Classification of Attacks on Big Data using Deep Learning Approach

Ms. Shalu Saini, Research Scholar, Department of Computer Science & Engineerng, BMU, Rohtak.

Dr. Priti Singla, Professor, Department of Computer Science & Engineering, BMU, Rohtak.

Abstract: The classification of attacks on big data using a deep learning approach is essential for safeguarding the integrity, confidentiality, and availability of large-scale data environments. MitM, DoS, and Brute Force represent significant threats to big data systems, and leveraging deep learning techniques can provide an effective means of detection and mitigation. Deep learning models can autonomously analyze massive datasets, adapt to evolving attack strategies, and identify intricate attack patterns. However, this approach is not without challenges. Deep learning for big data security requires access to diverse and labeled datasets, which can be scarce and challenging to obtain for specific attack types. Additionally, the resource-intensive nature of deep learning models may pose scalability and operational challenges for organizations with extensive big data infrastructure. Moreover, false positives and false negatives can be a concern, demanding continuous fine-tuning and adjustment to minimize misclassifications. Despite these challenges, the potential benefits of using deep learning in attack classification within big data environments are substantial, as they enable real-time threat detection and proactive security measures, ultimately reinforcing the overall security posture of these critical data systems.

Keywords: Big data security, MitM attack, DoS, Brute force, Deep learning

[1] Introduction

Securing big data using DL involves applying advanced ML techniques to detect and mitigate security threats and vulnerabilities in large-scale data environments. DL can be a valuable tool in enhancing the security of big data systems due to its ability to process and analyze vast amounts of data, adapt to evolving threats, and detect complex patterns. Here are some key ways to use deep learning for securing big data:

1. **Anomaly Detection:** DL models can be trained on historical data to learn normal patterns. When deployed, they can detect anomalies, which may indicate security breaches or unusual behavior in the data.
2. **Malware Detection:** Deep learning models can be used to identify malicious software or malware in big data. They can analyze code, network traffic, and behavioral patterns to detect known and novel threats.
3. **Intrusion Detection:** Deep learning can enhance intrusion detection systems (IDS) by continuously analyzing network traffic and identifying suspicious activities or known attack patterns. This helps in real-time threat identification.
4. **Natural Language Processing (NLP):** NLP models can be used to analyze text data and detect security threats, such as sentiment analysis for identifying cybersecurity threats in textual content, like forums or social media.
5. **Behavioral Analysis:** Deep learning models can monitor and analyze user and system behavior to identify deviations from normal patterns. This is particularly useful for detecting insider threats.
6. **Deep Packet Inspection:** For network security, deep learning models can perform deep packet inspection to analyze network traffic, flagging packets that exhibit suspicious behavior, potentially indicating attacks or data exfiltration.
7. **User and Entity Behavior Analytics (UEBA):** Deep learning can be employed in UEBA systems to monitor and detect unusual user and entity behavior, helping in the early identification of insider threats and compromised accounts.
8. **Secure Access Control:** Implementing deep learning-based authentication systems can enhance access control and reduce the risk of unauthorized access to big data resources.
9. **Data Loss Prevention (DLP):** Deep learning can assist in data loss prevention by analyzing data streams and identifying sensitive data that is being transmitted or accessed inappropriately.
10. **Dynamic Threat Intelligence:** Deep learning can be used to process and analyze threat intelligence data feeds to identify emerging threats and adjust security measures accordingly.
11. **Security Automation:** Implement deep learning models for automating security response, such as triggering alerts or taking predefined actions in response to detected threats.
12. **Data Privacy and Encryption:** Deep learning models can assist in data encryption and privacy protection by ensuring that sensitive data remains secure and identifying potential vulnerabilities.

Securing big data using deep learning is an ongoing process that requires continuous monitoring, model training, and adaptation to evolving threats.

1.1 Threats to security of Big Data

MitM, DoS, and Brute Force attacks are common security threats in the context of big data systems:

1. Man-in-the-Middle (MitM) Attack:

- Without the target's awareness, an attacker may secretly listen in on their conversation in a MitM attack. In a large data system, this can jeopardize the security of sent data.

- Implications for Big Data: MitM attack can lead to data interception, tampering, or even data theft. Attackers can potentially access sensitive information as it travels between components of the big data infrastructure.

