International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING

ISSN:2147-6799 www.ijisae.org Original Research Paper

Integrating User Attribute Influences and DL-Based Anonymization for Enhanced Privacy Protection in Medical Record Data Sharing for Publishing

Lingam Suman[1], **Dr. S. Venkata Lakshmi***[2]

Submitted:11/03/2024 **Revised**: 26/04/2024 **Accepted**: 03/05/2024

Abstract: In today's data-driven healthcare research landscape, sharing medical records for research is pivotal for advancing medical knowledge and patient care. However, ensuring individuals' privacy while maintaining data utility poses a significant challenge. To tackle this issue, this study proposes a novel Attribute Influence Anonymization using RVAE (AIARVAE) for enhancing both privacy and utility in medical records data sharing. The proposed model employs a preprocessing step to identify and filter Quasi-Identifiers (QIs) and Sensitive Attributes (SAs) from the dataset. Then quantify the susceptibility of QIs and measure the uncertainty of SAs using entropy. These metrics are then fed into a Recurrent Variational Auto-Encoder (RVAE) model, which replaces low-entropy SAs with sanitized values with the help of QI values. This approach mitigates the risk of explicit disclosure of private information while preserving data utility. By integrating attribute influences, the proposed model provides a comprehensive solution for safeguarding medical records data privacy during research sharing and promoting responsible and ethical data-driven healthcare research.

Keywords: *Privacy protection, Quasi-identifiers, Sensitive attributes, Sanitization, Utility*

1. Introduction

Due to the popularization of social networks (SNs), all SN service [16] providers collect and store user's data much like hospitals and banks. This data often contains information about the user's demographics, finances, SN activities and preferences, hobbies, community affiliations, medical status [25,26], and relationships with other online users. Releasing user's data is beneficial [18] for extracting accurate, timely, detailed, and multifaceted insights about the users with advanced data mining tools and pattern analysis applications [22,27]. Besides the individual's privacy problems [17,30], an adversary can infer the sensitive information [20,29] based on the users' QIs' values that can jeopardize the privacy of a specific community. Therefore, it is paramount that any sharing and mining of SN data must protect the privacy of the users' community. Privacy-preserving data publishing (PPDP) [19, 23 & 24] provides a set of tools, DL models, and methods to safeguard against the privacy threats that emerge from the mining and sharing of this data [15,31]. Most of the existing methods give equal weight to all attributes in data from a privacy and utility point of view. However, recent research has shown that each item within an attribute has a distinct impact on privacy and utility [11].

For example, a zip code allows locating someone more accurately than race and/or gender. Similarly, gender is more appropriate for making credit-related decisions, rather than age. Hence, quantifying the impacts of a user's attributes, and ensuring protection based on such statistics in the anonymization, is challenging [21,28]. To address these challenges, a novel model proposed AIARVAE. This model is designed to enhance both privacy and utility in the sharing of medical records for research. The AIARVAE model incorporates a multi-faceted approach that begins with a preprocessing step to identify and filter QIs and SAs from the dataset. Following this, the model quantifies the susceptibility of QIs and measures the uncertainty of SAs using entropy. These metrics provide a detailed understanding of the potential privacy risks associated with each attribute. These quantified metrics are then fed into an RVAE model. The RVAE uses this information to replace low-entropy SAs with sanitized values, leveraging the relationships and patterns identified within the QI values. This targeted approach mitigates the risk of explicit disclosure of private information, ensuring that sensitive attributes are effectively anonymized while maintaining the overall utility of the dataset. These contributions of the proposed AIARVAE are outlined as follows:

- To handle complex data patterns and relationships, leading to more effective anonymization outcomes RVAE model was used.

- To minimize unnecessary data alterations, thus preserving the semantic integrity and usefulness of the data the proposed model integrates attribute influences and focuses on the precise handling of QIs and SAs.

- To reduce the IL, the AIARVAE model identifies and

[1] Research Scholar, Dept. of CSE, GITAM (Deemed to be University), Visakhapatnam, Andhra Pradesh, INDIA
ORCID ID : 0009-0008-2399-6994
Email: ssuman2468@gmail.com
[2] Asst. Professor, Dept. of CSE, GITAM (Deemed to be University), Visakhapatnam, Andhra Pradesh, INDIA
ORCID ID : 0000-0003-0270-2594
** Corresponding Author Email: svlakshmi2014@gmail.com*

filters QIs and SAs, ensuring that only the most relevant attributes are anonymized.

- To lower computational complexity, AIARVAE focuses on high-risk QIs and low entropy SAs, the model prioritizes critical areas for anonymization, avoiding exhaustive processing of less relevant data.

The integration of attribute influences within the anonymization process allows the AIARVAE model to provide a comprehensive solution for safeguarding medical records data privacy during research sharing. The remaining portions of this study are arranged as follows: Section 2 reviews relevant work and highlights gaps in past research works. Section 3 goes over the suggested AIARVAE technique, dataset information, pre-processing, and the RVAE model. The research findings and analysis are presented in Section 4, while Section 5 examines the future extent and conclusion.

