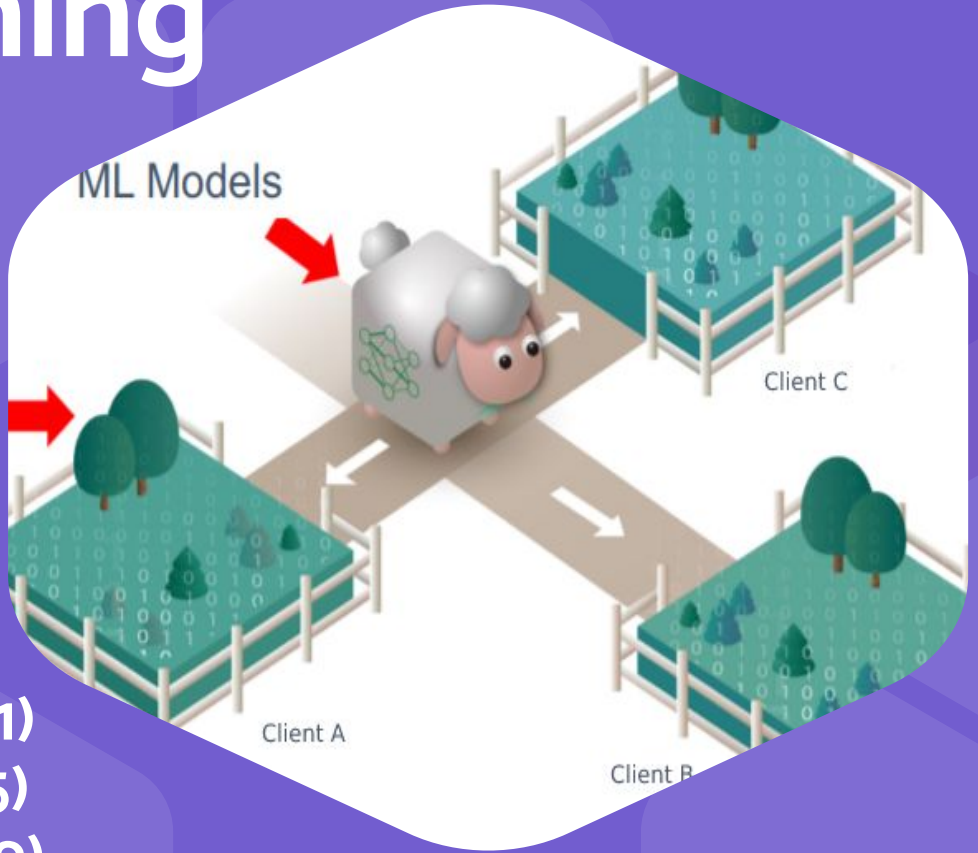
Federated Learning



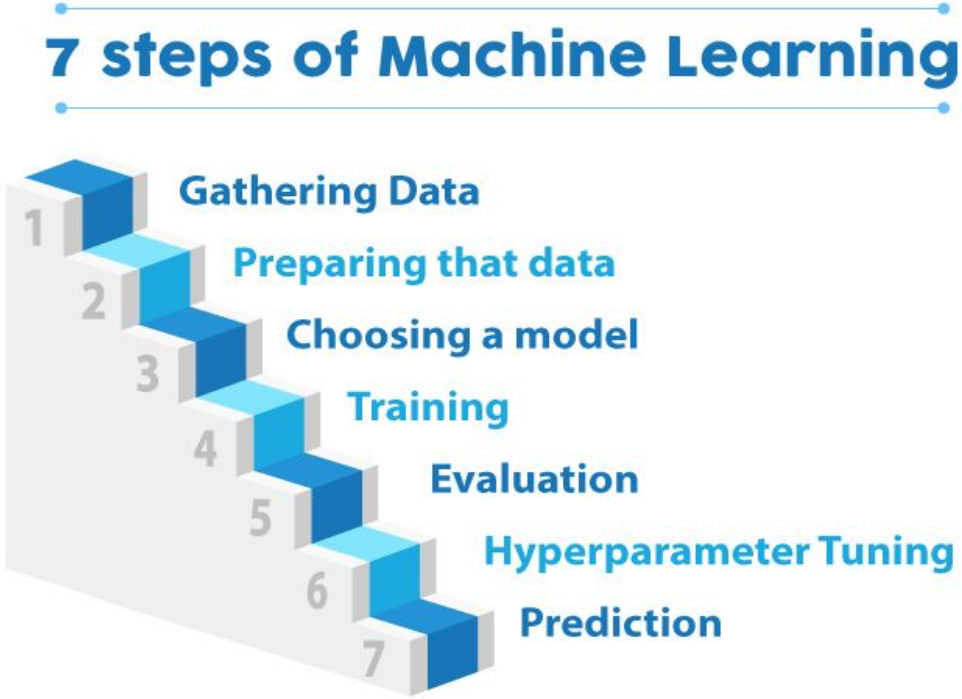
Team Members

Aaditya Mani Subedi (075BCT001)
Arpan Pokharel (075BCT015)
Saugat Kafley (075BCT099)

1 Background

The emergence of the need of Data Privacy

Contemporary Machine learning approach



Background

- Increasingly strict laws on Data Security and Data Protection.
- Growing concern on user privacy and data security
- Data exist in the form of isolated point in a system where data is kept and segregated from other parts of the architecture.

Data sharing among parties : Difficult, illegal or even Immoral

- Sensitive data between corporations cannot be sent directly such as
  - Medical reports data,
  - Research materials
- We need more control over data privacy and security
  - Corporate Security and Confidentiality Concerns
  - Data privacy concerns

Data sharing among parties : Difficult, illegal or even Immoral

