ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26mar858

Generative AI: For Cyber Based Fake Attacks Detection and Classification Using Deep Learing Techniques

Katikam Mahesh[1]

[1]Assistant Professor Department of CSE (JNTUH): CSE. B.V. Raju Institute of Technology Narsapur, Medak District, Near Hyderabad, Telangana

Publication Date: 2026/03/19

Abstract: With projections to rise from USD 2 billion in 2024 to USD 14.79 billion by 2034, the worldwide market for generative AI in cybersecurity is expanding quickly. Council for the Global Development of Skills (GSDC) .to save from various cyber-attacks need an effect technology to tackle it. In order to counter increasingly complex, automated threats that elude conventional detection techniques, generative artificial intelligence (GenAI) is crucial to current cybersecurity. By providing real-time anomaly detection, quick threat analysis, automatic vulnerability patching, and AI-driven phishing defines, it functions as a force multiplier and significantly improves security posture. Existing deep learning models such as ANN (Artificial Neural Network) and DNN (Deep neural networks).so to enhance performance of these with the help of Generative ai technique is GAN (Generative Adversarial Network) easy to detect fake and real images automatically

Keywords: *ANN (Artificial Neural Network), Generative ai Technique is GAN (Generative Adversarial Network), Generative Artificial Intelligence (GenAI), Deep Neural Networks(DNN).*

How to Cite: Katikam Mahesh (2026) Generative AI: For Cyber Based Fake Attacks Detection and Classification Using Deep Learing Techniques. *International Journal of Innovative Science and Research Technology,* 11(3), 1328-1333. https://doi.org/10.38124/ijisrt/26mar858

I. INTRODUCTION

Artificial intelligence that learns patterns from enormous existing datasets to produce new, creative content, such as literature, photos, music, or code, is known as generative AI. Generative AI uses deep learning models, such as Transformers, to create completely new, human-like outputs based on user cues, as contrast to traditional AI, which analyses or classifies data.

➢ *Generative AI Features:*

- *What it Produces:*

Text (ChatGPT), photos (DALL-E, Midjourney), video (Sora), audio, and computer code can all be produced by it.

- *How It Operates:*

Instead of just reproducing training data, it can forecast and produce new content by using foundation models trained on large amounts of data to discover structures and relationships.

- *Core Technologies:*

To simulate human creativity, it uses transformers and Generative Adversarial Networks (GANs).

- *Applications:*

AI chatbots are frequently used for customer support, automating creative jobs like graphic design or copywriting, and speeding up research in industries like software development and medical. By identifying intricate, non-linear patterns in data, Artificial Neural Networks (ANN) and Deep Neural Networks (DNN) are very successful in cybersecurity for identifying, categorizing, and stopping network threats. While ANNs have basic learning capabilities, DNNs employ several hidden layers to identify intrusions, malware. But detection and classification different cyber-attacks is very poor so enhance its performance generative ai is only key to improve its performance.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26mar858

