

Modified Deep Learning Approach Based on CPS to Secure the Data for IT Organization

Naveen Vemulapalli¹, Amith Arigapudi², Rohith Vallabhaneni³
^{1,2,3}Infylogy Corp

ABSTRACT: In the recent days, data's plays a significant role and its enormous quantity are extracted in IT Organization. Various devices and services generates huge data and it is processed to extract the knowledge data through data analytics. To process those data, which are extracted and needs Deep Learning (DL) model as it needs large number of input and its attributes. The deep learning models helps in classifying the data and process those data generated from IoT and those useful data will be useful to those consumers. As huge data are circulated and it contains some private information's to be secure from the third parties along with effective data processing. The modified deep learning approach based on Cyber Physical Systems (CPS) is propose to effectively process those IoT based data with enhanced data security. Based on deep learning model, the data access algorithm helps to process the raw data generated from IoT time to time and it helps to convert the data to knowledge data. Then the data are secured based on policy access control against various attacks like DoS and DDoS. Based on the performance analysis, the approach helps to provide effective classification of data and reliable data's to be maintained. Then security on IoT data to be verified against DoS and DDoS attacks on cyber physical system on real time applications.

KEYWORDS: Internet of Things, Deep Learning, Data Processing, Attacks and Security

I. INTRODUCTION

In the recent period, network-related activities have risen rapidly on a daily basis, and it has a large amount of confidential information. In this instance, the network is confronted with security-related problems as a result of intruders, or unauthorised users, who are involved in malicious and unwelcome actions. The main goal of intrusion attempts is to obtain sensitive information [1]. They are overcoming many attacks and unlawful activities that happen on the physical systems through cyber-attacks based on the various security-based network ways, and they continue to exist on it even after the methods are applied. Therefore, an efficient intrusion system based on a method linked to network intrusion is required to prevent those attacks. Numerous intrusion detection systems, such as Denial of Service (DoS) [2], Cross-Site Scripting (CSS) [3], and others, are employed to counteract the ongoing attacks.

In this case, the network will offer different services so that cyber-physical systems can connect with cyber-physical systems that use some real processors. Numerous Internet of Things (IoT)-based sensors are added, causing actuators and other components to be added, integrated with one another, and form a network structure [4]. The aforementioned will assist in the formation of a network so that they can interact through network-oriented connections such TCP/IP as well as numerous protocols linked to wireless networks. When this happens, the network's simple structure will make the system associated to cyber-physical susceptible to numerous

intrusions. Then the cyber-based systems will make the system safe and secure, even if system failure will happen and trigger based on attacks and the intruder will do more harm on the system even based on the network being properly accomplished and as people will depend on such systems. As a result, the aforementioned scenario will enable attacks to be prevented based on cyber-physical systems with efficient network intrusion systems [5]. The digital physical systems are then triggered to generate the resource and process the physical systems, allowing it to be controlled physically through the process of network computing and network communication via network device connections.

These network devices will make it possible to access resources remotely, have the ability to manage different network equipment and systems, and make it more crucial and vital in a variety of industrial settings [6]. When users rely more on cyber systems, and the systems are executed on the implementation of various security threats, it may result in severe damage to the objects of physical nature system, as well as a direct reflection on the users who rely on those cyber systems. Finally, an efficient network intrusion detection system will offer activity execution so as to prevent behaviours that could harm the system due to network attacks. The Internet of Things (IoT) is regarded as a crucial medium for enabling communication between IoT devices and the rest of the world. Through Internet of Things (IoT) devices, we are able to accomplish our daily tasks more easily. , smart E-bikes, autonomous vehicles, smart temperature sensors, and

other IoT gadgets with applications in many other sectors are just a few examples. Data generation from IoT devices and their supporting platforms will increase significantly as a result of these fantastic outcomes. These data must be sent over and stored in back-end storage facilities, including cloud infrastructure. To calculate massive volumes of data in addition to the typical classical big data analysis, several specialised computing talents are required. IoT hardware produces continuous streaming data in a raw format. These data cannot be stored or processed in their raw form [7]. To obtain the fullest, most meaningful information and knowledge from the raw data, these data should be processed and cleaned. Next, handle the heterogeneity of data from different IoT devices, depending on the application domain (e-health, smart automobiles, smart electrical grids, smart home management, etc.). According to a poll from a McKinsey report, IoT impact on commercial areas will increase over the next several years. The financial impact of IoT-based services will propel the development of numerous industries to a new level. In 2025, the annual economic impact of IoT will range between \$2.7 and \$6.2 trillion. Table 1 below displays how the Internet of Things will affect many industries in 2025 [8].

