Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 8, August : 2024

STRATEGIES FOR DETECTING DATA POISONING IN DISTRIBUTED MACHINE LEARNING

[1]KAKUMANU PUJITHA, [2] MRS. CH. DEEPTI

[1] PG Scholar in the department of MCA at QIS College of Engineering & Technology (AUTONOMOUS), Vengamukkapalem, Ongole- 523272, Prakasam Dt., AP., India.

[2] Assistant Professor in the department of CSE/MCA at QIS College of Engineering & Technology (AUTONOMOUS), Vengamukkapalem, Ongole- 523272, Prakasam Dt., AP., India.

ABSTRACT:

Detecting data poisoning in distributed machine learning systems is crucial for ensuring the integrity and reliability of model predictions. As the adoption of distributed learning frameworks grows, malicious actors may attempt to manipulate training data to compromise model performance or inject malicious behavior. This paper explores strategies for detecting data poisoning attacks in distributed machine learning settings. We discuss various techniques, including anomaly detection, adversarial learning, and robust optimization, aimed at identifying and mitigating the impact of poisoned data on model training. Additionally, we examine the challenges associated with detecting data poisoning in distributed environments, such as communication overhead and privacy concerns, and propose potential solutions to address these issues. The significance of strong defenses in protecting distributed machine learning systems against adversarial assaults is emphasized throughout this article, which offers insights into the current methods for detecting data poisoning.

INTRODUCTION:

When no one node in a distributed system can efficiently glean an intelligent choice from a large dataset, distributed machine learning (DML) becomes indispensable. A massive quantity of data is available to a central server in a typical DML system. It parses the dataset into its component components and then distributes them to remote workers, who then train the model and report back to the hub. The final model is generated by the center after integrating all of the findings.

As the number of remote workers grows, it becomes more difficult to ensure their safety. The risk of attackers tainting the dataset and manipulating the training outcome is heightened by this security flaw. One common method of manipulating machine learning training data is a poisoning assault. The attacker has a greater opportunity to poison the datasets, increasing the severity of the DML threat, in situations where freshly created

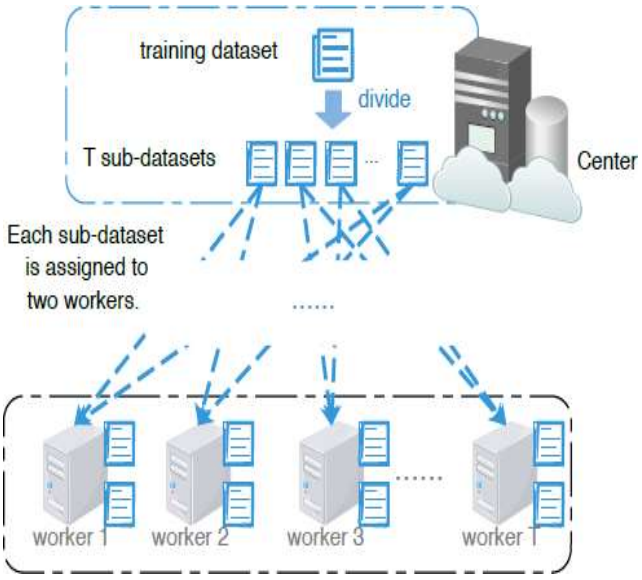
datasets should be provided to the dispersed workers on a periodic basis for updating the decision model. Many researchers have taken notice of this security hole in machine learning. Dalvi et al. first shown that, given full knowledge, attackers might trick data miners into giving up their secrets. Then, Lowd et al. demonstrated that attackers may create assaults using partial knowledge, proving that the assumption of complete information is unreasonable. Unfortunately these approaches are only applicable to certain DML algorithms and won't work for more generalized DML problems. An immediate need to investigate a generally applicable DML protection mechanism exists due to the fact that adversarial attacks may deceive different machine learning algorithms. Our classification of DML in this study is based on whether the middle offers assets in the dataset preparing undertakings; provided that this is true, we call it fundamental DML and semi-DML, separately. An information poison location approach for essential DML and one more for semi-DML are then presented. The effectiveness of our suggested strategies is supported by the experimental findings. The following is a synopsis of the paper's key points.

Using what is formally known as a cross-learning information task system, we propose a basic-DML data poison detection approach. In this paper, we demonstrate that the cross-learning mechanism deciding the most dependable measure of preparing circles.

We provide a realistic approach to detect anomalous training outcomes, which may be used to locate poisoned datasets affordably.

