Initial Plan

Generating Differentially Private Datasets Using Deep Learning

(in collaboration with: Office for National Statistics)

Module: 1819-CM3203 - 40 credits

Author: Wei Loon Teh

Supervisor: George Theodorakopoulos

Moderator: Luis Espinosa-Anke

1. Project Description

Government organisations, businesses, academia, members of the public and other decision-making bodies require access to a wide variety of administrative and survey data to make informed and accurate decisions. However, the collecting bodies are often unable to share sensitive data with third-parties without risking breaking the confidentiality and consent checks required to obtain this data.

Even aggregation methods or other functions that distort original data could still leak information through reconstruction attacks or model inversion attacks, such as described in (Fredrikson, Jha, & Tech, n.d.)

Therefore, researchers have proposed many methods for generating synthetic data to replace the raw data for the purposes of processing and analysis. A good synthetic dataset has two properties: it is representative of the original data and it provides strong guarantees about privacy.

This project is in collaboration with the Office for National Statistics – Data Science Campus, and it involves the application of the concept of differential privacy for the generation of synthetic data using deep learning techniques. It is a condition by our collaborator that the code from the project be open and available on github.

2. Aims and Objectives

The foremost goal of the project is to be able to create a privacy preserving synthetic dataset that is retains its utility and representativeness. The objectives to achieve are listed below

Essential Objectives

- Able to generate representative privacy preserving synthetic data
    - Initial approach idea is to use Generative Adversarial Neural Networks to produce data and Differential Privacy techniques to protect privacy
    - Time permitting: May explore other methods such as autoencoders.
- Evaluate and analyse the representativeness of the synthetic data
    - Test and tweak model to make sure the synthetic data is still useful for analytics
- Evaluate and analyse the privacy protection of synthetic data
    - Test that data retrieved is sufficiently obscured that original information is not leaked

3. Ethical Approval

After discussion with my supervisor, ethical approval is not needed. We will be using a public data set for the purposes of the project, more specifically the US Adult Census Dataset (available from https://archive.ics.uci.edu/ml/datasets/Adult) and as stated in the Ethics Procedure, publicly and lawfully published information such as in the Census is not subject to the Ethical Review process.

4. Work Plan

Week (Begins)	Work Progress	Milestone
Week 1 (27th Jan)	• Complete Initial Plan • Literature research on work done in the current field	Complete & Submit First Initial Plan
Week 2	• Obtain US Census bureau data for pre-processing and data exploration. • Learning and experimenting with Tensorflow, Keras and GANs • Learning and reading about Differential Privacy Techniques	Obtain first data & suitable libraries to process data
Week 3	• Train simple GAN without Differential Privacy techniques on the dataset to generate initial batches of synthetic data	First Synthetic Data
Week 4	• Analyse the representativeness of the data and improve on the GAN output by modifying GAN structure and training • Test for common data errors such as overfitting or underfitting	**First Review Meeting**

Week 5	• Introduce noise into the data using differential privacy techniques	First Privacy Preserving Synthetic Data
Week 6 Week 7 Week 8	• Develop platform based on earlier work for modifying GAN structure to test different models and automating portions of training and output • Experiment with different levels of noise and different neural net structures	Code platform that allows for training GANs with different structures
Week 9 Week 10	• Test and evaluate data privacy and representativeness of different models • Experiment with different levels of noise and different neural net structures • Explore other options and possibilities if time permits	Complete all Coding requirements **Second Review Meeting**
Week 11 Week 12 Week 13	• Final Report Write up • Contingencies	Final Report Drafts
Week 14 (5th May)	• Review final report, proofread and check references • Publish code to GitHub	Submit final report

5. References

1. Fredrikson, Matt, Jha, Somesh, & Ristenpart, Thomas. 2015. Model inversion attacks that exploit confidence information and basic countermeasures. Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security Pages 1322-1333 https://dl.acm.org/citation.cfm?id=2813677