

# PRESENTATION OUTLINE: Determining the susceptibility of MD-GAN networks to nodes containing datasets with different class distributions

Omid Davoudi  
School of Computer Science  
Carleton University  
Ottawa, Canada K1S 5B6  
*omiddavoudi@cmail.carleton.ca*

November 5, 2019

## 1 Distributed Machine Learning

- Very larger datasets
- More computational power needed
- Globalization of data
- Regulations preventing data movement

## 2 Distributed Machine Learning - Common Methods

- Parameter Server
- Federated Learning
- Gaia

## 3 Generative Adversarial Networks (GANs)

- Generative models
- Most commonly used for generating images
- Consist of two neural networks, a Generator and a Discriminator
- Training is tricky, convergence is hard to achieve

## 4 MD-GAN

- GAN architecture designed for distributed training
- Main idea is to share parameters for the generator, but have unsynchronized local discriminators
- Assumes each node has a dataset with similar class distributions
- Periodically switches discriminators between nodes to avoid overfitting on local datasets.

## 5 MD-GAN - Potential vulnerability

- Differences in class distributions between different datasets
- These differences always exist in real life
- Theoretically, they can result in non-convergence
- Switching discriminators between nodes with highly divergent class distributions could result in catastrophic forgetting

## 6 MD-GAN - Testing the extent of vulnerability

- Proposed test setup: Cluster with at least 4 nodes
- Distribute the data evenly among the nodes
- Train until convergence
- Create skewed datasets each with one class being overrepresented
- Repeat the experiment with increasingly more skewed datasets

## 7 MD-GAN - Vulnerability results

- Results go here: Convergence time (or lack of convergence) for each dataset

## 8 MD-GAN - Potential avenue for training speed-up

- Training of the discriminator could happen while the generator is being synchronized
- Training happens using the old generator
- Better utilization of computational power
- Could destabilize the training process or result in non-convergence
- Test done with non-skewed datasets

## **9 MD-GAN - Speed-up results**

- Results go here: Convergence time (or lack of convergence) for using this method compared to not using it

## **10 Conclusions**

- MD-GAN is valid/invalid for real world/skewed datasets
- MD-GAN speed-up did/did not give favourable results

## **11 References**

- References go here...