ILLINOIS TECH

Gaussian Differential Privacy

CS 528 Data Privacy and Security

By

Yuanyuan Sun (A20487775)

Kevan Dedania (A20522659)

Hemanth Vennelakanti (A20526563)

Under the Guidance of Prof. Binghui Wang

Research Project Contributions

General Contributions	
Research	Kevan Dedania, Yuanyuan Sun, Hemanth Vennelakanti
Discussion	Kevan Dedania, Yuanyuan Sun, Hemanth Vennelakanti
Presentation	Kevan Dedania, Yuanyuan Sun, Hemanth Vennelakanti
Code Implementation	Hemanth Vennelakanti: CODE: GDP Exploration between Utility and Privacy
	Yuanyuan Sun: CODE: GDP with Machine Learning

Report Contributions	
1. Introduction	Kevan Dedania
2. Project Background	Yuanyuan Sun
3. An Overview of Research Focus, Dataset and Statistical Techniques	Yuanyuan Sun
3.1 Research Focus	Yuanyuan Sun
3.2 Introduction and Analysis of the Data Set	Yuanyuan Sun
3.3 Statistical Techniques We Work On	Yuanyuan Sun
4. Introduction to Gaussian Differential Privacy	Kevan Dedania
4.1 Gaussian Mechanism	Kevan Dedania
4.2 Differential Privacy and Randomized Techniques	Kevan Dednia

4.3 Gaussian Differential Privacy Properties	Kevan Dedania
4.4 Trade Off Functions of GDP	Kevan Dedania
4.5 Benefits of GDP	Kevan Dedania
5. Utility and Privacy Metrics	Hemanth Vennelakanti
5.1 Gaussian Noise Calculation	Hemanth Vennelakanti
5.2 Relative Error and Utility	Hemanth Vennelakanti
5.3 Experimental Observations	Hemanth Vennelakanti
6. GDP with Machine Learning	Yuanyuan Sun
6.1 Introduction to the Third Research Topic	Yuanyuan Sun
6.2 The “Noise Data” Approach	Yuanyuan Sun
6.3 Classification Models We Use to Examine Utility	Yuanyuan Sun
6.4 Measuring Utility of Noise Data for ML	Yuanyuan Sun
6.5 Budget's Influence on ML Features and Performance	Yuanyuan Sun
6.6 Intuitive Privacy Analysis of GDP with Machine Learning	Yuanyuan Sun
7. Challenges in Our Project	Kevan Dedania Yuanyuan Sun Hemanth Vennelakanti

1. Introduction

Kevan Dedania (A20522659)

Gaussian Differential Privacy is a technique for protecting the privacy of individuals in a dataset. The core idea of Gaussian Differential Privacy is to add random noise to the data before releasing it. The amount of noise added depends on the sensitivity of the data. The amount of noise added is determined by the sensitivity of the query and desired level of privacy protection. Gaussian differential privacy is a method of achieving differential privacy that involves adding noise to a query result in a way that it follows a Gaussian (normal) distribution. This technique can be used to protect sensitive data in a hospital dataset while still allowing researchers to gain insights from the data.

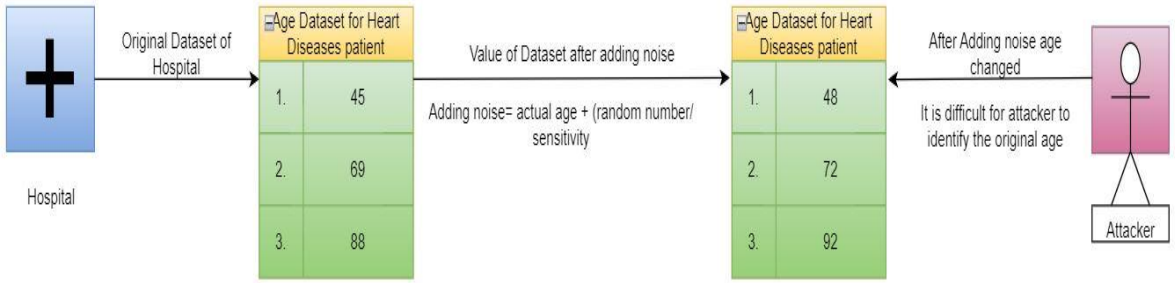


Figure 1 Age Dataset for Heart Diseases Patient

Let's consider an example to illustrate this concept in detail. Suppose we have a hospital dataset containing the age of heart disease patients. It might be possible that hospital need to release this data in front of the public. During that time, adding noise to the dataset changes the actual age of the heart disease patient. Adding noise makes it difficult for the attacker to identify the original age.

Formula for adding noise in the dataset

(Adding noise=actual age + [random number/sensitivity]).

To do this, they could use Gaussian differential privacy by first specifying a privacy parameter, epsilon (ε), which determines the level of privacy protection. The smaller the value of epsilon, the greater the privacy protection, but the more noise that needs to be added to the output.

Using the privacy parameter and sensitivity value, they can then determine the amount of noise to add to the query result. They would do this by generating a random value from a Gaussian distribution with mean zero and standard deviation sigma (σ), where the value of sigma is determined by the privacy parameter and sensitivity value. For example, if epsilon is 1 and the sensitivity is 2, then the value of sigma would be 1/epsilon* sensitivity = 2. Finally, by adding the random value generated from the Gaussian distribution to the true query result, to obtain the differentially private result. This process ensures that the output of the query is noisy enough to protect the privacy of the individual patients in the dataset.

Gaussian differential privacy is a useful technique for protecting sensitive data in a hospital dataset while still allowing statistical analysis to be performed on the dataset. By adding noise to the query results, this technique ensures that the privacy of individual patients is protected.

2. Project Background

Yuanyuan Sun (A20487775)

There are healthcare statisticians who work to research statistical information about the population based on certain conditions and then analyze the data to make

decisions on how to assign resources to the communities. This type of work is often done by public health officials, epidemiologists, and other professionals in the healthcare field. In order to make informed decisions about resource allocation, these healthcare statisticians would need to gather data from a variety of sources, including hospitals, clinics, and other healthcare providers. The data they collect could include information about the prevalence of certain health conditions, demographic information about the population, certain diseases, or other relevant factors. Once they have collected and analyzed the data, these healthcare statisticians can use their findings to make decisions about where to allocate resources in order to improve health outcomes for the community. This could include things like distributing vaccines, providing funding for community health programs, or investing in new healthcare infrastructure in areas that are underserved.

The healthcare statisticians are interested in obtaining heart disease data for the community from a specialist hospital, but the specialist hospital is not willing to provide the original data set. Instead, the specialist hospital will use Gaussian differential privacy to add random noise to the original data set, which will protect the privacy of individual patients while still allowing the staff members to analyze the data.

