

"A SECURE FEDERATED INTRUSION DETECTION MODEL WITH BLOCKCHAIN AND DEEP BIDIRECTIONAL (LSTM)"

Khaja Shaik Nazar Mohammed, Md Umair Mudassir, Syed Ali Nawaz Uddin, Dr. Md Zainlabuddin ^{1,2,3}B. E Student, Department of CSE, ISL College of Engineering, India.

⁴Associate Professor, Department of CSE, ISL College of Engineering, Hyderabad, India.

Abstract: Several specialists have suggested using Machine Learning (ML) methods to create Intrusion Detection Systems (IDMs) in order to enhance the security of computing and network resources. Machine Learning replicates human cognitive skills by modeling deduction, inference, extrapolation, and synthesis. The use of this technology might be employed to construct Intrusion Detection Systems (IDMs), thereby facilitating accurate detection of harmful network traffic, resulting in a reduction in false positive alarms. Support Vector Machines (SVM), Decision Trees (DT), Bayesian Networks, and Naïve Bayes are often used machine learning techniques for creating Intrusion Detection Systems (IDMs). Nevertheless, the aforementioned conventional machine learning (ML) techniques prioritize the labor-intensive process of designing features and are less effective when confronted with vast amounts of data that need categorization in a real-world setting. As the amount of the dataset increases, the accuracy of many categorization jobs decreases. To address the difficulties in analyzing large amounts of data that classical machine learning (ML) algorithms face, deep learning (DL) methods are used to enhance the efficiency of intelligent data mining (IDM), particularly in situations involving many classes.

An intrusion refers to any illegal action that results in damage to a computer system or network, with the potential to jeopardize the confidentiality, integrity, or availability of information. Intrusion detection models are often integrated into software programs to monitor networks or computers for malicious activity in order to maintain system security [10]. As a proactive solution, IDM detects and halts possible risks to a system or network before they may do harm, including both hostile insiders and external hackers.

INTRODUCTION

Cloud computing has arisen as a solution to the need for computation services that can be accessed and used in a similar way to utilities such as electricity and water. Cloud computing enables the delivery of IT services to remote places over the Internet. In this method, individuals make use of the computing capabilities offered by Cloud Service Providers (CSP) and are charged based on their actual use, rather than paying a fixed amount in advance or entering into long-term agreements. By purchasing and using computational power on a pay-per-use basis, the fees associated with acquiring capital goods may be transformed into ongoing operating costs. Shared resources in the realm of cloud computing include processing and storage resources, software applications, operating systems, and network infrastructures. Businesses and consumers may access cloud services via many means such as social media, email, web application hosting, and more. Cloud storage systems such as AWS, Microsoft Azure Blobs, and Google Cloud Storage are extensively used by many cloud vendors and may be



accessed from several devices. More than \$1 trillion has already been allocated to cloud-based computing, either via active investment or passive funding. The advantages of cloud computing technology include scalability, allocation of resources according to demand, reduced administrative burden, a flexible pricing model (pay-asyou-go), and simplified application development and deployment. Establishing large data centers entails exorbitant initial costs. Smaller organizations may try to achieve a similar scale effect by forming federations of computing and storage services. Cloud federation supports three primary interoperability elements: resource migration, resource redundancy, and the aggregation of complementary resources or services. It involves aggregating services from several providers into a single pool [6]. The main cloud service types are Infrastructure-as-a-Service (IaaS), Platform-as-a-Service (PaaS), and Software-as-a-Service (SaaS). IaaS offers virtualized resources for computing, storage, and networking. Platform-as-a-Service (PaaS) offers a framework for creating, developing, launching, and overseeing software applications. On the other hand, Software-as-a-Service (SaaS) allows end users to access programs without the need for physical installation on their machines. The numbers 2, 4, 7, and 8 are listed. With the progression of technology, there has been a noticeable increase in the frequency of hacking occurrences. The proliferation of security vulnerabilities has emerged as a significant obstacle in our intricate technological landscape. Cloud providers and consumers often report instances of attacks [9]. Identifying cyber risks in cloud computing is a costly and time-consuming process due to their growing complexity. Active Intrusion Detection Systems or Models (IDS/M) are crucial for preventing and securing Information Technology (IT) cloud infrastructure and operations [2].