Many of the innovations and technology that have profoundly impacted the development of man, his relationships with others, and the environment have their roots in the emergence of the internet. The capacity to wirelessly connect computer systems throughout the globe made communication, cooperation, and data access simple, serving as a tool for innovation and creativity. The International Telecommunication Union (ITU) links broadband internet penetration to employment and economic growth [9]. This is so because the internet is the enabling technology for the digital revolution, which gave rise to paradigms and platforms such as online banking, online education, online shopping, e-health, and e-government. Beyond the internet, this digital revolution is being propelled by developments in wireless communications, the spread of high-capacity mobile devices, the relatively reduced cost of computing hardware, alternative energy sources, and access to larger memories. Parallel/distributed computing, cloud computing, quantum computing, nanotechnology, microelectronics, and optoelectronics have all made significant contributions to these advancements. Internet connection was initially only available through computers, and then subsequently through smartphones. However, advancements in wireless sensors have made it possible to incorporate tiny and high-capacity sensors into hardware devices used for commonplace tasks, thereby establishing connectivity to the internet. This innovation known as the Internet-of-Things (IoT) has made it possible to expand the internet from a global network of

computers to a global network of computers and things, giving rise to terms such as Internet-of-battlefield-things (IoBT), Internet-of-vehicles (IoV), and Industrial Internet-of-Things (IIoT). IoT and wireless sensor networks (WSN) have pushed back the boundaries of research in industries like industry, transportation, healthcare, home automation, military warfare, entertainment, and security [2]–[8].

The proposed work is summarized as follows,

- The modified deep learning approach based on Cyber-Physical Systems (CPS) is proposed to efficiently analyse the IoT-based data with improved data security.
- Based on the deep learning model, the data access algorithm processes raw IoT information and converts it to knowledge data.
- Following that, the data are protected using policy-based access control against different assaults including DoS and DDoS.
- Performance analysis is a method that enables trustworthy data to be kept and successful data classification. The security of IoT data must therefore be checked against DoS and DDoS assaults on cyberphysical systems running in-the-moment applications.

II. LITERATURE SURVEY

CPS use the internet and WSN as intelligent systems that automate tasks that were once heavily reliant on human efforts. Although they have been variously defined by many writers [9], in essence, these refer to physical and engineering systems where a computational and communication core monitors, coordinates, controls, and integrates the operations. CPS offer a control action to the IoT's computational and networking characteristics. The level of automation in CPS exceeds that of connected devices due to the ability to use control feedback to tell a mechanism to take act according to physical parameters collected by sensors. Because of their enormous potential to revolutionise practically every sphere of human endeavour and solve real difficulties in our world [10], CPS has piqued the curiosity of industry, government, and academia. Whenever the Internet and CPS are used in vital infrastructure such as agriculture, health, military, transportation, home automation, and power systems, a tremendous volume of data is generated. This is due to the fact that the gadgets often have constant power and real-time connections. The notion of big data analytics (BDA) has thus been useful to IoT and CPS systems since it allows information to be accessed from data for making decisions such as fault prediction, diagnostics, and predictive maintenance [11].

