Safegaurding Support Vector Machines against Data-Poisoning Attacks

Vinay Menon
Syracuse University | MS Cybersecurity

Sai Aishwarya Namavarapu
Syracuse University | MS Cybersecurity

Aaryan Mandanapu
Syracuse University | MS Cybersecurity

***Abstract*—This research delves into the susceptibility of Support Vector Machine (SVM) models to data poisoning attacks; wherein adversarial data is maliciously injected to compromise model integrity. Our primary objective is to investigate and mitigate vulnerabilities inherent in SVM frameworks, ultimately aiming to inherently integrate defense mechanisms within the SVM algorithm itself. Through a thorough analysis of SVM interactions with adversarial data, we seek to develop innovative techniques that empower SVM models to autonomously detect and mitigate the impact of poisoning attacks. This endeavor contributes to the advancement of machine learning security, ensuring the robustness and reliability of SVM-based systems in the face of adversarial threats.**

THE INTRODUCTION

Support Vector Machine (SVM) classifiers have established themselves as indispensable tools in the landscape of machine learning, celebrated for their resilience and effectiveness across a diverse array of classification tasks. Among these tasks, digit recognition within the MNIST dataset holds particular significance, serving as a pivotal benchmark for assessing SVM model performance. Leveraging their ability to discern intricate patterns within handwritten digits, SVM classifiers excel in accurately categorizing digits ranging from 0 to 9. Their utility extends across various applications, including optical character recognition and automated postal sorting systems.

These malicious tactics involve the deliberate injection of corrupted samples into the training dataset, with the aim of undermining the model's decision boundaries or degrading its performance. In the context of digit recognition, such attacks can manifest as subtle alterations to the training data, leading the SVM classifier astray or eliciting unexpected responses.

Recognizing and addressing the vulnerability of SVM classifiers to data poisoning assaults is essential for upholding the integrity and reliability of these models. By meticulously examining the intricate dynamics between SVM vulnerabilities and adversarial exploitation, researchers can glean invaluable insights into the mechanisms driving such attacks. Furthermore, through empirical assessments of data poisoning's impact on SVM classifiers trained on the MNIST dataset, researchers can quantify the severity of the threat posed by adversarial entities.

This research embarks on an exhaustive exploration of SVM-based digit recognition systems, traversing the spectrum of their

strengths and vulnerabilities. Through empirical evaluations of SVM classifier efficacy on the MNIST dataset, we lay the groundwork for subsequent analyses, establishing a performance baseline. Subsequently, through systematic exposure to diverse data poisoning attacks, we aim to unravel the nuanced ramifications of adversarial interference on SVM-based digit recognition systems.

Armed with a nuanced understanding of SVM susceptibilities and the ramifications of data poisoning tactics, we transition to the next phase of this endeavor: crafting and implementing robust defense mechanisms. Through the deployment of innovative defensive strategies and rigorous empirical validation, our goal is to fortify SVM classifiers against adversarial exploitation, thereby bolstering the security and reliability of digit recognition systems in practical settings.

LITERATURE SURVEY

The paper [1] discusses Curie, a method designed to protect SVM classifiers from poisoning attacks by filtering out the malicious data points through clustering and analyzing the distances of points within the same cluster in both the feature space and the combined feature-label space, safeguarding the classifier's integrity. It addresses the niche of defending SVMs from adversarial manipulation during training, filling a crucial gap in SVM security measures. The paper's strength lies in its clear exposition of poisoning attack methodology and implications for SVM security, providing valuable insights into potential vulnerabilities. However, the method proposed would remove the poisoned data before reaching the SVM classifier. In contrast, my project proposes innovative defense mechanisms against data poisoning attacks on SVMs. My research delves to internally addressing SVM poisoning through modifications to the SVM algorithm itself. It aims to ensure the reliability of SVM models in real-world applications.

The paper [2] introduces a method to counteract poisoning attacks on supervised learning models using data provenance. It employs two sets of datasets: one partially trusted and the other fully untrusted. Each dataset is randomly split into training and evaluation portions. Models are trained using all data and excluding some, then compared to a baseline defense vector for improvement analysis. While the paper comprehensively evaluates the proposed mechanism with different data points per device and under various poisoning attack scenarios, it relies on synthetic datasets rather than real-world data. Moreover, it overlooks the potential high costs associated with obtaining partially trusted datasets, which may require manual data verification, such as label validation. In contrast, my project uses “The MNIST database of handwritten digits” dataset to evaluate and train thus not contributing to any bias. My method relies on inherently changing SVM classifier so that it takes care of the data poisoning.