Recently, blockchain technology has become prevalent in all aspects of Information and Communication Technology (ICT), and its use has grown. The rise in the value of cryptocurrencies and substantial venture capital investments in blockchain start-up firms have generated interest in the advancement of this technology [16], [17]. Blockchain primarily functions as a ledger, recording and sharing all transactions conducted by participating peers with all participants. As additional blocks are added over time, the size of this blockchain continues to grow. Commonly used cryptocurrency blockchains are accessible to the public, and transactions may be examined using online platforms like blockchain.com, enabling anybody to inquire about the blockchain transactions. Blockchain allows participants to engage in commercial transactions without the need for a trusted intermediary [17], [18].

LITERATURE REVIEW

Talaei Khoei and Kaabouch [11] devised an innovative approach focused on identifying outliers to detect zero-day intrusions. The main goal was to develop an advanced Intrusion Detection System (IDS) model capable of accurately detecting zero-day assaults with a high level of accuracy. The model's performance was assessed using the CICIDS2017 and NSL-KDD datasets, revealing detection accuracies of 90.01% for DoS (GoldenEye), 98.43% for DoS (Hulk), 90.01% for port scanning, and 99.67% for DDoS assaults. Sharafaldin et al. [38] introduced the Hierarchical Deep Learning System based on Big Data (BDHDLS). The BDHDLS utilizes behavioral and content features to understand the characteristics of network traffic and the information included in the payload. Each deep learning model in the Big Data High-Dimensional Learning System (BDHDLS) is designed to learn the unique data distribution inside a particular cluster. This technique enhances the detection



rate of intrusive assaults in comparison to previous methods that used just one learning model. The model underwent training using the Information Centre of Excellence for Tech Innovation (ISCX) Intrusion Detection Assessment dataset (ISCXIDS2012) and obtained a remarkable accuracy rate of 99.5%. Lee et al. [39] created a framework for adaptive ensemble learning. The authors devised a multi-tree technique by adjusting the quantity of training data and creating many decision trees. The researchers used several fundamental classifiers, including Decision Trees, Random Forests, K-Nearest Neighbors (KNN), and Deep Neural Networks (DNN), in order to create an ensemble adaptive voting method that improves the overall detection performance. The authors used the NSL-KDD dataset to authenticate and verify the effectiveness of their suggested methodology. The multi-tree approach achieved an accuracy of 84.2%, whilst the adaptive voting system achieved an accuracy of 85.2% based on their results. The researchers in [40] introduced a network intrusion detection model that utilizes the Convolutional Neural Network (CNN) approach. The model autonomously extracts the salient features of incursion samples, enabling accurate categorization. It utilizes convolution and pooling methods to enhance the extraction of feature correlations within the data. The evaluation using the KDD99 datasets demonstrated that the suggested model attained a precision rate of 99.23%. Ring et al. [41] introduced a network intrusion detection model based on a CNN-IDS. The CNN model was used to automatically extract data characteristics that had been decreased in dimensionality. The authors used a standard KDD-CUP99 dataset to assess the performance of the CNN model. Based on the results, the CNN model achieved an accuracy of 94%, a Detection Rate of 93%, and a False Alarm Rate of 0.5%. The authors in [42] created a novel Intrusion Detection System (IDS) that combines different classifier methods using decision tree and rule-based algorithms. These algorithms include the Reduced Error Pruning (REP) tree, Repeated Incremental Pruning (RIP) algorithm, and Forest by Penalizing Attributes (Forest PA) algorithm. The purpose of this IDS is to detect and classify suspicious activity. The first and second algorithms were used to classify benign network traffic by using the dataset's attributes as inputs. The third classifier takes into account the characteristics of the original dataset and uses the outcomes of the first and second classifiers as inputs. The suggested Intrusion Detection System (IDS) model achieved an accuracy of 96.67%, a detection rate of 94.48%, and a false positive rate of 1.15% on the CICIDS2017 dataset. The model was trained for 195.5 seconds and tested for 2.27 seconds. Chattopadhyay et al. [43] introduced an Intrusion Detection System (IDS) that utilizes Core Vector Machines (CVM), which is a classifier based on data mining. When compared to Support Vector Machine (SVM) and ensemble classifiers, it exhibits a reduced false positive rate and requires less processing resources. The classifier underwent training and testing using the KDDCup'99 dataset. The evaluation findings indicate a detection rate of 99% and a false positive rate of 27%. Shen et al. [45] introduced a Genetic approach (GA) and Support Vector Machine (SVM) based approach for detecting intrusions in a human-controlled Intrusion Detection System (IDS). By using the crossover and mutation probabilities of GA as a baseline, the authors enhanced both the effectiveness of the population search and the capacity of people to exchange information. The algorithm's convergence time and SVM training speed were improved.