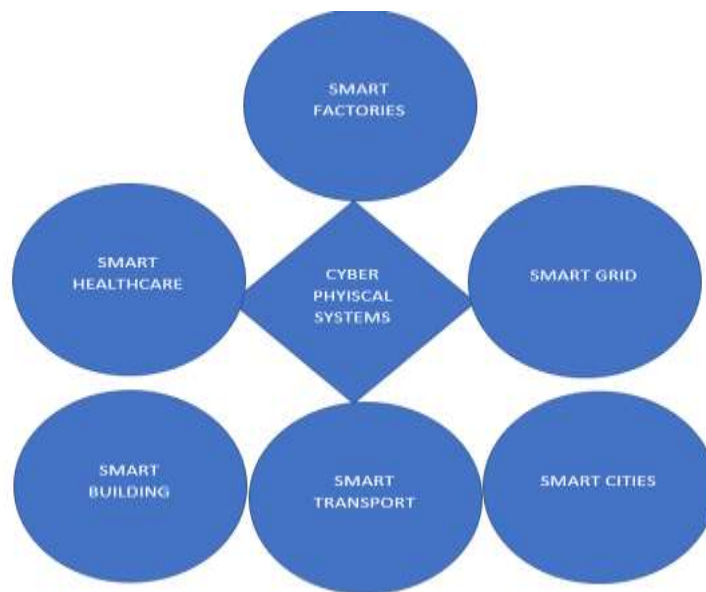


Fig 1. CPS Applications

The application of statistical techniques to identify hidden patterns, correlations, and insights from vast amounts of data to derive value from the created data is growing increasingly common as they add new functionality to the systems outlined above. CPS and other networked controlled systems are currently responsible for a vast volume of data produced around the world. Various types of applications that demonstrate the concept of CPS are depicted in Image 1. Despite all the buzz surrounding CPS, worries about safety and security prevent their practical adoption to address real-world problems [10]. The CPS has incredibly strict requirements, including the necessity to operate in real-time and sensitivity to network issues like latency. Furthermore, the impact of a failure on human life and infrastructure is greater when compared to typical information technology systems. The study of cybersecurity has taken centre stage in computer science and information technology. Security was initially ensured through the use of intrusion detection systems (IDS), firewalls, and techniques such as malware, adware, spyware, and ransomware to target information. Increased interconnectivity of sensors, actuators, and controls within CPS contributes significantly to the rise of cyber attacks, as it enlarges the attack surface and renders these systems susceptible to hostile actions. Recently, ML and artificial intelligence (AI) algorithms have been employed to increase the effectiveness of numerous systems [11]–[13]. The references to data and artificial intelligence as "new oil" and "new electricity" respectively [14] emphasise their impact on the contemporary world. In addition, cutting-edge technologies like IoT, CPS, BDA, and AI are driving the fourth industrial revolution [15], [16]. These advancements have increased research interest in the use of big data and data science techniques to protect systems from adversarial attacks. The application of AI and ML for cybersecurity began with their inclusion in intrusion detection systems (IDS).

In this area of study, malware and anomaly detection in information and communication systems were investigated. ML was applied to achieve cybersecurity in IoT systems after the success was noted [17]–[21]. Additionally, the development of deep learning (DL) and reinforcement learning (RL) has greatly aided the use of ML algorithms to tackle real-world problems that have proven difficult for shallow algorithms as well as the more well-known supervised and unsupervised algorithms to solve. Factors that support the use of DL in CPS include the high dimensionality of the data produced and the fact that the amount of data keeps growing [12]. Recently, researchers have coupled DL and RL to create deep reinforcement learning (DRL); this breakthrough has caused a significant revolution in CPS research and continues to show a huge promise to offer solutions to current and impending difficulties. Because of the requirement to constantly make decisions like lane changes and responding to traffic signs autonomously through image and pattern recognition, this revolution is prominent in vehicular CPS such as autonomous vehicles. [22]–[26]. However, there are growing worries regarding the usage of AI in cybersecurity. Recent studies have demonstrated that systems that rely on ML algorithms for security are also vulnerable to various adversarial attacks. Data-dependent machine learning algorithms use data generated by numerous sensors in networked systems to draw conclusions or make predictions. A common type of cyber attack in CPS and other systems is the creation of ways to tamper with data or input. As a result, the model is compelled to generate inaccurate results. Neural networks, especially deep neural networks (DNN) used to protect CPS systems, are especially prone to this. Furthermore, the idea of repurposing tactics designed to defend a system to attack it has been a subject of concern. The employment of AI and ML algorithms to protect systems can be utilised by adversaries to target those systems and launch adversarial attacks. As a

result of their increased sophistication, speed, and relative affordability as a result of utilising the defence systems' attempts to make them stronger and harder to stop or detect, it has recently been shown that such attacks have significant potential.

