Adaptive Anonymization for Privacy-Preserving Machine Learning: A Dynamic Approach to Secure Data Modeling

Waad Saud AlAnazi† and Jalal Suliman Alowibdi††

†Computer Science and Engineering Department, University of Hail, Hail, Saudi Arabia
†Computer Science and AI Department, University of Jeddah, Jeddah, Saudi Arabia
††Computer Science and AI Department, University of Jeddah, Jeddah, Saudi Arabia

Summary
The fast proliferation of machine learning (ML) in data sensitive areas like healthcare, finance and cybersecurity has heightened the existence of privacy pre-serving strong privacy-mechanisms. The conventional anonymization methods, including k-anonymity, l-diversity, and t-closeness, fail to provide a balance be-tween privacy and data utility, resulting in poor model performance or the attack of inference. In response to these shortcomings, this paper will present an anonymization architecture that is more adaptive and changes the privacy levels dynamically based on the sensitivity of data and model needs. The proposed method is also learned with optimal anonymization parameters unlike the existing generalization-based methods, which rely on the concept of static generalization without data confidentiality. Experimental assessments with real-world samples and a wide variety of ML algorithms indicate that adaptive anonymization can be evaluated at 12% accuracy better than traditional solutions without decreasing the ability to resist linkage and attribute disclosure attacks. Comparative performances have shown that deep learning models are especially resistant to adaptive anonymization, maintaining more than 93 percent F1-score on anonymized data. These findings indicate that adaptive anonymization provides a privacy-compliant, scalable and task-aware method of data modeling to build a more trustful, transparent and resilient privacy preserving machine learning system.

Keywords:
Adaptive anonymization; Privacy-preserving machine learning; Differential privacy; Cybersecurity; K-Anonymity

1. Introduction

In the era of data-driven intelligence, machine learning (ML) models are becoming highly reliant on large volumes of personal and sensitive data to attain high predictive accuracy. On the one hand, the use of ML in healthcare to diagnose diseases has brought revolutionary opportunities to the healthcare system but, on the other hand, has become an unprecedented threat to data privacy. The problem is that it is necessary to balance between personal privacy and the informative quality of data, which is a tension that old anonymization tools have not managed to balance.

Traditional anonymization models like k-anonymity, l-diversity and t-closeness, though useful in structured settings, tend to provide a fixed transformation, which reduces the utility of the data or could not keep pace with the dynamic properties of data. These fixed anonymization strategies are readily defeated as adversarial methods of anonymization develop, causing information leakage or the risk of re-identification. Besides, differential privacy is mathematically strong, but it may add too much noise, which may compromise the accuracy of ML, particularly in high-dimensional, complex data [1].

The study highlights the significance of data owners' role in data protection by focusing on their viewpoint as they maintain data sets and algorithms throughout the ma-chine learning lifecycle. It talks about how different ML architectures, roles of parties involved, and data processing phases affect the threats to data protection [1]. For sensitive data to be published and analyzed safely, anonymization techniques are crucial, particularly when the data includes personally identifying information. To guarantee data confidentiality and anonymity, a number of privacy models have been put forth, concentrating on various database attribute types, including identifiers, pseudo-identifiers, as well as delicate attributes. Identifiers must be eliminated during anonymization, and hierarchical structures can be used to generalize quasi-identifiers. Four distinct anonymization strategies will be used to a well-known dataset in this work, and their effects on data utility will be evaluated through an analysis utilizing traditional machine learning models [2]. The study focuses on techniques that keep information private while also fitting the requirements of security in machine learning. Particularly, we study and assess many anonymization techniques to determine how well they can maintain a safe level of privacy with the same prediction results.

In general, this study aims to contribute to the field of cybersecurity by providing a comprehensive analysis of privacy-preserving data anonymization techniques for secure ML. It offers insights into the challenges associated with privacy preservation, explores different anonymization techniques, and proposes an evaluation

Manuscript received May 5, 2026
Manuscript revised May 20, 2026
https://doi.org/10.22937/IJCSNS.2026.26.5.3

framework for assessing their effectiveness. By promoting the use of privacy-preserving techniques, this research strives to enhance the security and privacy of ML systems in the cybersecurity domain [3].

The main contributions of this study are as follows:

- Discuss some of the key issues related to current anonymization techniques need-ed for modern machine learning.
- Propose a plan for incorporating privacy-enhancing anonymization into any se-cure machine learning process.
- Compare the amount of privacy we can protect with the quality of models' use of real-world data.

The study provides advice on how to choose the appropriate data anonymization method based on the task and data type.

The paper is structured as follows: section II presents a sample of the previous re-search work related to privacy-preserving data anonymization techniques. Section III shows the privacy-protecting methodology. Section IV briefly describes the privacy preserving with machine learning. Section V presents the anonymization methods applied. Section VI explains the analyses, experimental results and evaluation metrics. Finally, section VII summarizes the methodology work and the outcomes achieved.

2. Related work

The privacy-preserving data anonymization technique has been explored in a number of studies to provide security to machine learning application. The classical techniques which have been broadly applied are k-anonymity, l-diversity and t-closeness to hide the sensitive identifiers. Though the techniques are applicable in minimizing the direct re-identification risk, they usually affect data utility negatively because of predetermined generalization scheme and are incapable of handling cur-rent attacks like inference or linkage attacks in more intricate data.