Figure 2.1

Gaussian differential privacy is a method for adding random noise to a data set in order to protect the privacy of individual patients. Essentially, the specialist hospital would apply a mathematical algorithm to the original data set that would add a certain amount of noise to each data point. This noise would be random and would not contain any personal information about individual patients, but it would still provide some statistical information about the data set as a whole. Once the specialist hospital has added noise to the data set, they can share it with the healthcare statisticians without compromising patient privacy. However, it's important to note that the amount of noise that is added to the data set can impact the accuracy of the analysis that the healthcare statisticians are able to do. If too much noise is added, the data may become too distorted to provide meaningful insights.

3. An Overview of Research Focus, Dataset and Statistical Techniques

Yuanyuan Sun (A20487775)

The healthcare statisticians are interested in obtaining heart disease data for the community from a specialist hospital, but the hospital is not willing to provide the original data set. Instead, the hospital will use Gaussian differential privacy to add random noise to the original data set, which will protect the privacy of individual patients while still allowing the staff members to analyze the data. Hence, we will perform some tasks to add Gaussian noise to the original heart disease dataset. It should be noted that the amount of noise that is added to the data set can impact the accuracy of the analysis. We will use some metrics and statistical techniques to do

comparisons among different heart disease noise datasets based on our research focus.

3.1 Research Focus

Yuanyuan Sun (A20487775)

In this research project, there are three research objectives/focus. One of our main objectives is to learn a new differential privacy mechanism called Gaussian differential privacy (GDP). GDP is a commonly used mechanism in privacy-preserving data analysis and has shown promising results in providing strong privacy guarantees while maintaining a reasonable level of utility. By exploring Gaussian differential privacy (GDP), we hope to gain a better understanding of how it works and how it can be applied in different contexts. The second focus of our research is exploring the utility and privacy of GDP. In order to protect sensitive information while still maintaining the utility of the data, we need to explore the relationship between privacy and utility within the privacy budget range we determined. By studying this relationship, we hope to identify strategies for achieving better privacy protection while minimizing the loss of utility. The third focus/topic in this research is that we aim to apply GDP with machine learning to see how it affects the accuracy of the classification models. Machine learning has become an essential tool for data analysis and decision-making, but it often requires access to sensitive data. By applying GDP to machine learning models, we hope to demonstrate that we can still achieve high accuracy while preserving privacy or gain insightful ideas and understanding regarding its effectiveness in preserving privacy while maintaining model accuracy by observing and analyzing the implementation of GDP with seven different classification models. This research will have important implications for the

development of privacy-preserving machine learning techniques and their adoption in real-world applications. Simplistically, our research focus is:

(1) Study and introduce a new differential privacy mechanism: Gaussian Differential Privacy (GDP).

(2) Explore GDP based on utility and privacy metrics.

(3) Apply GDP to seven classification models to gain insightful information about whether or not the noise data influences the performance of ML models significantly (verify the utility of the noise data for seven classification models). We can try to gain insights into the utility of different noise heart disease datasets and identify which budgets are most suitable for analyzing data with different levels of privacy protection we customize.

3.2 Introduction and Analysis of the Data Set

Yuanyuan Sun (A20487775)

This heart disease data set contains 1025 records of patients from 1988. There are four databases in this data set: Cleveland, Hungary, Switzerland, and Long Beach V. It contains 76 attributes, including the predicted attribute, but all published experiments refer to using a subset of 14 of them. In particular, the Cleveland database is the only one that has been used by ML researchers to this date. In our research, we use a data source that has 13 independent variables (also known as a predictor variable or feature) and one dependent variable (also known as a response variable or target variable). The dependent variable is "target" field. The "target" field refers to the presence of heart disease in the patient. It is integer-valued 0 = no disease and 1 = disease. It is convenient for researchers to do machine learning and statistical

analysis. We get the heart disease data set from the UCI Machine Learning Repository[1]. The UCI Machine Learning Repository is a collection of databases, domain theories, and data generators that are used by researchers in the field of machine learning for the empirical analysis of algorithms and data. The repository was created by the Center for Machine Learning and Intelligent Systems at the University of California, Irvine. The datasets in the UCI Machine Learning Repository cover a wide range of topics, including classification, regression, clustering, and recommender systems. Researchers can use these datasets to train and test machine learning algorithms as well as compare the performance of different algorithms. Here are the attribute names and their corresponding types in the heart disease data set we got from UCI:

ID	Attribute	Type	Meaning
1	age	numeric	Age of the patient in years
2	sex	binary	Sex of the patient (1 = male; 0 = female)
3	cp	ordinal	Type of chest pain the patient experienced (1 = typical angina; 2 = atypical angina; 3 = non-anginal pain; 4 = asymptomatic)
4	trestbps	numeric	Resting blood pressure of the patient (in mm Hg) upon admission to the hospital
5	chol	numeric	Serum cholesterol level of the patient (in mg/dl)
6	fbs	binary	Whether or not the patient's fasting blood sugar level was greater than 120 mg/dl (1 = true; 0 = false)
	restecg	ordinal	Results of the resting electrocardiogram (ECG) of the patient (0 = normal; 1 = having ST-T wave

ID	Attribute	Type	Meaning
7			abnormality; 2 = showing probable or definite left ventricular hypertrophy)
8	thalach	numeric	Maximum heart rate achieved by the patient during exercise
9	exang	binary	Whether or not the patient experienced exercise-induced angina (1 = yes; 0 = no)
10	oldpeak	numeric	ST depression induced by exercise relative to rest
11	slope	ordinal	Slope of the peak exercise ST segment (1 = upsloping; 2 = flat; 3 = downsloping)
12	ca	nominal	Number of major vessels (0-3) colored by fluoroscopy
13	thal	nominal	Blood disorder called thalassemia (3 = normal; 6 = fixed defect; 7 = reversible defect)
14	target	ordinal	integer-valued 0 = no disease and 1 = disease

Based on the attribute information, we can determine which variables are continuous and which are categorical. Continuous variables are those that can take on any numerical value within a certain range [2]. In this dataset, the following variables are continuous:

- age (numeric, Age of the patient in years)
- trestbps (numeric, resting blood pressure)
- thalach (numeric, maximum heart rate achieved)
- oldpeak (numeric, ST depression induced by exercise relative to rest)
- chol (Serum cholesterol level of the patient (in mg/dl))

Categorical variables are those that take on a limited number of values or categories [2]. In this dataset, the following variables are categorical:

- sex (binary: 0 = female, 1 = male)
- cp (chest pain type): categorical with 4 values (0 = typical angina, 1 = atypical angina, 2 = non-anginal pain, 3 = asymptomatic)
- fbs (fasting blood sugar): binary (1 = true, 0 = false)
- restecg (resting electrocardiographic results): categorical with 3 values (0 = normal, 1 = having ST-T wave abnormality, 2 = showing probable or definite left ventricular hypertrophy by Estes' criteria)
- exang (exercise induced angina): binary (1 = yes, 0 = no)
- slope (the slope of the peak exercise ST segment): categorical with 3 values (0 = upsloping, 1 = flat, 2 = downsloping)
- ca (number of major vessels colored by flourosopy): categorical with values 0, 1, 2, 3
- thal (thalassemia diagnosis): categorical with 3 values (0 = normal, 1 = fixed defect, 2 = reversible defect)

