significant value of stored data, which gives hackers several opportunities to make money, the banking industry is especially susceptible to these kinds of attacks [4].

Cyber-attack detection in banking platforms can be expensive and hard to operate, particularly for smaller businesses with limited funds and infrastructure, due to the possibility of false positives, the inability to respond to quickly emerging threats, and the need for substantial computational resources [5]. To lessen these issues, automated cyberattack detection based on

38 artificial intelligence (AI) was developed. Large datasets are analyzed by AI in automated
39 cyberattack detection to find suspicious trends, allowing for quicker and more precise detection.
40 In a subfield of artificial intelligence, machine learning algorithms are employed to help robots
41 learn from experiments more effectively [6].

2.   Related works

43 A Deep Belief Network (DBN) was developed for cyber-attack detection in industrial
44 automation using temporal patterns in network data to identify anomalies over time [7]. The
45 method achieved 93.2% accuracy with reduced false positives and an error rate of 6.8%.
46 However, its reliance on labeled data limits effectiveness against zero-day and unseen
47 attacks.Deep Neural Network (DNN) was proposed for real-time cyber-attack detection by
48 analyzing large-scale network traffic patterns [8]. It achieved 94.5% accuracy, 92.7% precision,
49 and 91.2% specificity, supported by image enhancement techniques. Despite strong performance,
50 high computational complexity makes deployment difficult on resource-constrained devices.

51 A Long Short-Term Memory (LSTM) framework was introduced for cyber-attack
52 detection using feature selection to reduce data dimensionality and improve accuracy [9]. The
53 model achieved 94.9% detection accuracy and 93.4% precision, effectively identifying
54 sophisticated attack patterns. However, the feature selection process may exclude critical
55 features, causing occasional misclassification.Recurrent Neural Network (RNN) was developed
56 for cyber-attack detection using temporal data [10]. It monitors variations in network behavior
57 over time, effectively capturing subtle and continuous attack patterns. By employing a patch-
58 based approach, the system enhanced detection performance, achieving an accuracy of 94.8%
59 and a specificity of 91.7%. However, it relies heavily on large volumes of sequential data,
60 making it less effective in cases where data is sparse or lacks continuity.

61 A Convolutional Neural Network (CNN) was designed to detect cyber-attacks in IoT
62 networks using a graph-based method, where devices and their communications are represented
63 as nodes and edges [11]. The model achieved an accuracy of 92.4% and a precision of 90.7%,
64 with computational efficiency that supports real-time applications. Nonetheless, as IoT networks
65 scale, the increased data complexity can reduce performance, making it more challenging to
66 maintain detection accuracy in large environments.Graph Neural Networks (GNN) was proposed
67 to identify cyberattacks on dispersed networks [12]. The distributed model protects data privacy
68 while assisting in the detection of cyberattacks in a variety of contexts. The complexity of the
69 approach is increased by communication overhead between federated systems, which make
70 scaling for real-time detection across several nodes challenging.

71 A verifiable dynamic access control mechanism with randomized ensemble SVM was
72 developed to address user revocation in IoT architectures [13]. The approach integrates
73 VOMAACS (Verifiable Outsourced Multi-Authority Access Control Scheme) with CPASBE
74 (Ciphertext Attribute Set-Based Encryption) to enhance data security. In CPASBE, sensitive

75 information remains with the data owner, ensuring confidentiality, preventing duplication, and
76 protecting against agreement breaches regardless of provider reliability. When the system
77 receives new or unknown attacks that match pre-existing signatures, its performance suffers due
78 to its dependence on signature-based detection [14]. When presented with noisy or insufficient
79 data, CNN's performance can decrease, decreasing the accuracy of detection in practical settings.
80 Occasionally, sophisticated attacks may be misclassified as a result of the process's unintentional
81 exclusion of crucial features. Scaling for real-time detection over numerous nodes is challenging
82 due to the approach's complexity [15]. To improve the security of Wireless Sensor Networks
83 (WSNs) by detecting and averting cyberattacks, an intelligent hybrid model that combines
84 artificial intelligence and machine learning was proposed. Singular Value Decomposition (SVD)
85 and Principal Component Analysis (PCA) are two feature reduction approaches used in the
86 study. The K-means clustering model enhanced information gain (KMC-IG) is also used for
87 feature extraction. Intrusion detection systems and network traffic classification are presented
88 with the introduction of the Synthetic Minority Excessively Technique for data balancing [16].