2. Denial of Service (DoS) Attack:

- The goal of a DoS assault is to make a target system, service, or network inaccessible or unresponsive to its intended users by flooding it with traffic. This has the potential to interfere with how a large data system usually operates.
- Implications for Big Data: In a big data environment, a DoS attack can lead to service outages, degradation of performance, and a loss of data availability. It can impact the timely processing and analysis of data.

3. Brute Force Attack:

- The goal of a brute force assault is to find the proper password or key combination by trying every conceivable combination. Brute force attacks can target authentication mechanisms within a big data system.
- Implications for Big Data: If a brute force attack is successful, it can lead to unauthorized access to the big data system or critical data. This poses a significant security risk, especially when dealing with sensitive or confidential information.

1.2 protection schemes

To protect big data systems against these types of attacks, several security measures can be implemented:

- Encryption: Encrypting data in transit and at rest can prevent MitM attacks by ensuring data confidentiality and integrity.
- Access Control: Implement strict access controls and authentication mechanisms to protect against brute force attacks.
- IDS: Utilize IDPS to detect and mitigate DoS attacks by identifying abnormal traffic patterns and responding in real-time.
- Network Segmentation: Isolate critical components of the big data system from the public network to reduce the attack surface and limit the impact of attacks.
- Monitoring and Logging: Keep a close eye on the system for any unusual or suspicious activity and record all important events for analysis after an occurrence.
- Patch and Update Management: Apply security fixes to software and systems on a regular basis to fix known vulnerabilities that hackers may use.
- Incident Response Plan: Create a solid incident response strategy that specifies what to do in the case of an attack, from preventing it to investigating it and recovering from it.

Protecting big data systems against MitM, DoS, and brute force attacks is essential for ensuring data integrity, availability, and confidentiality, which are fundamental for maintaining the reliability and security of these complex and valuable environments.

1.3 Need of deep learning based classification of Attacks over big data

The need for classifying MitM, DoS, and brute force attacks on big data using a deep learning approach is driven by several important considerations:

1. **Scale and Complexity of Big Data:** Big data environments process and store massive volumes of data, making them attractive targets for attackers. Traditional methods of attack detection and prevention may not scale to handle the complexity and size of big data systems.
2. **Real-Time Detection:** Big data systems often require real-time processing, and attacks can have immediate and severe consequences. Deep learning models can analyze data streams in real-time, allowing for faster detection and response to attacks.
3. **Advanced Attack Techniques:** Attackers are continually evolving their tactics. Deep learning models can adapt and learn from data, making them more capable of identifying both known and emerging attack patterns.
4. **Zero-Day Attacks:** Traditional rule-based systems are less effective at detecting new, previously unseen attack patterns. Deep learning models can identify anomalies and unknown attack vectors without relying on predefined rules.
5. **Anomaly Detection:** Big data attacks often involve subtle anomalies in data patterns. Deep learning models are well-suited for anomaly detection because they can learn the expected data distribution and identify deviations from it.
6. **Mitigation of False Positives:** Deep learning models can reduce false positives by learning to distinguish between normal fluctuations in data and actual attacks, improving the efficiency of security operations.
7. **Automation and Efficiency:** Automating the process of attack detection and classification using deep learning allows security teams to focus on responding to confirmed attacks, rather than sifting through large volumes of data and alerts.
8. **Multi-Vector Attacks:** Attackers often use a combination of techniques in a coordinated manner. Deep learning models can handle complex, multi-vector attacks by learning the relationships between various attack indicators.
9. **Continuous Learning:** Deep learning models can adapt and improve over time by continuously learning from new data. This adaptability is crucial for staying ahead of evolving attack strategies.
10. **Reduced Manual Rule Creation:** Deep learning reduces the reliance on manually creating rules for attack detection. This can save time and resources and enable security teams to respond more effectively to new and complex threats.
11. **Scalable and Resource-Efficient:** Deep learning models can be deployed at scale without significant resource overhead. This is important for big data environments that handle vast amounts of data and require efficient attack detection methods.

The need for classifying MitM, DoS, and brute force attacks on big data using a deep learning approach is driven by the complexity, scale, and evolving nature of big data systems and the attacks targeting them. Deep learning offers an adaptive and effective solution for identifying and responding to various types of attacks, enhancing the security and reliability of big data environments.

[2] Literature review

Juan Jose Mura Privacy and security concerns related to big data are a hot topic in 2014 [1]. Modern issues of privacy and security are the focus of this research. A wide variety of sources may include unstructured data, including social networking sites, sensors, scientific apps, surveillance systems, video and picture archives, medical records, financial transactions, internet search results, and system logs.

