[**https://sci-hub.ru/**](https://sci-hub.ru/) **find free paper by => doi number given**

1. **The History of Artificial Intelligence**
2. In the first half of the 20th century, science fiction familiarized the world with the concept of artificially intelligent robots.
3. By the 1950s, we had a generation of scientists, mathematicians, and philosophers with the concept of artificial intelligence (or AI) culturally assimilated in their minds.
4. However, while the basic proof of principle was there, there was still a long way to go before the end goals of natural language processing, abstract thinking, and self-recognition could be achieved.
5. The biggest was the lack of computational power to do anything substantial: computers simply couldn’t store enough information or process it fast enough.
6. It offers a bit of an explanation to the roller coaster of AI research; we saturate the capabilities of AI to the level of our current computational power (computer storage and processing speed), and then wait for Moore’s Law to catch up again.
7. In the long term, the goal is general intelligence, that is a machine that surpasses human cognitive abilities in all tasks.

<https://sitn.hms.harvard.edu/flash/2017/history-artificial-intelligence/>

This AI Model Can Generate Ultra-Short Summary of Scientific Papers

<https://medium.com/swlh/tldr-model-can-generate-single-sentence-summary-of-scientific-papers-fa5f55a55835>

Few topic article ([**Recommendation Algorithms**](http://sitn.hms.harvard.edu/flash/2017/recommended-machine-learning-helps-choose-consume-next/)) ([**Computers Dreaming**](http://sitn.hms.harvard.edu/flash/2017/psychosis-dreams-memory-ai/))

<https://sitn.hms.harvard.edu/special-edition-artificial-intelligence/>

### **Machine Learning and Data Security**

An ML-based approach to **malware** detection typically uses **Supervised Machine Learning** that derives a malware detection model from the vast array of labeled examples of malware (Trojan horses, spyware etc.). To create such a model, we first need to collect a training data set from available malware examples, cleanse, and prepare it for the training. Eventually, we can feed this data to the learning algorithm designed to identify recurring patterns and features in the training set. If we’ve done everything right, we can expect that the ML algorithm comes up with some abstract representation of malware patterns which can be thereon employed to detect new malware. A similar approach is already effectively used for spam detection based on the vast amounts of labeled spam examples.

<https://datascience.foundation/datatalk/machine-learning-and-data-security>

# **নিউরনের গল্প**

তো কী করা যায় এই নিউরাল নেটওয়ার্ক দিয়ে? আসলে কম্পিউটার দিয়ে আমরা সাধারণত যে প্রোগ্রামগুলো লিখি বা যেই সফটওয়্যারগুলো ডেভেলপ করি তাদের শেখার ক্ষমতা নাই। মানে সে মানুষের মত শিখতে পারে না। এখন আমরা যদি চাই কোন কম্পিউটারকে কিছু ইনপুট আর আউটপুট দিব। সে সেই ইনপুট, আউটপুটগুলা দেখে শিখবে। পরে আমি যখন তাকে নতুন কোন ইনপুট দিব সে আমাকে তখন সেই নতুন ইনপুটের জন্য একটা আউটপুট দিবে। ব্যাপারটা এরকম, আমি এমন একটা প্রোগ্রাম লিখব যাকে প্রথমে কিছু কুকুর আর বিড়ালের ছবি দিয়ে বলব কোনটা কুকুর আর কোনটা বিড়াল। সে দেখে দেখে শিখবে। তারপর তাকে যখন নতুন কোন বিড়াল বা কুকুরের ছবি দিব সে আমাদের বলে দিব সেটা কী? কুকুর নাকি বিড়াল? এই কাজ করতে গেলেই আমাদের নিউরাল নেটওয়ার্কের কাছে হাত পাততে হয়।  
<https://medium.com/%E0%A6%AA%E0%A7%8D%E0%A6%B0%E0%A7%8B%E0%A6%97%E0%A7%8D%E0%A6%B0%E0%A6%BE%E0%A6%AE%E0%A6%BF%E0%A6%82-%E0%A6%AA%E0%A6%BE%E0%A6%A4%E0%A6%BE/%E0%A6%A8%E0%A6%BF%E0%A6%89%E0%A6%B0%E0%A6%A8%E0%A7%87%E0%A6%B0-%E0%A6%97%E0%A6%B2%E0%A7%8D%E0%A6%AA-afab31782e3a>