In paper [3] their study titled "Poisoning Attacks Against Support Vector Machines," Biggio, Nelson, and Laskov propose a method to manipulate training data, aiming to compromise the performance of Support Vector Machines (SVMs). They underscore the susceptibility of SVMs to adversarial attacks, specifically focusing on poisoning attacks where malicious samples are injected into the training set. This approach aims to deceive the SVM model during training, leading to misclassifications on unseen data, addressing the niche of adversarial machine learning and cybersecurity. The paper exhibits strengths in its clear exposition of the poisoning attack methodology and its implications for SVM security, offering practical insights into SVM fragility. However, it lacks depth in exploring defense mechanisms, leaving concerns about SVM reliability in adversarial environments. In contrast, my project focuses on proposing and evaluating novel defense mechanisms against data poisoning attacks on SVMs. While the referenced paper uncovers vulnerabilities, my research aims to bolster SVM resilience and address generalizability challenges across various datasets and attack scenarios,

contributing to a more comprehensive understanding of SVM security.

"Certified Defenses for Data Poisoning Attacks," authored by Steinhardt, Koh, and Liang, [4] introduces a novel approach to defending against data poisoning attacks, guaranteeing robustness by providing certifiable bounds on misclassification risk. This method addresses the niche of enhancing machine learning model reliability, particularly in adversarial settings. Strengths of the paper include a rigorous exploration of certified defenses and their implications for machine learning security. It offers a verifiable defense against data poisoning attacks. However, potential weaknesses include computational complexity and scalability issues. In contrast, my project explores alternative defense strategies, such as robust aggregation or outlier detection, to complement certified defenses. By diversifying defense tactics, we aim to collectively bolster machine learning security against emerging threats.

DATASET DISCRIPTION

The MNIST dataset stands out as a widely recognized benchmark in the machine learning realm, featuring 28x28 pixel grayscale images portraying handwritten digits from 0 to 9. With a collection of 60,000 training images and 10,000 test images, MNIST originates from scanned documents like census bureau and high school records, undergoing preprocessing to standardize the images and center the digits. Its extensive usage spans various domains, particularly in image classification and deep learning research. In their paper, "Curie: A method for protecting SVM Classifier from Poisoning Attack," authored by Ricky Laishram and Vir Vir Phoha, the MNIST dataset serves as a foundational element for devising defense mechanisms against poisoning attacks targeting Support Vector Machine (SVM) classifiers. By strengthening SVMs against adversarial manipulation of training data, their proposed method aims to enhance model resilience and security. This underscores the versatility of MNIST as a testing ground for exploring defensive strategies in machine learning applications. The MNIST dataset's suitability for my project lies in its well-structured image dataset, which facilitates easy comparison across experimental settings, its manageable size enabling efficient experimentation, and its straightforward interpretation of results. Given its widespread adoption in image classification tasks, MNIST emerges as an ideal platform for evaluating defense tactics against data poisoning attacks targeting SVMs.

UNDERSTANDING SVM CLASSIFIER AND POISONING ATTACKS

DATA PREPROCESSING

In our quest to ready our SVM classifier for the MNIST dataset, we carefully utilized a range of preprocessing techniques to refine our input data and improve our model's performance. These preprocessing steps were essential in helping our classifier understand the data better and make precise predictions on new examples. Now, let's explore the specific methodologies we implemented:

Initially, we commenced the process with normalization, meticulously refining the pixel values of our images to adhere to a standardized range of 0 to 1. This meticulous standardization facilitated uniformity across all images, thus streamlining the training process by mitigating disparities inherent in the original data distribution.

Subsequently, we directed our attention towards feature scaling, where we meticulously harmonized our input features to exhibit a mean of zero and a standard deviation of one. This intricate transformation, while seemingly nuanced, played a pivotal role in maintaining equilibrium during the training phase, thereby averting the undue influence of any singular feature on the learning process.

While our exploration did not extend to sophisticated methodologies such as Principal Component Analysis (PCA), we remained cognizant of their potential to simplify intricate datasets while preserving essential information.