PROPOSED SYSTEM:

This project proposes a Secure Federated Intrusion Detection Model (SecFedIDM-V1). The model uses Bidirectional Long Short-Term Memory (LSTM), a deep learning algorithm, to detect and classify malicious

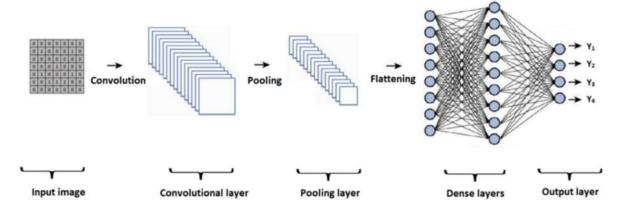


activities on the network by monitoring incoming network packets. The details of the packets classified as malicious are stored in a blockchain ledger in a private network where only certified nodes can access it, and the ledger serves as a signature database.

CNN Architecture, Process & Inputs

Architecture:

CNNs contain a combination of layers which transform an image into output the model can understand.

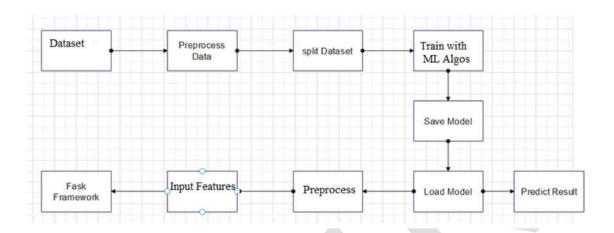


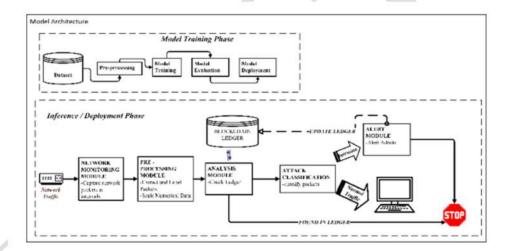
- Convolutional layer: creates a feature map by applying a filter that scans the image several pixels at a time
- Pooling layer: scales down the information generated by the convolutional layer to effectively store it
- Fully connected input layer: flattens the outputs into a single vector

SYSTEM ARCHITECTURE

A system architecture or systems architecture is the conceptual model that defines the structure, behavior, and more views of a system. An architecture description is a formal description and representation of a system. Organized in a way that supports reasoning about the structures and behaviors of the system.









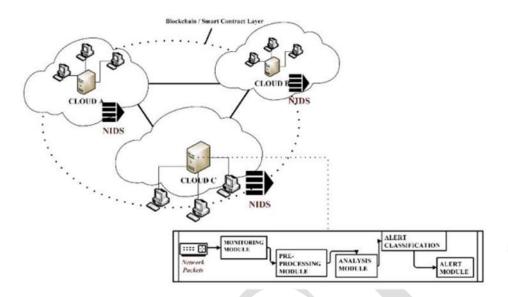
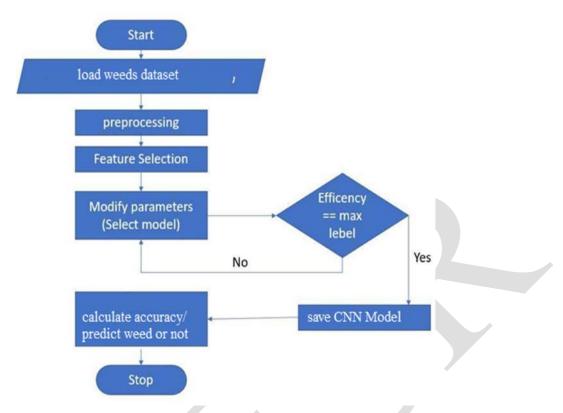


