

Research Proposal——Add privacy protection in Generative Adversial Networks(GAN)

Haihan Gao

October 7, 2022

Generative Model has been used to generate fake data from the distribution of real data. Some impressive work has been done with data generation such as Generative Adversarial Network(GAN) and Variation Autoencoder(VAE). An universal idea of building a generative model is to map real data(such as images and texts) into latent variables which represent the information or cluster characteristic in latent space. Then we use latent variables from encoder as the parameter of the target distribution. Sampling from the target distribution will generate new data which retains features from origin data as well as contains differences compared with origin input. In order to use Gradient descent to optimize deep model, we need to design loss function predefinitely. Generally the loss function is made up with two parts. The first one is difference with input, such as mse loss. Another one is KL divergence between latent distribution and a prior assumption. If we give the first part a high priority, we will find that our model tends to "copy" input data into output data. This may result into leakage of training data.

1 Scope of Work - 4 Questions

In this section the essence of the proposed work is described by answering four key questions.

What is the problem you want to address in your work? insert answer here

Why is it a problem? insert answer here

What is the solution you developed in your work? insert answer here

Why is it a solution? include answer here

2 Preliminary Table of Contents

In this section the table of contents for the proposed work is described.

1. **Section 1 Name** insert brief description
 - a) **Subsection 1 Name** insert brief description
 - b) **Subsection 2 Name** insert brief description
2. **Section 2 Name** insert brief description

3 Relevant Related Work

In this section, identified related work is described.

[gruba'how'2017] insert brief description

[zobel'writing'2015] insert brief description