# **BUET- Dissertation/Theses - Department of Computer Science and Engineering: Recent submissions**

Many topic idea you find from there **recent work** and thesis **Paper found**

<http://lib.buet.ac.bd:8080/xmlui/handle/123456789/60/recent-submissions>

The role of AI in cybersecurity and also a roadmap for implementing AI in cybersecurity

<https://www.capgemini.com/wp-content/uploads/2019/07/AI-in-Cybersecurity_Report_20190711_V06.pdf>

Basic Recommendation system Discuss

<https://www.youtube.com/watch?v=_hf_y-_sj5Y&list=PLZoTAELRMXVN7QGpcuN-Vg35Hgjp3htvi>

# **Artificial Intelligence for Detection, Estimation, and Compensation of Malicious Attacks in Nonlinear Cyber-Physical Systems and Industrial IoT**

<https://ieeexplore.ieee.org/document/8917652>

**Virus developers can also use AI. Those AI-powered viruses can potentially cause more damage than other viruses because they may be able to detect antivirus software, attack its code, and bypass it**

<https://nordvpn.com/blog/artificial-intelligence-in-cyber-security/>

# How to tell if you have malware

<https://nordvpn.com/blog/signs-of-malware/>

# What is machine learning?

<https://nordvpn.com/blog/machine-learning/>

* AI can also be used by [hackers](https://nordvpn.com/blog/what-is-a-hacker/) for malicious purposes and to initiate even more sophisticated and **large-scale attacks**. AI also can help to **identify and exploit vulnerabilities** more quickly and efficiently.

**Harnessing Artificial Intelligence Capabilities to Improve Cybersecurity**

**Abstract:**

Cybersecurity is a fast-evolving discipline that is always in the news over the last decade, as the number of threats rises and cybercriminals constantly endeavor to stay a step ahead of law enforcement. Over the years, although the original motives for carrying out cyberattacks largely remain unchanged, cybercriminals have become increasingly sophisticated with their techniques. Traditional cybersecurity solutions are becoming inadequate at detecting and mitigating emerging cyberattacks. Advances in cryptographic and Artificial Intelligence (AI) techniques (in particular, machine learning and deep learning) show promise in enabling cybersecurity experts to counter the ever-evolving threat posed by adversaries. Here, we explore AI’s potential in improving cybersecurity solutions, by identifying both its strengths and weaknesses. We also discuss future research opportunities associated with the development of AI techniques in the cybersecurity field across a range of application domains

<https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=8963730>

# 

**Mail to Sir And you also read**

# **Study of Artificial Neural Networks in Information Security Risk Assessment (2020)**

study may be Impact of security measures on the reduction of information security risk show. Conclusion: after presentation of the organization and review the data and information available and the resources to implement the research model of this study in order to improve results and compliance with the the coin real information below to copy the proposed change found Login to have consisted of real attacks Statistics information to the relevant organization in order to carry out the investigation after the observation and study of the statistics Vazhga attacks the nervous system sorted them by teaching the necessary research to the night of the the security risk information by using the algorithm of the themes of artificial intelligence 68 M to be in this way given the number 87789 in the neural network C of the conjugate graded 69 to 10 layers, and 5 percent were trained in control and test data.

# <http://journals.researchub.org/index.php/JMAS/article/view/861>

# **Intelligent Cyber Attack Detection and Classification for Network-Based Intrusion Detection Systems (2021)**

Intrusion Detection Systems (IDS) are important security mechanisms that can perform the timely detection of malicious events through the inspection of network traffic or host-based logs. Many machine learning techniques have proven to be successful at conducting anomaly detection throughout the years, but only a few considered the sequential nature of data. This work proposes a sequential approach and evaluates the performance of a Random Forest (RF), a Multi-Layer Perceptron (MLP), and a Long-Short Term Memory (LSTM) on the CIDDS-001 dataset.

