

# EFFECTS OF FEDERATED LEARNING ON THE PERFORMANCE OF GENERATIVE MODELS

Author: Alexandru Ojica  
a.n.ojica-1@student.tudelft.nl

Supervisor: Swier Garst  
Responsible Professor: David Tax

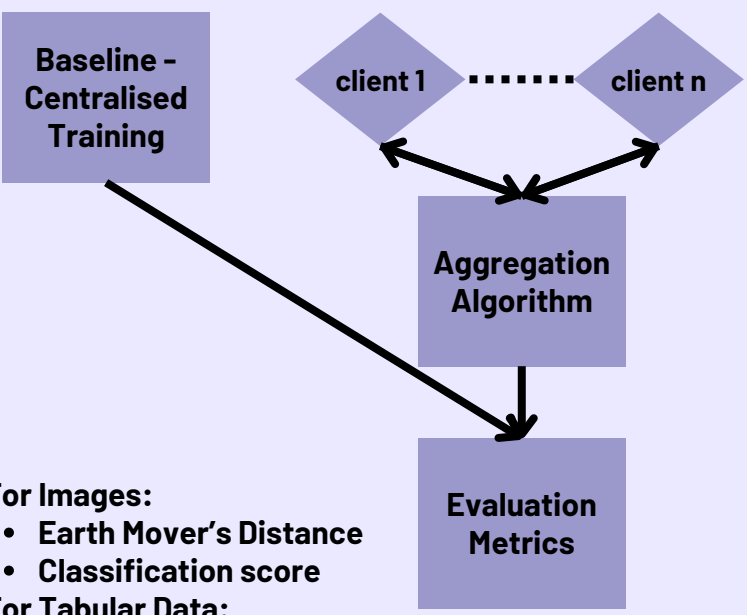
## 1. Introduction

- Federated Learning(FL) makes training ML models in a distributed manner.
- Non-IID data in FL leads to model divergence
- Data augmentation using generated data may solve this issue
- Research on generative models (GANs and VAEs) trained in a federated manner is needed

## 2. Methodology

Questions to be answered:

- How well do generative models perform when trained in a Federated Manner on image generation tasks?
- How well do generative models trained in a Federated Manner on Tabular Data generation tasks?