3.   Proposed Methodology

90 Cyber security is the practice of defending computers, servers, mobile devices, electronic
91 systems, networks, and data from malicious attacks. Because it guards against theft and
92 destruction to all types of data, cyber security is crucial. False positives from cyberattack
93 detection systems might overload security professionals with warnings that aren't harmful. They
94 might also find it difficult to identify complex, dynamic dangers in real time. In this case, a deep
95 learning-based solution that can identify and safeguard the system in an aberrant state is
96 suggested. Figure 1 depicts the suggested architecture for deep learning-based cyber security
97 threat identification.

98 To increase the quality of the data, the obtained data are first pre-processed. Feature
99 extraction is carried out using TF-IDF, followed by dimensionality reduction through Multi-
100 Kernel Principal Component Analysis. The processed data is then classified using an Extended
101 Physics-Informed Neural Network (EPINN) for cyber-attack detection. A Verifiable Ring
102 Signature-based Authentication Protocol ensures secure user authentication, while BFV
103 encryption secures data storage in the cloud.

**(i) Data Pre-Processing:**

105 In cybersecurity-oriented IoT environments, preprocessing plays a crucial role in
106 preparing raw network traffic, sensor logs, and user access records for reliable intrusion
107 detection and anomaly analysis. Since IoT data is often heterogeneous, large-scale, and prone to
108 noise or missing values, robust preprocessing ensures the integrity and accuracy of subsequent
109 detection models.

110 Data Cleaning is an essential step which focuses on addressing invalid, inconsistent, or
111 incomplete information. IoT traffic logs may contain missing values due to packet loss, device

112 malfunctions, or communication delays. These can be handled in two ways: (1) ignoring the
113 tuples with missing data, a viable option only in very large datasets, or (2) filling the missing
114 values by imputation techniques such as mean replacement, predictive estimation, or using
115 default values. This preserves data continuity, which is crucial for real-time threat analysis.

116   Another challenge is noisy data, which may arise from corrupted sensor readings,
117 unauthorized access attempts, or data entry errors in IoT networks. To address this, methods such
118 as the Binning Method are applied, where continuous data (e.g., traffic rates, packet counts) is
119 sorted and divided into equal-sized intervals to smooth out inconsistencies. This reduces false
120 alarms in anomaly detection.

121
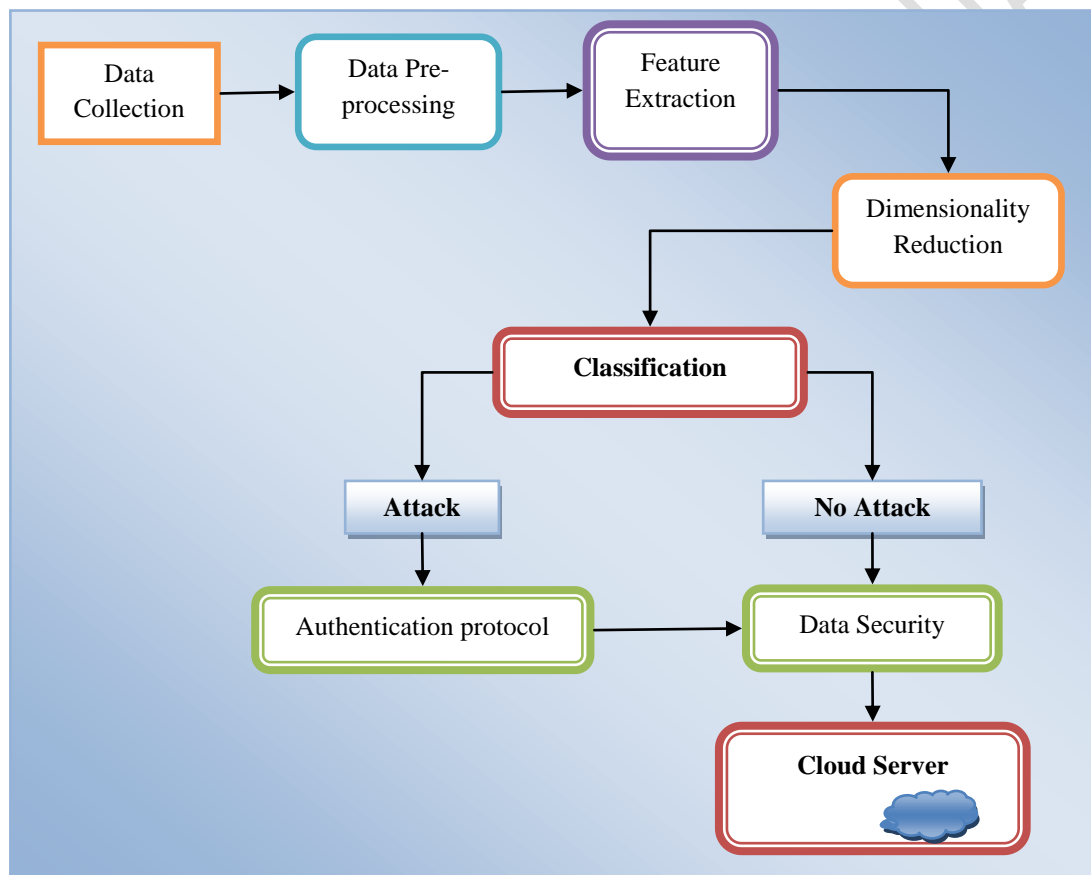
122   **Figure 1: Block diagram for Proposed Model**