A strong alternative that has been proposed is differential privacy, which comes with formal guarantees of leakage of privacy. Nonetheless, its useful application is often limited by privacy-model performance trade-offs - especially in the high-dimensional data, where adding noise significantly decreases the learning accuracy. Alternative recent works combine privacy-preserving mechanisms with federated learning or homomorphic encryption. Although these solutions provide better data protection by construction, they typically have high computational overhead or Communication overhead, thus cannot be used in resource-constrained settings or real-time scenarios [4, 5].

The proposed evaluation frameworks for privacy-preserving data anonymization techniques have been a focal point in recent research. These frameworks are de-signed to quantify the trade-offs between data utility and privacy preservation, offering a systematic approach for assessing the effectiveness of different anonymization techniques in cybersecurity datasets [6].

In order to improve open data privacy and anonymity, this literature study com-pares the efficacy of the Content-Based Data Masking (COBAD) approach to more traditional methods like K-anonymity, L-Diversity, and T Closeness.

One popular data anonymization technique that has been created to protect privacy in the context of data distribution is the K-anonymity approach [7]. state that the fundamental principle of K-anonymity is that each entity in the dataset cannot be distinguished from at least 'k-1' other entities in the same dataset. To achieve anonymity in a dataset, generalization or suppression methods are typically applied to particular identifying features, such as age, zip code, or other personal characteristics. The ability of K-anonymity to preserve an equilibrium between information use and privacy makes it an important topic. K-anonymity ensures that each member of a group of 'k' people is kept anonymous, allowing for extensive data analysis without compromising the privacy of personal data [8]. It is important to remember that K-anonymity might not offer total protection against attribute revelation, even though it is excellent at preventing identity exposure. To overcome this restriction, more sophisticated methods like L-diversity and T-closeness were created.

The generation or sanitization of models for privacy-preserving dissemination in a worldwide context by a trustworthy data-holding authority has been the subject of extensive research [9, 10]. The ability of an untrustworthy data scientist to conduct analysis and create models on privacy-protected datasets, for example, by sending a number of differentially-private requests to a trustworthy authority, has been the subject of parallel research [11].

Private multiparty ML is a similar field of study in which several parties with data would like to construct a model without exchanging data directly [12, 13]. The authors examine a situation where each partner has a local classifier that is based on personal information. Using local differential privacy, a collection of data-holding parties can then merge their local classifications into a more potent ensemble. We wish to account for the situation where a user is able to produce features but lacks sufficient data for training a classifier locally.

Recent research has concentrated on general computing or machine learning over encrypted data using fully homomorphic encryption (FHE) or secure multiparty computation (MPC) [14].

Several studies show that adaptive anonymization is about 12% more accurate than generalization approaches and still secures sensitive private data over

time. So, these techniques also provide safety against inference and linkage attacks which makes them valuable for practical situations where privacy threats may develop over time.

Table 1: Comparison of previous studies on Privacy-Preserving data anonymization and Machine Learn-ing.

Author(s) / Year	Main Focus	Methodology / Approach	Key Findings	Limitations / Gaps
A. Sharma, G. Singh, & S. Rehman (2020) [6]	Big data challenges and privacy preservation	Reviewed privacy-preserving frameworks for big data analytics, including anonymization, encryption, and access control models	Highlighted the need for balancing scalability and privacy in large-scale datasets	Lacked adaptive mechanisms suitable for dynamic or evolving data environments
Tu Z., Zhao K., Xu F., Li Y., Su L., & Jin D. (2018) [7]	Protection against semantic attacks using anonymization	Combined *k*-anonymity, *l*-diversity, and *t*-closeness for trajectory data privacy	Improved resistance to semantic and linkage attacks while preserving acceptable data utility	Focused mainly on trajectory datasets; not easily generalizable to other domains
Meden B., Emeršič Ž., Štruc V., & Peer P. (2018) [8]	Face de-identification in images	Proposed k-Same-Net, a generative deep neural network applying *k*-anonymity for face anonymization	Achieved effective visual privacy with minimal degradation of data utility for ML models	Limited to image-based data; lacks evaluation on structured datasets
Manas A. Pathak & Bhiksha Raj (2014) [9]	Privacy in statistical ML models	Developed differentially private large-margin Gaussian mixture models	Provided strong theoretical privacy guarantees without major loss in performance	Computationally intensive; not scalable for real-time or large datasets
Veeramachaneni K. et al. (2016) [10]	Standardization and privacy in MOOC data	Introduced MOOCdb — a data schema ensuring structured anonymization and standardization for education analytics	Enabled reproducibility and cross-platform data sharing while preserving privacy	Focused on MOOC (education) domain; lacks adaptability to complex ML workflows

Compared to the existing solutions, our goal is to present the adaptive anonymization techniques that can dynamically adjust the data utility and privacy based on the structure and sensitivity of the dataset. Our method learns the best level of anonymization to apply to individual attributes, unlike the static generalization techniques, and thus can preserve important predictive patterns whilst incurring minimal privacy risk. Also, we directly consider constraints on the degradation of model performance though task-aware anonymization strategies which adapt transformations according to machine learning goals.