2. Background Study

Majeed et al. [1] provided a comprehensive examination of Clustering-based Anonymization Mechanisms (CAMs) utilized in securely releasing personal data across diverse formats. They meticulously categorized existing CAMs according to ten distinct data styles, offering an analysis of their respective attributes such as strengths, weaknesses, and the clustering algorithms employed. Additionally, they addressed numerous unresolved challenges encountered by anonymization methods utilizing clustering principles. Lastly, the authors deliberated on potential avenues for future research, acknowledging the evolving landscape of privacy risks amidst ongoing technological advancements.

Peethambaran et al. [3] introduced a composite classifier model prioritizing privacy considerations to assess classification accuracy. Initially, they constructed a composite classifier using a high-dimensional dataset anonymized through k-anonymity, a privacy algorithm. They observed that the composite classifier's incorporation of various algorithmic characteristics effectively balanced classification accuracy, which could otherwise be impacted by the privacy model. Additionally, they assessed the model's performance in terms of execution time by employing the parallel computing framework Spark.

Cai et al. [4] introduced a strategy aimed at preserving privacy in interactive messaging. This approach utilizes the concept of user credibility, which evaluates individuals' reputations based on their social interactions. Additionally, the scheme employs the item response theory from psychometrics to evaluate the risk levels associated with users' interactive messages. To ensure privacy during message exchanges, the scheme implements message obfuscation through replication and the replacement of sensitive attributes. Experimental validation conducted on a dataset akin to a Facebook-style Social Network demonstrates the scheme's robustness, exhibiting both high availability and accuracy.

Ciampi et al. [6] presented a modular architecture capable of collecting clinical data from heterogeneous sources and transforming them into a standard format useful for secondary use for research, governance, and medical training purposes, as well as applied the necessary de-identification techniques to respect privacy and ethics law requirements. The adequate and efficient use of data from heterogeneous sources was obtained by making the representation of the processed information compliant and standard.

Shakeel et al. [10] proposed a novel approach to k-NDDP, which is the extension of k-NMF. k-NDDP is a degree anonymization method that extends the concept of k-anonymity and differential privacy based on Node DP for vertex degrees. The proposed approach provides a solution to the problem that reveals the individual behind any vertex of the social graph and causes identity disclosure. To defend against identity disclosures, the suggested method inserts the least number of dummy connections into the original graph while preventing the adversary from identifying the vertices and preserving as much graph information as possible.

2.1 Problem Statement and Scope

Table 1. Findings reveal several limitations in PPDP. Gangarde et al. [2] employ a semi-automated approach, missing out on the

Abbasi & Mohammadi [7] struggle with removing of infrequent

Table 1. Detailed analysis of the SOTA studies

Author	*Method Proposed*	*Strength*	*Weakness*
Gangarde et al. [2]	A novel privacy preservation model using the enhanced clustering mechanism (EC)	The reduction in Information Loss (IL) amounted to 23.23%, while there was an improvement of 25.35% in the Degree of Anonymization (DoA).	The model's limitation lies in its semi-automated approach to privacy preservation, as it doesn't explore leveraging DL for automated privacy protection.

Onesimu et al. [5]	Attribute-focused privacy preserving data publishing scheme (AFPP)	The classification accuracy showed a 13% improvement, while the IL was reduced by 12%.	A drawback of this approach is the exclusion of quasi-attributes from being treated as sensitive or semi-sensitive attributes.
Abbasi & Mohammadi [7]	Novel approach based on the clustering process using the K-means++ method	Reduce IL 1.5 times and execution time 3.5 times.	Remove infrequent data elements that could impede the process of discovering knowledge.
Bazai et al. [8]	Mondrian algorithm on Spark framework	Significantly decrease the essential expenses associated with re-computation, shuffling tasks, communication between processes, and cache administration.	The computational complexity is considerably high.
Kumar and Kumar [9]	Privacy Preserving Rewiring Algorithm (PPRA)	The PPRA algorithm assisted in balancing the trade-off between the degrees of privacy protection for individuals' IL.	The suggested approach entails significant computational demands and raises concerns regarding privacy.

data elements hindering knowledge discovery. Bazai et al. [8], and Kumar and Kumar [9] face computational complexity and privacy concerns. Recent studies stress the importance of considering individual items within attributes for privacy and utility. However, quantifying these impacts within CAMs, as highlighted by Majeed et al. [1], remains challenging. To address these issues, this work aims to develop a PPDP model for the medical domain. It integrates anonymization and DL to effectively quantify user attributes, enhancing data privacy while minimizing computational burdens and IL.