<https://www.mdpi.com/2076-3417/11/4/1674/htm>

**ZERO DAY ATTACK DETECTION USING MACHINE LEARNING TECHNIQUES**

zero-day attacks are cyber attack as opposed to software blemish that is unknown and has no patch. Based on the recently discovered types of zeroday attacks it has become evident that operating system protection is becoming less effectual attacks that are performed by exploiting the system vulnerabilities are more common, and cyber attacks are becoming more sophisticated and better at bypassing any organization defences. The traditional signature-based approach was not enough to detect the zero-day attacks. And yet the signature prediction approach was also been not seems to be as operative as it has to be. The key purpose of this project was to propose a solution based on the detection of zero-day malware by using machine learning and artificial intelligence.

<https://www.ijrar.org/papers/IJRAR19J1648.pdf>

# **Malware Detection and Prevention using Artificial Intelligence Techniques**

<https://ieeexplore.ieee.org/abstract/document/9671434>

**MUST Read**

# **Cybersecurity Tasks and Machine Learning**

Instead of looking at ML tasks and trying to apply them to cybersecurity, let’s look at the common cybersecurity tasks and machine learning opportunities. There are three dimensions (Why, What, and How).

The first dimension is a goal, or a task (e.g., detect threats, predict attacks, etc.). According to [Gartner’s PPDR model](https://www.gartner.com/document/3286317), all security tasks can be divided into five categories:

* prediction;
* prevention;
* detection;
* response;
* monitoring.

The second dimension is a technical layer and an answer to the “What” question (e.g., at which level to monitor issues). Here is the list of layers for this dimension:

* network (network traffic analysis and intrusion detection);
* endpoint (anti-malware);
* application (WAF or database firewalls);
* user (UBA);
* process (anti-fraud).

Each layer has different subcategories. For example, network security can be Wired,Wireless or Cloud. Restassured thatyou can’t apply the same algorithms with the same hyper parameters to both areas, at least in near future. The reason is the lack of data and algorithms to find better dependencies of the three areas so that it’s possible to change one algorithm to differentones.

The third dimension is a question of “How” (e.g., how to check security of a particular area):

* in transit in real time;
* at rest;
* historically;
* etc.

For example, if you are about endpoint protection, looking for the intrusion, you can monitor processes of an executable file, do static binary analysis, analyze the history of actions in this endpoint, etc.

Some tasks should be solved in three dimensions. Sometimes,there are no values in some dimensions for certain tasks. Approaches can be the same in one dimension. Nonetheless, each particular point of this three-dimensional space of cybersecurity tasks has its intricacies.

It’s difficult to detail them all so let’s focus on the most important dimension — technology layers. Look at the cybersecurity solution from this perspective

<https://towardsdatascience.com/machine-learning-for-cybersecurity-101-7822b802790b>

**Give to sir and you also read**

# **Cyber Threat Intelligence-Based Malicious URL Detection Model Using Ensemble Learning (2022)**

Web applications have become ubiquitous for many business sectors due to their platform independence and low operation cost. Billions of users are visiting these applications to accomplish their daily tasks. However, many of these applications are either vulnerable to web defacement attacks or created and managed by hackers such as fraudulent and phishing websites. Detecting malicious websites is essential to prevent the spreading of malware and protect end-users from being victims. However, most existing solutions rely on extracting features from the website’s content which can be harmful to the detection machines themselves and subject to obfuscations. Detecting malicious Uniform Resource Locators (URLs) is safer and more efficient than content analysis. However, the detection of malicious URLs is still not well addressed due to insufficient features and inaccurate classification. This study aims at improving the detection accuracy of malicious URL detection by designing and developing a cyber threat intelligence-based malicious URL detection model using two-stage ensemble learning. The cyber threat intelligence-based features are extracted from web searches to improve detection accuracy. Cybersecurity analysts and users reports around the globe can provide important information regarding malicious websites. Therefore, cyber threat intelligence-based (CTI) features extracted from Google searches and Whois websites are used to improve detection performance. The study also proposed a two-stage ensemble learning model that combines the random forest (RF) algorithm for preclassification with multilayer perceptron (MLP) for final decision making. The trained MLP classifier has replaced the majority voting scheme of the three trained random forest classifiers for decision making. The probabilistic output of the weak classifiers of the random forest was aggregated and used as input for the MLP classifier for adequate classification. Results show that the extracted CTI-based features with the two-stage classification outperform other studies’ detection models. The proposed CTI-based detection model achieved a 7.8% accuracy improvement and 6.7% reduction in false-positive rates compared with the traditional URL-based model.