S. Riaz 2020 [2] focuses on the present issues and future research prospects in connection to preserving privacy and security of big data are highlighted. Data kept in the cloud is not yet safeguarded against any and all attacks and invasions, whether they come from inside the system or from the outside. This is because a third party retains control over cloud servers. In recent years, data security has become more important as a result of the use of cloud storage for sensitive information. In this study, they survey state of art in data security and privacy, as well as the defining features of big data. For the sake of future study, the most pressing dangers, unanswered questions, and pressing concerns, as well as their effects on businesses, are explored. Data security frameworks and architectures have evolved over time in response to evolving threats, and this presentation reviews and analyses these developments to better inform practises for securing data in the cloud.

The challenges, issues, risks, and privacy concerns associated with big data were first presented by Taran Singh Bharati 2020 [3]. Within the context of big data, there are conflicts around issues of identity, transparency, and authority. Lambda and kappa are the names of the big data architectures. Protecting the privacy and security of the system against vulnerabilities, gaps, and attacks requires proactive and reactive measures. They would consider a system to be secure if it met their criteria of being access-controlled, integrated, genuine, and confidential. An adversary will use all opportunities to their advantage. Important data was protected by procedures and protocols designed to detect, prevent, and fix data security breaches. Software like anti-spam programs, anti-virus programs, firewalls, and internet security suites may help protect against these threats. There are numerous situations when even a little change to certain data might cause massive misunderstanding.

Vishal Joshi 2020 [4] laid forth several fundamental principles of privacy and security issues unique to big data, which is crucial if we are to redirect our efforts toward strengthening big data infrastructures. The article "Security/Privacy Issues and Challenges in Big Data" explains how people at various points in the big data ecosystem encounter problems while dealing with big data's day-to-day activities. This article explains the research done to tackle the most pressing problems with big data security and highlights some things to keep in mind while working with such a complex and potentially dangerous resource.

In his 2018 [5] study, R. Sumithra examines the security, privacy, and challenge concerns related to big data and cloud security in great depth. Big data enables effective decision-making by delivering data warehouse advantages with the added capability of analyzing data from distant file systems. There is a great ROI with big data. There was also a code of conduct for big data and a discussion of the legal issues surrounding intellectual property rights in this publication.

Professor Kamakshi, P. Dec 2014 [6] considered privacy issues in the age of big data are the focus of this study. Several privacy-related issues, as well as the advantages and applications of big data, are investigated.

In their study of current layered cloud designs, Tawalbeh, A., and S, Gokay [7] offered a solution for large data storage, P2P Cloud System utilization, and a hybrid mobile cloud computing model based on cloudlets. As a case study, this technique was subsequently applied to the analysis of patient data included inside healthcare systems.

According to research by Minit Arora et al. (2016) [8], companies use several strategies to mask client information for safety and privacy reasons. Making a solemn vow, either in writing or orally, was the most typical way to guarantee one's privacy and security. However, this approach was incorrect, as we can see from past events. Common low-level technological methods for enforcing privacy and security while sharing and aggregating data over dynamic, dispersed networks include passwords, restricted access, and two-factor authentication.

Additional evidence for this assertion may be found in the research and analysis of big data and cloud computing's varied degrees of security conducted by K.P. Mahaheswari, P. Ramya, and S. Nirmala Devi (2017) [9]. Some industries and parts of government were hit particularly hard by the difficulties posed by big data. The same security problems that plagued massive data systems and technologies also plagued cloud computing. Since cloud computing involves connecting several independent systems, it was imperative that these connections be as secure as possible.

For the purpose of keeping individual health records in a big data environment, J. L. Dhas, S. Maria Celestin Vigila, and C. Ezhil Star (2017) [10] created an architecture that takes privacy and security into consideration. There are a number of immediate concerns that arise from keeping medical information as massive data sets. Learning to keep your data secure on the cloud is one of them. The next step is figuring out how to spot the file and block unauthorised users from accessing sensitive medical data.