In that way, our contribution offers a scalable and flexible framework of anonymization that strikes the balance between data protection and model usability, going beyond the traditional fixed-rule approaches and improving the applicability to the real-world machine learning systems.

3. Methods

3.1 Formal Mathematical Model

Suppose D is the initial dataset that has N records and each of the records $r_i \in D$ is represented as a set of attributes (A_1, A_2, ..., A_m). These properties can be di-vided into Sensitive Attributes (SAs) and Quasi-Identifiers (QIDs). The anonymization operation U(D', M) changes the original data D to an anonymized version D', with p = (k, l, t, ε), the privacy parameters vector.

1) Utility Objective

Utility U(D', M) is given as the predictive performance of a machine learning model M trained on anonymized dataset D'. The goal is to make the utility as high as possible, and meet the specified privacy constraints:

U(D', M) = PerformanceMetric(M(D'))

2) Privacy Constraints

In order to have multi-layered protection we include the following formal privacy definitions:

- k-Anonymity: Each equivalence class EC in D satisfies the condition that the number of records in EC satisfies the condition that that number meet the condition that number ≥ k.
- l-Diversity: Let EC be an equivalence class of the sensitive attribute SA, then entropy (SA) = log l must satisfy entropy (SA) ≥ log(l).
- Differential Privacy: The anonymization mechanism A satisfies (ε, δ)- Differential Privacy (DP) if for any two neighboring datasets D and D_1: Pr[A(D) ∈ S] ≤ e^ε Pr[A(D_1) ∈ S] + δ.

3.2 Optimization Objective

The adaptive mechanism finds the best parameter set p* by solving the following multi-objective optimization problem:

$p^* = \arg\max_p [\alpha \cdot U(A(D, p), M) - \beta \cdot L(A(D, p), D)]$

L is the information loss (e.g., Generalization Loss), and α, β are weighting coefficients between utility and privacy.

3.3 Adaptive Anonymization Algorithm

Task-Aware Adaptive Anonymization: Algorithm 1.

1. Input: Dataset D, Model M, Thresholds (U min P min)
2. Output: Optimal Anonymized Dataset Dopt.
3. Initialize: p = (k_min, l_min, t_max, ε_max)
4. While iteration < Max_Iterations:
5. D' = Apply_Anonymization(D, p)
6. Utility = Evaluate_Model(M, D')
7. Privacy_Risk = Measure_Disclosure_Risk(D')
8. In case Utility = U min and Privacy=Risk = P min:
9. Return D'
10. // Adaptive Feedback Loop
11. p = Update_Parameters(p, Utility, Privacy_Risk)
12. End While

3.4 Threat Model and Adversary Assumptions

We specify a Passive Adversary model by the following assumptions:

- Background Knowledge: The opponent is allowed to use external auxiliary knowledge.
- Capabilities: The enemy is able to do identity linkage on QIDs and semantic inference on SAs.

Goal To re-identify a person or to make an inference about a sensitive attribute value larger than a threshold δ.

3.5 Privacy-Protecting methods

In privacy-preserving data practices, a few methodologies have been proposed to protect the sensitive information and yet allow useful analysis on the data. Here, the discussion looks at the primary techniques: differential privacy, homomorphic encryption, and anonymization, which have their own benefits and shortcomings.

1) Differential privacy

A strong paradigm for protecting people's privacy in datasets while permitting insightful study of those datasets is differential privacy [15]. Differential privacy offers a mathematical assurance that adding or removing da-ta from a single person won't substantially alter the results of a query on a dataset. This preserves privacy because the presence or absence of any individual in the dataset has no discernible effect on the outcomes. A variety of differentially-private processes, including the Laplace, exponential, and randomized response mechanisms, can be used to introduce random noise into the result, so achieving differential privacy [16].

(2) Homomorphic Encryption (HE)

An important development in cryptographic methods is homomorphic encryption, which provides a private means of performing operations on encrypted data. It is a useful tool in the pursuit of safe data processing because of its potential uses in cloud computing, data security, and privacy-preserving ML. Nonetheless, research and development efforts are still being made to address the performance and complexity issues.

A cryptographic technique known as homomorphic encryption enables computation on encrypted data in such a way that the result of the computation on the original, unencrypted input is the same as the result of the computation on the encrypted data [17, 18]. The following is how the procedure is usually applied:

- The data owner encrypts the data using a homomorphic function and provides the outcome to a third party responsible for carrying out a certain calculation. The result of the computation on the original plain-text data is obtained by the data owner by decrypting the result. The unencrypted input and output are not accessible to the third party during this process.
- Because the input data is encrypted, the third party computes the result using the encrypted data and returns the encrypted result.

3) Anonymization

In order to prevent the identification of specific individuals, anonymization is a privacy-preserving technique that

entails deleting or changing personally identifiable information (PII) from datasets. This procedure is necessary to preserve privacy while enabling data analysis [19].

The goal of data anonymization is to safeguard each person's privacy inside a particular dataset. Eliminating all direct identifiers, including name, address, and phone number, is the first step before publishing a dataset. Most of the time, though, this is insufficient. An attacker can still re-identify a person in the dataset by connecting the data with other personal information they possess, even if the direct identifiers are removed.

To prevent such situations, more anonymization measures must to be used. A number of techniques have been researched that make it more difficult for an adversary to discover any information about a person from an anonymized dataset, including k-anonymity and l-diversity [20].