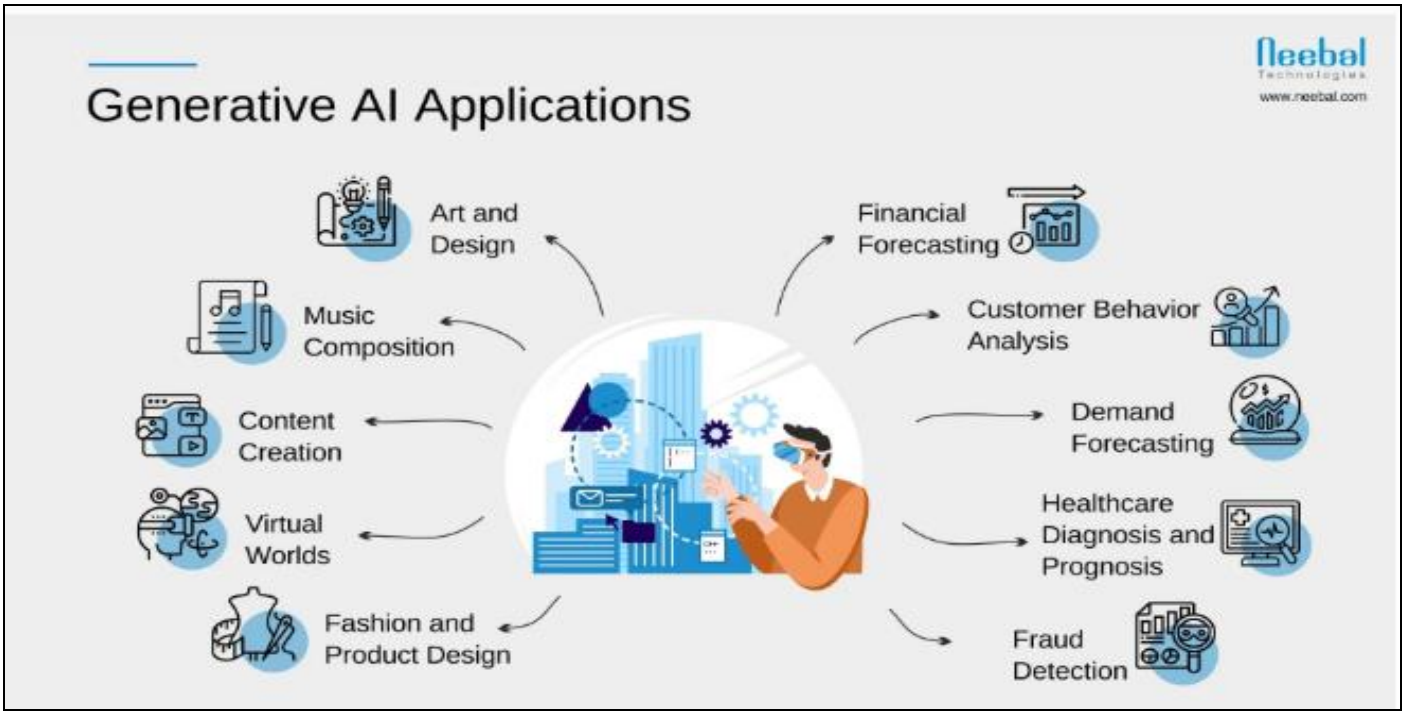

Fig 1 Generative AI Applications

In cybersecurity, artificial neural networks (ANNs) are employed for intrusion detection systems (IDS), real-time threat detection, and behavioural analysis to spot malicious trends.

In cybersecurity, Deep Neural Networks (DNNs) have become a potent, high-accuracy technique that frequently outperforms conventional machine learning techniques in tasks like threat identification and malware categorization. They can spot minute patterns in network traffic because they are particularly good at processing complicated, high-dimensional, and unstructured data.

According to specialists at Palo Alto Networks, generative AI is changing cybersecurity from a reactive to a proactive paradigm by spotting intricate patterns, automating threat detection, and producing realistic, synthetic data for testing. It helps with incident analysis and vulnerability management by enabling quicker, more intelligent responses to cyberthreats using natural language. Phishing detection, malware analysis, and AI-powered security co-pilots are important use cases, according to NTT Data.