123 **(ii) Feature Extraction:**

124   The process of turning unstructured IoT or cybersecurity data into valuable features that
125 serve as crucial information for analysis or modeling is known as feature extraction. By
126 choosing, transforming or producing variables that emphasize the most pertinent patterns, this
127 stage is essential to improving the performance of machine learning algorithms. Effective feature
128 extraction aids in the detection of harmful traffic patterns, illegal access attempts, and unusual
129 device behavior when communicating in cybersecurity-based IoT systems.

130   Term Frequency–Inverse Document Frequency (TF-IDF) is a popular method for feature
131 extraction in cybersecurity data that is based on text and logs.TF-IDF is a statistical measure that
132 evaluates the importance of a word (or token) in a dataset relative to its frequency across
133 multiple documents. TF-IDF can be used to extract discriminative features for anomaly or attack
134 detection in network traffic logs, system call traces, or access records in the context of IoT
135 cybersecurity. The term frequency measures how frequently a term occurs in a document, which
136 is given in equation 1,

$$TF(t,d) = \frac{f_{t,d}}{\sum f_{t',d}} \qquad (1)$$

138   Here, $f_{t,d}$ is the number of times term 't' appears in document 'd'. Inverse Document
139 Frequency (IDF) reduces the weight of common terms and increases the weight of rare terms
140 across the dataset and it is given in equation 2,

$$IDF(t) = \log \frac{N}{1+n_t} \qquad (2)$$

142 where 'N' is the total number of documents, and n_t is the number of documents containing term
143 t.

144   The terms that appear frequently in a particular document (or log entry) but infrequently
145 in other documents are thus emphasized by TF-IDF, which makes them effective markers of
146 anomalous activity or possible cyber threats.

147   Features like abnormal command sequences, infrequent access attempts, or abnormal
148 protocol usages that can indicate cyberattacks can be extracted using TF-IDF from cybersecurity-
149 based IoT. In order to increase detection accuracy and reduce false positives, these collected
150 features are subsequently input into classifiers (such as deep learning or ensemble models).

151   High-dimensional data is common in cybersecurity-based IoT systems due to the massive
152 volume of logs, network packets, and sensor readings. Therefore Principal Component Analysis
153 (PCA) is widely used as a dimensionality reduction technique.PCA transforms the original
154 correlated features into a smaller set of uncorrelated variables known as principal components
155 (PCs), while retaining most of the variance (information) in the data. This helps in reducing
156 noise, improving model efficiency, and focusing on the most discriminative patterns in IoT
157 cybersecurity data.