In our research, we utilize two datasets to conduct our analysis. These datasets differ by a single record. The first dataset is the original dataset, which consists of 14 variables - 13 independent variables and one dependent variable. The second dataset is the noise dataset (df_v2), which is created by removing a single record from the original dataset and adding specific noise to the independent variable values.

```
There are 3 people who are the oldest in the dataset, with an age of 77.
Number of records in the original dataset: 1025
By removing one record where age=77, the number of records in the new dataset (df_v2): 1024
+-----------+---------------------+
| Dataset   |   Number of Records |
+===========+=====================+
| df_v1     |                1025 |
+-----------+---------------------+
| df_v2     |                1024 |
+-----------+---------------------+
```

Figure 3.2.1

3.3 Statistical Techniques We Work On

Yuanyuan Sun (A20487775)

In our research, we will be utilizing several statistical measurements, such as mean (μ) and mean difference, variance and variance difference, and standard deviation. These measurements will help us understand the distribution of our data and identify any differences between our datasets.

Mean μ and Mean Difference

In statistics, the mean μ (pronounced "mu") is a measure of the central tendency of a set of numerical data. It is also known as the average and is calculated by summing all the values in the data set and dividing by the number of values. It is commonly used in descriptive statistics to give an idea of where the data is centered. We will use it to initially compare the difference between original independent variables and noise-independent variables. Mean-difference (diff_mean) = each mean of original independent variables - each mean of noise-independent variables. Comparing the mean difference on continuous/categorical independent variables between two data sets that differ in one record helps us understand how sensitive the model is to small changes in the data. The mean difference measures the average difference in the

values of the continuous/categorical independent variables between the two data sets. By comparing the mean difference between the original data set and the noise data set, we can observe how the addition of specific noise to the independent values affects the mean values of the variables. This can help us evaluate the impact of adding noise on the data's distribution and identify any potential bias that may have been introduced. Additionally, by comparing the mean difference between different privacy budgets, we can investigate the trade-off between privacy and model accuracy and find a good balance between them based on customized privacy budgets.

Variance and Variance Difference

Variance difference is a statistical measure that indicates the difference in spread or variability between two datasets. It is calculated by taking the difference between the variances of the two datasets. If the variance difference is close to zero, it suggests that the spreads of the datasets are similar. If the variance difference is large, it suggests that the spread of one dataset is more significant than the other. In machine learning, comparing the variance difference between two datasets can help us identify if one dataset has more variability or noise than the other. This information can help us determine which dataset is more suitable for training a model or which dataset may need further preprocessing to reduce noise or variability. It is important to note that variance differences alone may not be sufficient to draw a definitive conclusion about the differences between two datasets. It should be used in conjunction with other statistical measures and exploratory data analysis techniques to gain a more comprehensive understanding of the data.

Accuracy of Machine Learning Models

The accuracy of machine learning models is a statistical measure that is commonly used to evaluate the performance of a model. Accuracy is defined as the proportion of correct predictions made by a model out of the total number of predictions made. It is a metric that helps assess how well a model can generalize to new, unseen data [3]. Accuracy is often used in machine learning for binary classification problems, where there are two possible outcomes for each prediction. Using accuracy as a statistical metric to evaluate the performance of classification models on binary classification problems is a common and appropriate choice. In our research, we trained seven classification models on the original data set and different noise data sets, which is a binary classification problem. Using accuracy can help us compare the performance of the models on different datasets and assess their ability to generalize to new, unseen data. While accuracy is a useful statistical measure for evaluating the performance of machine learning models, it may not capture all aspects of model performance in certain scenarios, such as imbalanced datasets or when the costs of different types of errors vary. However, since our research focuses on observing and comparing the differences/influence after applying Gaussian differential privacy to train models, using accuracy as a performance metric is still useful. By comparing the accuracy of the models trained with and without Gaussian differential privacy, we can get a sense of how the privacy mechanism is impacting the models' predictive performance.

4. Introduction to Gaussian Differential Privacy

Kevan Dedania (A20522659)

Gaussian differential privacy is a technique used to protect the privacy of sensitive data in statistical analysis by adding noise to the output of a query. It belongs

to the family of differential privacy mechanisms that guarantee the privacy of individual data points while still allowing useful statistical analysis of the dataset as a whole. GDP is based on the idea that adding random noise to the data can help to obscure the individual data points while still preserving statistical properties of the data set.

The amount of noise added to the output of a query in GDP can be controlled by a parameter known as the privacy budget, which specifies the maximum amount of noise that may be added while still maintaining differential privacy. The privacy budget is frequently denoted by epsilon (ε) and is a measure of the mechanism's privacy protection strength.

GDP employs a Gaussian distribution to add noise to the output of a query, which is a probability distribution commonly used to model noise or errors in data. The quantity of noise introduced to the output is determined by the Gaussian distribution's standard deviation (sigma), which is proportional to the privacy budget (ε).

The larger the privacy budget, the more noise can be added to the report, increasing the likelihood that the output will be erroneous. If, on the other hand, the privacy budget is modest, the quantity of noise added to the output is also small, implying that the output is more likely to be correct.

In Gaussian differential privacy, the noise added to the output of a query is drawn from a Gaussian distribution. This ensures that the added noise is smooth and

doesn't significantly affect the utility of the data. The amount of noise added to the output is determined by the privacy budget, which is a parameter that represents the maximum amount of privacy that we're willing to sacrifice for the utility of the data.

The privacy budget controls the variance of the Gaussian distribution used to generate the noise. A larger privacy budget results in more noise being added to the output, providing stronger privacy protection at the expense of reduced utility. Conversely, a smaller privacy budget results in less noise being added to the output, providing weaker privacy protection but higher utility.

Gaussian Differential Privacy is a privacy-preserving approach that uses Gaussian noise to introduce randomization to a query's result while maintaining the privacy of individual data points. It is controlled by a privacy budget, which controls how much noise is added to the output, and this noise is generated using a Gaussian distribution with a standard deviation proportional to the privacy budget.

4.1 Gaussian Mechanism

Kevan Dedania (A20522659)

Probability of the mechanism outputting a result in S for the original dataset D should not be much larger than the probability of it outputting a result in S for the neighboring dataset D', where "not much larger" is defined by the parameter ε.

$$Pr[M(D) \in S] \leq Pr[M(D') \in S] \times e^{\varepsilon} + \delta$$

The parameter δ represents the maximum probability by which this condition may be violated (privacy loss parameter).

When δ = 0, the guarantee is simply called ε-DP.

4.2 Differential Privacy and Randomized Techniques

Kevan Dedania (A20522659)

The datasets are fixed, and the probabilities are calculated exclusively over the mechanism's randomness. The event E, in particular, can take any measurable set in the range of M. The randomized technique is required to achieve differentiated privacy. Consider the difficulties of privately revealing the average cholesterol level of people in the dataset S = (x1,..., xn), where xi corresponds to the individual's cholesterol level. A privacy-preserving mechanism may take the form.

$$M(S) = \frac{x_1 + \cdots + x_n}{n} + \text{noise}.$$

The level of noise must be high enough to disguise the features of any individual's cholesterol level, but not so high that it distorts the population average for accuracy purposes.

