Hasso-Plattner-Institute, Potsdam, Germany
Privacy Preserving Outlier Detection - Winter Semester 2023/24

Assignments #4-5: Anonymising Textual Data and De-Anonymisation – (100 marks)

January 10, 2024

1 General Goal

This assignment builds on the previous one. In this case, our goal is two-pronged the first is to anonymise a dataset containing textual data and the second is to find elements within a given dataset that enable de-anonymisations.

2 Dataset

Each task will indicate which or what type of dataset you should use.

3 Tasks...

3.1 Textual Data Anonymisation – 30 marks

As a first task, select a textual dataset of your choice. For example, you may select the Trip Advisor Reviews Dataset (https://www.kaggle.com/datasets/andrewmvd/trip-advisor-hotel-reviews) or the Tweets Emotions Dataset (https://www.kaggle.com/datasets/pashupatigupta/emotion-detection-from-text).

1. Do some research to determine what needs to be anonymised in the data and why.
2. Using a Natural Language Processing library (e.g. Python's spaCy), analyse the text to identify elements of personally identifiable information (PII).
3. Using the techniques you applied in Assignment #1, apply a masking or transformation mechanism to modify the detected PII elements and substitute with suitable replacements.

4. Analyse the text to determine what if any information can be obtained after the transformation process. What conclusions can you draw from this?

3.2 De-anonymising a dataset – 50 marks

Exchange datasets with your colleagues. You should then have a new anonymised dataset provided to you that is different from the one you have been working on for Assignments #1 and #2-3. (NOTE: Your proposed algorithm/model must be inspired by outlier detection. In other words you need to find features within the data that allow you to characterise a given data point as an outlier/anomaly in terms of how it can be used to violate the anonymity of the dataset generated from the anonymisation process.)

1. Using standard search mechanisms, determine if there are any elements within the dataset that you received, that allow for de-anonymistaions to occur. Make a note of what you find and explain the procedure you used.

2. Design a de-anonymisation algorithm and apply to both the received dataset and your dataset. Report on the following:

   (a) What are you able to discover using your de-anonymisation algorithm?

   (b) How your discoveries (with the anonymised algorithm) compare to the results from Q1?

   (c) Propose an alternative anonymisation approach and apply it to the original version (not anonymised) of the dataset received from your colleague. (**NOTE:** You must complete the tasks above before you request the original dataset in order to do this task.)

.

3.3 Experiments – 20 marks

Design experiments to test the following:

1. The utility of the data that you have generated using your proposed anonymisation scheme (algorithms) for Q2.c.

2. Analyse the new (anonymised) dataset for risks of de-anonymisation.

3. Propose a method of assessing the risk of disclosure (de-anonymisation) and use this metric to evaluate your anonymised datasets (from Assignments #1, and #2-3), the anonymised dataset received from your colleague, and your version of the anonymised dataset that you obtained in Q2.c.

4 Submissions

Once you have completed your assignment, please submit a 6-10 page report analysing your results. Code can be uploaded to Github or Gitlab and a link shared. Please note that you should aim to keep the same repository for all the assignments. Your report should contain enough details to ensure that your procedure is repeatable for grading purposes.