Figure 5. 1 System Architecture

3-Tier Architecture:

The three-tier software architecture (a three-layer architecture) emerged in the 1990s to overcome the limitations of the two-tier architecture. The third tier (middle tier server) is between the user interface (client) and the data management (server) components. This middle tier provides process management where business logic and rules are executed and can accommodate hundreds of users (as compared to only 100 users with the two tier architecture) by providing functions such as queuing, application execution, and database staging. FLOW CHART:





A use case diagram in the Unified Modeling Language (UML) is a type of behavioral diagram defined by and created from a Use-case analysis. Its purpose is to present a graphical overview of the functionality provided by a system in terms of actors, their goals (represented as use cases), and any dependencies between those use cases. The main purpose of a use case diagram is to show what system functions are performed for which actor. Roles of the actors in the system can be depicted.

implementation

To conduct studies and analyses of an operational and technological nature, and To promote the exchange and development of methods and tools for operational analysis as applied to defense problems.

System testing

Testing is the debugging program is one of the most critical aspects of the computer programming triggers, without programming that works, the system would never produce an output of which it was designed. Testing is best performed when user development is asked to assist in identifying all errors and bugs. The sample data are used for testing. It is not quantity but quality of the data used the matters of testing. Testing is aimed at ensuring that the system was accurately an efficiently before live operation commands.

Testing objectives:

The main objective of testing is to uncover a host of errors, systematically and with minimum effort and time. Stating formally, we can say, testing is a process of executing a program with intent of finding an error.



Output Screens

9.1 PROGRAM EXECUTION OUTPUT SCREEN



9.2 PACKET TRACKING INTERFACE

Packet Tracking Using Machine Learning Packet tracking is a critical component of network security, particularly in the context of intrusion detection systems (IDS). The objective of packet tracking is to monitor network traffic, analyze the packets transmitted over the network, and detect any anomalies or malicious activities. Leveraging machine learning for packet tracking enhances the ability to identify sophisticated and evolving cyber threats. Here, we outline the theoretical foundation of using machine learning for packet tracking

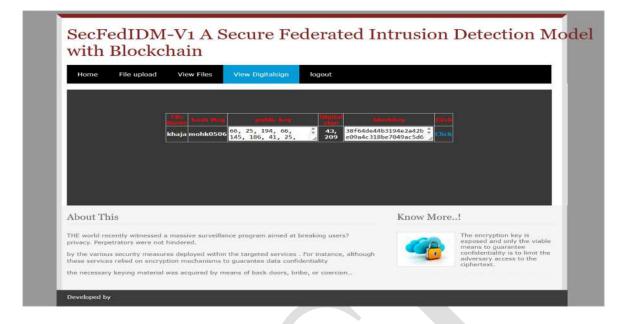
1. Introduction to Packet Tracking

Packet tracking involves capturing and analyzing the data packets that travel across a network. Each packet contains metadata and payload information, which can provide insights into the nature of the traffic and potential security threats. In traditional systems, rule-based methods are used to detect known threats, but these systems often fail to identify novel or evolving attacks.

1. Machine Learning for Anomaly Detection

Machine learning algorithms are well-suited for identifying patterns and anomalies in large datasets. In the context of packet tracking, machine learning can be employed to: Detect unusual patterns: By analyzing packet headers and payloads, machine learning models can identify deviations from normal traffic patterns, indicating potential intrusions. Classify packets: Supervised learning techniques can be used to classify packets as benign or malicious based on labeled training data. Predict future attacks: By learning from historical data, machine learning models can predict and preemptively identify potential future threats