As a result, the probability distributions of M(S) and M(S') are close to each other for all datasets S, S' that differ in only one individual record.

4.3 Gaussian Differential Privacy Properties

Kevan Dedania (A20522659)

Strong Privacy Protection: Gaussian differential privacy provides robust privacy

protection by adding noise to query results in a way that protects the private of individual records in the dataset. This means that even if an opponent knows some information about an individual in the dataset, they cannot deduce sensitive information about that individual based on the query results.

Differential privacy composition: Gaussian differential privacy is modular; it can be used in conjunction with differential privacy to give strong privacy protection, allowing for privacy-preserving data analysis that protects sensitive data at every step.

Flexibility in noise generation: The use of Gaussian differential privacy allows for more flexibility in the synthesis of noise to be added to query results. The noise can be created using a Gaussian distribution with a mean of zero and a standard deviation that depends on the query's sensitivity and the level of privacy protection sought. This adaptability enables the optimization of the noise added to query results in order to balance privacy protection and query accuracy.

4.4 Trade Off Functions of GDP

Kevan Dedania (A20522659)

Informally, all variations of differential privacy require that it be difficult to identify any pairs of neighboring datasets based on information released by a private a mechanism M. From the standpoint of an attacker, it is natural to formalize the concept of "indistinguishability" as a hypothesis testing issue for two neighboring datasets S and S'.

H0 : the underlying dataset is S versus H1 : the underlying dataset is S`.

The output of the mechanism M serves as the basis for performing the hypothesis testing problem. Differential privacy is most naturally defined through a hypothesis testing problem from the perspective of an attacker who aims to distinguish S from S′ based on the output of the mechanism.

4.5 Benefits of GDP

Kevan Dedania (A20522659)

Strong Privacy Protection: GDP offers a stronger privacy guarantee than other mechanisms such as randomized response By adding noise drawn from a Gaussian distribution, it offers greater protection against attacks such as reconstruction and membership inference attacks.

Flexibility: GDP is a flexible method that can be applied to a variety of datasets and applications, such as machine learning, data mining, and statistical research. It can be utilized in instances where the data has continuous values, giving it a versatile tool for protecting sensitive data privacy.

Differential Privacy Composition: GDP has the advantage of allowing for differential privacy composition, which means it can be applied to the same dataset numerous times, each time adding a different quantity of noise. This enables for a gradual trade-off between privacy protection and data utility, allowing for the desired level of privacy while keeping data utility.

5. Utility and Privacy Metrics

Hemanth Vennelakanti (A20526563)

Utility in the context of differential privacy refers to how well the output of the data analysis algorithm preserves the accuracy and usefulness of the original data. High Utility implies high accuracy and useful results. A differential privacy system's utility and privacy guarantees can be evaluated using below metrics.

1. Epsilon : Privacy loss parameter. Measures the maximum amount of privacy loss. Smaller the value of epsilon, Higher level privacy but may also result in lower level of utility.
2. Delta : This measures the probability that the privacy guarantee will be violated, even when epsilon is set to the desired value. A smaller delta value corresponds to a higher level of privacy.
3. Accuracy : This measures how closely the output of the data analysis algorithm matches the true underlying distribution of the data.
4. Sensitivity : This measures the maximum amount that the output of the algorithm can change when a single data point is added or removed from the dataset.

5.1 Gaussian Noise Calculation

Hemanth Vennelakanti (A20526563)

According to the Gaussian mechanism, for a function f(x) which returns a number, the following definition of F(x) satisfies (ε, δ) - differential privacy.

$$F(x) = f(x) + N(\sigma)$$

N(σ) is calculated based on the Normal Distribution's scale.

$$\sqrt{2\log(1.25/\delta)}\ \frac{\Delta_2 f}{\varepsilon}$$

Here $\Delta_2 f$ means sensitivity of the query, ε is Privacy loss parameter, δ is Failure probability of privacy violation.

5.2 Relative Error and Utility

Hemanth Vennelakanti (A20526563)

Utility metric of numerical query result is based on the absolute value of the relative Error, which is given as :

$$\tilde{E} = |x_i' - x_i|/|x_i|$$

where xi' is the perturbation value of xi

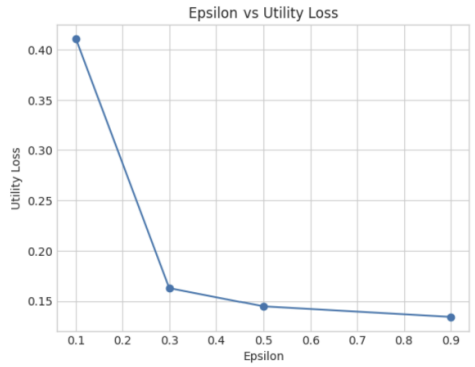
The utility metric of numerical query result for $f_i(D)$ is U = 1 - E

$$E = \frac{|f_i(D) + u_i - f_i(D)|}{|f_i(D)|}$$

For any query function f: D $\rightarrow R^K$, $f_i(D) + u_i$ is the perturbation value of $f_i(D)$.

5.3 Experimental Observations

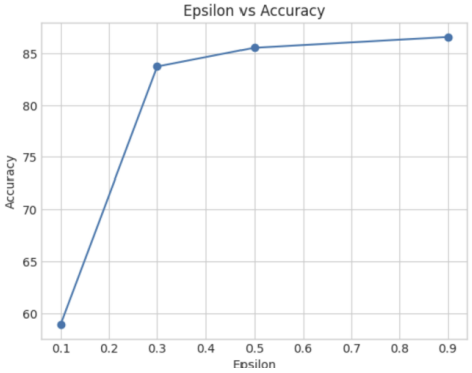
Hemanth Vennelakanti (A20526563)



epsilon	UtilityLoss
0.1	0.41
0.3	0.16
0.5	0.14
0.9	0.13

Figure 5.3.1

For different values of epsilon we tried to plot the utility loss. It can be clearly seen that Utility loss is decreasing for numerical query results and noise applied for increasing values of epsilon.



epsilon	Accuracy
0.1	58.92
0.3	83.73
0.5	85.54
0.9	86.58

Figure 5.3.2

From the accuracy table, for the highest value of epsilon (0.9), highest accuracy is achieved and accuracy is increasing for increasing values of epsilon. This is because, when epsilon is large, we have higher utility, less deviation from the original value, hence higher accuracy.