The danger of undermining machine learning algorithms used to improve cybersecurity in networked systems is a challenge that researchers must address. In other words, it must be acknowledged that machine learning algorithms cannot completely stop machines from accessing systems that need to be safeguarded. As may be shown from [27], [28], questions about the security of ML have drawn attention for more than ten years. Though numerous studies have concentrated on spam email filtering and IDS, few have delved into what the systems must do in the face of adversarial attacks [29]. The goal to ensure that ML-enhanced systems continue to provide the services they should in the face of hostile attack gave rise to the concept of resilient ML [30]. To improve the actual implementation of ML-based cyber security, notably in CPS, this study area needs to be thoroughly examined.

III. PROBLEM FORMULATION

The framework development for processing the data for each level of IoT device to assure security is the subject of this study. As data processing and data access are linked as one framework, proper data classification, data dependability, and efficacy on scalable data as IoT solutions are ensured. Here, a feature extraction approach is used, which helps to reduce the number of datasets needed to apply feature extraction and perform effectively. Sets of features are created to identify malware based on mistakes in improperly classified data in order to improve data accuracy. The machine learning algorithm is then used to calculate the improved computational speed and increased detecting speed.

IV. PROPOSED METHODOLOGY

The cloud audit and authenticator approach is propose based on certain file system, which makes the malicious user to get authenticate the data auditing verification as the existing

algorithm has low security based on entropy. The propose a data auditing approach integrates with file management and authenticator of data deduplication. The propose approach performs the authenticator process and new form of file tag, which helps guarantee the effective security based on the random generation of message key

Initially the encryption process is performed on a file, which gets identified based on the cloud.

Algorithm 1: Encryption Algorithm

Input: Data 'd' & File 'F_d'

Output: File Tag and Authenticator

The encryption process gets initiated on the file.

Identify the File 'F_d' on cloud.

If (File = exists in database)

{

File Found in the database

Else

File not found in the database

Initiate key generation process

To generate the key tag and authenticator.

Move the encrypted file into the cloud.

Move the file tag and authenticator into the cloud

}

Algorithm 2: User Process in cloud

Initially the file identity is sent to the cloud.

To check the malicious behaviour of the user to access the cloud.

First, the users → the ownership of the file access

Proof work is checked and sends to the user.

If (Proof of Work = Valid)

{

User can execute the file uploading operation

File gets tag in the cloud

Else

Users' operation to access the file gets refused

}

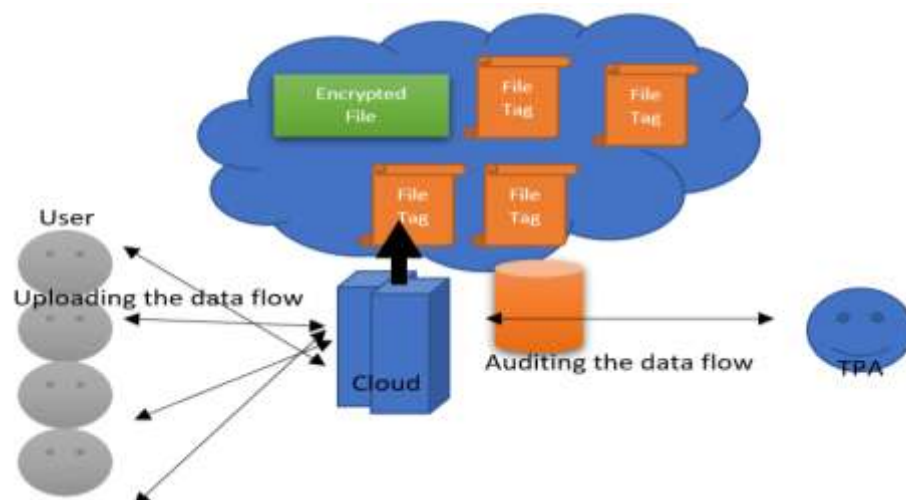


Fig. 2. Proposed Approach

User choose signing key to get verified and here secure digital signature algorithm is applied. The user is randomly selected based on the secret key and compute the public key. The user makes secret and publish the system information.