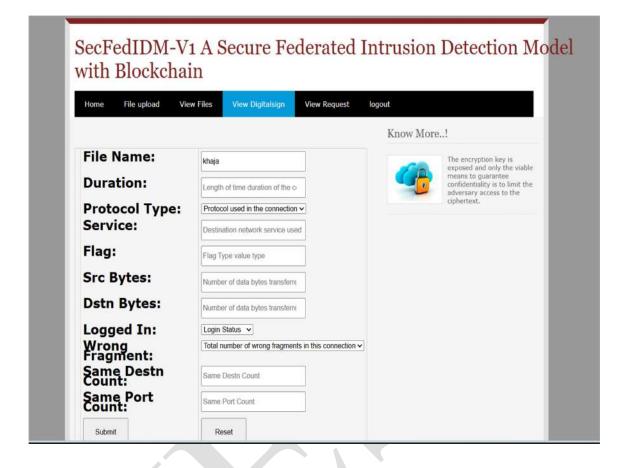


In modern network security, Intrusion Detection Systems (IDS) are crucial for identifying and mitigating unauthorized access and attacks. Traditional IDS methods often fall short in addressing the complexity and sophistication of contemporary cyber threats. This project presents an advanced approach by integrating machine learning, specifically

Deep Bidirectional Long Short-Term Memory (BiLSTM) networks, with blockchain technology in a federated learning environment to enhance intrusion detection capabilities. Packet tracking is a key component of network monitoring and security. It involves the continuous monitoring and analysis of data packets transmitted over a network to identify anomalies and potential security threats. Each data packet carries essential information such as source and destination addresses, protocol types, payload data, and timestamps. Analyzing these packets can reveal suspicious activities and potential intrusions. Machine learning, especially deep learning, offers powerful tools for recognizing patterns and detecting anomalies in large datasets.

In this project, a BiLSTM network is utilized for its superior ability to handle sequential data and capture temporal dependencies, which is crucial for analyzing network packets. Unlike standard Recurrent Neural Networks (RNNs), BiLSTMs process input data in both forward and backward directions, allowing the model to consider both past and future context in its predictions. This dual perspective enhances the model's capability to detect complex patterns and anomalies within packet data.







Federated learning is employed to enable the training of machine learning models across multiple decentralized devices without requiring the transfer of raw data to a central server. This approach significantly enhances privacy and security, as sensitive data remains local and only model updates are shared. Federated learning allows multiple organizations to collaboratively improve the intrusion detection model while maintaining data privacy and security to ensure the integrity and security of the federated learning process, blockchain technology is integrated.

Blockchain provides a transparent and tamper-proof ledger where each model update is recorded. This integration addresses potential trust and reliability issues in federated learning environments, ensuring that the collaborative learning process remains secure and trustworthy the packet tracking interface is a critical element of the ids in this project.

It provides a real-time view of network traffic and highlights detected anomalies. The interface tracks and displays key packet attributes, including packet size, time intervals, protocol types, source and destination addresses, and payload data. These attributes are essential for analyzing network traffic and identifying potential threats the anomaly detection process involves the continuous analysis of incoming packets using the bilstm model, which has been trained through federated learning. The process includes data preprocessing to extract relevant features and normalize the data, feature extraction to identify key attributes from each packet, and model inference where the bilstm network processes the extracted features.

The model compares these features to learned patterns of normal and malicious traffic, identifying packets that significantly deviate from normal patterns as potential threats. Detected anomalies trigger alerts, prompting further investigation or automated mitigation actions

CONCLUSION

In modern times, cloud federation has gained usage among Cloud Service Providers (CSP) as a means of sharing computational resources for better service delivery. However, with the increase in malicious activities and cyber-attacks, the need for secure transactions and communication between federated entities becomes critical. This study presents a BiLSTM-based IDS model for a federated cloud platform named SecFedIDM-V1. Firstly, we developed a cloud federation testbed using openstack. Secondly, we experimentally evolved an IDS



model using CIDDS dataset with BiLSTM RNN giving the best performance metrics for intrusion traffic detection on an OpenStack-based federated cloud testbed. Thirdly, in order to enhance the security of the proposed architecture, we integrated a blockchain to serve as a secure datastore for intrusion signatures. In order to make the model usable for cloud system administrators, we developed a cloud native web application that incorporates the proposed BiLSTM IDS model. As a proof of concept, we illustrated how different classes of malicious traffic could be detected with high precision from the cloud application. The web app could be deployed on other experimental or production-federated cloud platforms to detect malicious intrusions.