[3] Problem statement

Classifying MitM, DoS, and Brute Force attacks on big data using a deep learning approach presents several notable challenges and issues. First and foremost, acquiring labeled datasets for deep learning models can be a formidable obstacle, as obtaining diverse and representative data for specific attack types is often a complex and resource-intensive task. The resource demands of deep learning models pose another challenge, particularly in big data environments, where scalability, processing power, and memory requirements can strain existing infrastructure. Additionally, the interpretability of deep learning models can be a concern, making it difficult to explain the rationale behind the model's classifications and leading to a potential lack of trust in automated security systems. Furthermore, false positives and false negatives in attack classification are common challenges, necessitating continuous fine-tuning and adjustments to reduce misclassifications and enhance accuracy. The evolving nature of cyber threats means that deep learning models must be regularly updated to adapt to new attack strategies and vulnerabilities, requiring a commitment to ongoing maintenance and improvement. Despite these hurdles, the potential advantages of employing deep learning in attack classification within big data ecosystems, such as real-time threat detection and proactive security measures, make addressing these challenges essential for bolstering the security of these critical data systems.

[4] Proposed work

Classifying attacks on big data using a deep learning approach is a complex and valuable task in enhancing the security of big data systems. Deep learning models can automatically learn and identify attack patterns within large volumes of data. Here are the steps to classify attacks on big data using a deep learning approach:

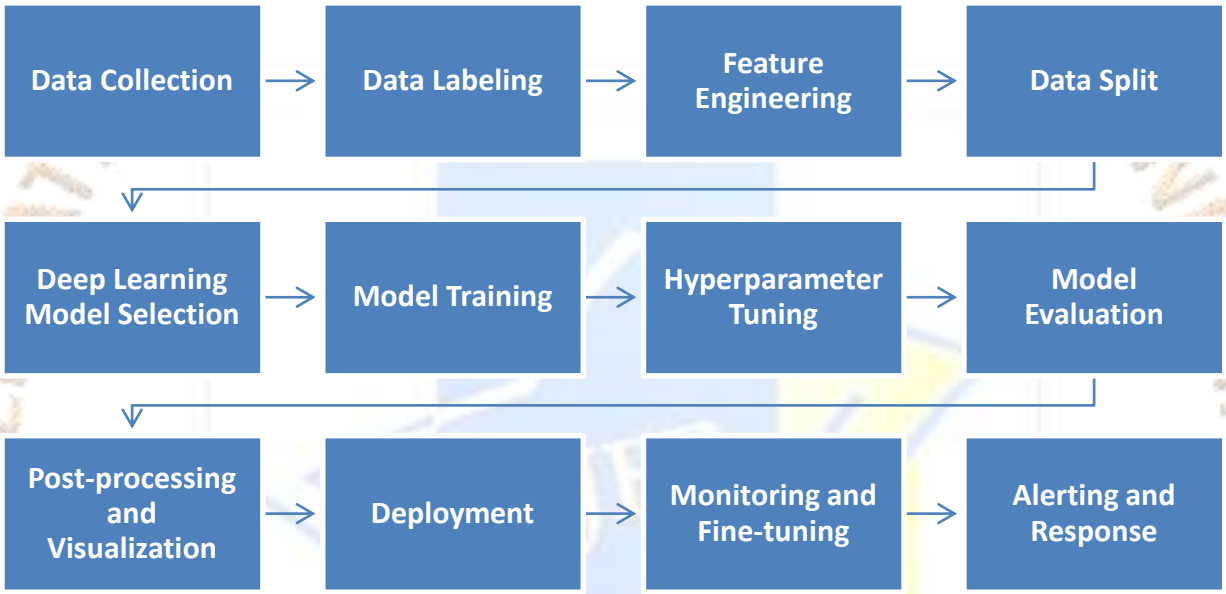

Fig 1 Proposed Research Methodology

1. **Data Collection:** Collect and preprocess a large dataset of network traffic, or logs. This dataset should include examples of various attacks, such as MitM, DoS, and brute force attacks.
2. **Data Labeling:** Annotate the dataset to indicate whether each data sample represents a normal or an attack instance. For attack samples, specify the type of attack, such as MitM, DoS, or brute force.
3. **Feature Engineering:** Identify useful data characteristics for feeding into the deep learning model. Network packet characteristics, system logs, and any other pertinent data that may differentiate between normal and attack patterns are examples of such attributes.
4. **Data Split:** Separate the dataset into three parts: training, validation, and testing. Perhaps 70% would go toward teaching, 15% toward validation, and 15% toward testing.
5. **Deep Learning Model Selection:** To complete the categorization job, choose a deep learning architecture that is suitable. Popular options include hybrid models, RNNs for sequential data, and CNNs for picture data.
6. **Model Training:** Backpropagation and gradient descent are used to train the deep learning model using the training dataset. The goal of training the model is to have it differentiate between typical and malicious patterns and to assign appropriate labels to each kind of assault.
7. **Hyperparameter Tuning:** Learn how to improve the deep learning model's performance on the validation dataset by experimenting with various hyperparameters, such as learning rate, batch size, and architecture.
8. **Model Evaluation:** Check the trained model's recall, accuracy, precision, F1-score, and other important metrics on the testing dataset. Making ensuring the model works properly with new data is the goal of this stage.
9. **Post-processing and Visualization:** After classifying attacks, apply post-processing techniques to refine the results and reduce false positives. Visualization tools can help interpret the model's decisions and aid in understanding the attack patterns.
10. **Deployment:** Once the model achieves satisfactory accuracy and performance, deploy it in your big data environment to continuously monitor for attacks in real-time.
11. **Monitoring and Fine-tuning:** Continuously monitor the system's performance in detecting attacks and adapt the model as new attack patterns emerge. Fine-tune the model as needed to maintain high accuracy.

