PRIVACY-PRESERVING DEEP-LEARNING

Jerald Yik, Tan Zhen Yuan, Zaw Maw Htun

Created for the purpose of Final Presentation for Renaissance Capstone Project RCP2020/21

ABSTRACT

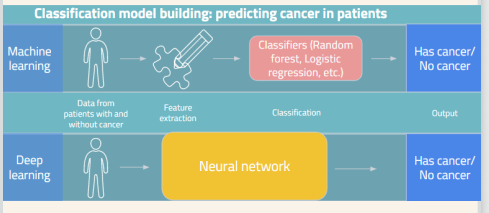
As concern about privacy grows with increasing usage of machine learning, federated learning (FL) appeals in its ability to preserve privacy across clients whilst training a central model. Differential privacy (DP) complements FL by obfuscating personally identifiable information (PII).
We investigated the effects of the different hyperparameters on accuracy, and ascertained the relationship between privacy and noise added.

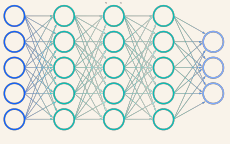
OBJECTIVE

1. Explore the effects of Federated Learning & Differential Privacy on accuracy
2. Ascertain the relationship between the privacy budget ε and the nature of noise added (standard deviation)

DEEP LEARNING

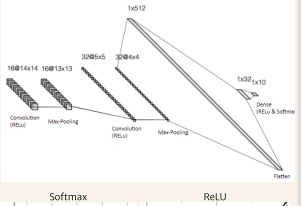
Deep learning is a subset of machine learning, with some crucial differences:

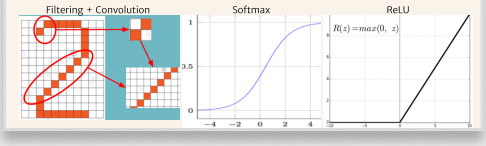




Models define key features by themselves within the layers of neural networks

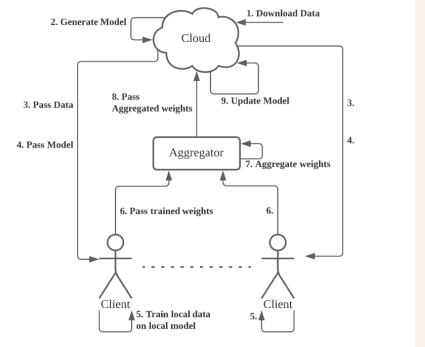
Convolutional Neural Network (CNN)
Utilises layers such as filtering, convolution, ReLU, softmax





FEDERATED LEARNING

The federated learning (FL) stakeholders in our project are grouped by the Central Server and the Clients, and the process follows this flow:



Central Server

Cloud: *Model creation, consolidation of weights, updating global model, retrieval & parsing of dataset*
Aggregator: *Consensus change to model weights*

Client

Trains local model using local data

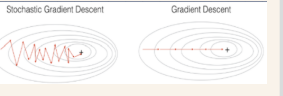
DIFFERENTIAL PRIVACY

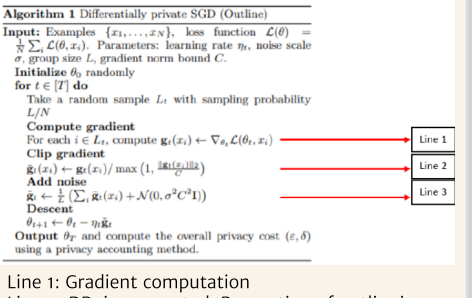
Helpful in preserving privacy by preventing attacks such as membership inference attack

Mechanism:
1. DP based on randomised function M with privacy budget ε:
$\Pr[M(D) \in S] \leq \exp(\varepsilon) \Pr[M(D') \in S] + \delta$
2. Noise n added (obfuscation) to query response f of a data D in function M:
$M(D) = f(D) + n$

Smaller ε gives better privacy preservation

Stochastic gradient descent:
Used to optimize parameters

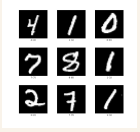




Line 1: Gradient computation
Line 2: DP-incorporated, Prevention of outlier in each epoch
Line 3: Noise addition (Gaussian)

IMPLEMENTATION, EXPERIMENTS & INSIGHTS

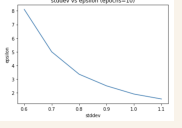
Dataset



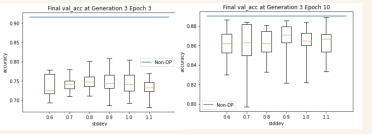
Services Used

colab, Amazon SageMaker, jupyter

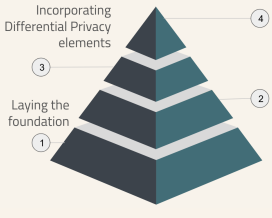
Inverse Relationship



Less accurate for **DP** & **less-trained** models

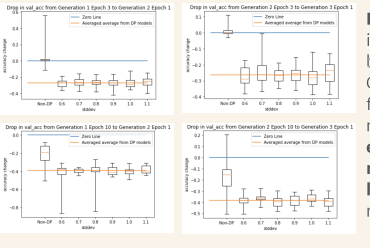


Implementation Timeline



3 Epochs per Gen:
10 Epochs per Gen:

Generation 1 to 2:
Generation 2 to 3:



Definite Drop in accuracy between Generations for **all** models, **except** for **non-DP, less-trained** models

FUTURE DEVELOPMENTS

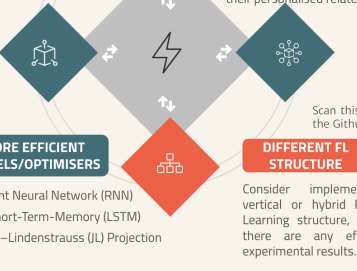
DIFFERENT HYPERPARAMETERS

Experiment with different hyperparameters to ascertain a relationship between Standard Deviation & Epsilon through experiments.

DECENTRALISED FEDERATED LEARNING

This will remove the need for a Learning Coordinator, and the client-side is responsible for training local models.

Clients can train on their local models on data not received from a centralised pool from the Cloud, but instead with their personalised related dataset.



Scan this QR Code to view the Github Repository

MORE EFFICIENT MODELS/OPTIMISERS

Recurrent Neural Network (RNN)
Long-Short-Term-Memory (LSTM)
Johnson–Lindenstrauss (JL) Projection

DIFFERENT FL STRUCTURE

Consider implementing a vertical or hybrid Federated Learning structure, to see if there are any effects on experimental results.