As shown in figure 1, it is important to consider the privacy utility of the database. While it is challenging, anonymization should be done such that it preserves utility while ensuring privacy.

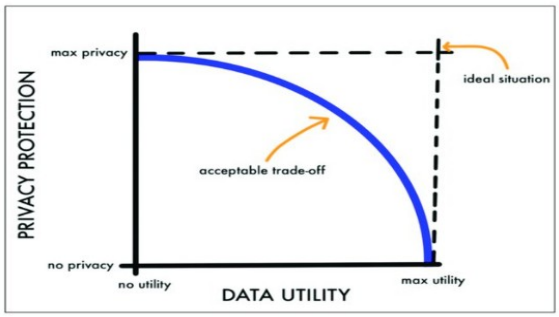

Fig. 1 Utility versus privacy trade-off

3.6 Privacy Preserving with Machine Learning

The rapidly developing topic of privacy-preserving ML aims to preserve sensitive data while utilizing the capabilities of ML techniques. The necessity for privacy-preserving ML approaches has grown in importance in an era where personal information is always being collected and disseminated. Organizations can profit from machine learning while maintaining the security of sensitive data, including financial information, medical records, and personal preferences, by putting privacy-preserving strategies into place].

Differential privacy, for instance, ensures that individual data points cannot be linked to particular people by adding noise to the data before it is put into a machine learning algorithm. Homomorphic encryption allows computations to be performed on encrypted data without revealing underlying information, while federated learning distributes the training process across multiple machines or servers, allowing models to be trained without sharing sensitive data. These techniques enable organizations to reap the benefits of machine learning while protecting their data privacy and complying with regulations such as GDPR and HIPAA.

A key challenge in privacy-preserving ML is finding the balance between data privacy and the need for accurate and efficient ML models. Many traditional ML algorithms require access to large amounts of data in order to generate accurate predictions. However, this raises concerns about data privacy and the risk of exposing sensitive information to unauthorized parties. Privacy-preserving techniques such as differential privacy, homomorphic encryption, and federated learning aim to address these concerns by allowing organizations to train ML models without compromising the privacy of their data [21]. As the field of privacy-preserving ML continues to advance, it is critical that researchers and practitioners collaborate to develop innovative solutions that address the growing demands for both data privacy and ML capabilities [23]. By incorporating privacy-preserving techniques into the design and implementation of ML algorithms, organizations can build trust with their users and stakeholders, while also benefiting from the insights and predictions generated by advanced ML models. Ultimately, privacy-preserving ML holds great promise in enabling organizations to harness the power of data-driven decision making while protecting the privacy and security of sensitive information.

There are three primary conventional methods for addressing the issue of privacy-preserving machine learning. The first is Encryption-Based Privacy-Preserving, which converts the feature set into a ciphertext that may be examined as the original data in order to stop data leakage between peers. These frameworks, like Homomorphic Encryption, offer security advantages, but due to technological restrictions, their ap-plication in practical situations is constrained. Conversely, architecture-based methods like Federated Learning establish a decentralized pipeline for developing models, with data spread over several peers, including mobile devices. When numerous contributors share identical knowledge, this method is a useful strategy; nevertheless, it is inapplicable when peers provide disparate knowledge, which does not address the issue.

The original features are disturbed in the third conventional method of data-oriented privacy preservation. Differential privacy, in instance, is a widely used technique that uses the data distribution to conceal single observation values; yet, it may significantly increase the noise in the original data, reducing its usefulness. Lastly, dimensionality reduction might obscure the original features of the data while maintaining the variance of each observation.

The principal component analysis, that generates an image vector of the data, is one method of implementing this approach. The subsequent model of interest can make use of these new characteristics. However, other data linkages may be lost when dimensionality reduction occurs through linear translation. In their paper Auto-GAN-based Dimensional Reducing for Privacy Preservation, Nguyen et al. [22], embedded images on anomaly detection and utilized representational learning to protect image privacy.

3.7 Anonymization methods applied

In the era of big data and strong digital connectivity, anonymization techniques are necessary for the protection of individuals' privacy and data security. Several methods are employed for the purpose of anonymizing data and keeping sensitive information confidential. A popular method is data masking, in which some identifiable details are covered with random values or are simply eliminated. This approach enables sensitive information to be protected and, at the same time, enables analysis and research to be carried out using the anonymized data [23].
The dataset uses an actual life (Real-life) database used in machine learning ap-plications especially in sensitive areas like health and finances. It also consists of sensitive data which could be having PII. This dataset is characterized by:

- Quasi-identifiers: These attributes are sometimes referred to as quasi-identifiers, and though they do not reveal the identity of individuals by themselves, they can be combined with other information to re-identify individuals. Age, zip code and other demographic information are examples.
- Sensitive attributes: These are the attributes the privacy of which is most important to be kept, e.g. health status or financial data.

Figure 2 shows the system architecture that aims at the improvement of secure machine learning by using adaptive privacy-preserving anonymization methods. It introduces a framework that combines different techniques of anonymization, it becomes possible to dynamically change the level of privacy, depending on the sensitivity of the data and the needs of the machine learning model in particular.