12. **Alerting and Response:** Integrate the model with an alerting system to notify administrators or initiate automated responses when attacks are detected. This can help mitigate the impact of attacks in real-time.

By following these steps, you can use a deep learning approach to classify various types of attacks on big data. This approach can provide an effective and adaptive way to enhance the security of big data systems by automatically identifying and responding to threats.

[5] Result and discussion

In this section, there are 3 subsections named as conventional approach, proposed work and comparative analysis of accuracy. There are confusion matrix contains 4 categories for testing operation. It calculates accuracy parameters such as precision, recall, and f-score as shown in section 5.3.

5.1 Conventional Approach

Testing operation made using conventional model and 1000 signal have been considered for each class. Confusion matrix for conventional has been shown in table 1

Table 1 Confusion matrix

	Normal	MitM	DoS	Brute force
Normal	911	19	22	18
MitM	23	914	17	29
DoS	27	18	937	17
Brute force	39	49	24	936

Following table 2 is presenting the accuracy, recall, precision and f1 score for proposed work considering table 1.

TP: 3698 and Overall Accuracy: 92.45%

Table 2 Accuracy chart for conventional

Class	n (truth)	n (classified)	Accuracy	Precision	Recall	F1 Score
1	1000	970	96.3%	0.94	0.91	0.92
2	1000	983	96.13%	0.93	0.91	0.92
3	1000	999	96.88%	0.94	0.94	0.94
4	1000	1048	95.6%	0.89	0.94	0.91

5.2 Proposed approach

Testing operation made using proposed model and 1000 signal have been considered for each class. Confusion matrix for proposed has been shown in table 3

Table 3 Confusion matrix

	Normal	MitM	DoS	Brute force
Normal	949	11	13	11
MitM	15	956	11	17
DoS	18	10	959	9
Brute force	23	17	17	963

Following table 4 is presenting the accuracy, recall, precision and f1 score for proposed work considering table 3.

TP: 3827 and Overall Accuracy: 95.68%

Table 4 Accuracy chat for Proposed

Class	n (truth)	n (classified)	Accuracy	Precision	Recall	F1 Score
1	1000	984	97.85%	0.96	0.95	0.96
2	1000	999	97.83%	0.96	0.96	0.96
3	1000	996	98.05%	0.96	0.96	0.96
4	1000	1021	97.63%	0.94	0.96	0.95

5.3 Comparative analysis

Considering table 2 and table 4, accuracy, precision, recall, F1-score have table have been obtained.

1. Accuracy

Comparison of accuracy in case of conventional and proposed accuracy is shown in table 5.

Table 5 Comparison graph of accuracy

Class	Conventional Accuracy	Proposed accuracy
1	97.85%	96.3%
2	97.83%	96.13%
3	98.05%	96.88%
4	97.63%	95.6%

Considering table 5, comparison of accuracy has been made for conventional and proposed work.