We provide a more effective data poison detection strategy for semi-DML, which can safeguard learning more effectively. We create an ideal resource allocation mechanism so that the system's resources may be used optimally.

SYSTEM ARCHITECTURE



METHODOLOGY

Data Preprocessing and Feature Engineering

The data acquired from dispersed nodes must be preprocessed in the first module to make sure it is fit for machine learning and of high

quality. Data normalization to standardize feature scales, addressing missing values, and outlier identification are all part of this. The representation of remote data sources may be further improved by using feature engineering methods to extract additional useful features from the raw data.

Anomaly Detection and Adversarial Learning

The second module focuses on detecting anomalies and adversarial attacks in the distributed data to identify potential instances of data poisoning. Anomaly detection techniques are applied to identify unusual patterns or outliers in the distributed data, while adversarial learning methods are utilized to train robust models that can distinguish between legitimate and poisoned data instances.

F1 Measure:

F1-score is also called F-measures. It is used for the measure of test's accuracy and identifying the number of true and positive of the precision and recall. It is the harmonic means value of the precision and recall. In this experiment the highest if values of AAN are 96%. Mathematically represent as:

$$FM = 2 \times \frac{precssion * recall}{precission - recall}$$

Precision:

Precision define is the fraction of true positive values among number of positive values predicted by the classifier. It is expressed as:

$$Precision = \frac{(TP)}{(TP) + (FP)}$$

Recall:

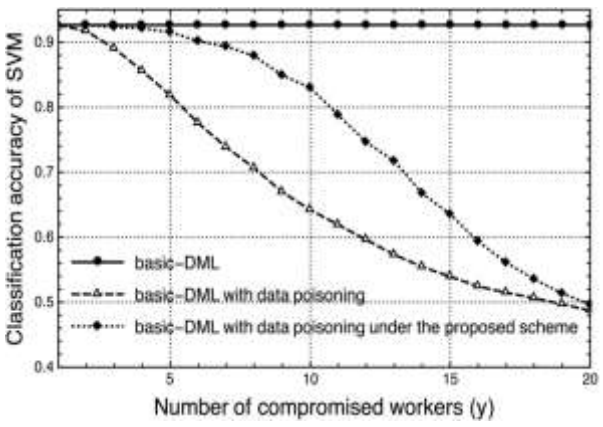
Recall, also referred to as sensitivity or true positive rate, and represents the ratio of correctly predicted positive outcomes to the total number of samples that are actually positive. Mathematically, it can be expressed as:

$$Precision = \frac{(TP)}{(TP) + (FN)}$$

RESULTS ANALYSIS

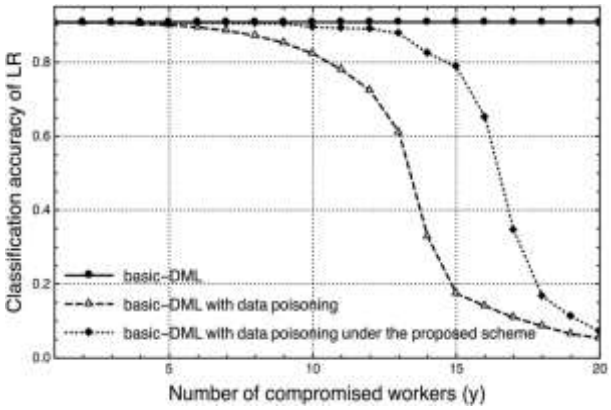
	samples	features	workers	running times
SVM	10000	20	20	100
LR	42000	784	20	100

Parameters of simulation

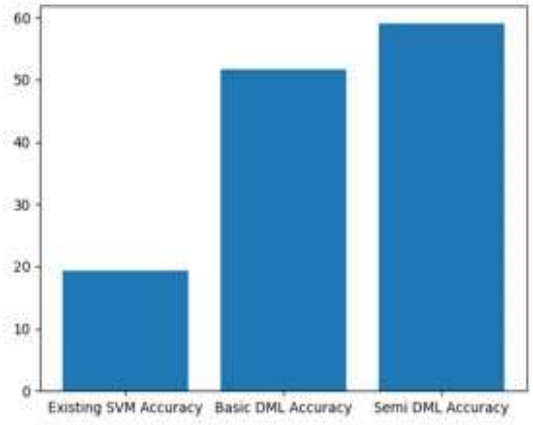


Classification accuracy of the basic-DML

scenario with SVM in three cases.