6. GDP with Machine Learning

Yuanyuan Sun (A20487775)

Differential privacy is a privacy-preserving technique that aims to protect individuals' sensitive information while still allowing data to be used for statistical analysis and machine learning to train models on a dataset without revealing any individual's sensitive information. Differential privacy can help achieve this by adding noise to the data such that the output of the machine-learning algorithm is not significantly affected. Gaussian differential privacy is a specific type of differential privacy that adds Gaussian noise to the data to preserve privacy. This technique is often used in machine learning because it allows models to be trained on sensitive data without revealing any specific individual's information. The amount of noise added to the data is controlled by a parameter called the "privacy budget", which determines the level of privacy protection offered. In machine learning, the privacy budget is often used to control the amount of noise added to the model parameters during training. By controlling the privacy budget, it is possible to balance the trade-off between privacy and accuracy and achieve an appropriate level of privacy while still obtaining useful insights from the data. However, the third research topic is based on how noise data influences the performance of the seven classification models. We focus on the noise influence. So, we don’t apply GDP to models by adding the amount of noise to the model parameters. On the contrary, we add Gaussian noise to the original data to train the models and observe their performance.

6.1 Introduction to the Third Research Topic

Yuanyuan Sun (A20487775)

In the previous research, we explored the relationship between utility and privacy with different privacy budgets when the privacy budget was between 0 and 1. The third research topic is to explore how a specific range of values of epsilon (between 0 and 1) influences the performance of seven classification machine learning models when the sensitivity is set to 1 and the delta is set to 1e-5. Based on this idea, we try to prove the utility by applying the different noise data to the seven classification machine learning models. We want to see if the seven different binary classification models whether or not still have statistical information after we add Gaussian noise to the independent variables. Based on prior research, we determined to work on the privacy budget epsilon $\in$ (0, 1), which is e=0.1, e=0.3, e=0.5, and e=0.9 to be added to the original heart disease dataset before sharing it with healthcare statisticians. The healthcare statisticians will then collect and analyze the noisy dataset using seven different classification machine learning models to make informed decisions about resource allocation that can help improve the health outcomes of the community. To use Gaussian differential privacy in machine learning, we try to follow these general steps to verify if the privacy budget influences the performance of the machine learning model:

1) Determine the specific privacy budget to be added to the original heart disease data (independent variables).
2) Split the noise heart disease dataset into two parts, a training set and a test set.
3) Train the machine learning model on the noisy training set.
4) Evaluate the performance of the trained model on the test set using appropriate

evaluation metrics such as accuracy, precision, recall, or F1 score. Here we will use the most common metric, which is called ‘accuracy’.

5) Repeat steps 1-4 for different privacy budgets to examine the effect of the privacy budget on the performance of the model.
6) Compare the performance of the models trained with different privacy budgets to determine if there is a significant difference in performance.

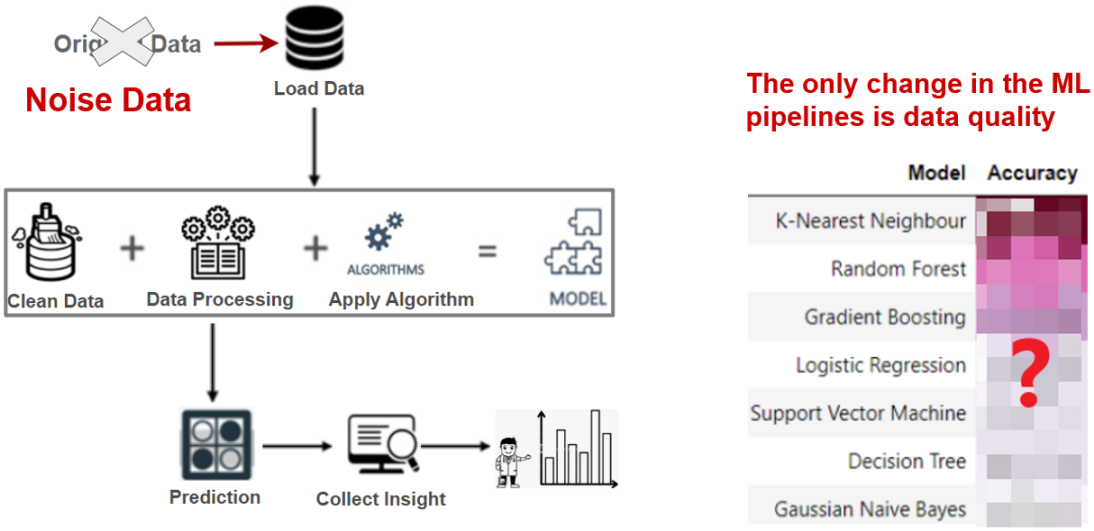


Figure 6.1 Model training process

By following these steps, we can check if the privacy budget influences the performance of the machine learning model significantly.

6.2 The “Noise Data” Approach

Yuanyuan Sun (A20487775)

We defined an approach called the "noisy data" approach: adding noise directly to the independent variables instead of adding noise to the output of the machine learning models or models’ parameters when using the Gaussian mechanism

for differential privacy. In our research, we adopt the “noisy data” approach, where Gaussian noise is added to the original data set based on the different types of features before it is used for training the 7 different classification models. Applying Gaussian differential privacy to add noise to feature values can be an effective approach for protecting the privacy of individuals in a dataset while preserving the utility of the data for machine learning.

This approach has several advantages, including simplicity, efficiency, and flexibility. It can be easily integrated into the machine-learning pipeline without requiring significant changes to the existing algorithms or models. It is simple to implement because it involves adding noise directly to the data, and it is efficient because it can be applied to large datasets without significantly impacting the computational resources required for machine learning. Moreover, the noisy data approach is flexible because we can adjust the amount of noise added to the data based on the level of privacy protection required. For instance, we can control the amount of noise added to the data by adjusting the standard deviation of the Gaussian noise added to the data. Although the noisy data approach can be effective in preserving privacy, it can also lead to a loss of accuracy in the machine learning model. This is because adding noise to the data can make it more difficult for the model to learn the underlying patterns in the data.

6.3 Classification Models We Use to Examine Utility

Yuanyuan Sun (A20487775)

We choose to use the 7 different classification models, which are K-Nearest Neighbour, Random Forest, Gradient Boosting, Logistic Regression, Support Vector Machine, Decision Tree, Gaussian Naive Bayes. Each of these models has its own

strengths and weaknesses, which can help us identify which influence the privacy budget could bring by comparison.

1) K-Nearest Neighbour is a simple yet powerful classification algorithm that classifies data points based on the class of their nearest neighbors in the feature space.
2) Random Forest is an ensemble learning algorithm that combines multiple decision trees to improve the accuracy of classification.
3) Gradient Boosting is another ensemble learning algorithm that builds multiple models in a sequential manner, where each model learns from the mistakes of the previous models.
4) Logistic Regression is a statistical method for analyzing a dataset in which there are one or more independent variables that determine an outcome. It is often used for binary classification problems.
5) Support Vector Machine is a classification algorithm that finds the hyperplane that maximally separates data points belonging to different classes.
6) Decision Tree is a simple yet powerful classification algorithm that recursively partitions the data based on the most informative features.
7) Gaussian Naive Bayes is a classification algorithm that is based on Bayes' theorem and assumes that the features are conditionally independent given the class variable.