3. Methods and Materials

3.1. Dataset

The proposed model uses the publicly available dataset [14]. In the dataset, each 15 column provides specific information about the patient, their admission, and the healthcare services provided. The overall framework of AIARVAE is shown in Fig. 1.

3.2 Preprocessing

The initial dataset,DS includes explicit identifiers (EI), quasi-identifiers (QIs), sensitive attributes (SA), and non-sensitive attributes (NSA). Upon inputting DS, the proposed algorithm categorizes attributes into these four types, standard practice in PPDP, and subsequently removes EI and NSA. This results in DS containing only QIs and SA, denoted as $DS\{QI, SA\}$. Here, the set $QI = \{qi_1, qi_2, \dots qi_n\}$ represents the QIs in DS, with each qi_n representing a specific type of QI like race or sex. SA can either be a single type (e.g., disease) or multiple types (e.g., disease and salary) depending on the scenario. For this study, DS contains a single SA with n distinct values $SA = \{sa_1, sa_2, \dots sa_n\}$. Thus, each user, U_x in DS possesses two attribute types: $QI_{U_x}$and SA_{U_x}. Each user is represented as a tuple, denoted as T. The original user dataset may include outliers, missing QI values, and incomplete tuples. Hence, the algorithm pre-processes the data before anonymization, removing outliers via $min - max$ analysis and eliminating records with missing values. Redundant records are discarded using similarity-based analysis. Additionally, the algorithm formats the data and enriches it if necessary to meet processing model requirements. Through data preprocessing, a refined dataset containing comprehensive user information is obtained.

3.3 Finding Susceptibility of QI

Several studies have highlighted potential attacks targeting QIs with numerous unique values, emphasizing the importance of quantifying the significance of these QIs. However, most existing research overlooks the susceptibility of QIs in terms of their potential to infer SAs from QI values, which can inadvertently reveal private information about user communities rather than individual users. Such vulnerable QIs pose a risk to the privacy of user communities rather than the identity or attributes of individual users. Susceptible QIs are characterized by having many similar values, enabling attackers to group users based on these values and infer their associated SAs. In this study, AIARVAE approach begins by identifying susceptible QIs from dataset DS and assessing their importance using Random Forest (RF) to mitigate the risk of multiple user identifications. QIs with few unique values but high frequency of each value has minimal impact on prediction accuracy, making them highly susceptible to community privacy risks [11]. The importance of each QI is determined using Eq. (1):

$$QII(qi_n) = \frac{\sum_{y\in T} I(z_y=\hat{z}_y)}{|T|} - \frac{\sum_{y\in T} I(z_y=\hat{z}_{y,\pi_n})}{|T|} \quad (1)$$

Here by $\hat{z}_y$ is the predicted SA value for yth observation before permutation and by $\hat{z}_{y,\pi_n}$is the predicted SA value for yth observation after qi_n values permutation. The QII for each QI is then calculated as the mean x_{qi_n} importance from all trees using Eq. (2):

$$x_{qi_n} = \frac{\sum_{b=1}^{mtree} QII(qi_n)}{mtree} \quad (2)$$

Here x_{qi_n} gives the mean score from all trees. The standard deviation s_{qi_n} and susceptibility weight w_{qi_n} can be computed using Eq. 3 and 4, respectively.

$$s_{qi_n} = \sqrt{\frac{1}{mtree-1}\sum_{b=1}^{mtree}(QII(qi_n) - x_{qi_n})^2} \quad (3)$$

$$w_{qi_n} = \frac{x_{qi_n}}{s_{qi_n}} \quad (4)$$

Using this approach, it's able to calculate the weights of all QIs and classify them of high, medium, and low susceptibility. Special consideration is given to highly susceptible QIs due to their potential to facilitate unique identifications of multiple users, particularly within Equivalence Classes (ECs) characterized by low entropy values.

3.4 Finding similar Users and Creating EC

To safeguard the privacy of user communities effectively and mitigate IL during the anonymization process, users within EC must possess highly similar attributes. In this study, rank users based on their QI values using the cosine similarity (CS) metric. This measure, ranging from 0 to 1, provides a straightforward and dependable indication of similarity between users. A CS value of 1 indicates exact similarity between two users, while a value of 0 signifies dissimilarity. The similarity between two distinct users A_1 and B_1 can be computed using Eq. 5:

$$CS(A_1, B_1) = \frac{\sum_{x=1}^{q} A_{1x}\times B_{1x}}{\sqrt{(\sum_{x=1}^{q} A_{1x})^2}\times\sqrt{(\sum_{x=1}^{q} B_{1x})^2}} \quad (5)$$

In this context, x represents the QIs of users A_1 and B_1, and q denotes the total number of QIs. Utilizing Eq. 5 enables the computation of similarity between all n users, resulting in the generation of a matrix M containing users exhibiting high degrees of similarity. Subsequently, matrix M is divided into a set C consisting of distinct ECs, denoted as $C = \{C_1, C_2, \ldots, C_n\}$, guided by the privacy parameter p. The data owner selects the value of p, which can be any integer greater than 1. The number of ECs, nc for a group of n highly similar users can be determined using Eq. 6.

$$nc = \frac{n}{p} \quad (6)$$

These ECs will be used for further processing before the generalization of QI's values.