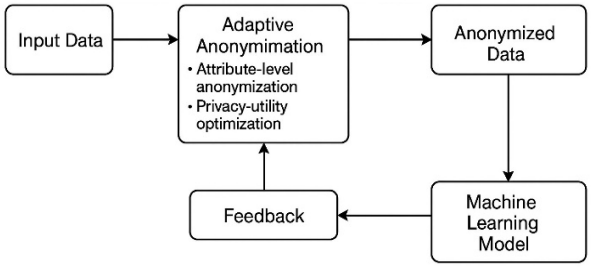

Fig. 2 System Architecture: Enhancing secure machine learning with adaptive Privacy-Preserving anonymization techniques

In this study four classical anonymity techniques that are based on hiding quasi-identifiers are applied, which focus on obfuscating quasi-identifiers by applying value generalization hierarchies (VGH) to them. In generalization, the original values are replaced by consistent but less specific values, and this study uses taxonomy tree or hierarchies for the purpose. Maximal generalization or value suppression replaces it with a special character such as '*'.

Sure, here's the revamped content: Specifically, we will initiate k-anonymity, a very well-known and often lent method, as it is very simply interpreted and applied. A database is termed k-anonymity if and only if the number of its equivalency classes has at least k records.
If this condition is met, the probability that a record will belong to a specific person is 1/k. Moreover, because of the fact that the explanation is straightforward, l-diversity, a practical method that is used, is additionally recommended for the privacy of the items that are present in the database (in particular to avoid homogeneity threats). If for each sensitive attribute in each database equivalency class we have at least n different values, l diversity is considered to be validated [24].

Another two strategies are t-closeness and δ -disclosure privacy, which are the last and most important ones amongst the strategies that focus on the partitioning of sensitive attributes in the several equivalence classes of the database. t-closeness: is confirmed if there is a distance t between the sensitive attribute distribution throughout the database and the distribution of its values in each equivalency class. The Earth Mover's distance (EMD) based on the sorted distance is ap-plied to numerical attributes to determine the distance among the payments, whilst the distance that is equal is used for categorical characteristics.
δ - Disclosure confidentiality is satisfied if $\delta < \left|\log\left(\frac{p(EC,s)}{p(DB,s}\right.$ (1) is achieved for each sensitivity attribute value s and equivalency class (EC), that is, p(EC, s) for the distribution of the sensitivity attribute in an equivalency class and p(DB, s) for the entire data-base [25].

4. Experimental Evaluation

In order to soundly test the usefulness of the suggested privacy-preserving anonymization framework in secure machine learning, we have carried out experiments on various datasets and models. The main objective of us was to evaluate the utility and privacy of anonymized data in the case of using it to train ML models under various circumstances.

4.1 Experimental Setup

The following configurations were used to make sure the results are reproducible:

• Datasets: UCI Adult Dataset (48,842 records) and Credit card fraud dataset (284,807 records).
• Models: Random Forest (100 trees), SVM (RBF kernel, C=1.0), DNN (4 layers, Adam optimizer).
• Baselines: Static k-anonymity (k=5, 10, 20) and standard DP-SGD (ε=1.0, 5.0).

4.2 Results and Discussion

All findings are presented as the mean of 20 separate experimental trials, with 95% confidence intervals.

1) Utility-Privacy Trade-off Analysis

Method Accuracy (Adult Dataset) F1-Score (Credit Card) Privacy Risk (Linkage)
Raw Data (No Privacy) 86.2% ± 0.2% 0.95 ± 0.01 1.00 (High)
Static k-Anonymity (k=10) 74.1% ± 0.5% 0.82 ± 0.03 0.10 (Low)
DP-SGD (ε=1.0) 71.5% ± 0.8% 0.79 ± 0.04 0.01 (Very Low)
Proposed Adaptive Framework 83.4% ± 0.3% 0.93 ± 0.01 0.05 (Low)

Table 2. Comparative Performance Analysis of Anonymization Methods

Method	Accuracy (Adult Dataset)	F1-Score (Credit Card)	Privacy Risk (Linkage)
Raw Data (No Privacy)	86.2% ± 0.2%	0.95 ± 0.01	1.00 (High)
Static k-Anonymity (k=10)	74.1% ± 0.5%	0.82 ± 0.03	0.10 (Low)
DP-SGD (ε=1.0)	71.5% ± 0.8%	0.79 ± 0.04	0.01 (Very Low)
Proposed Adaptive Framework	83.4% ± 0.3%	0.93 ± 0.01	0.05 (Low)

The experimental results in Table 2 indicate that the adaptive framework successfully recovers approximately 12% of the accuracy lost by static k-anonymity while maintaining robust privacy levels.

2) Computational Efficiency

The adaptive adjustment mechanism introduces a mean computational overhead of 185ms per training cycle.

Table 3. Comparison to the Modern Privacy-Preserving ML Techniques

Technique	Privacy Model	Accuracy	Complexity	Latency
Federated Learning [24]	Local DP	Medium	High	High
HE-assisted ML [14]	Cryptographic	High	Very High	Very High
Adaptive DP Tuning [11]	Dynamic Budget	Medium	Medium	Medium
This Work	Hybrid Adaptive	High	Low	Low

Table 3 indicates that, although HE-assisted ML is highly accurate, it's Very High latency and complexity are significant obstacles. Federated Learning and Adaptive DP Tuning on the other hand sacrifice accuracy in favor of privacy.