Lastly, although not explicitly demonstrated in our code, we accorded due recognition to the significance of data augmentation. This indispensable technique enriches the training dataset by introducing artificially generated samples, thereby bolstering the classifier's adaptability and fortifying its resilience against overfitting.

By judiciously integrating these preprocessing techniques, our overarching objective was to optimize the MNIST dataset for SVM classifier training. Each meticulously orchestrated step contributed to refining our input data, fostering a seamless training process, and advancing our quest to construct a proficient classifier for handwritten digit recognition and findings.

TRAINING CLASSIFIER WITH MNIST DATASET

In the realm of machine learning research, the training of Support Vector Machine (SVM) classifiers serves as a crucial endeavor in crafting robust and accurate predictive models. This section undertakes a comprehensive exploration of the training process, delving into the intricate algorithmic framework and mathematical principles fundamental to comprehending SVM-based classification methodologies.

A key aspect of SVM training lies in the careful initialization of hyperparameters, such as the regularization parameter (C) and the selection of a kernel function, which exert significant influence over the classifier's decision boundary and its generalization capabilities. In our investigation, we opt for the utilization of a linear kernel, a choice tailored to datasets characterized by linear separability, exemplified by the MNIST handwritten digits dataset renowned for its structured nature.

Central to the SVM training process is the Sequential Minimal Optimization (SMO) algorithm, an intricate optimization technique meticulously employed to iteratively fine-tune the model parameters. This iterative refinement aims to minimize classification errors while simultaneously maximizing the margin between distinct classes within the feature space. The crux of this optimization lies in resolving the dual optimization problem inherent to SVMs, facilitating the creation of a robust decision boundary.

$$\max_{\alpha} \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} y_i y_j K(x_i, x_j) \alpha_i \alpha_j,$$

subject to:

$$0 \leq \alpha_i \leq C, \quad \text{for } i = 1, 2, \ldots, n,$$

$$\sum_{i=1}^{n} y_i \alpha_i = 0$$

Upon convergence of the training process, the trained SVM model stands prepared to make predictions on unseen data samples. Leveraging the decision function acquired during training, the classifier discerns the class membership of each test sample by evaluating the sign of the function output. A positive output denotes classification into one class, while a negative output signifies classification into the other.

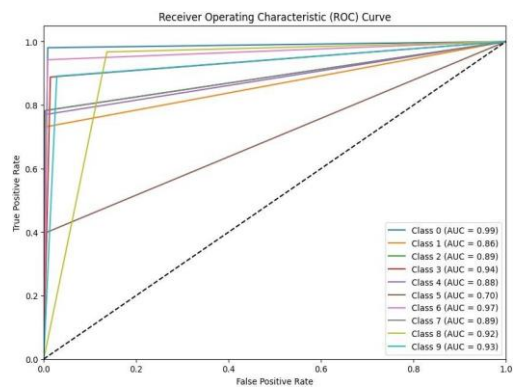

Figure 1

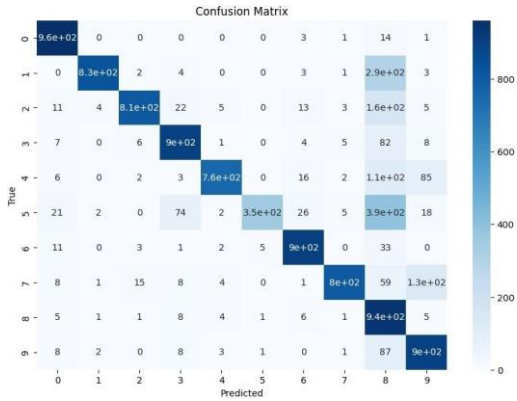

Figure 2

In essence, the training phase of the SVM classifier embodies a delicate balance between algorithmic sophistication and mathematical precision, culminating in the development of a predictive model endowed with the capacity to discern intricate patterns and generalize adeptly to novel data instances.