Algorithm 1: Similar user finding and EC creation

Input: Dataset $DS\{QI, SA\}$, Each user has $QI = \{qi_1, qi_2, \ldots qi_n\}$ and $SA = \{sa_1, sa_2, \ldots sa_n\}$

Output: Set of ECs $C = \{C_1, C_2, \ldots, C_n\}$

```
Start
Initialize, set $C = \emptyset$
        for $x = 1$ to $N$ do
            for $y = 1$ to $N$ do
                Compute $CS(A_i, B_j)$ value for each users depend upon $QI$ from Eq. (5).
                Add the $CS(A_i, B_j)$ values in Matrix $M$
                Repeat for all users.
            End for
        End for
        Sort the $CS$ values in descending order for find highly similar users
        Calculate the values of $nc$ by Eq. (6)
        Create ECs $C = \{C_1, C_2, \ldots, C_{nc}\}$
End
```

3.5 Find Entropy of Sensitive Attribute

The entropy of SA values within each EC is assessed to measure the degree of uncertainty regarding potential explicit disclosures. Particular focus is directed towards ECs exhibiting low entropy values. The subsequent process entails computing the entropy E for each EC, denoted as C_x, following the outlined procedure. First, identify the distinct SA category values present in C_x. Then calculate the total occurrences of each category value $F(SA, C_x) = \{f_{sa_1}, f_{sa_2}, \ldots f_{sa_n}\}$. Calculate the proportion r of each SA's category value in an EC. The proportion (r_n) of nth SA value given its frequency f_{sa_1} can be computed using Eq. 7.

$$r_n = \frac{f_{sa_1}}{p} \quad (7)$$

The entropy E of an EC C_x can be computed using Eq. 8.

$$E(C_x) = -\sum_{x=1}^{sa} r_n \log_2 r_n \quad (8)$$

The value, E, falls within the range of 0 to 1, denoted as $E[0,1]$. An E value of 0 indicates that there is no uncertainty within the EC regarding SA values, as all users share the same SA value. Conversely, an E value of 1 signifies a high level of uncertainty for attackers, with SA values distributed evenly among users on average. A lower E value simplifies SA disclosure, making sensitive user information more readily apparent, whereas an E value of 1 is optimal for safeguarding user privacy. In AIARVAE approach, incorporate E values into the data anonymization process to enhance both the privacy protection of user communities and the utility of anonymous data.

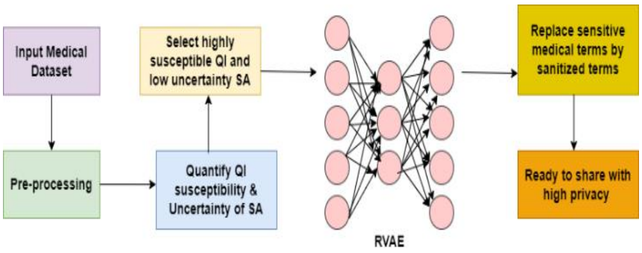
Input Medical Dataset
Pre-processing
Quantify QI susceptibility & Uncertainty of SA
Select highly susceptible QI and low uncertainty SA
RVAE
Replace sensitive medical terms by sanitized terms
Ready to share with high privacy

Fig. 1. Overview of AIARVAE model

3.6 Sanitization by RVAE

The generalization for each QI is based on its susceptibility weight, and the sanitization for SA is based on its corresponding class's entropy value. Data generalization is used to replace the original QI values with less specific but semantically consistent values. By utilizing the entropy and susceptibility statistics jointly to effectively preserve both users' community privacy and anonymous data utility. For example, higher levels of generalization are preferred when there exists a risk that would uniquely identify a multiple of users and their SA's disclosure.

In the proposed AIARVAE model, RVAE takes the SA as the input and reconstructs that to desensitized data to protect the SA. The sanitization data can be used to train the language model while maintaining high utility. The Proposed AIARVAE model is based on the RVAE. It mainly involves two modules: the encoder and decoder, all single-layer LSTM to adapt the original autoencoder to data [12]. The RVAE can effectively approximate inference with the directed probabilistic models. Given the observed SA from the $E(C_x)$, contains a string of words and can be denoted by $SA = \{sa_1, sa_2, \dots sa_n\}$, n is several texts. The goal of the model is to estimate the parameters θ while minimizing marginal log-likelihood.

$$log\, d_\theta\; (E(C_x)) = \sum_{n=1}^{N} log \int_u^{\cdot} d(u)\, d_\theta(sa_n|u)du$$
(9)