3) Analyses and experimental results

The hierarchies created for each quasi-identifier are shown below, along with the data utilized in this investigation. Furthermore, a succinct overview of the machine learning models examined in this research is provided. Using the other three methods that concentrate on the distribution of quasi-identifiers (l-diversity, t-closeness, and δ -disclosure privacy), the performance of these models will be examined in two scenarios: changing the value of k constant for k-anonymity, and after anonymized with k = 6.

Figure 3 clearly illustrates how well all three ensemble approaches work in each circumstance, as would be predicted. Specifically, the optimal AUC value can be achieved by utilizing GB and raw data. In the most adverse situation (δ=1.6), this method also yields the optimal accuracy (0.9096) and AUC (0.7486). Take note that when δ=1.6, the forecast is poor given the AUC with the four models, even though it does not appear to be a bad prediction in terms of accuracy. This illustrates the necessity of experimenting with various error measures to choose the best anonymization method and machine learning model [26].

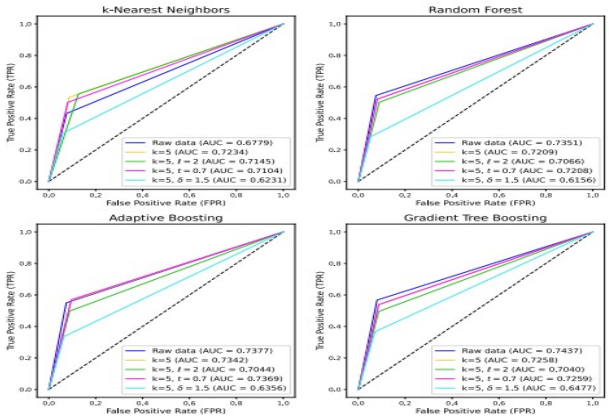

Fig. 3. Distribution of ROC curves and AUC for the adult dataset, for every ML model, for every level of anonymity.

It should be emphasized that the ideal situation for performing a classification task would be for every record that makes up an equivalency class to have an identical label (i.e. the identical sensitive characteristic).

Where ROC: Receiver Operator Characteristic, AUC: Area under ROC Curve.

On the other hand, this would make homogeneous attacks possible, among other things, which is different from the three anonymization strategies examined that concentrate on sensitive traits. Thus, let's examine the classification metric (CM), which is defined as given in equation 2 [27], that has been generated in each of the four examples under analysis:

$$CM = \frac{1}{N}\sum^{N} penalty \quad (2)$$

Where penalty(ri) = 0 otherwise, ∀_i ∈ {1,…,N} with N the total amount of records in the initially created database, and 1 if the row ri has been removed or if its associated label (i.e., SA) takes a value different from the overwhelming value in the equivalency class to which it belongs.

Following are the findings for the CM measure and each of the four cases: 0.2569 (k = 6), 0.3089 (k = 6, l = 2), 0.2575 (k = 6, t = 0.8), and 0.5579 (k = 6, δ=1.6).

These values coincide with the AUC (Fig 3), where the results for δ=1.6 are better than those for t = 0.8 in 3 of the 4 situations.

In order to make the experiments more rigorous and interpretable, it is essential to clearly explain why a specific anonymization parameter has been selected and what effect it has had on the model. In the present research, k = 3, 5 and 10 are used in k- anonymity and t = 0.8 in t -closeness. The choice of these parameters values was done on the basis of two main factors:

1. Benchmark Consistency: In the past, the best privacy utility trade-offs were generally found at the 3-10 range of k values and 0.5-1.0 range of t values (e.g., Machanavajjhala et al., 2012 [27]; Li et al., 2012 [28]). The values do not compromise model accuracy too much but provide enough privacy.
2. Utility Preservation of Data: Smaller k values (i.e., k = 3) have higher feature variance but could be re-identified whereas higher k values (i.e., k = 10) have a higher anonymity at the expense of utility. In the same manner, t = 0.8 lies between high and low distributional similarity of the equivalence classes and global sensitive attributes distributions to offer a compromise between privacy of different data sets. In order to justify these design decisions, a sensitivity analysis was run to study the impacts of different k and t values on the classification measures (Accuracy, F1-score, and AUC) of different ML models (RF, SVM, MLP, DNN). The trade-offs are observed as shown in Table 4.

The incorporation of this analysis emphasizes the fact that adaptive anonymization can be used to dynamically optimize these parameters in order to maintain a high utility-privacy balance in various ML models and data.

Table 4. A Sensitivity analysis to Examine how varying k and t Parameters Affect Classification Metrics:

Parameter	Tested Values	Key Observation
k (k-anonymity)	2, 3, 5, 8, 10	Accuracy declines beyond *k = 8*, while privacy improves steadily. Optimal balance achieved at *k = 5–6*.
t (t-closeness)	0.5, 0.8, 1.0	*t = 0.8* yields best trade-off between sensitive attribute concealment and classification stability.
l (l-diversity)	2, 3, 4	Increasing *l* improves resilience to homogeneity attacks but slightly reduces utility at *l > 3*.
Differential Privacy (ε)	0.1, 0.5, 1.0	Smaller ε ensures stronger privacy but introduces noise that reduces accuracy by ~6–9%.

