International Journal of Applied Engineering & Technology

A SYSTEMATIC STUDY OF IMPLEMENTATION OF PROGNOSTIC APPROACH TO ENHANCE DATA SECURITY USING COGNITIVE BEHAVIOR OF MACHINE LEARNING

Mr. Prabhanjan Chaudhari[a], Dr. Guddi Singh[b], Dr. Amit Bhusari[c]
[a]Research Scholar, Kalinga University, New Raipur, Chhattisgarh
[b]Assistant Professor, Department of Computer Science and Engineering, Kalinga University, New Raipur, Chhattisgarh
[c]Assistant Professor, Department of MCA, Trinity Academy of Engineering, Pune, Maharashtra

ABSTRACT
This study introduces the significance of data science and machine learning in the modern era and highlights the growing importance of understanding cognitive behavior within machine learning systems. This paper explains the role of data science, emphasizing its multidisciplinary nature. It provides an overview of Machine Learning and Big data used for training models to make predictions and decisions based on data using Prognostic approach like in health care sector. The review explores practical applications of machine learning with cognitive behavior, recommendation systems, and autonomous systems. It discusses usability and ethical implications associated with cognitive behavior in machine learning. Emphasizing the significance of understanding and harnessing cognitive behavior within machine learning for advancing data science and AI applications. The objective of to connect the concept of prognostic approach used in health care and medical domain in cyber threats detection.

Keywords: Prognostic approach, Machine Learning, Big Data, Cognitive Behavior, Data Security.

1. INTRODUCTION
There are two blue-chips: Machine Learning and Big Data of today's IT sector. Large volumes of data are reviewed and information extracted using big data storage. Machine learning, on the other hand, refers to a computer's ability to learn and develop without being explicitly taught. Automatic data processing and decision-making algorithms are the pillars of machine learning that learn from their past experiences and improve at each stage of their job. "Evolve via learning," to put it another way. To track growth and changes in data flow in the apropos of Big Data, Machine Learning is applied in order to provide constantly evolving and relevant insights. Machine learning algorithms explain and detect patterns in the incoming data, they are then translated into other languages with actionable insights that can be incorporated into business processes. Many more decision-making processes were then automated by use of the algorithms. [1]

Decision trees and neural networks are used in combination with machine learning methods for these reasons. Many sectors have seen amazing development as a result of the dominating mix of Machine Learning and big data. One of these industries is the e-commerce industry. Integrating statistical models with data assists financial analysts in determining the solutions for various financial crises and to decide the remedies to overcome from it and for future perfectness.

Financial analysts may use predictive analytics to track and exchange critical information about the various economic problems. They automatically retain data on their daily transactions, payments and linked systems, allowing customers to remotely access and manage the financial transactions using the concept of cognitive behavior of Machine Learning [2].

2. Various studies

2.1 Prognostics Approach
Prognostics is a discipline of engineering concerned with predicting when a system or component may cease to function as intended. This loss of performance is usually triggered by the point at which the system is no longer usable to accomplish the desired outcomes. After then, the projected term is transformed to Remaining Useful

Life (RUL), which is very important perception in contingency planning. To anticipate future performance, prognostics examine the degree of a system's departure or degradation from its projected normal operating circumstances. Prognostics is the study of failure processes, monitoring and identifying potential problems before they become major problems. If a detailed understanding of the non-effective mechanisms that are the probable causes to generate poverty state is obtained, an effective prognostic solution may be applied. And, finally, system failures are achieved. As a consequence, gaining prior information regarding a product's possible problems is crucial (including the location, mode, cause, and method). This information is required in order to establish the parameters of the system that need to be kept an eye on. Prognostics may be useful in condition-based maintenance. Health and prognostics (PHM) are a term that is used interchangeably with other terms in the transportation industry, such as vehicle health management (VHM) and engine health management (EHM), is a discipline that links failure mechanism research with system lifecycle management. The three kinds of technical techniques to construct models in prognostics include techniques based on data, models, and hybrid approaches. [3]

2.2 Data Security

You might think of data security as a set of procedures and rules that are meant to secure your key information technology infrastructure (IT). Everything from databases to accounts to networks was scanned for evidence. Controls, apps, and procedures must be used in conjunction to evaluate the value of diverse datasets and execute the most appropriate security rules.

While implementing operative data security measures, the compassion of diverse datasets along with regulatory acquiescence must be considered. Security measures like data and perimeter protection aren't a one-size-fits-all strategy for keeping hackers at bay. When it comes to safeguarding and managing information, one of the most important tools is data security.