Algorithm 3: Key and Authenticator Generation

Input: User, Third Party Authenticator

Output: Update the key Generation

Initially,

User choose and compute the secret tag key

To compute the authenticator by the user

User choose and compute the public tag key

To compute the re-key auditing by the user

User to compute the re-key process auditing → Third Party Authenticator

User choose and fix the file tag → Public Information

Third Party Authenticator → File Tag (Proof Verification Algorithm)

File Tag → Public Tag key

Subsequent user → File Tag Recompilation

Compute the authenticator to check the data integrity in cloud

User to upload the data into the cloud

User Processing Algorithm is applied by the user to compute the requested upload and it is sent to the cloud.

If (File = Cloud)

{

 User = User Processing Algorithm

}

When the user does not need upload the duplicate file as the cloud checks for that particular file to access. The cloud selects randomly by sending the PoW to cloud.

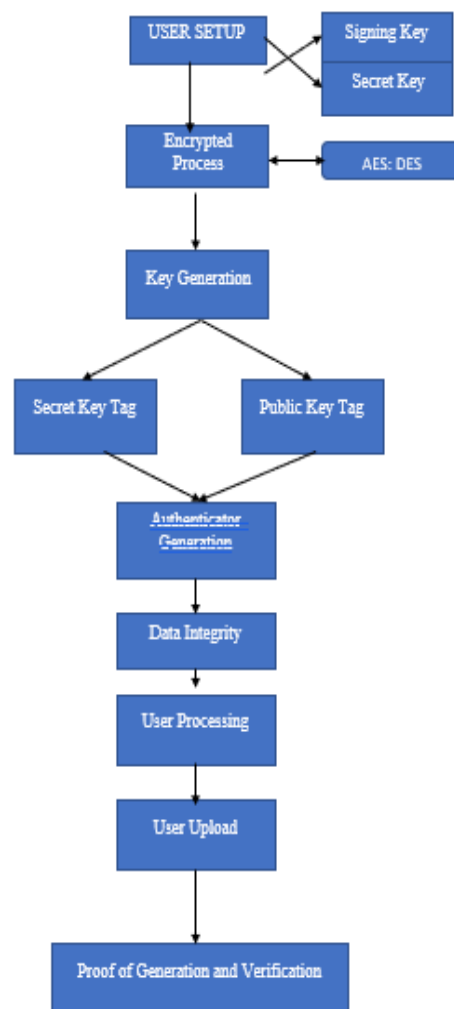


Fig. 3. Flow Chart for the Propose Work

V. PERFORMANCE ANALYSIS

In the performance analysis, FL based Secure Approach is proposed to secure the data communication and perform data pre-processing with training process based on DL/ML

algorithm. In Table 1, the data loss is calculated based on the process of training phase for the proposed system and it performs a gradual increase in the value based on the variation in the iteration from 0 to 500.

Table 1. Number of Iteration Vs. Data Loss

S.No.	Number of Iteration	Data Loss (Training Process)
1	0	0.00075
2	100	0.00096
3	200	0.00156
4	300	0.00321
5	400	0.00386
6	500	0.01056

In the Table 2, the data accuracy is determined based on the training process by varying the number of iteration from 0 to 500. The data accuracy of the proposed FLSA gets decreases gradually and the average value of 92.5. In Table 3, the data

accuracy is determined based on the training process by varying the number of iterations 0 to 500 based on the testing process. The data accuracy of the proposed FLSA gets decreases gradually and the average value of 0.925.