Here $d_\theta(u)$ is the prior distribution of a latent variable u, where u is sampled from a multivariate diagonal Gaussian distribution. Because of the integration inside the marginal log-likelihood, the equation is un-differentiable and cannot directly use the gradient descent method to optimize the parameter θ so it can be inverted to the evidence lower bound of the marginal log-likelihood by using an approximation posterior distribution $c_\emptyset(u|sa)$ of $d_\theta(sa|u)$.

$$log\, d_\theta\;(sa) \geq E_{c_\emptyset}(u|sa)[log\, d_\theta\;(sa|u)] - D_{KL}(c_\emptyset(u|sa)||d(u)) \qquad (10)$$

AIARVAE framework used the encoder which includes a single-layer LSTM combining with two fully-connected layers to predict the posterior distribution $c_\emptyset(u|sa)$. More concretely, the posterior distribution $c_\emptyset(sa|u)$ can be assumed as a multivariate diagonal Gaussian distribution.

$$c_\emptyset(u|sa) = \mathcal{N}(u; \mu_\emptyset(h), \sigma_\emptyset(h)) \qquad (11)$$

The function $\mu_\emptyset$ and $\sigma_\emptyset$ both are linear layers to predict the mean and variance of the multivariate diagonal Gaussian distribution according to the hidden state vector h which is the final state output of the LSTM encoder that maps a sequential text input of $SA = \{sa_1, sa_2, \dots sa_n\}$. The operation of sampling latent variable u from $c_\emptyset(u|sa)$is non-continuous resulting that the gradient cannot be computed and passed through. So we refine the sampling operation by $u = \mu_\emptyset(sa) + \beta \sum_\emptyset^{1/2} sa$, where the β is sampled from $\mathcal{N}(0, I)$.

The decoder module is also a LSTM layer which maps the latent variable u as the initial hidden states to the text sequence sample input and generates a new sanitization text for SA while modelling the distribution of $d_\theta(sa|u)$ relies on the latent variable u. Sanitization converts the most important phrases (SA) to the least important ones, such as “Breast Cancer”, can be masked by “Tumour.” The idea is extensively utilized to achieve safety measures. Suppose the data security is breached during the transition from the authenticated person to the cloud or someone. In that case, the adversaries cannot get the real un-sanitized data. Thus, there is a trade-off between privacy and utility. The more is the generalization, the less is the utility [13]. So here only the low entropy SA are sanitized and high susceptibility QI are generalized. Lastly, the medical terms are replaced by sanitized terms and shared. Tables 2 and 3 show the original and anonymized medical records of the patients.

$$Sanitized_{\forall Sensa_n \in SenSA} = Sensa_n$$
(12)

Table 2. Original medical record

	Quasi Identifiers	***Sensitive Attribute***
ID	***Age***	***Medical Condition***
247	20	Cancer

892	34	Covid
349	29	Breast Cancer
572	25	Covid
167	37	Insomnia

Table 3. Anonymous medical record

	Quasi Identifiers	***Sensitive Attribute***
ECs	***Age***	***Medical Condition***
C1	20-26	Tumor
	25-33	Tumor
	30-38	Virus
C2	20-29	Virus
	32-42	Virus

The model can be trained with the stochastic gradient descent and minimize the Loss function, where $\mathcal{N}$ is the batch size.

$$Loss(\, E(C_x); \emptyset, \theta) = \sum_{i=1}^{N} E_{c_{\emptyset}}\,(u|sa)[log\; d_{\theta}\;(sa|u)] - \sum_{i=1}^{N} D_{KL}(c_{\emptyset}(u|sa)||d(u)) \quad (13)$$

The first term of the equation is to encourage the model to reconstruct the original text input. The second term of the equation uses the Kullback-Leibler (KL) divergence, which can evaluate the similarity of the distributions.

4. Result and Discussion

In this section, the performance metrics used for the evaluation of the proposed privacy-preserving algorithms are presented. All experiments were conducted on a PC of an Intel Core i5-3320M CPU at a 2.60GHz clock speed and with 8GB of RAM. The effectiveness of the proposed AIARVAE was compared using four parameters with existing models: EC [2], AFPP [5], Mondrian [8], and PPRA [9].

4.1 Information Loss

To assess anonymous utility, information loss (IL) was utilized as a metric. Fig. 2 illustrates that the proposed algorithm results in lower ILs compared to existing methods. Typically, datasets with a larger number of tuples tend to have higher ILs. This aligns with the theoretical understanding that datasets with a high volume of records are more susceptible to alterations in data semantics, leading to decreased utility.