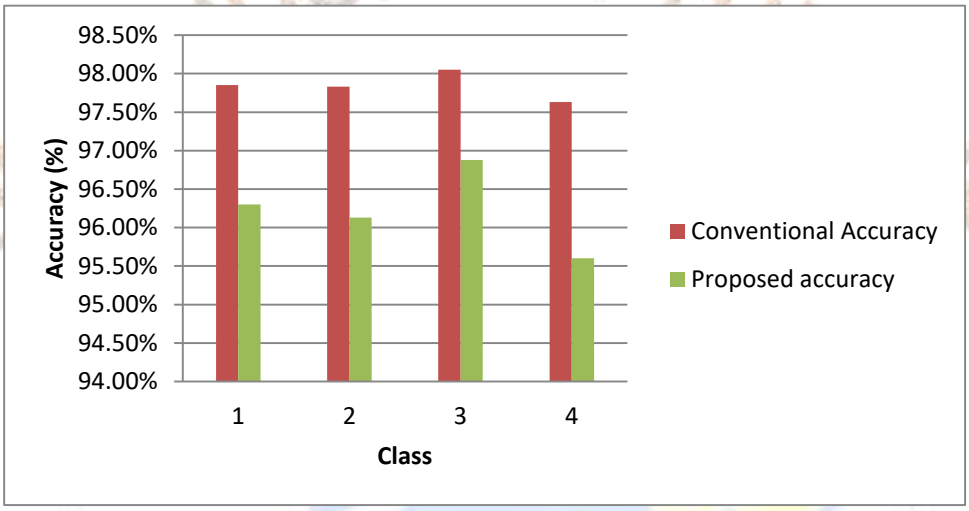

Fig 2 Comparison graph of accuracy

2. Precision

Comparative analysis of precision in case of conventional and present precision is shown in table 6.

Table 6 Comparison graph of Precision

Class	Conventional Precision	Proposed Precision
1	0.94	0.96
2	0.93	0.96
3	0.94	0.96
4	0.89	0.94

Considering table 6, comparison of precision has been made for conventional and proposed work.

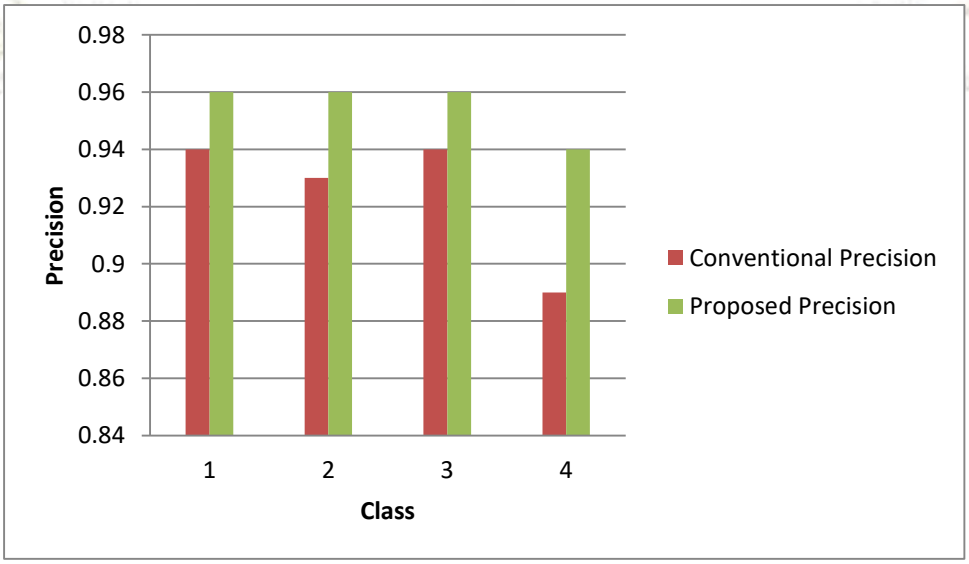

Fig 3 Comparison graph of Precision

3. Recall value

Comparison of recall value in case of conventional and proposed recall value is shown in table 7.

Table 7 Comparison graph of Recall

Class	Conventional Recall value	Proposed Recall value
1	0.92	0.95
2	0.92	0.96
3	0.94	0.96
4	0.91	0.96

Considering table 7, comparison of recall value has been made for conventional and proposed work.