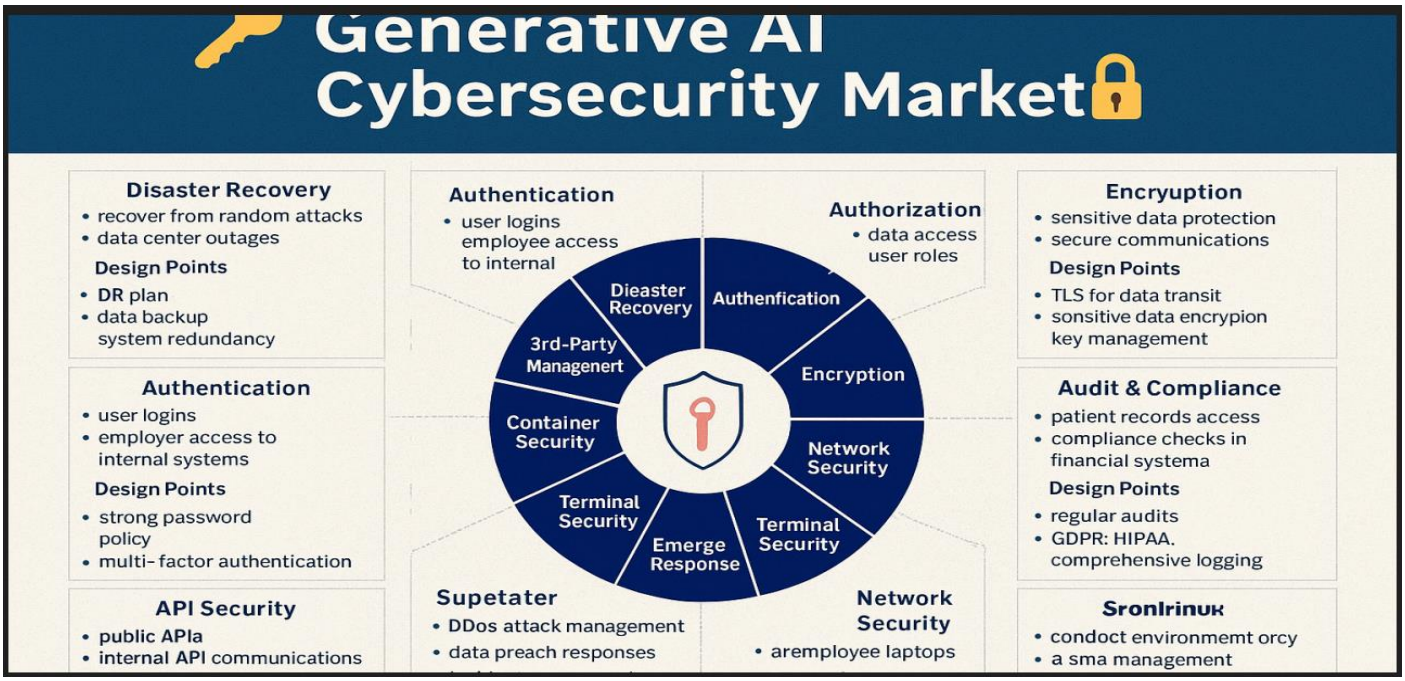

Fig 2 Generative for Cyber Security

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26mar858

By acting as a dual-purpose tool that simulates sophisticated attacks to train more resilient systems and strengthens defences through realistic synthetic data generation, Generative Adversarial Networks (GANs) are quickly revolutionizing cybersecurity. GANs, a type of deep learning, are very useful for handling imbalanced data in anomaly detection and malware classification because they employ a generator-discriminator architecture to produce synthetic data that is almost identical to real data.

II. LITERATURE STUDY

According to recent academic research, Generative AI (GenAI) in cybersecurity is a "double-edged sword" that greatly improves both offensive sophistication and defensive resilience [1]. Threat Detection & Anomaly Spotting: In order to identify patterns suggestive of malware, phishing, or zero-day attacks, large language models (LLMs) and generative adversarial networks (GANs) examine massive amounts of network data in real-time [2]. Incident Response & Automation: By automating repetitive processes like patch management, vulnerability assessments, and compliance checks, GenAI frees up human teams to concentrate on strategic choices [3]. Synthetic Data for Training: Without utilizing sensitive real-world data, GANs generate synthetic datasets that mimic uncommon attack scenarios, assisting in the training of more resilient intrusion detection systems (IDS) [4]. Vulnerability Remediation: During the Software Development Life Cycle (SDLC), tools like Sentinel One Purple AI and Microsoft Security Copilot help with secure code development and automated patching. Automated and Polymorphic Malware [5]: To avoid detection by conventional antivirus software, attackers employ GenAI to generate self-mutating malware that dynamically modifies its signature. Advanced Phishing: Because LLMs create highly customized, context-aware emails that imitate particular business tones, it is far more difficult to identify them using conventional keyword filters. Deepfakes & Misinformation: Social engineering techniques including "virtual kidnapping" schemes and fraudulent executive impersonation employ GAN-driven audio and video manipulation. Adversarial AI Attacks: By introducing malicious noise into inputs or contaminating training data, adversaries utilize GenAI to identify "blind spots" in defensive machine learning models. Frameworks like ISO 27001 and NIST CSF 2.0 frequently lack precise rules for the particular risks associated with GenAI, like algorithmic bias and rapid insertion.