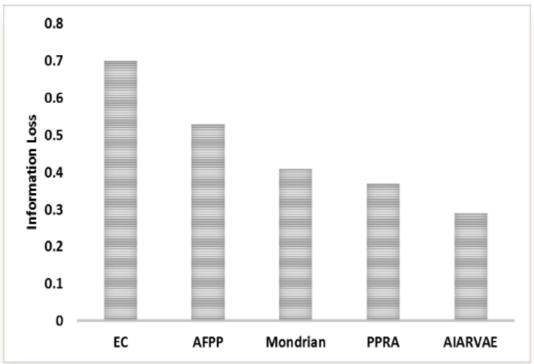

Fig. 2. IL Comparison

However, AFPP has a notable drawback, it excludes QI from being treated as sensitive or semi-sensitive, which can compromise privacy preservation. AIARVAE achieves the lowest IL, with a reduction of 8% compared to PPRA, 12% compared to Mondrian, 24% compared to AFPP, and a substantial 41% compared to EC. This method excels in balancing privacy and utility by effectively handling QIs and generalizing highly susceptible QIs within ECs with low entropy values.

4.2 Computational Complexity

Fig. 3 compares the computational complexity of various anonymization methods across different record sizes, measured in ms. EC method shows consistently high computational complexity, starting at 30ms for 100 records and increasing to 65ms for 1000 records. This indicates an increase in complexity as the record size grows. A major drawback of the EC model is its semi-automated approach to privacy preservation, as it doesn't leverage DL for automated privacy protection, which limits its efficiency and effectiveness.

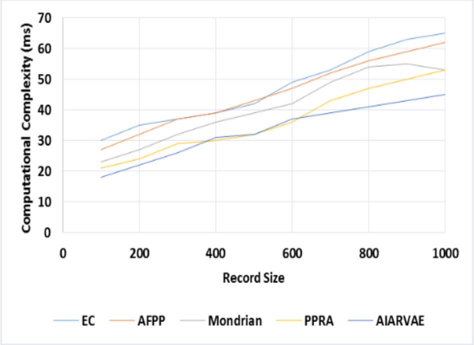

Fig. 3. Computational Complexity Comparison

AIARVAE consistently exhibits the lowest computational complexity, starting at 18 ms for 100 records and reaching 45 ms for 1000 records. The DL-based RVAE model is highly efficient in handling large datasets. Its recurrent nature allows for the processing of sequential data, making it particularly suited for the temporal aspects often present in medical records. Additionally, RVAE are designed to handle high-dimensional data effectively, which reduces the complexity involved in the data anonymization process.

4.3 Utility Preservation

Fig. 4. compares the preservation of utility among different anonymization models. The EC model shows the lowest preservation of utility at 45.97%, indicating significant data degradation and loss of utility. The AFPP model improves utility preservation to 57.82%, but it still falls short because it fails to treat QIs. The Mondrian method further enhances utility preservation to 63.17%, because of its more balanced approach in handling data generalization. However, it still lags due to its higher computational complexity. The PPRA model achieves 69.26% yet it encounters challenges related to computational efficiency and comprehensive privacy protection.

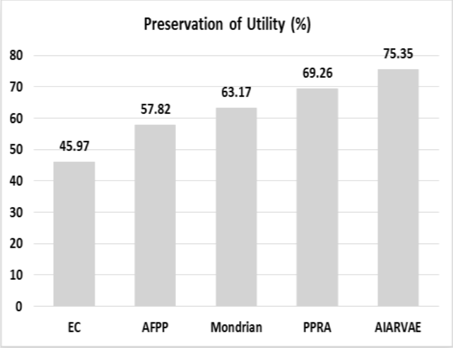

Fig. 4. Preservation of Utility

In contrast, the AIARVAE model leads with the highest preservation of utility at 75.35%. This superior performance is attributable to its preprocessing steps, which effectively filter and quantify QI and SAs. The integration of these metrics into the DL framework allows for precise sanitization, particularly focusing on low-entropy SAs, thereby minimizing explicit disclosure risks and preserving the utility of the data.

4.4 Recall, Precision, and F1 Score

From Fig. 5, the comparison of different methods based on their Precision, recall, and F1 score reveals significant insights into their effectiveness in maintaining data utility and privacy. The EC method demonstrates the lowest performance, with precision at 82.13%, recall at 80.16%, and an F1 score of 83.4. The lower scores indicate that the EC method struggles to effectively balance privacy and utility, leading to less accurate and comprehensive anonymization.