There is a high demand of Data security in both public and private sector businesses due to increase in various types of cyber-attacks. To begin with, companies have a protection against unauthorized access is a legal and moral obligation for companies. Payment Card Industry Data Security Standard (PCI DSS) can be used to protect customers private and confidential data like user id and password.

There are the chances that a data leakage or hacking may cause of financial loss of the customer. After assessing the damage, you'll need to determine which company processes were unsuccessful and what has to be changed. [4]

2.3 Types of Data Security

Controls for gaining access to restricted areas

Physical and digital access restrictions might be part of a data security plan that restricts vital systems and information. All computers and gadgets must be checked to make sure they are working properly password-protected, and that only authorized staff have access to the physical space.

i. Authentication

Accurately verifying the identity of a user prior to providing them with access to data is called Authentication. There are a variety of methods for gaining access to a system.

ii. Recovery & Backups

In data security this is another part which may result in a system failure, catastrophe, corrupted, or breach of data. To restore from a backup, you'll need a different kind of storage medium, for example HDD, SSD or Cloud Storage etc.

iii. Data Erasure

Disposing Data after regular intervals in an appropriate manner. Data Erasure may cause due to internal factors like software employees working in that organization.

iv. Data Masking
Data masking software conceals information by using proxy characters to mask letters & numbers. The data can be hidden, so that it cannot be access. The data converts in original form only when a legitimate user receives it.

v. Data Resiliency
If you have thorough data security, your systems should have that kind of technology updated withstand or recover from any type of failures. Resilience can be built in any of these two forms hardware & software to safeguard that security is not compromised by unforeseen circumstances like natural disasters or power outages.

vi. Encryption
A network security algorithm transforms written characters into an unintelligible format utilizing encryption keys. Only users with the proper keys & authorization can unlock & access the information.

2.4 Data Science
Predictive analytics and machine learning algorithms, as well as the creation of new technologies for mining massive amounts of data (Big Data) are all cornerstones of data science, which is an interdisciplinary study.

Begin with describing the data science life cycle and improving data science project management. The first phase Capture is a step in the data science pipeline that comprises gathering data, extracting for analysis.

In the maintenance stage, data warehousing, data cleaning, data processing, data staging, and data architecture are all part of the job. [5]

A few of the strategies used to analyze data include data mining, data categorization and clustering, data modeling, and communicating insights create usable data.

The next, and equally crucial, phase is data analysis. Among other things, data scientists here work text mining and qualitative analysis, as well as exploratory and confirmation work. When done correctly, data science today cannot be categorized as "cookie-cutter."

At the end of the process, the data scientist discusses his or her results. This includes data using a variety of business intelligence tools and techniques, as well as assisting enterprises, policymakers, and others in making better choices.

By 2020, the world's data will have grown to 40 zettabytes (40 trillion gigabytes). The amount of data accessible is growing at a breakneck speed. Over 90% of this vast amount of data is produced in the previous two years at any one time, according to firms like IBM and SINTEF. In the latest survey published by various agencies noted that daily, Internet end users are generating near about 2.0 to 3.0 quintillion bytes of data of various types. Every individual on the planet will generate an average of 146,880 GB of data per day by the year 2020, and 165 zettabytes per year by 2025.

This shows that data science still has a lot of work to do and a lot more to learn. According to The Guardian, just around 0.5 percent of the data was examined in 2012.

Normal data analysis may be used to understand data from only one source or a little quantity of data. Data science instruments and tools, on the other hand, are critical for making sense of large amounts of data and data from a variety of sources. This is shown through a look at some of the commercial applications of data science, which gives an impressive start to the field of data science.

2.5 Cybersecurity
The area of cybersecurity is concerned with the protection of data and systems. Due to the limits of cybersecurity solutions, this is a critical responsibility that will only get more difficult.

The purpose of cybersecurity is to protect systems against hacker attacks and to secure them. Hackers, on the other hand, always had the upper hand owing to their strategy of reacting to assaults after they occurred. As a result of reactionary efforts, cybersecurity is slower than the dangers it faces.

Web application firewalls (WAFs) are ways of detecting malicious code and determining the appropriate course of action for a firewall. Rule-based WAFs and signature-based WAFs are two types of WAFs.