158 **(iii) Classification:**

159   Classification is the process of assigning data points to predefined categories based on
160 patterns learned from labelled data. Extended Physics-Informed Neural Networks (EPINN)
161 integrates physical laws into neural networks for solving complex scientific problems, improving
162 accuracy in simulations.

163   Extended Physics-Informed Neural Networks (EPINN) incorporates physical laws and
164   domain knowledge into neural network architectures, improving accuracy and generalization in
165   solving complex, physics-based problems. The Structure Diagram of the Physics Informed
166   Neural Network is shown in Figure 2 is shown below.

167
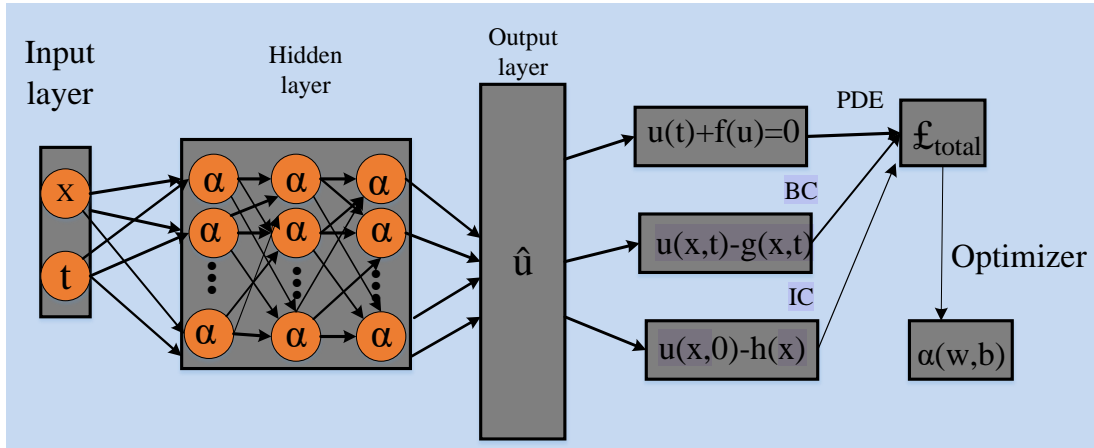
168   **Figure 2: Structure Diagram of the EPINN**

169   A Structure Diagram of a EPINN visually represents how neural networks integrate
170   physical laws into their architecture. It typically includes components such as input layers for
171   initial conditions, hidden layers where the physics-based constraints are enforced, and output
172   layers that produce predictions while satisfying governing equations. This framework enables the
173   model to learn both data-driven patterns and the underlying physics, enhancing accuracy and
174   generalization in scientific applications.

175   To approximate the solution of the partial differential equation using spatial-temporal
176   coordinates $u(x,t)$ as the input, Fully linked feed forward neural networks with multiple hidden
177   layers are used by EPINN are represented as shown below in equation 3,

178            $$(na_d^i(L_d(X^{d-1}))_i), d = 1,2\,\ldots.,D-1, i = 1,2\,\ldots\ldots N_d \tag{3}$$

179   Once classification process is done then the attacked packets are re-registered with the
180   cloud server using Signature based Authentication Protocol (SAP) and the normal packets are
181   directly send for the encryption processfor secure storage. VRSAP is usedto verify a user's
182   identity before allowing them access to systems or networks and it is important for protecting
183   systems, data, and applications from attacks.

184   Following user verification, the data is safely stored on a cloud server using
185   Brakerski/Fan-Vercauteren (BFV) cryptography. The BFV scheme, a lattice-based holomorphic
186   encryption technique, is perfect for data security and privacy-preserving computations since it
187   allows safe addition and multiplication operations on encrypted data.

188   **4.   Results and Discussion**

189   To improve network security, the proposed homomorphic encryption based
190 authentication protocol based on BFV and EPINN have been created for the detection of
191 cybersecurity attacks.With an Intel Core i5 CPU, an Nvidia GeForce GTX 1650 GPU, and 16GB
192 of RAM, it was designed using Python 3.8.

193   The Kaggle website is the original source of the data. This dataset has been used
194 extensively to assess anomaly detection and is based on data collected during the DARPA'98
195 IDS evaluation program.