- Data Privacy tampering.

Movement for Data Protection

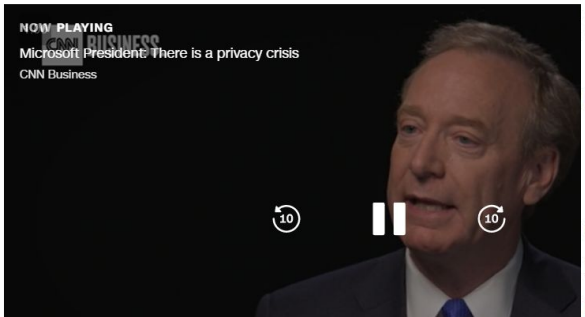
CNN BUSINESS Markets **Tech** Media Success Perspectives Videos

Facebook finally rolls out privacy tool for your browsing history

By Kaya Yurieff, CNN Business
Updated 1839 GMT (0239 HKT) August 20, 2019

Top Microsoft exec says online privacy has reached 'a crisis point'

By Clare Duffy, CNN Business
Updated 174



Google strengthens Chrome's privacy controls

Frederic Lardinois

Image Credits: Phillip Waterman / Gett

Google today announced a major new
long run, introduce significant changes
users' privacy across the web.

Challenges for prevailing AI

- Data is present in isolated forms and fragmented.
- Non-iid (Independent and identically distributed)
- Unbalanced data
- Data can be malicious and outdated.

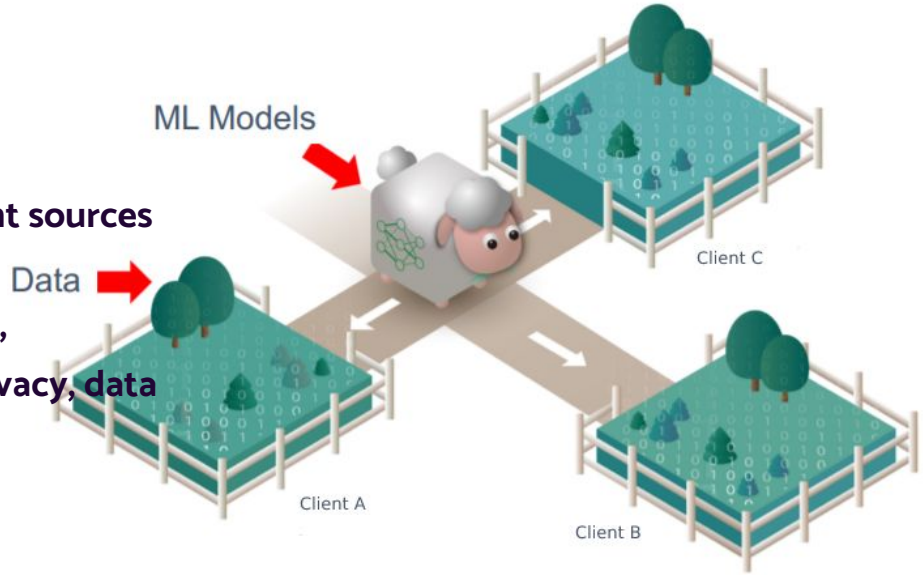
2

Federated Learning

A Potential solution!