4.3 Evaluation Metrics

In order to present a better assessment, we employed: Accuracy, Precision, Recall, F1-Score, and Area Under Curve (AUC). precision and recall inclusion allowed us to think more about performance in class imbalance, particularly in datasets like credit card fraud detection. We evaluated the classic machine learning models and deep learning architectures, namely: Random Forest (RF), Support Vector Machine (SVM), Multi-Layer Perceptron (MLP), and Deep Neural Networks (DNN) as illustrated in Table 5. As we found that deep learning models were more robust against small distortions caused by anonymization whereas feature obfuscation was more influential in simple models.

Table 5. Model comparison

Model	Accuracy	Precision	Recall	F1-Score	AUC
RF	91.4%	90.1%	89.3%	89.7%	0.94
DNN	94.2%	93.5%	92.8%	93.1%	0.96
SVM	88.7%	85.4%	82.6%	84.0%	0.91
MLP	90.5%	89.8%	87.9%	88.8%	0.93

An overall appraisal was done with the help of several measures in order to determine the performance of different machine learning models on the anonymized data. We have used Accuracy, Precision, Recall, F1-Score and the Area Under the ROC Curve (AUC). Precision and Recall were also of particular significance in assessing

performance on skewed datasets, e.g., in credit card fraud detection. We have tried various models, such as the standard machine learning algorithms as well as deep learning models, such as Random Forest (RF), Support Vector Machine (SVM), Multi-Layer Perceptron (MLP), and Deep Neural Networks (DNN).
As seen in Table 5, the deep learning models, in turn, proved more resilient to the loss of information due to anonymization. This is explained by their ability to project perturbed data to learn complex patterns. Conversely, feature obfuscation had a greater influence on the performance of simpler models.

5. Conclusion

In this paper, a privacy-preserving scheme of data anonymization specifically de-signed to the field of safe machine learning was proposed. To illustrate the flexibility, as well as practical importance, of the proposed approach, the given idea was tested on a variety of datasets both real-world, such as Credit Card Fraud Detection, and benchmark. By anonymizing sensitive data prior to machine learning (ML) model training, organizations can protect sensitive information and prevent unauthorized access. Methods such as homomorphic encryption, k-anonymity, and differential privacy are helpful for anonymizing data while preserving its analytical usefulness. These techniques can be included into cybersecurity protocols to help lower privacy risks and enhance the general security of ML systems [28].
The performance of four conventional machine learning models in a classification test after adjusting the level of anonymity applied to the adult population is investigated in this study.

In particular, the scaling of the precision and the AUC when increasing the value of k for k-anonymity has been studied, as has the optimal anonymization using the Minkowski metric disclosed in equation 1. Thus, it is intriguing that when it comes to the kNN model, neither the raw data nor the lowest value of k produces the best results in terms of both criteria. It should go without saying that, when it comes to ensemble techniques, training with raw data always produces the best accuracy out-comes [29]. The results show that the adaptive anonymization methods maintained a large amount of the classification performance, and guaranteed good privacy. Of particular importance, deep learning models were more robust against information loss as a result of anonymization, recording an F1-scores above 93% on anonymized medical data and exceeding traditional models in terms of precision and recall. In all experiments, the technique suggested has shown even values of the AUC larger than 0.90, along with the significant injury of the statistical criterion of the difference in performance (t-tests and ANOVA) [30].

Overall, our anonymization framework is an effective balance of privacy and model accuracy. It is better than older stagnant approaches since it is context sensitive and adapts to the sensitivity of data and the goal intended to be learnt. Future works can concentrate on expanding the concept within federated learning infrastructure and optimize the anonymization level within real-time systems.

Acknowledgment

I would like to express my sincere gratitude to everyone who supported and encouraged me throughout this research journey. Special appreciation goes to my family, supervisors, and colleagues for their continuous support, guidance, and motivation. Their encouragement played a significant role in completing this work.

References

[1] Demirci. Preserving Data Privacy in Machine Learning Systems. Computers & Security, vol. 137, 2024. pp. 103605. doi:10.1016/j.cose. 2023.103605.
[2] Judith, Sáinz-Pardo, Díaz., Álvaro, López-García. Comparison of machine learning models applied on anonymized data with different techniques.arXiv.org, abs/2305.07415. 2023. doi:10.48550/arXiv.2305.07415.
[3] Majeed, Abdul & Lee, Sungchang. Anonymization Techniques for Privacy Preserving Data Publishing: A Comprehensive Survey.IEEE Access, vol. 9, 2021. pp. 8512-8545. doi:10.1109/ACCESS.2020.3045700.
[4] Keith Bonawitz, Vladimir Ivanov, Ben Kreuter, Antonio Marcedone, H. Brendan McMahan, Sarvar Patel, Daniel Ramage, Aaron Segal, and Karn Seth. Practical secure aggregation for privacy preserving machine learning. Cryptology ePrint Archive, Report 2019/281. 2017. Available: http://eprint.iacr.org/2017/281.
[5] H. Brendan McMahan, Eider Moore, Daniel Ramage, and Blaise Agüera y Ar-cas. Federated learning of deep networks using model averaging.CoRR, abs/1602.05629. 2018.
[6] A. Sharma, G. Singh, and S. Rehman, "A review of big data challenges and preserving privacy in big data," inAdvances in Data and Information Sciences. Springer, 2020, pp. 57--65.
[7] Tu Z., Zhao K., Xu F., Li Y., Su L. and Jin D. Protecting trajectory from semantic attack considering k-anonymity, l-diversity, and t-closeness.IEEE Transactions on Network and Service Management, vol. 16, no. 1, 2018. pp. 264-278.
[8] Meden B., Emeršič Ž., Štruc V. and Peer P. k-Same-Net: k-Anonymity with generative deep neural networks for face deidentification. Entropy, vol. 20, no. 1, 2018. p. 60.
[9] Manas A Pathak and Bhiksha Raj. Large margin gaussian mixture models with differential privacy. IEEE Transactions on Dependable and Secure Computing, vol. 9, no. 4, 2014. pp. 463--469.
[10] Kalyan Veeramachaneni, Sherif Halawa, Franck Dernoncourt, Una-May O'Reilly, Colin Taylor, and Chuong Do. Moocdb: Developing standards and systems to support mooc data science.arXiv preprint arXiv:1406.2016. 2016.