196
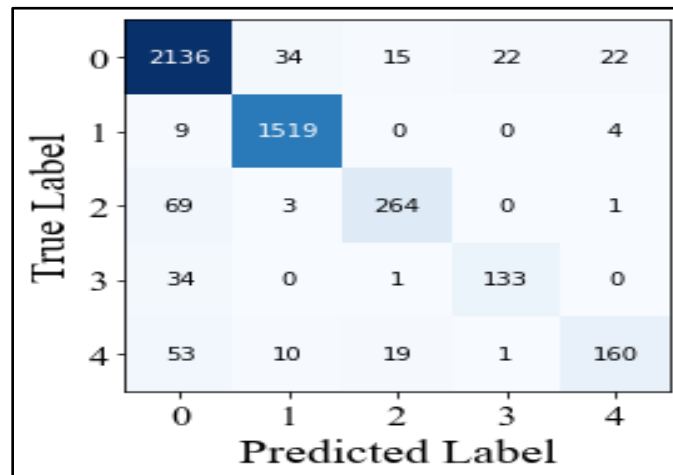
Figure 3: Confusion Matrix for Proposed Model

198 The confusion matrix in Figure 6 provides a comprehensive evaluation of the proposed
199 classification method for cyber-attack detection. It is used to validate the accuracy of the model
200 by comparing predicted outputs with the actual class labels. These results demonstrate the
201 effectiveness of the proposed method in distinguishing between normal and malicious traffic. By
202 accurately classifying different attack categories, the model proves its robustness and
203 applicability in real-world cybersecurity-based IoT environments where timely and precise
204 detection of various attack types is critical for system security.

205
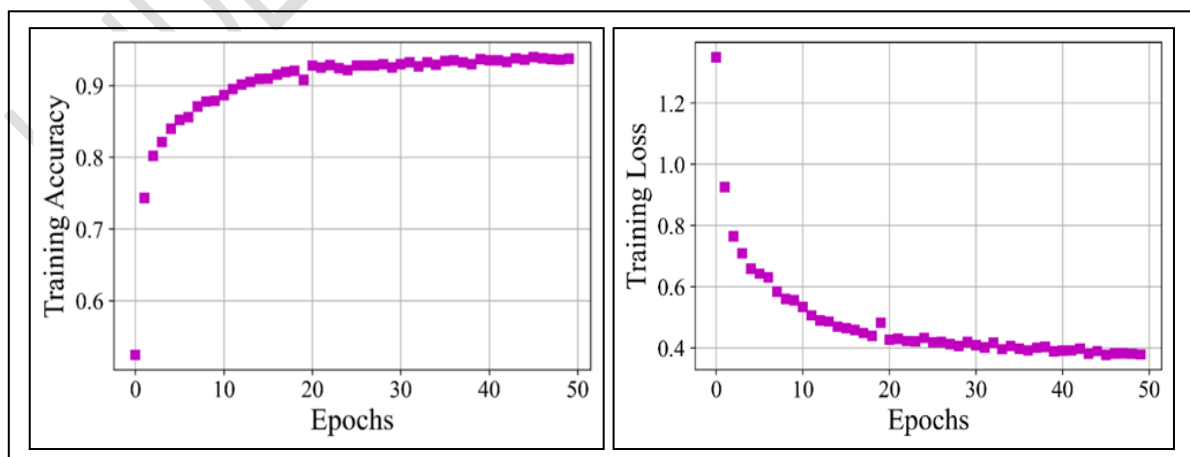
206 **Figure 4: Evaluation of Training Accuracy and Loss with varying Epochs**

207 The figure 4 shows the evaluation of training accuracy and training loss by different
208 numbers of epochs. The model gains more knowledge from the training data as the number of
209 epochs rises and both the training and validation datasets show an improvement in prediction
210 accuracy. This is because the model has more chances to minimize the loss function by
211 modifying its weights and biases.