III. RESERCH METHODOLOGY

Generative Adversarial Networks (GANs) have emerged as a result of the evolution of cyber threats and their increasing randomness[6]. Threat simulation is a compelling use for GANs, which were initially intended for image production. By enabling enterprises to proactively anticipate and counter future cyber threats, this integration opens up a new frontier in cybersecurity. Gans are deep neural network frameworks that can learn from training data and produce new data with the same properties as the training data, according to Deep AI. For instance, a generative adversarial network trained on images of human faces can produce completely fake faces that appear genuine.

The inventive generator and the critical discriminator are the two neural networks that make up this duelling artist. The generator functions as the creative brain, transforming random noise into completely original works of art, such as creating realistic faces or surreal landscapes from scratch[7]. The discriminator, on the other hand, assumes the role of a perceptive critic and learns to distinguish between actual examples and those the generator imagines. The discriminator becomes more adept at distinguishing between the real and fake as they train.

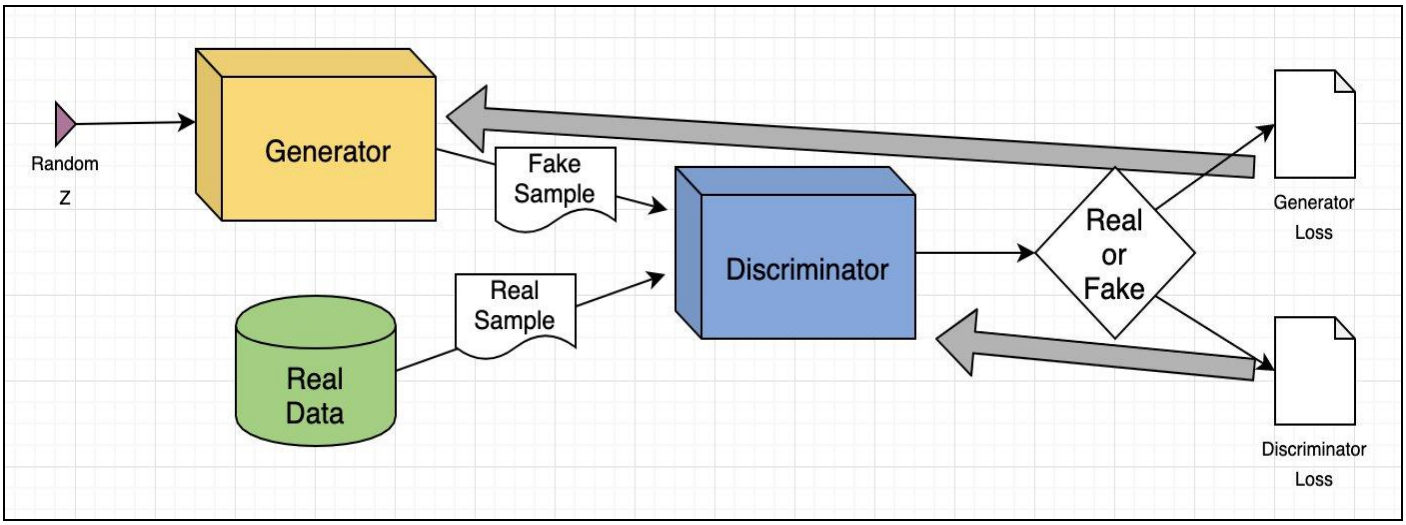

Fig 3 GAN(Generative Adversial Network) for Classification of Real or Fake Attacks

Adversarial training is a fascinating mechanism that makes all of this GAN magic possible. It's a creative dance in which the discriminator hones its abilities to distinguish between the real and the fake while the generator tries to create objects that appear exactly like reality. Because of this ongoing artistic struggle, the generator produces outputs that are so convincing that even a discriminator finds it difficult to distinguish between what is and is not real.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26mar858

➢ *Algorithm*

- GAN Algorithm:
- Start:
- Input: The CIFAR-10 dataset
- Outcome: Classification of fake or Real Images