The two systems, as seen in the diagram above, are tough, and fresh attacks must be pre-programmed. Signature-based detection looks for signs that an attack is about to happen. These signatures must be obtained ahead of time, and the approach is hampered significantly by assailants who have yet to be seen. Each preloaded code sample must be examined by a signature-based detector in search of the best match. As a consequence, reaction times are sluggish, and false positives are possible.

A distinct method is used in rule-based detection. Instead, then hunting for codes one by one, the technique focuses on the hack's consequences. When we speak about "rules," we're just referring to suspicious behavior that a hacker may do that clean code wouldn't. This technique is faster since it eliminates options based on the code's effect rather than the code itself having to read through each signature.

However, having instances of harmful code to examine is still necessary. And it is from this reactive posture that their overall plan emerges. You'll see the abbreviation FUD, which stands for 'fear, uncertainty, and doubt,' wherever you look online. Many people in the cybersecurity profession, according to my study, are wary of this idea being their guiding light. This comes down to security personnel operating in the dark and striking at their opponents blindly.

2.6 Data Security in Data Science

When these two fields, data science, and security, come together, cybersecurity gets a powerful tool to combat invasions. Data science serves as the eyes and ears for the sword of cybersecurity.

Cybersecurity data science (CSDS) is a method for detecting hostile assaults on digital infrastructures that are based on science. It is a data-driven strategy that employs machine learning methods to detect threats.

An important element of machine learning in cybersecurity is anomaly detection. Attacks are often carried out via code that deviates from the norm or performs activities that are deemed unusual. Using data science approaches to aid cybersecurity is as simple as creating a machine learning model to identify an abnormality.

Machine learning may also be used for penetration testing. Machine learning is a good test for firewalls that safeguard data and data structures because of its automation and ability to adapt from prior experiences.

Data scientists provide cybersecurity employees with information that helps them better understand how to defend against assaults.

2.7 Big Data

"Big Data" is the collection of all data created at an unprecedented pace throughout the world. Structured or unstructured data might be present. A strongly knowledge-oriented economy is responsible for a large portion of today's corporate firms' success.

Deciphering the numerous patterns and making meaning of this data by discovering previously undiscovered relationships within the large sea of data becomes a crucial and very satisfying activity. Big Data must be transformed into Business Marketing Intelligence that businesses can use immediately. Improved information contributes to better decision-making and a better approach to storiform, regardless of their size, location or market share, or client segmentation, may benefit from this strategy. Massive amounts of data may be processed with Hadoop.

Big Data has a set of properties that are specified by the 4Vs:

1. Volume:
Because organizations may gather a vast quantity of data, the volume of information or data becomes a significant aspect of Big Data analytics.

2. Velocity: The velocity at which latest or new information is created, owing to our reliance on the internet, sensors, and machine-to-machine interactions, is also critical for parsing Big Data promptly.

3. Variety: The data created is entirely heterogeneous in the sense that it may be in a variety of forms such as to get the most value out of Big Data obtained.

4. Veracity: Before understanding and utilizing Big Data for business requirements, it's critical to establish it's possible that the data isn't trustworthy.

2.8 Study of Cognition
As an example, Youngsters create their own mental architecture to fit into their cognitive surroundings. They actively create their own interpretations and meanings. The following describes the cognitive development process from infancy to adolescence.

Early life: The infant plans and synchronizes their bodily motions with their sensory experiences, such as hearing and seeing. They pick up on it fast that objects they see, even though they are no longer in their immediate surroundings, do not truly vanish. They have an active pattern-scanning ability and a developing memory in ways that modern neuroscience is still investigating.

Early Childhood: As a youngster gains experiences, their mental life expands. They have mental images of many different objects in the world.

In Cognitive Development, there are two very important processes. One is Critical Thinking and the second is creative learning [25]

2.9 Machine learning through a cognitive approach
As mentioned in the above explanation we can trend the machine through a cognitive approach and such a trained system can help in predictive analysis to find the vulnerabilities in the system like various threats and possibilities of attacks.

3. Related Work
Andrea K. Bowe (2022) [6] In comparison to its use in the clinical context, ML has gotten far less attention when applied to public health issues. Addressing differences in early childhood cognitive development is one such challenge. This is a complicated public health problem with roots in the social determinants of health, which is made worse by inequality & characterized by intergenerational transmission. Early life offers a window of opportunity for early intervention to enhance cognitive development because it is the time when neuroplasticity is at its peak. Unfortunately, many people will miss this window, and intervention might not take place at all or might happen only when overt indicators of cognitive impairment appear. In this study, we investigate the potential benefits of data science is useful in the early detection of children at risk for poor cognitive outcomes, an area where there appears to be a paucity of research. We compare and contrast ML approaches with conventional statistical methods, give instances of how ML has been applied thus far in the study of neurodevelopmental diseases, and present a review of the benefits and drawbacks of using it at the population level. The review finishes by outlining possible possibilities for this field's future research.