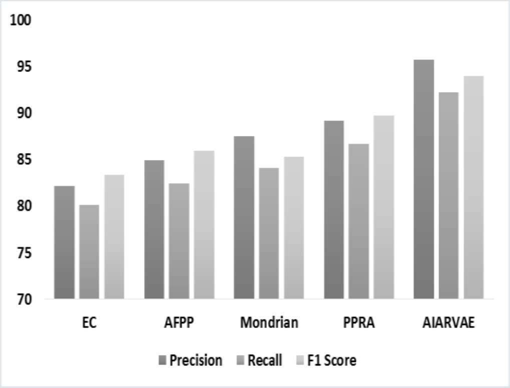

Fig. 5. Precision, Recall & F1 Score

The AFPP model with precision at 84.97%, recall at 82.39%, and an F1 score of 85.9. AFPP's improvement is due to its better handling of attributes, although it still falls short because it does not treat QI as sensitive, which affects its overall accuracy and completeness. The AIARVAE model begins with a targeted preprocessing step that accurately identifies and filters QIs and SAs. This ensures that only the most relevant attributes are considered for anonymization, leading to more precise handling of data and thus enhancing precision. By quantifying the susceptibility of QIs and measuring the uncertainty of SAs using entropy, the model effectively identifies which attributes are most at risk. This careful analysis ensures that the model focuses on the attributes that most impact privacy, thereby enhancing recall by ensuring that all relevant instances of sensitive data are properly anonymized. Because of all these, the proposed model obtains higher precision, recall, and F1 measures than other existing methods.

5. Conclusion and Future Work

In the rapidly evolving field of healthcare research, the need to share medical records while preserving patient privacy is critical. The proposed AIARVAE model addresses this need by providing an innovative solution that enhances both privacy and data utility. By integrating a preprocessing step to identify and filter QIs and SAs, and quantifying their susceptibility and uncertainty using entropy, the model ensures a focused and effective anonymization process. The use of a RVAE allows for sophisticated handling of data that maintains the integrity and utility of the dataset. The AIARVAE model leads with the highest preservation of utility at 75.35% and achieves the lowest IL, with a reduction of 8%. This approach not only mitigates the risk of explicit disclosure of sensitive information but also preserves the essential characteristics of the data, making it suitable for meaningful research. In the future, expanding the model to handle more diverse and complex datasets, including multimodal data that combines textual, numerical, and image data, could broaden its applicability across different domains beyond healthcare.

References

[1] Majeed, A., Khan, S., & Hwang, S. O. (2022). Toward privacy preservation using clustering-based Anonymization: Recent advances and future research outlook. IEEE Access, 10, 53066-53097. https://doi.org/10.1109/access.2022.3175219

[2] Gangarde, R., Sharma, A., & Pawar, A. (2023). Enhanced clustering based OSN privacy preservation to ensure K-anonymity, T-closeness, L-diversity, and balanced privacy utility. Computers, Materials & Continua, 75(1), 2171-2190. https://doi.org/10.32604/cmc.2023.035559

[3] Peethambaran, G., Naikodi, C., & Suresh, L. (2020). An ensemble learning approach for privacy–quality–Efficiency trade-off in data analytics. 2020 International Conference on Smart Electronics and Communication (ICOSEC). https://doi.org/10.1109/icosec49089.2020.9215250

[4] Cai, Y., Zhang, S., Xia, H., Fan, Y., & Zhang, H. (2020). A privacy-preserving scheme for interactive messaging over online social networks. IEEE Internet of Things Journal, 7(8), 6817-6827. https://doi.org/10.1109/jiot.2020.2986341

[5] Onesimu, J. A., J, K., Eunice, J., Pomplun, M., & Dang, H. (2022). Privacy preserving attribute-focused Anonymization scheme for healthcare data publishing. IEEE Access, 10, 86979-86997. https://doi.org/10.1109/access.2022.3199433

[6] Ciampi, M., Sicuranza, M., & Silvestri, S. (2022). A privacy-preserving and standard-based architecture for secondary use of clinical data. Information, 13(2), 87. https://doi.org/10.3390/info13020087

[7] Abbasi, A., & Mohammadi, B. (2021). A clustering-based anonymization approach for privacy-preserving in the healthcare cloud. Concurrency and Computation: Practice and Experience, 34(1). https://doi.org/10.1002/cpe.6487

[8] Bazai, S. U., Jang-Jaccard, J., & Alavizadeh, H. (2021). A novel hybrid approach for multi-dimensional data Anonymization for Apache spark. ACM Transactions on Privacy and Security, 25(1), 1-25. https://doi.org/10.1145/3484945

[9] Kumar, S., & Kumar, P. (2023). Privacy preserving in online social networks using fuzzy rewiring. IEEE Transactions on Engineering Management, 70(6), 2071-2079. https://doi.org/10.1109/tem.2021.3072812

[10] Shakeel, S., Anjum, A., Asheralieva, A., & Alam, M. (2021). K-NDDP: An efficient Anonymization model for social network data release. Electronics, 10(19), 2440. https://doi.org/10.3390/electronics10192440

[11] Majeed, A., & Lee, S. (2020). Attribute susceptibility and entropy based data anonymization to improve users community privacy and utility in publishing data. Applied Intelligence, 50(8), 2555-2574. https://doi.org/10.1007/s10489-020-01656-w

[12] Wang, Y., Meng, X., & Liu, X. (2023). Differentially private recurrent variational Autoencoder for text privacy preservation. Mobile Networks and Applications. https://doi.org/10.1007/s11036-023-02096-9