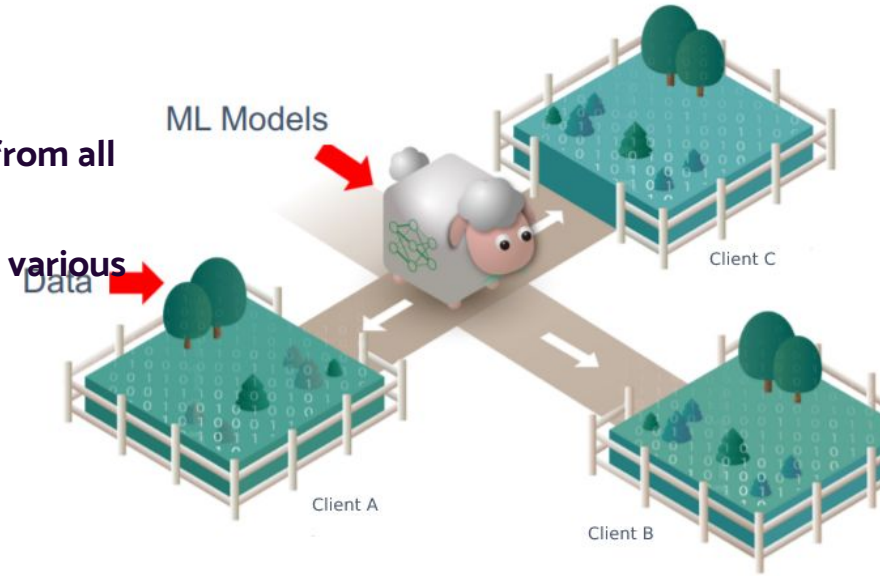
Let’s take a Simple Example.

- Our Interpretation
  - Model == sheep
  - Data == grass
- Originally, one need to purchase grass from different sources to feed sheep
  - Companies gather lots of data to train models,
  - where many challenges exist, such as user privacy, data security and regulations.



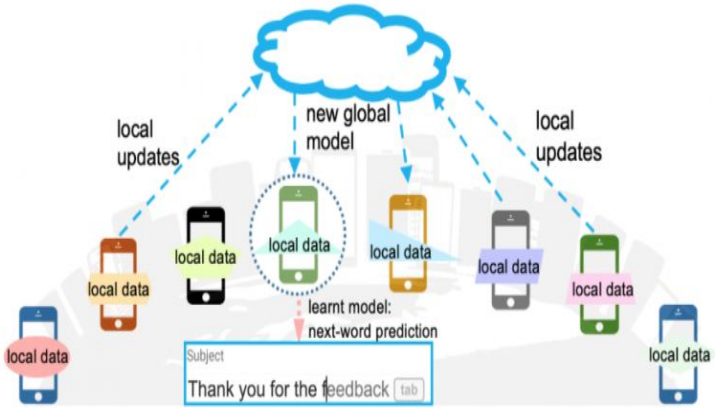
Let’s take a Simple Example.

- Federated Learning provides an alternative:
- Sheeps are led to different farms and can thus eat grass from all places without having to move the grass. ---
    - Federated learning models gather knowledge from various sources of data without having to observe



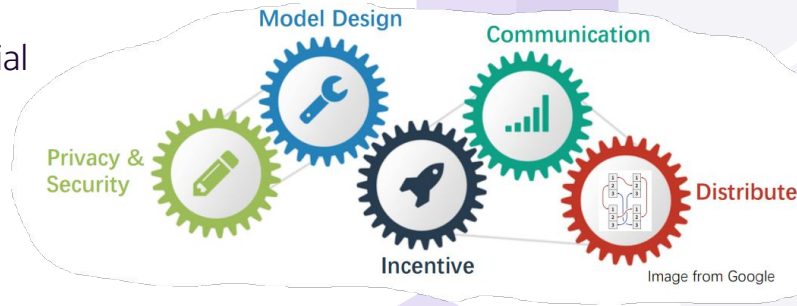
Federated Learning

- **Definition:**
  - Multiple parties, each of which owns some data, collaborate to jointly train a machine learning model.
  - During training, no data held by each party will leave that party
  - Only the trained (results) are transferred not the *data
  - The performance of the resulting model should be a good approximation of the ideal model, built with all data transferred to a single party