Afreen Khan (2022) [7] Alzheimer's disease (AD) kills brain cells. It is one of the main factors contributing to cognitive decline and memory loss in senior people all over the world. The healthcare community is more interested in early detection & simplifying diagnostic procedures. AD research has made substantial use of ML methods or other multivariate data exploration tools. This study's main objective is to demonstrate an automated categorization system for retrieving information patterns. We suggested a five-step machine learning pipeline. This machine learning pipeline is made up of a classifier system, data transformation, or feature selection methods that are integrated into an experimental and data analysis architecture. The Random Forest (RF) classifier's performance measures demonstrated the accurate output in classification.

Alessandro Allegra (2022) [8] Using ML algorithms & deep learning techniques, artificial intelligence has radically changed the landscape of oncology research. ML and Deep learning are the subsets of artificial intelligence are primarily represented by ANN, which are algorithms acts like a brain. AI is used in the diagnosis of multiple myeloma & present the most important DP & ML experiments conducted in the field. One of the most prevalent hematological cancers in the world, multiple myeloma is also one of the hardest to treat due to its high relapse and chemo resistance rates. The detection of new markers for their fast discovery & therapy selection, as well as a better assessment of its relapse & survival, are predicted to be among the future ways to combat this tumor with a poor prognosis.

Rutvij H. Jhaveri (2022) [9] In present era of Industry 5.0 revolution is one in which enormous amounts of data are being exchanged digitally. DL & ML Techniques have lately gained widespread recognition & adoption by a number of real-time engineering applications due to their outstanding performance. Developing automated & intelligent applications that can manage data in areas like health, cyber-security, & intelligent transportation systems requires a solid understanding of machine learning. The field of ML encompasses a variety of techniques, such as reinforcement learning, semi-supervised, unsupervised, & supervised algorithms. Recent research advances our understanding of how various machine learning techniques can be used in practical contexts like cyber security or intelligent system of transportation. This research reflects the main objective of the research and the challenges that ML methods face when handling practical applications.

Thanh Thi Nguyen (2021) [10] In recent years Cyber Systems are highly vulnerable. Due to complexity of Cyberattacks highly responsive, adaptable and scalable defense measures are needed like deep reinforcement learning (DRL) for effective protections. DRL provides high-dimensional cyber protection challenges in combination of deep learning with classical RL. This study reflects DRL strategies created for cyber security is presented. We discuss a variety of important topics, such as DRL-based cybersecurity methodologies for cyber-physical systems, self-contained intrusion detection systems, or multiagent DRL-based game theory simulation for cyberattack defense tactics. On DRL-based cyber security, extensive discussions & future research prospects are also provided.

Kosrat Dlshad Ahmed (2021) [11] IoT devices & connection offer people a better user experience or raise the standard of service from a variety of angles. In this regard, recent technical advancements & management of the necessary elements for performance delivery must be ensured. IoT applications are becoming increasingly important, and this has created tremendous opportunities for management & development. Users' attention has recently been drawn to cybersecurity and protecting user privacy. More and more people are connecting as a result of the social media platforms' rising popularity. People now require a more secure environment to connect due to the expanding opportunities for connectivity. In this paper, various aspects of cybersecurity based on DL models are discussed. These include understanding the concept of security & privacy, examining ML concepts, and helping to design & manage cybersecurity. Effective deep learning models like MLP, CNN, LSTP, and a hybrid model of CNN & LSTP have been examined to show that an understanding of cybersecurity in IoT networks has been achieved. Future studies prospects have also been suggested to aid in the learning process.

Julián Darío Miranda-Calle (2021) [12] Exploratory analysis firms are helpful in locating, managing and securing data from various cyber-attacks. It promotes the development of a plan for security measures that can aid to safeguard data, keep an eye on threats, and keep an eye on their organization's networks for any security breaches. The goal of this empirical tests is to demonstrate how data science is used to analyze data & provide a additional in-depth understanding of the most common cybersecurity attacks, the maximum and frequently used logical ports, discernible patterns, & attack trends.