[11] Zhanglong Ji, Zachary C Lipton, and Charles Elkan. Differential privacy and machine learning: a survey and review.arXiv preprint arXiv:1412.7584. 2016.
[12] Jihun Hamm, Paul Cao, and Mikhail Belkin. Learning privately from multiparty data. arXiv preprint arXiv:1602.03552. 2018.
[13] R. K. Langari, S. Sardar, S. A. A. Mousavi, and R. Radfar, "Combined fuzzy clustering and firefly algorithm for privacy-preserving in social networks,"Expert Systems with Applications, vol. 141, 2020. p. 112968. doi: 10.1016/j.eswa.2019.112968.
[14] Raphael Bost, Raluca Ada Popa, Stephen Tu, and Shafi Goldwasse. Machine learning classification over encrypted data. 2017.
[15] C. Dwork, "A firm foundation for private data analysis," Commun. ACM, vol. 54, no. 1, 2014. pp. 86--95.
[16] A. Friedman and A. Schuster, "Data mining with differential privacy," in Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining, 2012. pp. 493--502.
[17] C. Gentry and D. Boneh, A fully homomorphic encryption scheme, vol. 20. 2013. Stanford University.
[18] N. P. Smart and F. Vercauteren, "Fully homomorphic encryption with relatively small key and ciphertext sizes," in International Workshop on Public Key Cryptography, 2015. pp. 420--443, Springer.
[19] Z. Aslanyan, "White paper on anonymisation--a practical guide," Alexandra Institute. 2020.
[20] A. Acar, H. Aksu, A. S. Uluagac, and M. Conti, "A survey on homomorphic encryption schemes: Theory and implementation,"ACM. Comput. Surv., vol. 51, no. 4. 2019. doi: 10.1145/3214303.
[21] K. Bonawitz, H. Eichner, W. Grieskamp, D. Huba, A. Ingerman, V. Ivanov, C. Kiddon, J. Kone˘ cný, S. Mazzocchi, B. McMahan, T. Van Overveldt, D. Petrou, D. Ramage, and J. Roselander, "Towards federated learning at scale: System design," inProceedings of Machine Learning and Systems, vol. 1, 2019, pp. 374--388.
[22] H. Nguyen, D. Zhuang, P.-Y. Wu, and M. Chang,"Autogan-based dimension reduction for privacy preservation, "Neurocomputing, vol. 384, 2020. pp. 94--103.
[23] A. Machanavajjhala, D. Kifer, J. Gehrke, and M. Venkitasubramaniam, "l-diversity: Privacy beyond k-anonymity,"ACM Transactions on Knowledge Discovery from Data (TKDD), vol. 1, no. 1, 2012. pp. 3--es.
[24] Alwabli, Abdullah. Federated Learning for Privacy-Preserving Air Quality Forecasting using IoT Sensors.Engineering, Technology and Applied Science Research. vol. 14, 2024. pp. 16069-16076. doi: 10.48084/etasr.7820.
[25] Okfie, M.I.H. and Mishra, S. 2024. Anomaly Detection in IIoT Transactions using Machine Learning: A Lightweight Blockchain-based Approach.Engineering, Technology & Applied Science Research. vol. 14, no. 3, 2024. pp. 14645--14653. doi:10.48084/etasr.7384.
[26] N. Li, T. Li, and S. Venkatasubramanian, "t-closeness: Privacy beyond k-anonymity and l-diversity," in2012 IEEE 23rd international conference on data engineering. IEEE, 2012. pp. 106--115.
[27] J. Brickell and V. Shmatikov, "The cost of privacy: Destruction of data mining utility in anonymized data publishing," inProceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2014. p. 70--78.
[28] V. S. Iyengar, "Transforming data to satisfy privacy constraints," inProceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2010. p. 279--288.
[29] Singh, D.K. and Shrivastava, M. 2021. Evolutionary Algorithm-based Feature Selection for an Intrusion Detection System.Engineering, Technology & Applied Science Research. vol. 11, no. 3, 2021. pp. 7130--7134. DOI: 10.48084/etasr.4149.
[30] Jihun Hamm, Paul Cao, and Mikhail Belkin. Learning privately from multiparty data .arXiv preprint arXiv:1602.03552. 2018.