Homework 2

Name:

1. Laplace Mechanism

Using the same dataset as Homework 1: https://archive.ics.uci.edu/ml/datasets/Adult. Query the average age of the records with age greater than 25. Inject Laplacian noise to the query result (average age) to ensure 0.5-differential privacy and 1-differential privacy.

Tasks:

(a) In case of $\epsilon = 0.5$, generate 1,000 (random) results for the query (average age of the records with age greater than 25) over the original dataset, and generate 1,000 results for the query over each of three other datasets: (**10 points: 10 points for sensitivity and 10 points for implementation of noise**)

- removing the record with the oldest age
- removing any record with age 26
- removing any record with the youngest age

Hint: for the correct laplace noise satisfying the ϵ-DP, please have the discussion of the global sensitivity and corresponding noise amount in your report. Correct sensitivity and noise will have the full credits.

(b) In each of the above four groups of 1,000 noisy results, round each value to two decimal places. Then, discuss and empirically validate whether the last three groups of results (3,000 results in total) and the first group of 1,000 results are 0.5-*indistinguishable*. (**20 points**)

Guidance: To validate the 0.5-indistinguishability, consider the noisy query results from the original dataset D and the neighboring datasets D'_1, D'_2, D'_3 (obtained by removing the record with the oldest age / age 26 / youngest age).

1 **Discretize the output space:** Divide the entire result range into n equal-length intervals (bins) O_i, where $i \in [1, n]$.

2 **Estimate the probability distributions:** For each dataset, calculate the empirical probability of each interval:

$$Pr[\mathcal{A}(D) = O_i] = \frac{\#\text{ of results in } O_i}{1000}.$$

Similarly, obtain $Pr[\mathcal{A}(D'_1) = O_i]$, $Pr[\mathcal{A}(D'_2) = O_i]$, and $Pr[\mathcal{A}(D'_3) = O_i]$.

3 **Compute the empirical privacy ratio:** For each pair of neighboring datasets (e.g., D vs. D_1'), compute

$$\max_i \frac{Pr[\mathcal{A}(D) = O_i]}{Pr[\mathcal{A}(D_1') = O_i]}.$$

If this maximum ratio is smaller than $e^{0.5}$, the two datasets can be considered approximately 0.5-indistinguishable.

4 **Problems:** Based on your above calculation and observation, does the result satisfy the $\epsilon = 0.5$ privacy bound. If it not, what is the possible reason? Please show your above calculation results and your problem discussion in the report.

(c) Repeat all the above for $\epsilon = 1$ (another 4 groups of 1,000 results), compare all the 8 groups of 1,000 results (8,000 in total) and discuss the utility. (**10 points**) You can design a metric such as L_1 distance to calculate the error between the noisy result with true result to see the relationship between the ϵ and utility. Please have your utility error for different epsilon in the report.

2. Differentially Private Classification (50 points)

Considering the application of classifying the "iris plant" using the following dataset:

- The full dataset and description are available at: https://archive.ics.uci.edu/ml/datasets/Iris.
- Four attributes and three classes (using iris.data).
    1. sepal length in cm
    2. sepal width in cm
    3. petal length in cm
    4. petal width in cm
    5. class: Iris Setosa, Iris Versicolour, and Iris Virginica
- It is a small-scale dataset (150 records). Consider records (1-10, 51-60, 101-110) as testing data for prediction, and the remaining 120 records as training data.

Tasks:

(a) Build a Naive Bayes Classifier to predict the classes for records 1-10, 51-60, and 101-110 (both training and prediction). See more information about Naive Bayes Classifier: https://en.wikipedia.org/wiki/Naive_Bayes_classifier. (**10 points**) For this non-private classifier, please show the screenshot of running and results of training and testing in your report.

(b) Design and implement a differentially private algorithm (satisfying ϵ-differential privacy) to train the Naive Bayes Classifier for prediction. (**25 points**)

Hints: you can consider the Laplace Mechanism and allocate budgets to ensure ϵ-differential privacy for the algorithm. The algorithm consists of many queries, and you can think about sequential composition and parallel composition for the queries. You need to show the algorithm, the composition analysis and privacy budget allocation in your report.

(c) Set $\epsilon = 0.5, 1, 2, 4, 8$, and 16. Then, calculate the precision and recall of the prediction results (of records 1-10, 51-60, and 101-110) generated by the differentially private classifier. Note that the true results of the 30 records are given in the original dataset for benchmarking. (**15 points**) Please have this results in your report.

Submission Part: (1) a report including the analysis, discussion, screenshots of the procedures and results, and (2) source code files.