Shiv Hari Tewari (2021) [13] Data science has become a key driver of cybersecurity's numerous recent technological & operational improvements. The extraction of system security event patterns or insights from existing data of security breaches & development of a data-driven model are necessary for automating and improving a security system. The study & analysis of actual events using a variety of scientific methodologies & ML algorithms, are known as data science. Here, the analyst briefly discusses data science, its evolution, & applications in cloud security. She also discusses how cybersecurity data science came to be, the advantages it offers, and the procedures involved, such as gathering data from pertinent cybersecurity sources & combining it with analytics to produce more effective security solutions. Compared to conventional cybersecurity computing, cybersecurity data science provides more intelligent, actionable computing. The researcher then discussed a variety of potential problems that could occur from the widespread application of CSDS, as well as how ML & DL might be applied to it and the many kinds of algorithms that might be employed. Because of this, the research also looks at how a system that depends on data-driven intelligent decision-making might safeguard our system from both known and unknown cyber threats in addition to looking at the history of data science and its present uses in cybersecurity.

Charu Virmani (2020) [14] Technology has completely taken over our lives with the exponential development in technical knowledge in recent decades, but with the deployment of computer-aided technological systems in many areas of our day-to-day lives, the possible risks & threats have also surfaced, aiming at the many security aspects including confidentiality, integrity, authentication, authorization, so on. Worldwide efforts have been made by computer scientists to develop solutions to these looming issues. Attackers have developed intricate attacks over time that are challenging to understand and even more challenging to defend against. Organizations handle enormous amounts of data every second, which gave rise to the idea of "Big Data," making their systems more capable and clever of defending against unheard-of threats in real time. In this article, a study on the use of ML algorithms in cyber security is presented.

Iqbal H. Sarker, A. S. M. Kayes (2020) [15] In the computing world, data science is the force behind the recent dynamic changes in cybersecurity science. The secret to making a security system automated & intelligent is to extract patterns or insights related to security incidents from cybersecurity data and construct appropriate data-driven models. As compare to conventional ones, recent technological trends in the filed of cybersecurity data science provides scientific and effective ways to the computing process to be made more intellectual & lawful. Then, we go over and briefly highlight several related research problems and future directions. We also offer a multi-layered framework for cybersecurity modeling that is based on ML. Our overall objective is to address cybersecurity data science & pertinent methodologies while also emphasizing how data-driven intelligent decision making can be used to defend systems from cyberattacks.

Iqbal H. Sarker (2020) [16] A intelligent intrusion detection system based on data-driven can be created using AI, especially MLtechniques. To implement this intrusion detection tree ("IntruDTree") is depend on machine-learning-based security model in this paper. This model ranks security features in terms of their importance before developing a tree-based generalized intrusion detection model based on the features that have been determined to be most crucial. This model reduces the computing density of the model by decreasing the feature measurements, making it effective in calculation of accuracy for test cases that have not yet been known. Finally, tests were run using cybersecurity datasets to test the performance of our IntruDTree model, and the precision, recall, fscore, accuracy, & ROC scores were calculated. To get result this model also relate the outcomes of the IntruDTree model with others like the naive Bayes classifier, logistic regression, SVM, & KNN algorithms.

Ouissem Ben Fredj (2020) [17] Available systems in the market are not sufficient to render the present detection systems of cyber-attacks as cyber-attacks are exponentially increasing. The present detection systems are inadequate & necessitating the development of more pertinent prediction models and methodologies. Latest attack prediction algorithms are not that much sufficient to predict attack due to number of attacks are exponentially increased and many types of attacks, now ethical hackers have this is unresolved challenge. Due to their unmatched high performance in numerous prediction-based disciplines like in deep learning techniques. There is an exponential increment in cyber-attacks. Here, we tried to study how to investigates the use of deep learning algorithms for anticipating cybersecurity physical attack in this setting. LSTM (Long Short-Term Memory), RNN (Recurrent Neural Network), & MLP (Multilayer Perceptron) algorithms can be predict the type of physical attack that may harm existing cyber system. A newly launched dataset named CTF is used to validate the proposed models, with positive results, especially for the LSTM model with an f-measure above 93%.