By evaluating the performance of these 7 classification machine learning models on the noisy heart disease dataset with different levels of epsilon, we can try to gain insights into the utility of different levels of privacy and identify which models are most suitable for analyzing data with different levels of privacy protection. This can help us make informed decisions about the privacy budget to use for the dataset, and

can also inform the development of new machine-learning models that are more effective in privacy-preserving settings.

6.4 Measuring Utility of Noise Data for ML

Yuanyuan Sun (A20487775)

Machine learning model performance is a measure of how well a machine learning model is able to make accurate predictions on new, unseen data. It is an important aspect of machine learning and is used to evaluate the effectiveness of a model. Model performance can be evaluated using various metrics, depending on the type of problem being solved and the nature of the data. Common performance metrics include accuracy, precision, recall, F1 score, ROC curve, AUC score, and confusion matrix. The goal of model performance is to create a model that performs well on new data, which is known as generalization. In our research, we use a common and intuitive metric, which is ‘Accuracy’. As we discussed in sub-section 3.3, accuracy is often used in machine learning for binary classification problems. Using accuracy as a statistical metric to evaluate the performance of classification models on binary classification problems is a common and appropriate choice.

By comparing the differences among the accuracies of the 7 classification models trained on the original data and trained on the noisy data, we can get an idea of how much the noise has affected the model's performance. Would the privacy budget we choose result in the original data set losing its utility for machine learning training? In the third research topic, we use two intuitive and standard rules to do the comparison:

Rule 1. After applying GDP, if the differences in accuracies are changed within 20%

and accuracies are still greater than 80%, we consider the models trained on noise data still have enough statistical information to provide meaningful insights.

Rule 2. After applying GDP, if the differences in accuracies are changed by more than 20%, we consider the models trained on noise data do not have statistical information to provide meaningful insights. Healthcare statisticians cannot make use of it to make decisions.

Note: The rules only work for this research based on the accuracy results (the accuracy of all the seven models is greater than 80%) on the original dataset.

6.5 Budget's Influence on ML Features and Performance

Yuanyuan Sun (A20487775)

The heart disease dataset contains 13 independent variables, which can be classified into two types: continuous features and categorical features. Machine learning models are trained on independent variables. Here is the performance that was trained on the original data set:

Model	Accuracy
K-Nearest Neighbour	95.609756
Random Forest	88.780488
Gradient Boosting	86.829268
Logistic Regression	83.902439
Support Vector Machine	83.902439
Decision Tree	83.902439
Gaussian Naive Bayes	82.439024

Figure 6.5.1 Model Performance on Original Data

When training machine learning models on the heart disease dataset, it is important to handle these different types of features appropriately, for example, by normalizing the continuous features to have similar scales and encoding the categorical features as numerical values, such as one-hot encoding. We also notice that the privacy budget can influence the types of features to different degrees. Adding noise to the data can affect the continuous features and categorical features differently. For example, adding too much noise to the continuous features can cause the distribution of the data to become distorted, which can negatively impact the accuracy of the machine-learning models. On the other hand, adding noise to the categorical features may have less impact on the distribution of the data, but it can still affect the model's performance if the noise changes the meaning of the categorical values. Therefore, determining the specific privacy budget requires considering the specific characteristics of the data and the machine learning model being used. It may be necessary to adjust the budget for each feature type separately to achieve the desired balance between privacy and model performance. In subsection 3.2, we've classified independent variables into two types:

(1) continuous features

(2) categorical features

When we add noise to categorical features, we avoid adding noise to dummy variables and the corresponding features ('cp', 'thal', 'slope') which highly influence the 7 models' accuracy. Since the independent variables can be classified into two types: continuous features and categorical features. We have two sets of experiments regarding GDP with machine learning.

Experiment 1. Only apply GDP to add noise to continuous features among the 13 features and then train 7 different classification machine learning models on this noise data.

e=0.1	mean_diff
age	-10.617353
trestbps	-10.599774
thalach	-10.642743
oldpeak	-10.593524
chol	-10.708173

e=0.3	mean_diff
age	2.173167
trestbps	2.190745
thalach	2.147776
oldpeak	2.196995
chol	2.082346

e=0.5	mean_diff
age	-1.462767
trestbps	-1.445189
thalach	-1.488158
oldpeak	-1.438939
chol	-1.553587

e=0.9	mean_diff
age	-4.012632
trestbps	-3.995054
thalach	-4.038023
oldpeak	-3.988804
chol	-4.103452

Figure 6.5.2 The mean difference between the original and noise continuous features

By applying the GDP mechanism, we have added Gaussian noise to the original continuous feature values. The amount of noise added is controlled by the privacy budget, epsilon. We can clearly see the difference/change regarding the mean on each continuous feature by adopting different privacy budgets (epsilon $\in$ (0, 1)).

e=0.1	variance_diff
age	-0.417799
trestbps	0.257163
thalach	0.354893
oldpeak	0.000226
chol	-0.689636

e=0.3	variance_diff
age	-0.417799
trestbps	0.257163
thalach	0.354893
oldpeak	0.000226
chol	-0.689636

e=0.5	variance_diff
age	-0.417799
trestbps	0.257163
thalach	0.354893
oldpeak	0.000226
chol	-0.689636

e=0.9	variance_diff
age	-0.417799
trestbps	0.257163
thalach	0.354893
oldpeak	0.000226
chol	-0.689636

Figure 6.5.3 The variance difference between the original and noise continuous features

After only adding Gaussian noise to the values of continuous features, we can clearly see there is no obvious difference regarding the variance on each continuous feature by adopting different privacy budgets (epsilon $\in$ (0, 1)). Actually, based on the formula we used to add noise to the continuous features (Gaussian mechanism), the variance of the noise should be proportional to the privacy budget (epsilon). Therefore, we expect that the variance of the added noise will increase as we increase

the value of epsilon. The code implementation results in the opposite.

e=0.1	Model	Accuracy
	K-Nearest Neighbour	95.609756
	Random Forest	88.780488
	Gradient Boosting	86.829268
	Logistic Regression	83.902439
	Support Vector Machine	83.902439
	Decision Tree	83.902439
	Gaussian Naive Bayes	82.439024

e=0.3	Model	Accuracy
	K-Nearest Neighbour	95.609756
	Random Forest	88.780488
	Gradient Boosting	86.829268
	Logistic Regression	83.902439
	Support Vector Machine	83.902439
	Decision Tree	83.902439
	Gaussian Naive Bayes	82.439024

e=0.5	Model	Accuracy
	K-Nearest Neighbour	95.609756
	Random Forest	88.780488
	Gradient Boosting	86.829268
	Logistic Regression	83.902439
	Support Vector Machine	83.902439
	Decision Tree	83.902439
	Gaussian Naive Bayes	82.439024

e=0.9	Model	Accuracy
	K-Nearest Neighbour	95.609756
	Random Forest	88.780488
	Gradient Boosting	86.829268
	Logistic Regression	83.902439
	Support Vector Machine	83.902439
	Decision Tree	83.902439
	Gaussian Naive Bayes	82.439024

Figure 6.5.4 Model Performance on Continuous Noise Data

After only adding Gaussian noise to the values of continuous features, we can clearly see there is no obvious difference regarding the accuracy after training 7 classification models on continuous noise data by adopting different privacy budgets (epsilon $\in$ (0, 1)). According to rule 1, after applying GDP, if the differences in accuracies are changed within 20% and accuracies are still greater than 80%, we consider the models trained on noise data still has enough statistical information to provide meaningful insights. Models' performance does not significantly change after applying GDP with the specific epsilon to the continuous features.