- *GAN can be Implemented with the Following Steps:*

✓ Importing required libraries
✓ Building a simple generator network
✓ Building a simple discriminator
✓ Building a GAN by stacking the Generator and Discriminator
✓ Plotting the Generated images
✓ Training method for GAN
✓ Loading and processing MNIST data
✓ Training the GAN

- Stop:

➢ *Dataset*

The 60000 32x32 colour images in the CIFAR-10 dataset are divided into 10 classes, each with 6000 images. Ten thousand test photos and fifty thousand training images are available. Five training batches and one test batch, each containing 10,000 photos, make up the dataset. There are precisely 1000 randomly chosen photos from each class in the test batch. The remaining photographs are arranged randomly in the training batches, albeit certain training batches could include more images from one class than another. There are precisely 5000 photos from each class in the training batches[8].

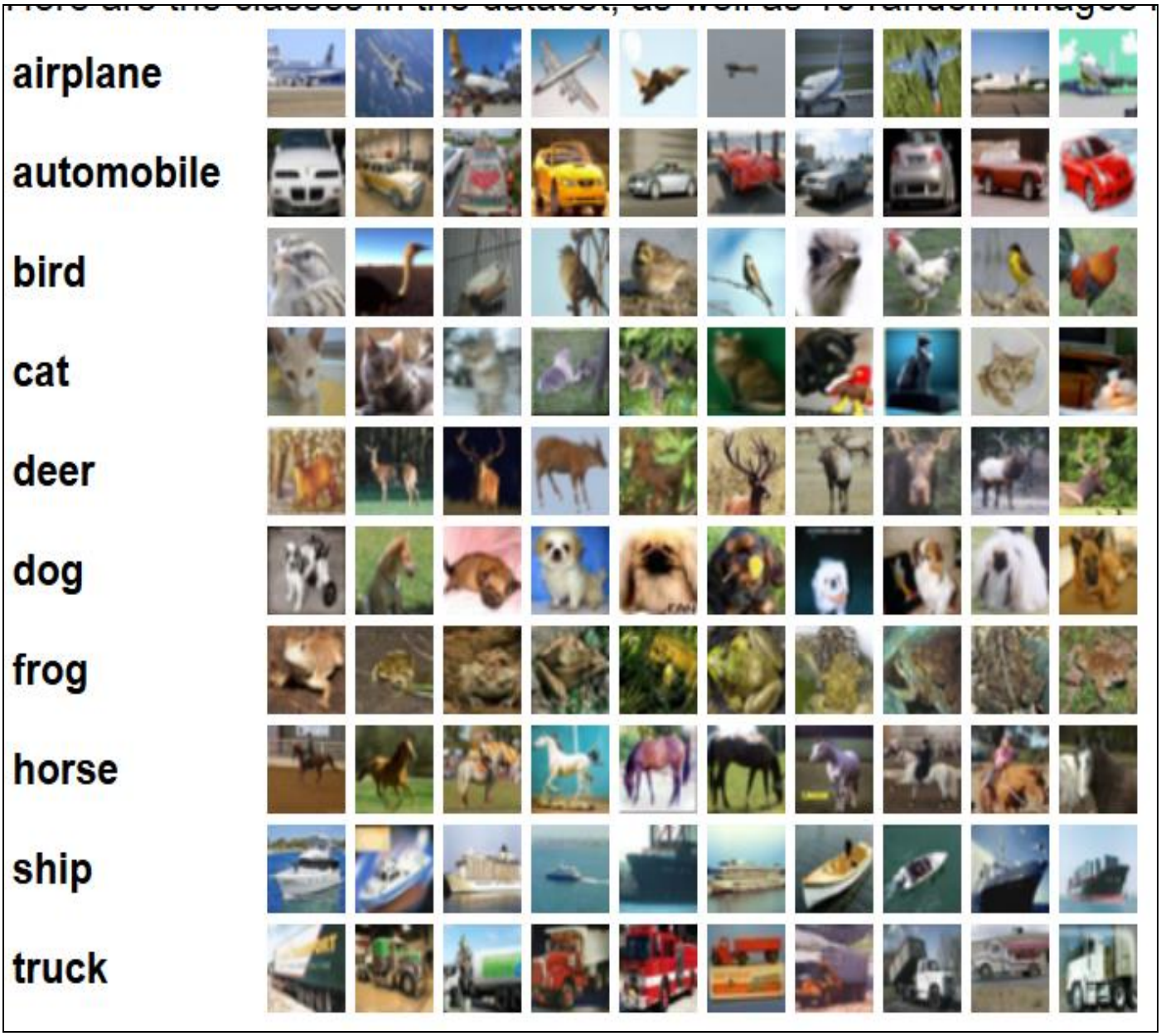

Fig 4 CIFAR-10 Datset Features.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26mar858