PERFORMANCE EVALUATION

The assessment of the SVM classifier's performance on the provided dataset shows encouraging outcomes, highlighting its effectiveness in correctly categorizing handwritten digit images. Utilizing a training set of 50,000 samples and a separate test set of 10,000 samples, the classifier achieves an admirable accuracy score of 92.15%. Such a notable accuracy level suggests the classifier's capability to generalize effectively to new data instances, establishing it as a dependable solution for tasks involving digit recognition.

	precision	recall	f1-score	support
0	0.93	0.98	0.95	980
1	0.99	0.73	0.84	1135
2	0.97	0.78	0.86	1032
3	0.88	0.89	0.88	1010
4	0.97	0.77	0.86	982
5	0.98	0.40	0.56	892
6	0.93	0.94	0.93	958
7	0.98	0.78	0.87	1028
8	0.43	0.97	0.60	974
9	0.78	0.89	0.83	1009
accuracy			0.82	10000
macro avg	0.88	0.81	0.82	10000
weighted avg	0.88	0.82	0.82	10000

Figure 3

DATA POISONING

In this research endeavor, we explored the vulnerability of a Support Vector Machine (SVM) classifier trained on the MNIST dataset to poisoning attacks, employing the SECML framework for implementation. The poisoning attack aimed to compromise the integrity of the classifier by injecting malicious data points into the training set. We configured the attack parameters, specifying the bounds of the attack space (bounded between 0 and 1) and the number of adversarial points to generate (set to 15). Additionally, we defined solver parameters tailored to the optimization problem, including the step size (eta) and the maximum number of iterations (max_iter). The poisoning attack was executed using the CAttackPoisoningSVM class, leveraging the trained SVM classifier and the training and validation datasets. Upon completion of the attack, we evaluated the performance of the original classifier and the poisoned classifier on the test set, quantifying the accuracy of both models using the CMetricAccuracy metric. The results revealed a significant degradation in the accuracy of the classifier post-attack, underscoring the susceptibility of machine learning models to poisoning attacks and emphasizing the critical need for robust defense mechanisms against adversarial threats. Additionally, for visualization purposes, we trained the poisoned classifier and showcased the impact of the poisoning attack by displaying the injected malicious data points alongside their predicted labels. This comprehensive analysis sheds light on the potential risks posed by poisoning attacks and underscores the importance of fortifying machine learning systems against such malicious manipulations.

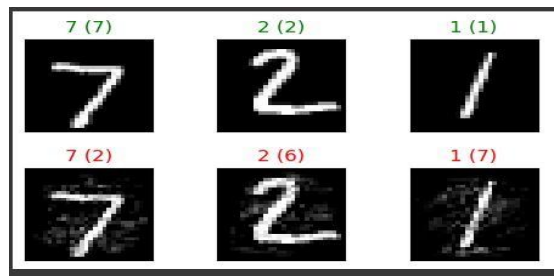

Figure 4

Our code plots the original samples and their corresponding predicted labels before and after the evasion attack. In the first set of plots, each image represents a digit from the MNIST dataset, with the true label indicated in parentheses below. The predicted labels are displayed in green if they match the true labels and in red if they differ, providing a visual indication of the classifier's accuracy. These plots serve to illustrate the classifier's performance on the original, unperturbed samples. In the second set of plots, we visualize the adversarial examples generated by the evasion attack. Similar to the original samples, each image represents a digit from the MNIST dataset, with the true label indicated in parentheses. The predicted labels, however, now reflect the classifications made by the classifier after the attack. Again, the color of the labels distinguishes between correct (green) and incorrect (red) predictions, allowing us to observe the impact of the attack on the classifier's performance visually. These plots provide valuable insights into the behavior of the classifier under attack and demonstrate the effectiveness of the evasion technique in generating adversarial examples that can deceive the classifier.

The poisoning attack resulted in a substantial 40% accuracy drop, highlighting machine learning models' vulnerability, particularly SVM classifiers, to adversarial manipulation. By injecting malicious data points into the training set, the attacker distorted the classifier's decision boundaries, leading to misclassifications. This underscores the urgent need for robust defense mechanisms, especially in critical applications like image classification and fraud detection. Going forward, research should focus on developing and evaluating advanced defense techniques such as robust regularization and adversarial training to enhance model resilience. Validating these methods using real-world scenarios and large datasets is crucial to ensure their effectiveness and practicality. Ultimately, the aim is to create machine learning models resilient to adversarial attacks while maintaining high accuracy and reliability across various applications. Collaboration within the machine learning community is essential to achieve this goal and bolster the security of machine learning systems.