<https://www.mdpi.com/1424-8220/22/9/3373/htm>

# **Optimized URL Feature Selection Based on Genetic-Algorithm-Embedded Deep Learning for Phishing Website Detection (2022) (good one)**

# Deep learning models for phishing URL classification based on character- and word-level URL features achieve the best performance in terms of accuracy. Various improvements have been proposed through deep learning parameters, including the structure and learning strategy. However, the existing deep learning approach shows a degradation in recall according to the nature of a phishing attack that is immediately discarded after being reported. An additional optimization process that can minimize the false negatives by selecting the core features of phishing URLs is a promising avenue of improvement. To search the optimal URL feature set and to fully exploit it, we propose a combined searching and learning strategy that effectively models the URL classifier for recall. By incorporating the deep-learning-based URL classifier with the genetic algorithm to search the optimal feature set that minimizing the false negatives, an optimized classifier that guarantees the best performance was obtained. Extensive experiments on three real-world datasets consisting of 222,541 URLs showed the highest recall among the deep learning models. We demonstrated the superiority of the method by 10-fold cross-validation and confirmed that the recall improved compared to the latest deep learning method. In particular, the accuracy and recall were improved by 4.13%p and 7.07%p, respectively, compared to the convolutional–recurrent neural network in which the feature selection optimization was omitted.

<https://www.mdpi.com/2079-9292/11/7/1090/htm>

# **An Enhanced Intrusion Detection Model Based on Improved kNN in WSNs (2022)**

Aiming at the intrusion detection problem of the wireless sensor network (WSN), considering the combined characteristics of the wireless sensor network, we consider setting up a corresponding intrusion detection system on the edge side through edge computing. An intrusion detection system (IDS), as a proactive network security protection technology, provides an effective defense system for the WSN. In this paper, we propose a WSN intelligent intrusion detection model, through the introduction of the k-Nearest Neighbor algorithm (kNN) in machine learning and the introduction of the arithmetic optimization algorithm (AOA) in evolutionary calculation, to form an edge intelligence framework that specifically performs the intrusion detection when the WSN encounters a DoS attack. In order to enhance the accuracy of the model, we use a parallel strategy to enhance the communication between the populations and use the Lévy flight strategy to adjust the optimization. The proposed PL-AOA algorithm performs well in the benchmark function test and effectively guarantees the improvement of the kNN classifier. We use Matlab2018b to conduct simulation experiments based on the WSN-DS data set and our model achieves 99% ACC, with a nearly 10% improvement compared with the original kNN when performing DoS intrusion detection. The experimental results show that the proposed intrusion detection model has good effects and practical application significance.

In the future, we will focus on developing an unsupervised or semi-supervised WSN intrusion detection model, such as k-means optimized by an evolutionary algorithm, and so on. These models will not only target a particular type of DoS attack, but also strive to cover Sybil attacks, routing attacks, and other possible attacks.

<https://www.mdpi.com/2079-9292/10/15/1854/htm>

# **Usability and Security Testing of Online Links: A Framework for Click-Through Rate Prediction Using Deep Learning**

Many internet applications, such as internet advertising and recommendation systems, rely on click-through rate (CTR) prediction to anticipate the possibility that a user would click on an ad or product, which is key for understanding human online behaviour. However, online systems are prone to click on fraud attacks. We propose a Human-Centric Cyber Security (HCCS) model that additionally includes AI techniques targeted at the key elements of user, usage, and usability. As a case study, we analyse a CTR prediction task, using deep learning methods (factorization machines) to predict online fraud through clickbait. The results of experiments on a real-world benchmark Avazu dataset show that the proposed approach outpaces (AUC is 0.8062) other CTR forecasting approaches, demonstrating the viability of the proposed framework.