IV. DNN:DEEP NUERAL NETWORKS

➢ *DNN having Three Imortent Layers as Shown in Below Table*

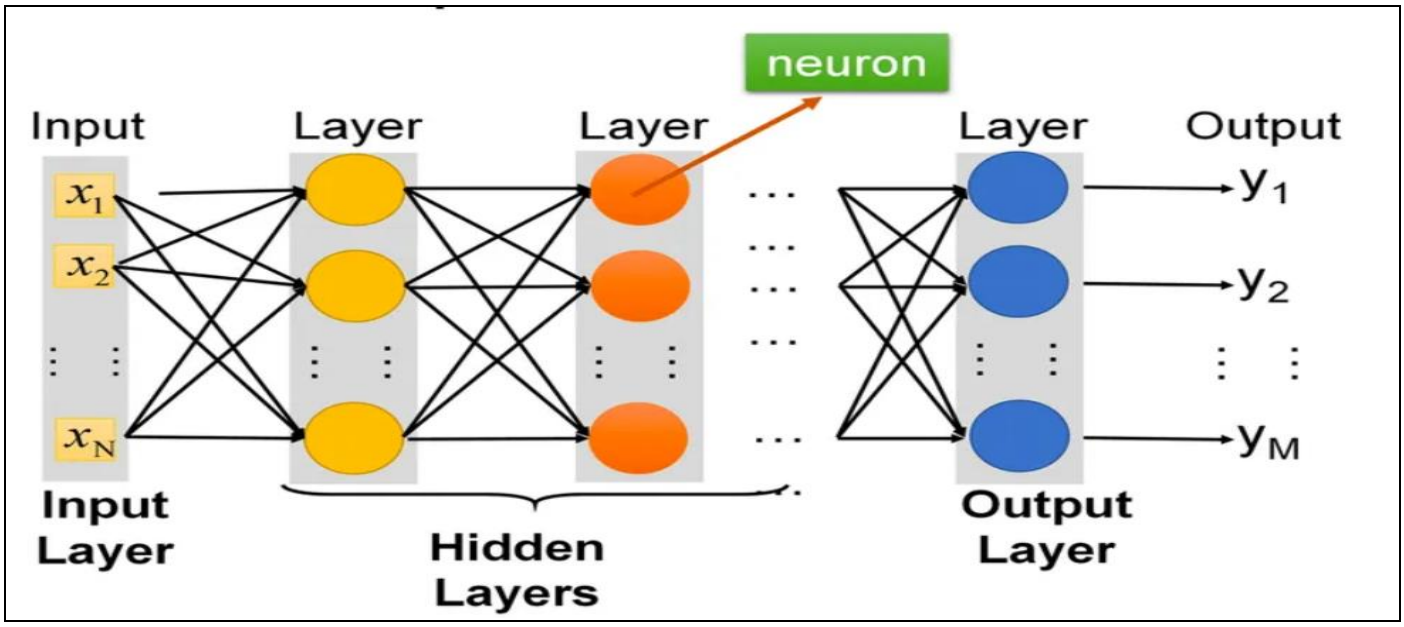

Fig 5 Deep Neural Network for Detection of Cyber Attacks

- Input Layer: The trip begins with the input layer, which is where data enters the network. This might be a picture, word, or any other type of data that the network must comprehend.

- Hidden Layers: Several hidden layers exist between the input and output layers. Each buried layer is made up of virtual neurons that are linked together. These neurons process information from the layer before them and forward it to the next layer. The more hidden layers there are in the network, the more complicated patterns it can learn [9].

- Neurons and Connections: Neurons in each layer are linked to neurons in the following layer via weighed connections. These weights govern how significantly one neuron's input affects the final result of another. When the network undergoes training, these weights are adjusted to increase its performance.