Muhammad Usman (2020) [18] Network-based systems experience fresh cyberattacks on a daily basis in our technology-based era. Traditional cybersecurity strategies rely on outdated threat-knowledge databases, which must be updated daily to combat the latest cyber threats & safeguard underlying network-based systems. Data created by sensitive real-time applications must be properly managed and processed in addition to being updated in threat knowledge databases. A valuable tool for managing and utilizing the created data to extract relevant information has arisen in recent years. These computing platforms are built on representation learning techniques. Strong cybersecurity solutions can be created to safeguard the underlying network-based systems and enable delicate real-time applications if these platforms are used effectively. Here, several cyber threats, real-world examples, associated with the various international organizations are used to process & evaluate the collected data, we examine various computing systems based on various representation and learning methods. We highlight a number of well-liked datasets that have been made available by reputable international organizations that can be utilized to train representation learning algorithms to anticipate and identify threats. In conclusion of this discussion numerous constraints, difficulties, and datasets that need to be taken into account when applying these efforts to the development of cybersecurity systems.

K. Rajeshkumar et al. (2020) [19] We must identify the threats that big data technology poses to the security measures in place today. Like earlier technological advancements, it is currently advancing. As a result of current technology advancements like cloud, network-connected smartphones, & omnipresent digital conversion of massive volumes of all forms of data, there are new potential risks to sensitive data. Big data demands a greater degree of accountability due to its increased susceptibility. Nearly 90% of the data produced during the past two years has been produced in that time. The hardest stage in any type of data processing is safeguarding confidential information from unauthorized access. A group of tactics & processes known as data leakage detection can be applied to effectively address the problem of considerable data leakage. Most of the data available today is unstructured. To obtain relevant information from huge data, we must develop better analytical techniques. We now have more security measures, but they are challenging to put into practice for significant volumes of data. We must protect both user information and essential information through the utilization of big data security methods. Sensitive information about the patient, various coding patterns, and a collection of traits must be protected using a machine learning technique. Systems that use machine learning have a wide range of library features to protect private client data. We provide the Secure Pattern-Based Data Sensitivity Framework to protect such sensitive data from huge data using machine learning (PBDSF). The proposed system uses HDFS to evaluate large amounts of data, find the most important information, or securely transform sensitive data.

Lidong Wang, Alexander (2019) [20] defines Big Data may be used to enhance & personalize patient care, increase provider-patient relationships, or reduce medical expenses. This article covers Healthcare data, big data in healthcare systems, and the applications & advantages of big data analytics in healthcare.

Andreas Holzinger (2018) [21] AI and Deep learning techniques outperform humans at some tasks. In addition to requiring large amounts of high-quality data, powerful computers, and engineering work, such approaches also have some drawbacks. They are fetching more opaque, and even if we comprehend the underlying mathematical concepts of such models, they still lack explicit declarative knowledge. For instance, words are translated into high-dimensional vectors that are incomprehensible to humans. Context-adaptive methods, or systems that create contextual explanatory models for groups of real-world occurrences, are what we will need in the future. The aim of explainable Artificial Intelligence, a discipline that is not new because the difficulty of explanation is older than AI itself.

Dongxia Zhang (2018) [22] The power system development trend that is receiving the most attention globally is the use of "smart grids." AI has moved into a new stage of growth, & AI 2.0 is advancing quickly. This page discusses the idea and current state of the three methods mentioned above, summarizes how they might be used in smart grids.

Giovanni Apruzzese (2018) [23] ML is used in many different fields, where it outperforms conventional rule-based methods. With the intention of assisting or possibly taking the place of the initial security analysts, these techniques are being included into cyber detection systems. Although complete automation of detection and analysis is a desirable objective, machine learning's effectiveness in cyber security must be carefully assessed. We give a study of machine learning methods used for spam, malware, & intrusion detection that is aimed at security professionals. The two objectives are to evaluate the solutions' current maturity & pinpoint the major drawbacks that prohibit the instant implementation of ML cyber detection techniques. Our findings are supported by a thorough literature analysis, experimentation on actual enterprise systems, or network traffic.

Dr.S.V.Achuta Rao (2017) [24] The field of ML has developed over the past few decades from an effort by a small group of computer enthusiasts to explore the idea of computers learning to play games and from a branch of mathematics (statistics) that rarely considered computational approaches to an independent research area that has not only established the foundation for mathematical principles of learning procedures but has also created a number of commonly utilised algorithms. This essay aims to compare the three most prominent machine learning algorithms based on certain fundamental ideas, as well as to explain the concept & evolution of machine learning. The performance of each technique in respect of training time, prediction time, and prediction accuracy has been documented & compared using the Sentiment140 dataset.