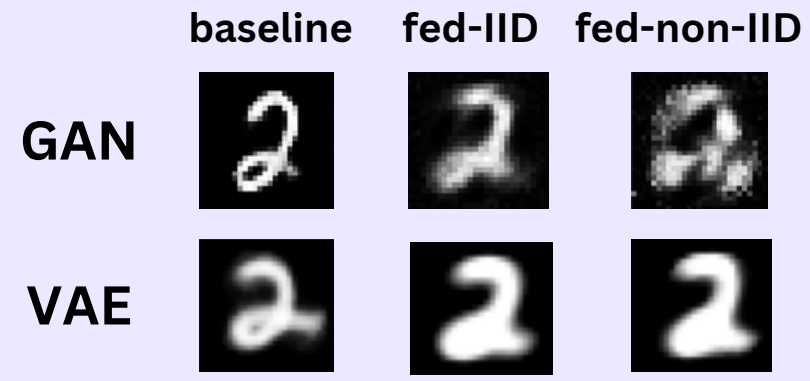


- For Images:**
- Earth Mover's Distance
  - Classification score

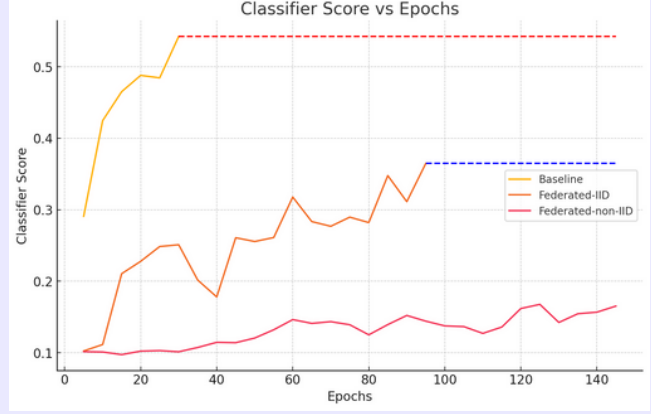
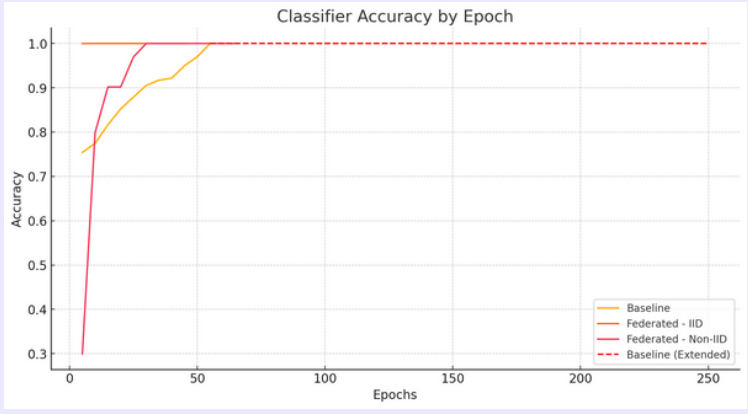
- For Tabular Data:**
- Resemblance - how well synthetic data distribution resembles real data distribution
  - Discriminability - how well a classifier distinguishes between real and synthetic data
  - Downstream Utility - how well a ML model performs on synthetic data
  - Privacy - how well is the synthetic data protected

## 3. Results

### MNIST

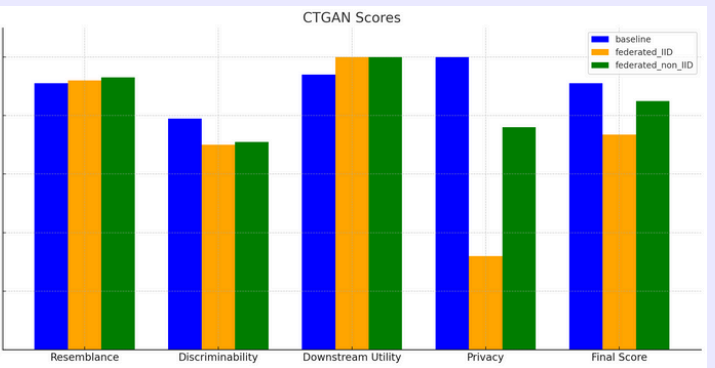
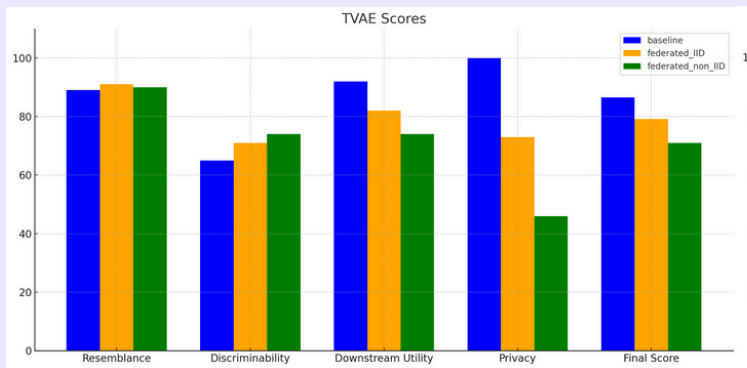


### CIFAR-10



### VAE- MNIST

### GAN - CIFAR-10



### TVAE

### CTGAN

## 5. Conclusion

- Federated Learning with non-IID data has a big impact on the convergence of GANs, especially on conditional models
- VAEs converge faster and are more stable in Federated Settings
- When it comes to tabular data, GANs are overall better suited for Federated Learning with non-IID data while VAEs are more stable in Federated setups with IID data

## 6. Future work

- Investigate how different architectures of GANs and VAEs perform under federated setups
- Investigate how Diffusion Models perform when used in Federated Learning