- Activation Function: After calculating the weighted aggregate of inputs for each neuron, the data is sent through a process called the activation function. This function adds non-linearity to the network, allowing it to record complicated data interactions. REL (Rectified Linear Unit) and sigmoid are two common activation functions.

- Output Layer: The network's prediction or classification is produced by the output layer, which is fed by the last hidden layer. For example, if the network is trained to recognize animals in images, the output layer may include neurons representing various animals (e.g., cats, dogs, birds).

- Training: Training a DNN entails providing it with an extensive collection of inputs as well as the proper outputs (labels). By evaluating the predictions to the actual labels, the algorithm changes its weights as well as its biases. This procedure is often carried out utilizing optimization methods such as gradient descent.

V. CONCLUSION

Effective technology is required to combat different types of cyberattacks. Generative artificial intelligence (GenAI) is essential to modern cybersecurity in order to combat more sophisticated, automated threats that defy traditional detection methods. It acts as a force multiplier and greatly enhances security posture by offering real-time anomaly detection, rapid threat analysis, automatic vulnerability patching, and AI-driven phishing definitions. Current deep learning models include DNNs (Deep Neural Networks) and ANNs (Artificial Neural Networks). GAN (Generative Adversarial Network) makes it simple to automatically distinguish between actual and fraudulent photos in order to improve their performance. DNN-GAN with GEN AI which drives to improve more Attacks detect and classify, in future detect and classify newly attacks with performance evolution in metric such as Accuracy.

REFERENCES

[1]. M.Sladic, V. Valeros, C. Catania, S. Garcia, LLM in the shell: generative honeypots, in: 2024 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW), IEEE Computer Society, Los Alamitos, CA, USA, 2024, pp. 430–435, https://doi.org/10.1109/EuroSPW61312.2024.00054

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26mar858

[2]. W. Tann, Y. Liu, J.H. Sim, C.M. Seah, E.-C. Chang, Using Large Language Models for Cybersecurity Capture-The-Flag Challenges and Certification Questions, 2023 arXiv preprint arXiv:2308.10443.

[3]. O.G. Lira, A. Marroquin, M.A. To, Harnessing the advanced capabilities of LLM for adaptive intrusion detection systems, in: L. Barolli (Ed.), Advanced Informatio

[4]. H. Lai, M. Nissim, A survey on automatic generation of figurative language: from rule-based systems to Large Language Models, ACM Comput. Surv. 56 (10) (2024) 1–34, https://doi.org/10.1145/3654795

[5]. .A. Ferrag, M. Ndhlovu, N. Tihanyi, L.C. Cordeiro, M. Debbah, T. Lestable, N.S. Thandi, Revolutionizing cyber threat detection with Large Language Models: a privacy-preserving BERT-based lightweight model for IoT/IIoT devices, IEEE Access 12 (2024) 23733–23750,https://doi.org/10.1109/ACCESS.2024.3363469

[6]. Z. Liu, A review of advancements and applications of Pre-Trained Language Models in cybersecurity, in: 2024 12th International Symposium on Digital Forensics and Security (ISDFS), 2024, pp. 1–10, https://doi.org/10.1109/ ISDFS60797.2024.10527236.

[7]. S. Jamal, H. Wimmer, I.H. Sarker, An improved transformer-based model for detecting phishing, spam and ham emails: a large language model approach. Security and Privacy, 2024 e402, https://doi.org/10.1002/spy2.402.

[8]. A. Fan, B. Gok kaya, M. Harman, M. Lyubarskiy, S. Sengupta, S. Yoo, J.M. Zhang, Large Language models for software engineering: survey and open problems, in: 2023 IEEE/ACM International Conference on Software Engineering: Future of Software Engineering (ICSE-FoSE), IEEE Computer Society, Los Alamitos, CA, USA, 2023, pp. 31–53, https://doi.org/10.1109/ICSE-FoSE59343.2023.00008.

[9]. J. Wu, W. Gan, Z. Chen, S. Wan, P.S. Yu, Multimodal Large Language Models: A Survey, 2023 arrive preprint arXiv:2311.13165.