Future scope:

In the future, we hope to extend the architecture as well as the web application by covering novel and higher number of network attack classes..

REFERENCES

- 1. Ijteba Sultana, Dr. Mohd Abdul Bari, Dr.Sanjay," Routing Performance Analysis of Infrastructure less Wireless Networks with Intermediate Bottleneck Nodes", International Journal of Intelligent Systems and Applications in Engineering, ISSN no: 2147-6799 IJISAE, Vol 12 issue 3, 2024, Nov 2023
- 2. Md. Zainlabuddin, "Wearable sensor-based edge computing framework for cardiac arrhythmia detection and acute stroke prediction", Journal of Sensor, Volume 2023.
- **3.** Md. Zainlabuddin, "Security Enhancement in Data Propagation for Wireless Network", Journal of Sensor, ISSN: 2237-0722 Vol. 11 No. 4 (2021).
- 4. Dr MD Zainlabuddin, "CLUSTER BASED MOBILITY MANAGEMENT ALGORITHMS FOR WIRELESS MESH NETWORKS", Journal of Research Administration, ISSN:1539-1590 | EISSN:2573-7104 , Vol. 5 No. 2, (2023)
- 5. Vaishnavi Lakadaram, "Content Management of Website Using Full Stack Technologies", Industrial Engineering Journal, ISSN: 0970-2555 Volume 15 Issue 11 October 2022
- 6. Dr. Mohammed Abdul Bari, Arul Raj Natraj Rajgopal, Dr.P. Swetha, "Analysing AWSDevOps CI/CD Serverless Pipeline Lambda Function's Throughput in Relation to Other Solution", International Journal of Intelligent Systems and Applications in Engineering, JISAE, ISSN:2147-6799, Nov 2023, 12(4s), 519–526
- 7. Ijteba Sultana, Mohd Abdul Bari and Sanjay," Impact of Intermediate per Nodes on the QoS Provision in Wireless Infrastructure less Networks", Journal of Physics: Conference Series, Conf. Ser. 1998 012029, CONSILIO Aug 2021
- **8.** M.A.Bari, Sunjay Kalkal, Shahanawaj Ahamad," A Comparative Study and Performance Analysis of Routing Algorithms", in 3rd International Conference ICCIDM, Springer 978- 981-10-3874-7_3 Dec (2016)
- 9. Mohammed Rahmat Ali,: BIOMETRIC: AN e-AUTHENTICATION SYSTEM TRENDS AND FUTURE APLLICATION", International Journal of Scientific Research in Engineering (IJSRE), Volume1, Issue 7, July 2017