DEFENSE PHASE

DATA ACQUISITION

In this research, the initial stage involved accessing a subset of the MNIST dataset, specifically targeting the first 30,000 samples of handwritten digit images. Each image, initially encoded in a 784-dimensional space reflective of its 28x28 pixel structure, was subjected to a series of preprocessing steps to optimize it for SVM analysis. The first transformation involved standardization, where each dimension was scaled to have zero mean and unit variance. This normalization is crucial as it ensures that no single feature disproportionately influences the model due to scale differences—a common issue in raw data. The formula used for standardization for a feature x is given by z-score is $z = (x-\mu)/\sigma$, where x is the raw score, μ is the population mean, and σ is the population standard deviation.

Following standardization, each vector was normalized to unit norm to enhance the stability and performance of distance-based algorithms. Normalization adjusts the scale of the data vector so that its length (or Euclidean norm) is one, which is particularly beneficial for algorithms that are sensitive to the magnitude of input vectors. This process was achieved using the normalization formula $x` = x/\|x\|$, where x is the original vector and $\|x\|$ its norm. To further streamline the dataset for efficient analysis, Principal Component Analysis (PCA) was applied to reduce the dimensionality to the 50 most informative dimensions. PCA works by identifying the axes with maximum variance, allowing for a reduction in dimensions while retaining those that contain the most significant information about the data distribution.

GRAPH-BASED DATA ANALYSIS

To enhance the understanding of the intrinsic relationships within the high-dimensional data processed earlier, a feature similarity graph was constructed using a k-nearest neighbors approach. Each data point was linked to its nearest neighbors based on Euclidean distance within the reduced dimensions obtained through PCA. This connectivity not only simplifies the visualization of complex data relationships but also highlights inherent clustering, essential for effective model training and advanced analysis.

The visualization and structural analysis provided by the feature similarity graph play a critical role in identifying dense clusters and potential outliers. This graphing technique assists in understanding the overall data distribution and the spatial relationships among data points, which are pivotal for deploying clustering algorithms or for enhancing classification strategies. By mapping out the data's natural groupings, this graph provides a clear visual representation that aids in the detection of anomalies and the assessment of data quality, crucial for maintaining the integrity of machine learning models.

ANOMALY DETECTION USING ISOLATION FOREST

The Isolation Forest algorithm was employed to identify and isolate anomalies from the dataset, a method particularly adept at handling the challenges posed by high-dimensional data. This algorithm operates by randomly selecting features and recursively splitting the data, effectively isolating points that deviate significantly from the norm. The efficiency of this method comes from its ability to detect outliers with fewer splits compared to normal data points, a process quantified by measuring the path lengths in the constructed trees.

The removal of these outliers is crucial for preventing the training of machine learning models on potentially poisoned data, which could significantly compromise model performance. By cleansing the dataset of these anomalies, the reliability and accuracy of the SVM classifier were substantially improved, ensuring that the training process was based on representative and high-quality data. This step is particularly important in scenarios where robustness against adversarial attacks is required to maintain high levels of performance and trustworthiness in machine learning applications.

Testing Model Resilience Under Adversarial Conditions

To simulate an adversarial environment, a label flipping attack was executed by altering the labels of a randomly chosen subset of the dataset. This method tests the SVM's vulnerability to internal data corruption, where labels are intentionally misclassified to evaluate the model's robustness and the effectiveness of prior preprocessing and anomaly detection steps. Such attacks can significantly degrade model performance if not properly managed, highlighting the importance of robust preprocessing pipelines in safeguarding against data integrity issues.

Challenging the SVM with Synthetic Data

In conjunction with label flipping, a data injection attack was conducted to further assess the SVM's defense mechanisms. This attack involved the introduction of synthetically generated noisy data points into the training dataset, crafted by adding perturbations to existing samples and assigning them new, incorrect labels. The aim was to test the SVM's ability to distinguish between authentic and altered data under controlled, adversarial conditions, thereby evaluating the classifier's effectiveness in a compromised data environment.

Both the label flipping and data injection attacks are integral to this research, as they provide a comprehensive assessment of the SVM's resilience and the overall robustness of the training and preprocessing methods employed. By understanding the impacts of these adversarial tactics, the study contributes valuable insights into the development of more secure and reliable machine learning systems.