Processes involved in FL

- **Model Design and hyperparameter tuning.** e.g CNN architecture.
- **Distributed learning algorithm,** e.g. Client selection, tackling non-IID.
- **Communication optimization,** e.g. alleviating the influence of network delay, model/gradient compression.
- **Security and privacy,** e.g. Homomorphic Encryption (HE), Differential Privacy (DP).
- **Incentive mechanism,** e.g. motivating organizations from different industries.

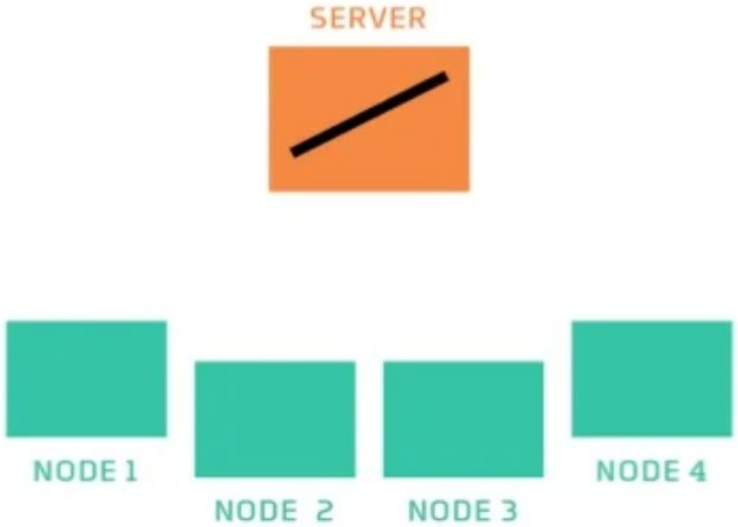


3

How does Federated Learning Work?

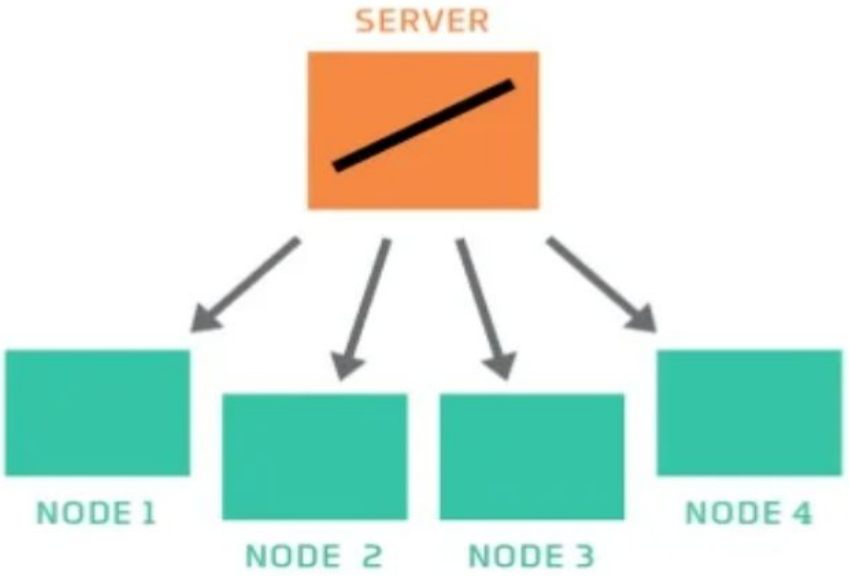
How Federated Learning works?