- 10. Mohammed Rahmat Ali,: BYOD.... A systematic approach for analyzing and visualizing the type of data and information breaches with cyber security", NEUROQUANTOLOGY, Volume20, Issue 15, November 2022
- 11. Mohammed Rahmat Ali, Computer Forensics -An Introduction of New Face to the Digital World, International Journal on Recent and Innovation Trends in Computing and Communication, ISSN: 2321-8169-453 456, Volume: 5 Issue: 7
- 12. Mohammed Rahmat Ali, Digital Forensics and Artificial Intelligence ... A Study, International Journal of Innovative Science and Research Technology, ISSN:2456-2165, Volume: 5 Issue:12.
- 13. Mohammed Rahmat Ali, Usage of Technology in Small and Medium Scale Business, International Journal of Advanced Research in Science & Technology (IJARST), ISSN:2581-9429, Volume: 7 Issue:1, July 2020.
- 14. Mohammed Rahmat Ali, Internet of Things (IOT) Basics An Introduction to the New Digital World, International Journal on Recent and Innovation Trends in Computing and Communication, ISSN: 2321-8169-32-36, Volume: 5 Issue: 10
- 15. Mohammed Rahmat Ali, Internet of things (IOT) and information retrieval: an introduction, International Journal of Engineering and Innovative Technology (IJEIT), ISSN: 2277-3754, Volume: 7 Issue: 4, October 2017.
- 16. Mohammed Rahmat Ali, How Internet of Things (IOT) Will Affect the Future A Study, International Journal on Future Revolution in Computer Science & Communication Engineering, ISSN: 2454- 424874 77, Volume: 3 Issue: 10, October 2017.
- 17. Mohammed Rahmat Ali, ECO Friendly Advancements in computer Science Engineering and Technology, International Journal on Scientific Research in Engineering (IJSRE), Volume: 1 Issue: 1, January 2017
- 18. Ijteba Sultana, Dr. Mohd Abdul Bari, Dr. Sanjay, "Routing Quality of Service for Multipath Manets, International Journal of Intelligent Systems and Applications in Engineering", JISAE, ISSN:2147- 6799, 2024, 12(5s), 08–16;
- 19. Mr. Pathan Ahmed Khan, Dr. M.A Bar i,: Impact Of Emergence With Robotics At Educational Institution And Emerging Challenges", International Journal of Multidisciplinary Engineering in Current Research(IJMEC), ISSN: 2456-4265, Volume 6, Issue 12, December 2021, Page 43-46
- 20. Shahanawaj Ahamad, Mohammed Abdul Bari, Big Data Processing Model for Smart City Design: A Systematic Review ", VOL 2021: ISSUE 08 IS SN: 0011-9342; Design Engineering (Toronto) Elsevier SCI Oct: 021
- 21. Syed Shehriyar Ali, Mohammed Sarfaraz Shaikh, Syed Safi Uddin, Dr. Mohammed Abdul Bari, "Saas Product Comparison and Reviews Using Nlp", Journal of Engineering Science (JES), ISSN NO:0377- 9254, Vol 13, Issue 05, MAY/2022.
- 22. Mohammed Abdul Bari, Shahnawaz Ahamad, Mohammed Rahmat Ali," Smartphone Security and Protection Practices", International Journal of Engineering and Applied Computer Science (IJEACS); ISBN: 9798799755577 Volume: 03, Issue: 01, December 2021 (International Journal, UK) Pages 1-6



- 23. .A.Bari& Shahanawaj Ahamad, "Managing Knowledge in Development of Agile Software", in International Journal of Advanced Computer Science & Applications (IJACSA), ISSN: 2156-5570, Vol. 2, No. 4, pp. 72-76, New York, U.S.A., April 2011
- 24. Imreena Ali (Ph.D), Naila Fathima, Prof. P.V.Sudha ,"Deep Learning for Large-Scale Traffic-Sign Detection and Recognition", Journal of Chemical Health Risks, ISSN:2251-6727/ JCHR (2023) 13(3), 1238-1253
- 25. Imreena, Mohammed Ahmed Hussain, Mohammed Waseem Akram" An Automatic Advisor for Refactoring Software Clones Based on Machine Learning", Mathematical Statistician and Engineering ApplicationsVol. 72 No. 1 (2023)
- 26. Mrs Imreena Ali Rubeena, Qudsiya Fatima Fatimunisa "Pay as You Decrypt Using FEPOD Scheme and Blockchain", Mathematical Statistician and Engineering Applications: https://doi.org/10.17762/msea.v72i1.2369 Vol. 72 No. 1 (2023)
- 27. Imreena Ali , Vishnuvardhan, B.Sudhakar," Proficient Caching Intended For Virtual Machines In Cloud Computing", International Journal Of Reviews On Recent Electronics And Computer Science , ISSN 2321-5461,IJRRECS/October 2013/Volume-1/Issue-6/1481-1486.
- 28. Heena Yasmin, A Systematic Approach for Authentic and Integrity of Dissemination Data in Networks by Using Secure DiDrip, INTERNATIONAL JOURNAL OF PROFESSIONAL ENGINEERING STUDIES, Volume VI /Issue 5 / SEP 2016
- **29.** Heena Yasmin, Cyber-Attack Detection in a Network, Mathematical Statistician and Engineering Applications, ISSN:2094-0343, Vol.72 No.1(2023)
- **30.** Heena Yasmin, Emerging Continuous Integration Continuous Delivery (CI/CD) For Small Teams, Mathematical Statistician and Engineering Applications, ISSN:2094-0343, Vol.72 No.1(2023).