[13] Moqurrab, S. A., Tariq, N., Anjum, A., Asheralieva, A., Malik, S. U., Malik, H., Pervaiz, H., & Gill, S. S. (2022). A deep learning-based privacy-preserving model for smart healthcare in Internet of Medical Things using fog computing. Wireless Personal Communications, 126(3), 2379-2401. https://doi.org/10.1007/s11277-021-09323-0

[14] Patil, P. (2024, May 8). Healthcare dataset. Kaggle: Your Machine Learning and Data Science Community. https://www.kaggle.com/datasets/prasad22/healthcare-dataset

[15] Majeed, A., & Hwang, S. O. (2023). Quantifying the vulnerability of attributes for effective privacy preservation using machine learning. IEEE Access, 11, 4400-4411. https://doi.org/10.1109/access.2023.3235016

[16] Manjula, G. S., & Meyyappan, T. (2023). Two-Phase Privacy Preserving Big Data Hybrid Clustering for Multi-Party Data Sharing. International Journal of Intelligent Systems and Applications in Engineering, 11(9s), 501-510.

[17] Udita, M., Ritu, N., & Amandeep. (2023). Secure and Compatible Integration of Cloud-Based ERP Solution: A Review. International Journal of Intelligent Systems and Applications in Engineering, 11(9), 695-707.

[18] Kavitha, G., Kavitha, K., & Sujatha, B. (2024). A Hybrid Multi-Client Filter Based Feature Clustering and Privacy Preserving Classification Framework on High Dimensional Databases. International Journal of Intelligent Systems and Applications in Engineering, 12(8), 93-107.

[19] Ge, Y., Wang, H., Cao, J., Zhang, Y., & Jiang, X. (2024). Privacy-preserving data publishing: An information-driven distributed genetic algorithm. World Wide Web, 27(1). https://doi.org/10.1007/s11280-024-01241-y

[20] Canbay, Y., Sagiroglu, S., & Vural, Y. (2022). A new

utility-aware anonymization model for privacy preserving data publishing. Concurrency and Computation: Practice and Experience, 34(10). https://doi.org/10.1002/cpe.6808

[21] Majeed, A. (2023). Attribute-centric and synthetic data based privacy preserving methods: A systematic review. Journal of Cybersecurity and Privacy, 3(3), 638-661. https://doi.org/10.3390/jcp3030030

[22] Kulkarni, Y. R., Jagdale, B., & Sugave, S. R. (2023). Optimized key generation-based privacy preserving data mining model for secure data publishing. Advances in Engineering Software, 175, 103332. https://doi.org/10.1016/j.advengsoft.2022.103332

[23] Kim, J. W. (2021). Efficiently supporting online privacy-preserving data publishing in a distributed computing environment. Applied Sciences, 11(22), 10740. https://doi.org/10.3390/app112210740

[24] Ekaputra, F. J., Ekelhart, A., Mayer, R., Miksa, T., Šarčević, T., Tsepelakis, S., & Waltersdorfer, L. (2024). Semantic-enabled architecture for auditable privacy-preserving data analysis. Semantic Web, 15(3), 675-708. https://doi.org/10.3233/sw-212883

[25] Andrew, J., Eunice, R. J., & Karthikeyan, J. (2023). An anonymization-based privacy-preserving data collection protocol for digital health data. Frontiers in Public Health, 11. https://doi.org/10.3389/fpubh.2023.1125011

[26] S.Venkata Lakshmi and Valli Kumari Vatsavayi, 2016. Query optimization using clustering and Genetic Algorithm for Distributed Databases. International Conference on Computer Communication and Informatics (ICCCI). IEEE.

[27] S Venkata Lakshmi and Valli Kumari Vatsavayi, April 2017. Teacher-Learner & Multi-Objective Genetic Algorithm Based Query Optimization Approach For Heterogeneous Distributed Database Systems. Journal of Theoretical and Applied Information Technology.

[28] Sunita A Yadwad, Dr V. Valli Kumari and Dr S Venkata Lakshmi, 2021. Service Outages Prediction through Logs and Tickets Analysis, International Journal of Advanced Computer Science and Applications (IJACSA).

[29] Simhadri Madhuri and Dr. S. Venkata Lakshmi (2023), A Machine Learning based Normalized Fuzzy Subset Linked Model In Networks for Intrusion Detection. Soft Computing.

[30] Simhadri Madhuri and Dr. S. Venkata Lakshmi (2023), Trusted Node Feedback Based Clustering Model For Detection Of Malicious Nodes In The Network. Journal of Theoretical and Applied Information Technology (JATIT), Vol.101. Issue No. 7.

[31] Improving Data Transmission Rate with Self Healing Activation Model for Intrusion Detection with Enhanced Quality of Service. International Journal on Recent and Innovation Trends in Computing and Communication (Q4), Volume: 11 Issue: 9s, 233-243. https://doi.org/10.17762/ijritcc.v11i9s.7417