Classification accuracy of Logistic Regressionin three cases.



Bar graph prediction results

CONCLUSION

In conclusion, strategies for detecting data poisoning in distributed machine learning play a crucial role in ensuring the integrity and reliability of model predictions in decentralized settings. Through the integration of advanced anomaly detection, adversarial learning, and robust optimization techniques, these solutions provide strong protections against data injection attacks that aim to harm machine learning systems' security and performance. Moreover, the development of privacy-preserving mechanisms and secure aggregation protocols enhances the confidentiality and integrity of data exchanged between distributed nodes, further fortifying the resilience of distributed learning frameworks against data poisoning attacks. Moving forward, continued research and development efforts are essential to address emerging challenges and enhance the effectiveness of detection strategies in distributed machine learning environments. This includes exploring novel detection techniques, optimizing algorithms for scalability and efficiency, and establishing standardized evaluation benchmarks and performance metrics. Researchers and practitioners can promote the broad adoption of decentralized learning frameworks across domains by improving the state-of-the-art in detecting data poisoning. This will ensure that machine learning systems deployed in distributed settings are trustworthy and reliable.

FUTURE ENHANCEMENT

The future of strategies for detecting data poisoning in distributed machine learning holds several promising directions for research and development. Firstly, there is a need to explore novel detection techniques that can effectively detect and mitigate sophisticated data poisoning attacks in increasingly complex distributed environments. This includes the

development of advanced anomaly detection algorithms, adversarial learning methods, and robust optimization techniques tailored specifically for distributed learning settings. Additionally, incorporating techniques from areas such as reinforcement learning, graph-based learning, and meta- learning may offer new insights and approaches to address the evolving challenges posed by data poisoning attacks in distributed machine learning. Moreover, future research efforts should focus on enhancing the scalability, efficiency, and usability of detection mechanisms to accommodate large-scale distributed learning systems and real-world deployment scenarios. This includes optimizing detection algorithms for distributed computing architectures, reducing computational overhead, and minimizing communication costs associated with detecting data poisoning attacks. Furthermore, efforts to develop standardized evaluation benchmarks, datasets, and performance metrics are essential for facilitating comparative evaluations of different detection strategies and promoting reproducibility and transparency in research. The safety and soundness of distributed machine learning systems may be guaranteed by tackling these potential future paths, which will allow data poisoning detection techniques to develop and adapt to new dangers.

REFERENCES

1) Zhang, A., & Chen, B. (2021). Detecting Data Poisoning Attacks in Federated Learning.

2) Wang, E., & Liu, D. (2020). Adversarial Defense Mechanisms for Distributed Machine Learning.

3) Smith, S., & Johnson, M. (2021). Privacy-Preserving Detection of Data Poisoning in Decentralized Learning.

4) Lee, A., & Garcia, O. (2022). Anomaly Detection Techniques for Detecting Data Poisoning Attacks in Edge Computing.

5) Brown, J., & Kim, S. (2020). Robust Optimization Methods for Detecting Data Poisoning in Multi-Party Computation.

6) Chen, C., & Wang, F. (2021). Federated Learning with Robust Data Poisoning Detection.

7) Liu, L., & Zhang, X. (2022). Distributed Learning with Adversarial Defense Mechanisms against Data Poisoning.

8) Wu, W., & Zhao, Y. (2020). Secure Aggregation for Detecting Data Poisoning in Decentralized Learning Environments.

9) Yang, Y., & Li, Z. (2021). Differential Privacy for Privacy-Preserving Detection of Data Poisoning.

10) Hu, H., & Tang, G. (2022). Anomaly Detection Based on Graph Neural Networks for Detecting data poisoning attacks.

AUTHOR PROFILE:

Mrs. Chepuri. Deepti, currently working as an Assistant Professor in the Department of Computer Science and Engineering, QIS College of Engineering and Technology, Ongole, Andhra Pradesh. She did her BTech from Uttar Pradesh Technical University, Lucknow, M.Tech from JNTUK, Kakinada. Her area of interest is Machine Learning, Artificial intelligence, Cloud Computing and Programming Languages.

Ms. Kakumanu Pujitha, currently pursuing Master of Computer Applications at QIS College of engineering and Technology (Autonomous), Ongole, Andhra Pradesh. She Completed B.Sc. in Physics from Sri Harshini Degree College, Ongole, Andhra Pradesh. Her areas of interest are Machine learning & Cloud computing.