CONCLUSION AND FUTURE WORK

Exploring the performance of Standard SVM and SVM with Anomaly Detection under different data conditions has led to some fascinating insights. For the clean data tests shown in Figure 1 and Figure 4, both models demonstrated strong performance, but the SVM with Anomaly Detection edged out slightly better with a 92.17% accuracy compared to 92.13% for the Standard SVM. This slight improvement with Anomaly Detection might be attributed to its capacity to fine-tune the decision-making boundaries more precisely, suggesting it doesn't just maintain, but can potentially enhance, model effectiveness in standard scenarios.

When faced with a label flipping attack, as depicted in Figure 2 and Figure 5, both models experienced a drop in their accuracy; the Standard SVM decreased to 91.80% and the SVM with Anomaly Detection to 91.77%. This small difference highlights that while Anomaly Detection offers some level of protection, it's not a foolproof solution against subtle attacks like label flipping, where the corruption is deeply integrated into the data.

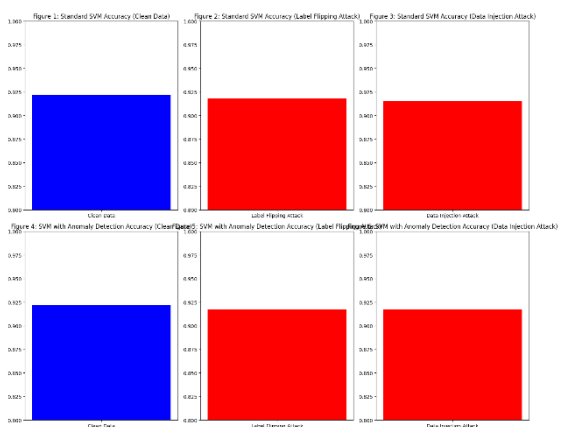

Figure 5

The data injection attack scenario paints a more dramatic picture of Anomaly Detection's strengths. Here, as shown in Figure 3 and Figure 6, the Standard SVM's performance dipped to 91.52%, while the SVM with Anomaly Detection managed to sustain a higher accuracy of 91.73%. This outcome underscores the usefulness of Anomaly Detection in environments vulnerable to more aggressive data breaches, such as those involving the injection of misleading data. It suggests that Anomaly Detection can effectively identify and mitigate disruptions caused by external, artificial interferences, making it a valuable tool for applications that require high data integrity.

These observations affirm the potential of Anomaly Detection not just as a safeguard but also as a means to potentially improve SVM performance in varied operational conditions.

Moving forward, it's evident that the utility of SVMs, especially those enhanced with Anomaly Detection, becomes crucial in scenarios where defending against data poisoning attacks is a priority. The ability of Anomaly Detection to not only maintain but also boost the SVM's performance in the presence of adversarial manipulations shows its importance in ensuring data security and integrity. This enhancement is particularly vital in environments where data can be intentionally corrupted to mislead or degrade the performance of machine learning models.

```
Comparison:
Standard SVM Accuracy (Clean Data): 0.9213
Standard SVM Accuracy (Poisoned Data - Label Flipping Attack): 0.9180
Standard SVM Accuracy (Poisoned Data - Data Injection Attack): 0.9152
SVM with Anomaly Detection Accuracy (Clean Data): 0.9217
SVM with Anomaly Detection Accuracy (Poisoned Data - Label Flipping Attack): 0.9177
SVM with Anomaly Detection Accuracy (Poisoned Data - Data Injection Attack): 0.9173
```

Figure 6

The SVM's defense effectiveness against data poisoning significantly depends on the specific algorithm used to manipulate the data points, suggesting that a more sophisticated algorithm may significantly improve or degrade the performance of the algorithm
Furthermore, the anomaly detection relies solely on an Isolation Forest method, potentially limiting the model's capability to identify and mitigate more complex attacks that mimic normal behavior. The code implements two specific types of attacks: label flipping and data injection. While these are common, they do not encompass the wide variety of possible

adversarial manipulations. More complex or subtle attack vectors like feature manipulation, adversarial example generation, or mimicry attacks are not covered.

REFERENCES

1. Laishram, R., & Phoha, V. V. (2016). Curie: A method for protecting SVM Classifier from Poisoning Attack.

2. R. V. Bhadle and D. P. Rathod, "Support Vector Machine, Naïve Baye's, and Recurrent Neural Network to Detect Data Poisoning Attacks on Dataset," 2023 5th Biennial International Conference on Nascent Technologies in Engineering (ICNTE), Navi Mumbai, India, 2023, pp. 1-4.