Table 2. Number of Iteration Vs. Data Accuracy (Training Process)

S.No.	Number of Iteration	Data Accuracy (Training Process)
1	0	95.5
2	100	91.8
3	200	89.8
4	300	93.6
5	400	90.6
6	500	90.1

Table 3. Number of Iteration Vs. Data Accuracy (Testing Process)

S.No.	Number of Iteration	Data Accuracy (Testing Process)
1	0	0.925
2	100	0.922
3	200	0.924
4	300	0.892
5	400	0.886
6	500	0.869

VI. CONCLUSION

To create a framework for data processing in each level of IoT devices and to offer data security. Data processing can take place at both the fog/edge and cloud levels of computing. Multiple Access Control mechanisms will be implemented at each level to ensure the security of the data. Data processing and data access control will be merged in this framework to deliver quality data classification, dependable, efficient, and scalable IoT security solutions. The feature selection approach aids in reducing the number of featured datasets and improves performance. Then, a featured set was able to more accurately identify malware with data misclassification-based mistakes. After that, use a machine learning technique to improve the detection and computation rates.

REFERENCES

1. Robert Mitchell and Ing-Ray Chen, “A survey of intrusion detection techniques for cyber-physical systems,” ACM Computing Surveys, Vol. 6, Issue. 4, pp. 1-26, 2014.
2. Li Zhao and Wei Li, “Co-Design of Dual Security Control and Communication for Nonlinear CPS Under DoS Attack,” IEEE Access, Vol. 8, pp. 19271 – 19285, 2020.
3. Hui Lin, Jia Hu, Jianfeng Ma, Li Xu, Zhengxin Yu, “A Secure Collaborative Spectrum Sensing Strategy in Cyber-Physical Systems,” IEEE Access, Vol. 5, pp. 27679 – 27690, 2017.
4. Haris Isakovic, Denise Ratasich, Christian Hirsch, Michael Platzer, Bernhard Wally, Thomas Rausch, Dejan Nickovic, Willibald Krenn, Gerti Kappel, Schahram Dustdar and Radu Grosu,” International Workshop on Design, Modeling, and Evaluation of Cyber Physical Systems Workshop on Embedded Systems and Cyber-Physical Systems Education CyPhy 2018, WESE 2018: Cyber Physical Systems. Model-Based Design, pp. 206-213, 2018.
5. Lei Wang; Zhaoyang Qu; Yang Li; Kewei Hu; Jian Sun; Kai Xue; Mingshi Cu, “Method for Extracting Patterns of Coordinated Network Attacks on Electric