Experiment 2. Only apply GDP to add noise to categorical features among the 13 features and then train 7 different classification machine learning models on this noise data.

e=0.1	mean_diff
sex	-0.573064
fbs	-0.573064
restecg	-0.572087
exang	-0.574040
ca	-0.575017

e=0.3	mean_diff
sex	4.069638
fbs	4.069638
restecg	4.070614
exang	4.068661
ca	4.067685

e=0.5	mean_diff
sex	-3.518385
fbs	-3.518385
restecg	-3.517409
exang	-3.519362
ca	-3.520338

e=0.9	mean_diff
sex	10.238434
fbs	10.238434
restecg	10.239411
exang	10.237457
ca	10.236481

Figure 6.5.5 The mean difference between the original and noise categorical features

By applying the GDP mechanism, we have added Gaussian noise to the original categorical feature values. The amount of noise added is controlled by the privacy budget, epsilon. We can see the difference/change in the mean for each categorical feature by adopting different privacy budgets (epsilon $\in$ (0, 1)).

e=0.1	variance_diff
sex	0.000117
fbs	0.000102
restecg	-0.000002
exang	-0.000212
ca	-0.003897

e=0.3	variance_diff
sex	0.000117
fbs	0.000102
restecg	-0.000002
exang	-0.000212
ca	-0.003897

e=0.5	variance_diff
sex	0.000117
fbs	0.000102
restecg	-0.000002
exang	-0.000212
ca	-0.003897

e=0.9	variance_diff
sex	0.000117
fbs	0.000102
restecg	-0.000002
exang	-0.000212
ca	-0.003897

Figure 6.5.6 The variance difference between the original and noise categorical features

After only adding Gaussian noise to the values of categorical features, we can clearly see there is no obvious difference regarding the variance of each categorical feature by adopting different privacy budgets (epsilon $\in$ (0, 1)). Actually, based on the formula we used to add noise to the categorical features (Gaussian mechanism), the variance of the noise should be proportional to the privacy budget (epsilon). Therefore, we expect that the variance of the added noise will increase as we increase the value of epsilon. The code implementation results in the opposite.

e=0.1	Model	Accuracy
	K-Nearest Neighbour	95.609756
	Random Forest	88.780488
	Gradient Boosting	86.829268
	Logistic Regression	83.902439
	Support Vector Machine	83.902439
	Decision Tree	83.902439
	Gaussian Naive Bayes	82.439024

e=0.3	Model	Accuracy
	K-Nearest Neighbour	95.609756
	Random Forest	88.780488
	Gradient Boosting	86.829268
	Logistic Regression	83.902439
	Support Vector Machine	83.902439
	Decision Tree	83.902439
	Gaussian Naive Bayes	82.439024

e=0.5	Model	Accuracy
	K-Nearest Neighbour	95.609756
	Random Forest	88.780488
	Gradient Boosting	86.829268
	Logistic Regression	83.902439
	Support Vector Machine	83.902439
	Decision Tree	83.902439
	Gaussian Naive Bayes	82.439024

e=0.9	Model	Accuracy
	K-Nearest Neighbour	95.609756
	Random Forest	88.780488
	Gradient Boosting	86.829268
	Logistic Regression	83.902439
	Support Vector Machine	83.902439
	Decision Tree	83.902439
	Gaussian Naive Bayes	82.439024

Figure 6.5.7 Model Performance on Categorical Noise Data

After only adding Gaussian noise to the values of categorical features, we can clearly see there is no obvious difference regarding accuracy after training seven classification models on categorical noise data by adopting different privacy budgets (epsilon $\in$ (0, 1)). According to rule 1, after applying GDP, if the differences in accuracies are changed within 20% and accuracies are still greater than 80%, we consider the models trained on noise data still have enough statistical information to provide meaningful insights. The models' performance does not significantly change after applying GDP with the specific epsilon to the categorical features.

6.5.1 Analysis of Noise Features Influence on Models' Performance

No matter whether we add noise to continuous features or categorical features, we don't want our noise data to lose statistical information. Based on the two sets of experiments. We can get the following conclusions:

1) The specific privacy budget (epsilon $\in$ (0, 1)) did not compromise statistical information on this data set.

2) Privacy budget's impact on continuous and categorical features deemed negligible in these 7 models.

We added different Gaussian noise to continuous and categorical features and compared the difference by mean of each feature after adopting different privacy budgets (epsilon $\in$ (0, 1)), the performance of 7 classification models was not obviously changed after adopting Gaussian differential privacy. Here are several possible reasons why the performance of the classification models was not significantly affected after adding Gaussian noise to the data:

1) Appropriate level of the privacy budget: The amount of noise added to the data is controlled by the privacy budget, which determines the level of privacy protection offered. If the privacy budget is set too high, the noise added to the data may be too large, resulting in a significant reduction in the utility of the data. On the other hand, if the privacy budget is set too low, the privacy guarantees may not be sufficient, especially if the epsilon values are customized between 0 and 1. It is highly possible that the privacy budgets we chose have been set at an appropriate level of privacy budget that allows for effective privacy protection while still preserving the utility of the data. As people know that the standard of privacy is determined by real-world problems, it's complicated, and ever-changing according to business requirements or user requirements. The practical research in our project is that we can measure the noise data's utility by models' performance.

2) Low sensitivity: The amount of noise added to the data depends on the sensitivity of the data, which is a measure of how much the output of the machine learning model can vary when one individual's data is removed. If the sensitivity of the data is low, adding Gaussian noise may not have a significant impact on the performance of the machine learning models. In the code implementation, I avoid adding noise to sensitive variables (dummy variables) and the corresponding features ('cp', 'thal', 'slope') which highly influence the 7 models' accuracy.

3) Small amount of noise: Gaussian differential privacy adds noise to the data by sampling from a Gaussian distribution with a mean zero and a standard

deviation determined by the privacy budget. If the standard deviation of the noise is small, the impact on the data may be minimal, resulting in little effect on the performance of the machine learning models.

4) Robust machine learning models: It is possible that the machine learning models you used are robust to noisy data. Robust machine learning models are less affected by noise and can still perform well even when the data is perturbed.

5) Appropriate choice of features: The choice of features can also have an impact on the performance of the machine learning models when adding noise. If the features selected are robust to noise or not sensitive to the privacy concerns, the impact of adding Gaussian noise may be minimal.