4. Design

4.1 Analyzing Big Data

Large data analytics is a field that examines and extracts information from large amounts of data in the business or data environment to develop appropriate conclusions. These insights might be used to forecast the company's future or predict it. This also helps with the formation of a historical pattern. The analysis of big data necessitates the use of skilled individuals with domain expertise in statistics and engineering. The data is massive, and analysis necessitates suitable determination and competence.

Large data analytics is a field that analyses and extracts information from large amounts of data in the business or data environment to develop appropriate conclusions. These insights might be used to foretell the company's future or predict what will happen next. This contributes to the formation of a historical pattern as well. It assists a company in comprehending the data they have and applying as a consequence, their firm will become more efficient, making more money and attracting more satisfied consumers.

Analytical professionals in the fields of big data analysis, predictive modeling, statistics, and more may use big data analytics tools to analyze the growing amount of structured as well as unstructured data. such a massive volume of data. Various data operations, including data mining, text mining, predictive analysis, forecasting, and so on, may be conducted using these tools; of It is important to note that these procedures are carried out in isolation and are part of high-performance analytics. With analytical tools and software for Big Data, a company may analyze a significant quantity of data to take future business decisions.

4.2 Analyzing Big Data Specific to Financial Data

Thanks to digitalization, technology like ML, AI, big data, and the cloud has revolutionized the financial institutions that how to compete in the market. These technologies are being used by large organizations to carry out digital transformations, fulfill the needs of clients, and enhance profit and loss. Even though most companies keep track of latest trends and important information, they aren't always sure how best to use it since the information is unstructured or undocumented internally. (Ruchi Verma 2020)

Businesses must adapt to the fast changes in the banking sector in a planned and complete way. Unstructured and high-volume data may be tapped for competitive advantages, new market opportunities, and more if financial businesses have the right technology in place to fulfill the difficult analytical needs of digital transformation.

4.3 Big data has revolutionized finance

Financial firms like Banks need of improvements those are allowing for easy, customized, and secure business solutions. This will help in big data analytics has been able to transform not just individual firm operations as well as the whole financial services sector.

A. Real time Stock market updates.

There are revolutionary changes in Trade and investment due to machine learning. Instead of only focusing on stock prices, big data may now be used to examine political and socioeconomic affects. Stock prices might be affected by this. With the help of machine learning, analysts can quickly gather and analyze relevant data, helping them to make well-informed judgments.

B. Fraud detection and prevention

In the fight against fraud, machine learning, which relies on vast amounts of data, is an invaluable tool. Credit card security risks have been reduced thanks to analytics that track buying habits. To protect their customers, banks may now quickly block their credit cards and transactions if they suspect that their credit card information has been compromised.

C. Accurate Risk Assessment

Major financial decisions, such as investments and loans, are being made using machine learning. The economy, client segmentation, and firm capital are all factors that may benefit from predictive analytics to spot potential risks such as erroneous investments or payments.

5. CONCLUSION

Machine learning through a cognitive approach can train the machine and such a trained machine can help to find the vulnerabilities in the system like various threats & possibilities of attacks.

The prognostic approach helps in any engineering field for predictive analysis. Here for the first time, we are implementing the concept of the prognostic approach in computer science for predictive analysis to find vulnerabilities in big data specific to financial data to predict possibilities of frauds.

REFERENCES

[1]. Bowe, A. K., Lightbody, G., Staines, A., & Murray, D. M., Big data, machine learning, and population health: predicting cognitive outcomes in childhood. *Pediatric Research*, 1-8, 2022.

[2]. Khan, A., & Zubair, S., An improved multi-modal based machine learning approach for the prognosis of Alzheimer's disease. Journal of King Saud University-Computer and Information Sciences, 34(6), 2688-2706, 2022.

[3]. Allegra, A., Tonacci, A., Sciaccotta, R., Genovese, S., Musolino, C., Pioggia, G., & Gangemi, S., Machine learning and deep learning applications in multiple myeloma diagnosis, prognosis, and treatment selection. Cancers, 14(3), 606, 2022.

[4]. Jhaveri, R. H., Revathi, A., Ramana, K., Raut, R., & Dhanaraj, R. K., "A review on machine learning strategies for real-world engineering applications", Mobile Information Systems, 2022.