- Power CPS Based on Temporal–Topological Correlation,” IEEE Access, Vol. 8, pp. 57260 – 57272, 2020.
6. Fei Pan; Zhibo Pang; Hong Wen; Michele Luvisotto; Ming Xiao; Run-Fa Liao; Jie Chen, “Threshold-Free Physical Layer Authentication Based on Machine Learning for Industrial Wireless CPS,” IEEE Transactions on Industrial Informatics, Vol. 15, Issue. 12, pp. 6481 – 6491, 2019.
7. Bo Cheng; Jingyi Zhang; Gerhard P. Hancke; Stamatis Karnouskos; Armando Walter Colombo, “Industrial Cyberphysical Systems: Realizing Cloud-Based Big Data Infrastructures,” IEEE Industrial Electronics Magazine, Vol. 12, Issue. 1, pp. 25 – 35, 2018.
8. Ling Li, “China's manufacturing locus in 2025: With a comparison of “Made-in-China 2025” and “Industry 4.0”, Technological Forecasting and Social Change, Elsevier, Vol. 135, pp. 66 – 74, 2018.
9. Antoine Bagula, Olasupo Ajayi and Hloniphani Maluleke, “Cyber Physical Systems Dependability Using CPS-IOT Monitoring,” Sensors, Vol. 21, Issue. 8, 2021.
10. Emmanuel Oyekanlu, “Fault-Tolerant Real-Time Collaborative Network Edge Analytics for Industrial IoT and Cyber Physical Systems with Communication Network Diversity,” IEEE 4th International Conference on Collaboration and Internet Computing (CIC), 2018.
11. Lukas Gressl, Michael Krisper, Christian Steger and Ulrich Neffe, “Towards an Automated Exploration of Secure IoT/CPS Design-Variants,” International Conference on Computer Safety, Reliability, and Security SAFECOMP 2020: Computer Safety, Reliability, and Security, pp. 372 – 386, 2020.
12. Mengru Tu , Ming K. Lim , Ming-Fang Yang, “IoT-based production logistics and supply chain system – Part 2: IoT-based cyber-physical system: a framework and evaluation,” Industrial Management & Data Systems, Vol. 118, Issue. 1, 2018.
13. Aslahi-Shahri BM, et al. A hybrid method consisting of GA and SVM for intrusion detection system. Neural Comput Appl 2016.
14. Sun S, Ye Z, Yan L, Su J, Wang R. Wrapper feature selection based on lightning attachment procedure optimization and support vector machine for intrusion detection. In: Proceedings of the IEEE 4th international symposium on wireless systems within the international conferences on intelligent data acquisition and advanced computing systems, IDAACS-SWS; 2018. p. 41–6.
15. Salman Raza , Shangguang Wang et al., A Survey on Vehicular Edge Computing: Architecture, Applications, Technical Issues, and Future Directions, Hindawi Wireless Communications and Mobile Computing Volume 2019.
16. H.V. Jagadish, B.C. Ooi, K.-L. Tan, C. Yu, R. Zhang, Idistance: an adaptive b⁺-tree based indexing method for nearest neighbor search, ACM Trans. Database Syst. (TODS) 30 (2) (2005) 364–397.
17. Issam El-Naqa, Yongyi Yang et al., “A Support Vector Machine Approach for Detection of Microcalcifications “IEEE TRANSACTIONS ON MEDICAL IMAGING, VOL. 21, NO. 12, DECEMBER 2002
18. Yunus Kologlu, Hasan Birinci et al., “ A Multiple Linear Regression Approach For Estimating the Market Value of Football Players in Forward Position”, 2018
19. Benjamin A Goldstein et al., “An Application of Random Forests to a genome-wide association dataset: Methodological considerations & new findings”,2010
20. Liangzhi Li, Kaoru Ota et.al,“When Weather Matters: IoT-Based Electrical Load Forecasting for Smart Grid” ,2017
21. S. Yi, Z. Hao, Z. Qin, and Q. Li, “Fog computing: Platform and applications,” in Proc. 3rd IEEE Workshop Hot Topics Web Syst. Technol. (HotWeb), Washington, DC, USA, 2015, pp. 73–78.
22. H.V. Jagadish, B.C. Ooi, K.-L. Tan, C. Yu, R. Zhang, Idistance: an adaptive b⁺-tree based indexing method for nearest neighbor search, ACM Trans. Database Syst. (TODS) 30 (2) (2005) 364–397
23. “Deep Learning for IoT Big Data and Streaming Analytics- A survey”, 2018, Mehdi Mohammadi et.al,
24. K. Kambatla, “Trends in big data analytics,” J. Parallel Distrib. Comput., vol. 74, no. 7, pp. 2561_2573, 2014
25. L. Wang, R. Ranjan, Processing distributed internet of things data in clouds, IEEE Cloud Comput. 2 (1) (2015) 76–80.
26. A. McCallum, K. Nigam, et al., A comparison of event models for naive bayes textclassification, in: AAAI-98 Workshop on Learning for Text Categorization, vol. 752, Citeseerwhw, 1998, pp. 41–48.
27. H. Zhang, The optimality of naive bayes, AA, 1 (2) (2004) 3.
28. H. Albreiki, L. Alqassem, K. Salah, M. H. Rehman, D. Svetinovic “Decentralized Access Control for IoT Data Using Blockchain and Trusted Oracles “ IEEE International Conference on Industrial Internet, 2019, Orlando, FL, USA..
29. Oscar Novo “Blockchain Meets IoT: an Architecture for Scalable Access Management in IoT”, IEEE Internet of Things Journal (Volume: 5 , Issue: 2 , April 2018, pp:1184-1195).
30. Y. Andaloussi, M. D. E Ouadghiri, Y. Maurel, J. M. Bonnin, H. Chaoui “Access control in IoT environments: Feasible scenarios”, Elsevier B.V., Procedia Computer Science 130 (2018), pp-1031–1036.