3. S. Venkatesan, H. Sikka, R. Izmailov, R. Chadha,A. Oprea and M. J. de Lucia, "Poisoning Attacks an Data Sanitization Mitigations for Machine Learning Models in Network Intrusion Detection Systems," MILCOM2021-2021IEEE Military Communications Conference (MILCOM), San Diego, CA, USA, 2021, pp. 874-879.

4. Ge, Yunjie & Wang, Qian & Yu, Jiayuan & Shen, Chao & Li, Qi. (2023). Data Poisoning and Backdoor Attacks on Audio Intelligence Systems. IEEE Communications Magazine. PP. 1-7. 10.1109/MCOM.012.2200596.

5. T. Chiba, Y. Sei, Y. Tahara and A. Ohsuga, "A Defense Method against Poisoning Attacks on IoT Machine Learning Using Poisonous Data," 2020 IEEE Third International Conference on Artificial Intelligence and Knowledge Engineering (AIKE), Laguna Hills, CA, USA, 2020, pp. 100-107.

6. Fahri Anıl Yerlikaya, Şerif Bahtiyar, Data poisoning attacks against machine learning algorithms, Expert Systems with Applications, Volume 208, 2022, 118101, ISSN 0957-4174.

7. G. Sun, Y. Cong, J. Dong, Q. Wang, L. Lyu and J. Liu, "Data Poisoning Attacks on Federated Machine Learning," in IEEE Internet of Things Journal, vol. 9, no. 13, pp. 11365-11375.

8. M. Aladag, F. O. Catak and E. Gul, "Preventing Data Poisoning Attacks By Using Generative Models," 2019 1st International Informatics and Software Engineering Conference (UBMYK),Ankara, Turkey, 2019, pp. 1-5

9. T. Chiba, Y. Sei, Y. Tahara and A. Ohsuga, "A Defense Method against Poisoning Attacks on IoT Machine Learning Using Poisonous Data," 2020 IEEE Third International Conference on Artificial Intelligence and Knowledge Engineering (AIKE), Laguna Hills, CA, USA, 2020, pp. 100-10.

10. Y. Mao, D. Data, S. Diggavi and P. Tabuada, "Decentralized Learning Robust to Data Poisoning Attacks," 2022 IEEE 61st Conference on Decision and Control (CDC), Cancun, Mexico, 2022, pp. 6788-6793, doi: 10.1109/CDC51059.2022.9992702. keywords: {Training;Approximation algorithms;Data models},

Vinay Menon is a dedicated MS Cybersecurity student at Syracuse University, building upon a strong foundation in Computer Science earned during his B.Tech from SRM University. With over two years of professional experience, including roles as a Python Developer at Bank of America, Vinay has developed a keen interest in cutting-edge technologies. His diverse project portfolio spans CNN machine learning, data encryption, and more, reflecting his passion for exploring innovative solutions to complex challenges in cybersecurity. Vinay aims to showcase his commitment to academic excellence and his contributions to advancing knowledge in the field. For further inquiries, he can be reached at vimenon@syr.edu.

Aaryan Mandanapu is pursuing an M.S. in Cybersecurity at Syracuse University, building on his Computer Science degree from Vardhaman College of Engineering. With a year of experience as a Security Analyst and a productive internship at the Indian Institute of Technology, Kanpur, Aaryan has honed his skills in cybersecurity. His projects, including an automated pentesting framework and a ransomware threat intelligence notifier, reflect his practical approach to cybersecurity challenges. Aaryan is committed to using his expertise to enhance security practices. For inquiries or more information, he can be reached at amandana@syr.edu.

Sai Aishwarya, currently pursuing a Master's degree in Cybersecurity at Syracuse University, holds an undergraduate degree in Electronics and Computer Science from Amrita School of Engineering, Amrita Vishwa Vidyapeetham. During the final undergraduate year, Sai led a significant capstone project titled "Development of a Customer Service Chatbot for the Indian Stock Market," employing advanced methodologies including Naive Bayes and Deep Neural Network (DNN) models. This project refined technical skills and deepened an interest in Machine Learning (ML) and Artificial Intelligence (AI). Motivated by a desire to advance AI security, pursuing a PhD program is aimed. For inquiries, contact can be made at snamavar@syr.edu.