<https://www.mdpi.com/2079-9292/11/3/400/htm>

New links:

Recently, researchers showed that phishing attacks can be performed by employing a deep neural network-based phishing URL generating system called DeepPhish. To prevent this kind of attack, we design an ensemble machine learning-based detection system called PhishHaven to identify AI-generated as well as human-crafted phishing URLs.To speed up the ensemble-based machine learning models, PhishHaven employs a multi-threading approach to execute the classification in parallel, leading to real-time detection.

<https://ieeexplore.ieee.org/document/9082616>

Here we present a comparative study between classical machine learning technique - logistic regression using bigram, deep learning techniques like convolution neural network(CNN) and CNN long short-term memory(CNN-LSTM) as architectures used to detect malicious URLs. On comparison CNN-LSTM gave the best accuracy of about 98% for the classification of phishing URLs

<https://ieeexplore.ieee.org/document/8494159>

In this paper, therefore we applied four machine learning algorithms to detect malicious URLs. The experimental results show that the best performance is achieved by Random Forest algorithm with an accuracy of 96% and 95% during two test phases.

<https://ieeexplore.ieee.org/document/9213792>

**New**

# **PWDGAN: Generating Adversarial Malicious URL Examples for Deceiving Black-Box Phishing Website Detector using GANs**

**In this article, we build a model based on generative adversarial network (GAN) – a deep learning-based framework to conduct black-box attacks using Phishtank and Alexa datasets that try to evade and bypass ML-based phishing detectors. We apply PWDGAN to carry out the attacks against the ML-based black-box classifiers and then deceive successfully these detectors with the rate of detecting malicious samples value of approximately 0% after 500 training epochs. The results of the paper demonstrate the effectiveness of GAN adoption in creating new patterns that can evade and bypass phishing detectors. These newly generated patterns can serve as material for future research in phishing website detection and improve the ability to detect novel anomaly attacks.**

[**https://ieeexplore.ieee.org/abstract/document/9690540**](https://ieeexplore.ieee.org/abstract/document/9690540)

# Machine learning based phishing detection from URLs

These products cannot prevent all of the phishing attacks. In this paper, a real-time anti-phishing system, which uses seven different classification algorithms and natural language processing (NLP) based features, is proposed. The system has the following distinguishing properties from other studies in the literature: language independence, use of a huge size of phishing and legitimate data, real-time execution, detection of new websites, independence from third-party services and use of feature-rich classifiers. For measuring the performance of the system, a new dataset is constructed, and the experimental results are tested on it. According to the experimental and comparative results from the implemented classification algorithms, Random Forest algorithm with only NLP based features gives the best performance

<https://www.researchgate.net/publication/344952543_Machine_learning_based_phishing_detection_from_URLs>

# **Classifying phishing URLs using recurrent neural networks**

explored the use of URLs as input for machine learning models applied for phishing site prediction. In this way, we compared a feature-engineering approach followed by a random forest classifier against a novel method based on recurrent neural networks. We determined that the recurrent neural network approach provides an accuracy rate of 98.7% even without the need of manual feature creation, beating by 5% the random forest method. This means it is a scalable and fast-acting proactive detection system that does not require full content analysis.

<https://ieeexplore.ieee.org/abstract/document/7945048>

Phishing Website Detection using Machine Learning Algorithms

<https://www.ijcaonline.org/archives/volume181/number23/mahajan-2018-ijca-918026.pdf>

Literature review for proposal:

[1] <https://link.springer.com/article/10.1007/s10115-022-01672-x>

[2] <https://ieeexplore.ieee.org/document/9213792>

[3]<https://www.researchgate.net/publication/351929313_Phishing_website_detection_using_machine_learning_and_deep_learning_techniques>

[4]<https://www.researchgate.net/publication/351946485_An_Optimized_Stacking_Ensemble_Model_for_Phishing_Websites_Detection>

[5] <https://norma.ncirl.ie/4509/1/sharadrajendraparmar.pdf>