6.6 Intuitive Privacy Analysis of GDP with Machine Learning

Yuanyuan Sun (A20487775)

In differential privacy, the privacy budget parameter (epsilon) controls the level of privacy protection provided to the data. Based on prior research, the privacy budgets in our research are customized between 0 and 1 (epsilon $\in$ (0, 1)). After applying different privacy budgets and comparing noise data with the original data, we see the changes in the heart disease data set such as Figure 6.6.1 and Figure 6.6.2 show. By comparing the continuous noise data with different values of epsilon (0.1 and 0.9) to the original continuous data (Figure 6.6.1), we observe that the original continuous data has obviously changed. This is because adding noise is a fundamental

aspect of differential privacy, which intentionally perturbs the data to protect privacy. Intuitively, we learned that continuous data has a certain privacy that loses the truth (it does not compromise statistical information).

After Applying GDP when e = 0.1:

age	trestbps	thalach	oldpeak	chol
41.405108	114.405108	157.405108	-9.594892	201.405108
42.405108	129.405108	144.405108	-7.494892	192.405108
59.405108	134.405108	114.405108	-7.994892	163.405108
50.405108	137.405108	150.405108	-10.594892	192.405108
51.405108	127.405108	95.405108	-8.694892	283.405108

Original Continuous Data:

age	trestbps	thalach	oldpeak	chol
52	125	168	1.000000	212
53	140	155	3.100000	203
70	145	125	2.600000	174
61	148	161	0.000000	203
62	138	106	1.900000	294

After Applying GDP when e=0.9:

age	trestbps	thalach	oldpeak	chol
48.009829	121.009829	164.009829	-2.990171	208.009829
49.009829	136.009829	151.009829	-0.890171	199.009829
66.009829	141.009829	121.009829	-1.390171	170.009829
57.009829	144.009829	157.009829	-3.990171	199.009829
58.009829	134.009829	102.009829	-2.090171	290.009829

Original Continuous Data:

age	trestbps	thalach	oldpeak	chol
52	125	168	1.000000	212
53	140	155	3.100000	203
70	145	125	2.600000	174
61	148	161	0.000000	203
62	138	106	1.900000	294

Figure 6.6.1 Continuous noise data v.s. original continuous data when e=0.1

By comparing the categorical noise data with different values of epsilon (0.1 and 0.9) to the original categorical data (Figure 6.6.2), we observe that the original categorical data has obviously changed. Intuitively, we learned that categorical data has a certain privacy that loses the truth (it does not compromise statistical information).

After Applying GDP when e = 0.1:

sex	fbs	restecg	exang	ca
0.426936	-0.573064	0.426936	-0.573064	1.426936
0.426936	0.426936	-0.573064	0.426936	-0.573064
0.426936	-0.573064	0.426936	0.426936	-0.573064
0.426936	-0.573064	0.426936	-0.573064	0.426936
-0.573064	0.426936	0.426936	-0.573064	2.426936

Original Categorical Data:

sex	fbs	restecg	exang	ca
1	0	1	0	2
1	1	0	1	0
1	0	1	1	0
1	0	1	0	1
0	1	1	0	3

After Applying GDP when e = 0.9:

sex	fbs	restecg	exang	ca
11.238434	10.238434	11.238434	10.238434	12.238434
11.238434	11.238434	10.238434	11.238434	10.238434
11.238434	10.238434	11.238434	11.238434	10.238434
11.238434	10.238434	11.238434	10.238434	11.238434
10.238434	11.238434	11.238434	10.238434	13.238434

Original Categorical Data:

sex	fbs	restecg	exang	ca
1	0	1	0	2
1	1	0	1	0
1	0	1	1	0
1	0	1	0	1
0	1	1	0	3

Figure 6.6.2 Categorical noise data v.s. original categorical data when e=0.9

Since the performance of the 7 classification models has been trained on heart disease data by Gaussian differential privacy with no obvious change/difference, the privacy budget's impact on continuous and categorical features can be deemed negligible when the privacy budgets are customized between 0 and 1 (epsilon $\in$ (0, 1)). Based on this research result, it's better to utilize a larger privacy budget in this customized range of epsilon since it does not compromise statistical information in this heart disease dataset. This heart disease data set can have better privacy with a larger privacy budget and good utility when we use epsilon $\in$ (0, 1). It's important to note that the analysis and conclusion are based on the observation and experiments of the specific data set (heart disease) and seven binary classification machine learning models with Gaussian differential privacy. It may not hold true for other datasets or different use cases, and the appropriate level of privacy budget should be determined based on specific privacy requirements and trade-offs with data utility. Moreover, it's important to consider that while the performance of the classification models may not have been significantly impacted, the addition of noise still alters the data and potentially affects downstream applications beyond the scope of this research. Therefore, it's important to carefully evaluate the trade-off between privacy and utility in each specific use case and to consider the potential impact of privacy-preserving

techniques on the entire data pipeline.

7. Challenges in Our Project

Challenge 1

The accuracy of machine learning models is critical in ensuring their effectiveness in solving real-world problems. In the case of the heart disease dataset, the original data presented some challenges in achieving accurate results due to factors such as incomplete data, outliers, and irrelevant features. As a result, the accuracy of the 7 classification machine learning models on the original dataset was below 50%, making it difficult to meaningfully compare the performance of the models before and after applying Gaussian differential privacy. To address this issue, we performed feature engineering to clean and preprocess the data, which led to a significant improvement in the accuracy of the models. This improvement enabled meaningful comparisons to be made between the performance of the models before and after applying Gaussian differential privacy. (*Yuanyuan Sun (A20487775)*)

Challenge 2

One of the challenges in GDP is choosing an appropriate privacy budget that balances privacy protection with data utility. A small privacy budget may provide strong privacy protection. Adding noise can improve privacy, but it can also reduce the accuracy of the statistical analysis and finding the right balance between privacy and accuracy is a challenging task. *(Kevan Dedania (A20522659))*

Challenge 3

Initially we tried to work on analyzing the optimal budget and based on that

we have decided to train the Machine Learning Models however after understanding the Differential privacy we removed the concept of optimal budget and trained models on different values of epsilon. *(Hemanth Vennelakanti (A20526563))*

Reference

[1] https://archive.ics.uci.edu/ml/datasets/heart+disease

[2] Cai, J., Luo, J., Wang, S., & Yang, S. (2018). Feature selection in machine learning: A new perspective. Neurocomputing, 300, 70–79. https://doi.org/10.1016/j.neucom.2017.11.077

[3] Yin, M., Wortman Vaughan, J., & Wallach, H. (2019). Understanding the effect of accuracy on trust in machine learning models. Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, 1–12. https://doi.org/10.1145/3290605.3300509

[4] Hai Liu, Changgen Peng Balancing Privacy-Utility of Differential Privacy Mechanism: A Collaborative Perspective.

[5]Gaussian Differential Privacy

https://academic.oup.com/jrsssb/article/84/1/3/7056089