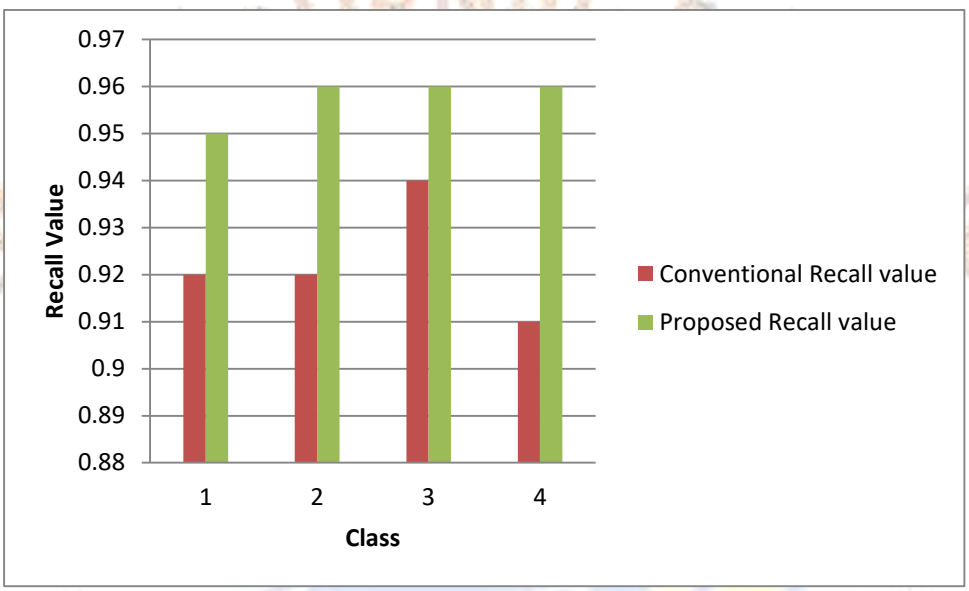

Fig 4 Comparison graph of Recall

4. F1-score

Comparison of f1-score in case of conventional and proposed f1-score is shown in table 5.

Table 5 Comparison graph of f1-score

Class	Conventional F1-Score	Proposed F1-Score
1	0.91	0.96
2	0.91	0.96
3	0.92	0.96
4	0.91	0.95

Considering table 5, comparison of f1-score has been made for conventional and proposed work.

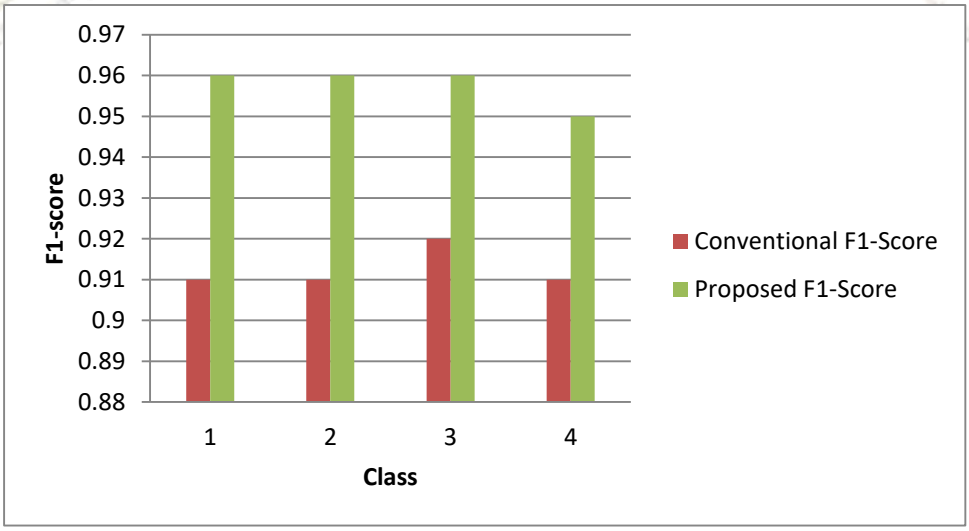

Fig 5 Comparison graph of f1-score

[6] Conclusion

It is concluded that proposed work has demonstrated improved performance compared to conventional approaches, particularly in terms of accuracy, precision, recall, and F-score. The utilization of optimization techniques to filter data is a significant contributing factor to these enhancements. By applying data filtering methods, you can reduce noise and irrelevant information, which in turn enhances the quality and relevance of the data being processed by your deep learning classification model. Furthermore, the subsequent improvement in the performance of the deep learning classification model is a positive outcome. Filtering the data not only streamlines the input but also helps the model focus on the most relevant and discriminative features, resulting in more accurate predictions. The use of optimization techniques and data preprocessing is a valuable strategy for enhancing the effectiveness of machine learning and deep learning models, as it enables them to make more informed and precise decisions. This can be particularly advantageous in the context of big data, where data quality, dimensionality, and complexity can present substantial challenges. It's essential to continue refining and fine-tuning these techniques to achieve even better results and to stay ahead of the evolving demands of big data analysis and classification.