[5]. Yu Xue, Nan Wei, Junyang Han, Chishe Wang, Moayad Aloqaily, "Design and Implementation of Enterprise Recruitment Mini Program", Journal of Cyber Security, Vol.3, No.3, pp. 125-132, 2021, DOI:10.32604/jcs.2021.016647, 2021

[6]. Jean Piagets, “Theory Of Cognitive Development”, Paul Main, June 11, 2021

[7]. Yeshi Nidup, “Awareness about the Online Security Threat and Ways to Secure the Youths”, Journal of Cyber Security, Vol.3, No.3, pp. 133-148, DOI:10.32604/jcs.2021.024136, 2021.

[8]. Prabhanjan Chaudhari, and Dr. Amit Bhusari. "Transfer Optimistic Outcome-based Learning for Mature Behavior of Machine in Deep Learning." 2020 IEEE International Conference on Advent Trends in Multidisciplinary Research and Innovation (ICATMRI). IEEE, 2020.

[9]. Nguyen, T. T., & Reddi, V. J., “Deep reinforcement learning for cyber security”, IEEE Transactions on Neural Networks and Learning Systems, 2021.

[10]. Ahmed, K. D., & Askar, S., “Deep learning models for cyber security in IoT networks: A review”, International Journal of Science and Business, 5(3), 61-70, 2021.

[11]. Miranda-Calle, J. D., Reddy C, V., Dhawan, P., & Churi, P., “Exploratory data analysis for cybersecurity”, World Journal of Engineering, 18(5), 734-749, 2021.

[12]. Shiv Hari Tewari on “Necessity of data science for enhanced Cybersecurity”, International Journal of Data Science and Big Data Analytics, Volume 1, Issue 1, ISSN: 2710-2599, 2021.

[13]. Virmani, C., Choudhary, T., Pillai, A., & Rani, M., “Applications of machine learning in cyber security”, In Handbook of research on machine and deep learning applications for cyber security (pp. 83-103). IGI Global, 2020.

[14]. Sarker, I. H., Kayes, A. S. M., Badsha, S., Alqahtani, H., Watters, P., & Ng, A., “Cybersecurity data science: an overview from machine learning perspective”, Journal of Big data, 7, 1-29, 2020.

[15]. Sarker, I. H., Abushark, Y. B., Alsolami, F., & Khan, A. I., “Intrudtree: a machine learning based cyber security intrusion detection model”, Symmetry, 12(5), 754, 2020.

[16]. Ben Fredj, O., Mihoub, A., Krichen, M., Cheikhrouhou, O., & Derhab, A., “Cyber Security attack prediction: a deep learning approach”, In 13th International Conference on Security of Information and Networks (pp. 1-6), 2020.

[17]. Usman, M., Jan, M. A., He, X., & Chen, J., “A survey on representation learning efforts in cybersecurity domain”, ACM Computing Surveys (CSUR), 52(6), 1-28, 2020.

[18]. K. Rajeshkumar, S. Dhanasekaran, V. Vasudevan, on “Exploration of Big Data Security Framework using Machine Learning”, International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249 – 8958 (Online), Volume-9 Issue-5, 2020.

[19]. Lidong Wang, Cheryl Ann Alexander, “Big Data Analytics in Healthcare Systems”, International Journal of Mathematical, Engineering and Management Sciences, Vol. 4, No. 1, 17–26, ISSN: 2455-7749, 2019.

[20]. Holzinger, A., “From machine learning to explainable AI”. World symposium on digital intelligence for systems and machines (DISA) (pp. 55-66). IEEE, 2018.

[21]. Zhang, D., Han, X., & Deng, C., “Review on the research and practice of deep learning and reinforcement learning in smart grids”, CSEE Journal of Power and Energy Systems, 4(3), 362-370, 2018.

[22]. Apruzzese, G., Colajanni, M., Ferretti, L., Guido, A., & Marchetti, M., “On the effectiveness of machine and deep learning for cyber security”, 10th international conference on cyber Conflict (CyCon) (pp. 371-390). IEEE., 2018.

[23]. Dr.S.V.Achuta Rao, “A survey on machine learning: concept, algorithms and applications”, International Journal of Innovative Research in Computer and Communication Engineering, 5(2), 1301-1309. 2017.

[24]. Rashmi N, Uma KM, Jayalakshmi K, Vinodkumar K P, “Big Data Security Challenges: Dealing with too many issues”; International Journal of Recent Development in Engineering and Technology, Volume 3, Issue 2, August 2014.

[25]. Pecht, Michael G., “Prognostics and Health Management of Electronics”, Wiley. ISBN 978-0-470-27802-4, 2008.