212 **5.   Conclusion**

213 The proposed optimized deep learning system with authentication based on homomorphic
214 encryption demonstrates significant potential for improving cyberattack detection in networks
215 enabled by the Internet of Things. The model achieves excellent accuracy with low false positive
216 rates by combining efficient preprocessing, TF-IDF-based feature extraction, PCA-driven
217 dimensionality reduction, and EPINN classification. Strong data security and user privacy are
218 also guaranteed in cloud environments via BFV encryption and a signature-based authentication
219 scheme. Therefore, this strategy provides a proactive and dependable protection mechanism,
220 enhancing cybersecurity resistance to changing threats.

221 **References:**

222 1. Elsisi, M., & Tran, M. Q. (2021). Development of an IoT architecture based on a deep
223   neural network against cyber attacks for automated guided vehicles. *Sensors*, *21*(24),
224   8467.
225 2. Chang, K., & Huang, H. (2023). Exploring the management of multi-sectoral
226   cybersecurity information-sharing networks. *Government Information Quarterly*, *40*(4),
227   101870.
228 3. Kumar, V., & Sinha, D. (2021). A robust intelligent zero-day cyber-attack detection
229   technique, Complex Intell. In *Syst* (Vol. 7, No. 5, pp. 2211-2234).
230 4. Sridhar, S., &Govindarasu, M. (2017). Model-based attack detection and mitigation for
231   automatic generation control. IEEE Transactions on Smart Grid, 5(2), 580-591.
232 5. Balta, E. C., Pease, M., Moyne, J., Barton, K., & Tilbury, D. M. (2023). Digital twin-
233   based cyber-attack detection framework for cyber-physical manufacturing systems. IEEE
234   Transactions on Automation Science and Engineering, 21(2), 1695-1712.
235 6. Zhang, F., Kodituwakku, H. A. D. E., Hines, J. W., & Coble, J. (2019). Multilayer data-
236   driven cyber-attack detection system for industrial control systems based on network,
237   system, and process data. *IEEE Transactions on Industrial Informatics*, *15*(7), 4362-
238   4369.
239 7. Jatothu, M. S., & Devi, D. S. L. (2024). Evolving Deep Belief Network for Cyber-Attack
240   Detection in Industrial Automation and Control Systems. Journal of Cybersecurity and
241   Privacy, 5(1), 1-15.

8. Vijayakumar, K. P., Pradeep, K., Balasundaram, A., &Prusty, M. R. (2023). Enhanced cyber attack detection process for internet of health things (IoHT) devices using deep neural network. *Processes*, *11*(4), 1072.

9. Zarzycki, K., Chaber, P., Cabaj, K., Ławryńczuk, M., Marusak, P., Nebeluk, R., ... & Wojtulewicz, A. (2023). Forgery cyber-attack supported by LSTM neural network: an experimental case study. Sensors, 23(15), 6778.

10. Udas, P. B., Roy, K. S., Karim, M. E., & Ullah, S. M. A. (2023, February). Attention-based RNN architecture for detecting multi-step cyber-attack using PSO metaheuristic. In 2023 International Conference on Electrical, Computer and Communication Engineering (ECCE) (pp. 1-6). IEEE.

11. Nedeljkovic, D., & Jakovljevic, Z. (2022). CNN based method for the development of cyber-attacks detection algorithms in industrial control systems. Computers & Security, 114, 102585.

12. Friji, H., Olivereau, A., &Sarkiss, M. (2023, May). Efficient network representation for GNN-based intrusion detection. In International conference on applied cryptography and network security (pp. 532-554). Cham: Springer Nature Switzerland.

13. Kumar, R. A., & Vinuthna, K. (2021). Randomized Ensemble SVM based Deep learning with Verifiable dynamic access control using user revocation in IoT architecture. Sādhanā, 46(4), 229.

14. Stojanović, B., Hofer-Schmitz, K., & Kleb, U. (2020). APT datasets and attack modeling for automated detection methods: A review. Computers & Security, 92, 101734.

15. Burnap, P., Anthi, E., Reineckea, P., Williams, L., Cao, F., Aldmoura, R., & Jones, K. (2024). Mapping automated cyber attack intelligence to context-based impact on system-level goals. Journal of Cybersecurity and Privacy, 4(2), 340-356.

16. Behiry, M. H., & Aly, M. (2024). Cyberattack detection in wireless sensor networks using a hybrid feature reduction technique with AI and machine learning methods. Journal of Big Data, 11(1), 16.