[7] Future scope

The future scope of classifying MitM, DoS, and brute force on big data using a deep learning approach is both promising and pivotal in the ever-evolving landscape of cybersecurity. As big data environments continue to expand in scale and complexity, the importance of robust security measures cannot be overstated. Deep learning's ability to process and study vast volum of data, classify intricate attack pattern, and adapt to emerging threats positions the forefront of security solutions. In the coming years, we can expect advancements in deep learning models specifically tailored to big data security. These models will not only offer improved accuracy and real-time threat detection but also enhanced adaptability to the ever-changing tactics employed by cybercriminals. Furthermore, the development of more comprehensive and diverse labeled datasets will be crucial in training and refining these models, ensuring their effectiveness in identifying both known and novel attack strategies. While the future holds great promise, several challenges persist. The acquisition of high-quality, real-world data for training and testing deep learning models remains a hurdle. Keeping these models up to date with evolving attack techniques and vulnerabilities is an ongoing task. Additionally, addressing the interpretability and transparency of deep learning models is essential for building trust in automated security systems and for regulatory compliance. Scalability and resource efficiency will also be key considerations, as big data environments grow in size and complexity. Striking the right balance between model complexity and computational feasibility will be vital for practical implementation. Mitigating false positives and false negatives in attack classification will continue to be a challenge, demanding constant optimization to reduce classification errors and enhance overall system performance. In conclusion, the future of attack classification in big data using deep learning is poised to play a pivotal role in fortifying the security of large-scale data systems. Ongoing research, innovation, and collaboration will be essential to address the challenges and seize the opportunities that lie ahead, ensuring that big data remains a valuable asset without compromising its integrity and confidentiality.

Reference

[1] J. Moura, “Security and Privacy Issues of Big Data,” Handbook of research on trends and future directions in big data and web intelligence., no. 20-52, 2015.
[2] S. Riaz, A. H. Khan, M. Haroon, S. Latif, and S. Bhatti, “Big data security and privacy: Current challenges and future research perspective in cloud environment,” Proc. 2020 Int. Conf. Inf. Manag. Technol. ICIMTech 2020, no. August, pp. 977–982, 2020, doi: 10.1109/ICIMTech50083.2020.9211239.
[3] T. S. Bharati, “Challenges, issues, security and privacy of big data,” Int. J. Sci. Technol. Res., vol. 9, no. 2, pp. 1482–1486, 2020.
[4] M. V. Joshi, ”Security/Privacy Issues and Challenges in Big Data,” International Research Journal of Engineering and Technology (IRJET), vol. 07, no. 06, 2020.
[5] R. Sumithra, ”Security, Privacy Issues and Challenges in Big Data and Cloud,” Special Issue based on Proceedings of 4th International Conference on Cyber Security (ICCS), 2018.
[6] P. Kamakshi, ”SURVEY ON BIG DATA AND RELATED PRIVACY ISSUES,” International Journal of Research in Engineering and Tech- nology, vol. 03, no. 12, Dec 2014.
[7] L. A. T. a. G. Saldamli, ”Reconsidering big data security and privacy in cloud and mobile cloud systems,” Journal of King Saud University – Computer and Information Science, 2019.
[8] M. A. a. D. H. Bahuguna, ”Big Data Security – The Big Challenge,” International Journal of Scientific Engineering Research, vol. 7, no. 12, Dec 2016.
[9] P. a. S. D. K.P.Maheswari, ”STUDY AND ANALYSES OF SECURITY LEVELS IN BIG DATA AND CLOUD COMPUTING,” International Journal of Innvative Research in Science and Engineering, vol. 3, no. 02, 2017.
[10] S. M. C. V. a. C. E. S. J.L. Joneston Dhas, ”A Framework on Security and PrivacyPreserving for Storage of Health Information Using Big Data,” International Science Press, 2017.
[11] “https://techvidvan.com/tutorials/big-data-applications/”.
[12] EU. Opinion 05/2014 on anonymisation techniques. ARTICLE 29 DATA PROTECTION WORKING PARTY, 2014.
[13] P. V. J. S. Batcha, ”THE FIELD OF BIG DATA FOR SECURITY INTELLIGENCE,” IJCRT, vol. 6, no. 2018, 2 April 2018.
[14] M. Parihar, ”Big Data Security and Privacy,” International Journal of Engineering Research Technology, 07 July 2021.
[15] R. V. Sitalakshmi Venkatraman, ”Big data security challenges and strategies,” vol. 4, no. 3.