- **The server has untrained model.**



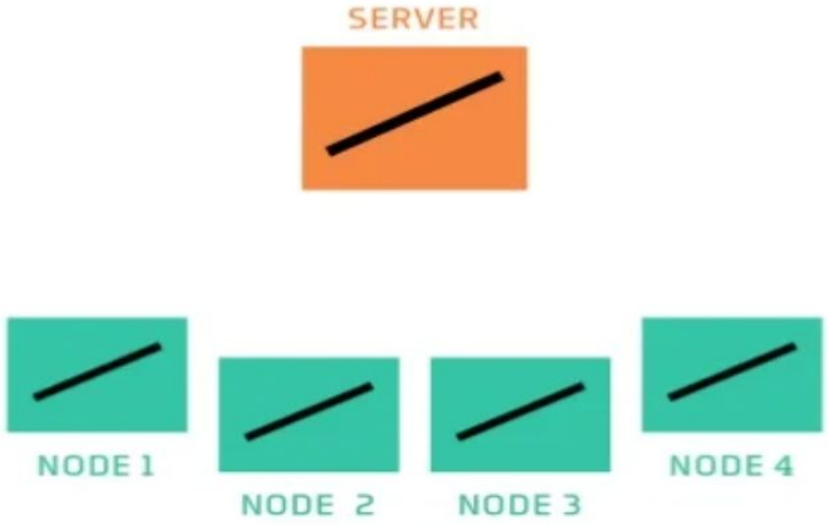
How Federated Learning works?

- **The server sends a copy of model to Nodes.**



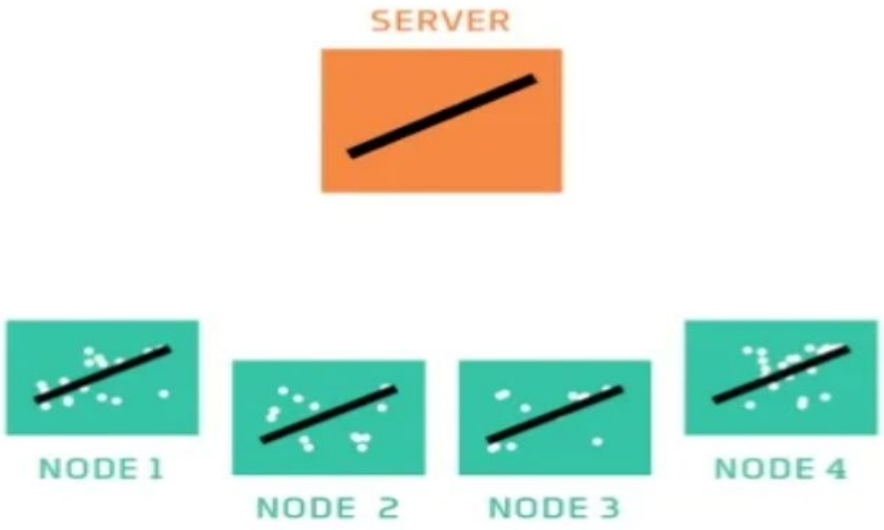
How Federated Learning works?

- **The Nodes now also have untrained model.**



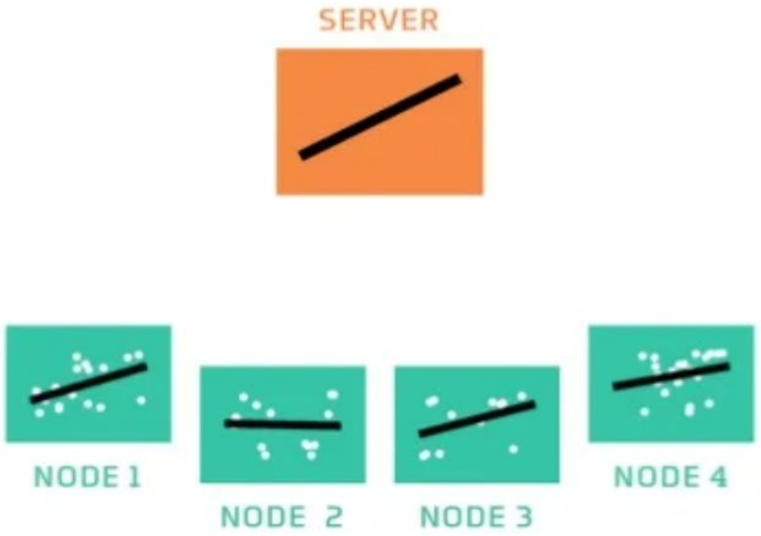
How Federated Learning works?

- **The Nodes have data to train their model.**



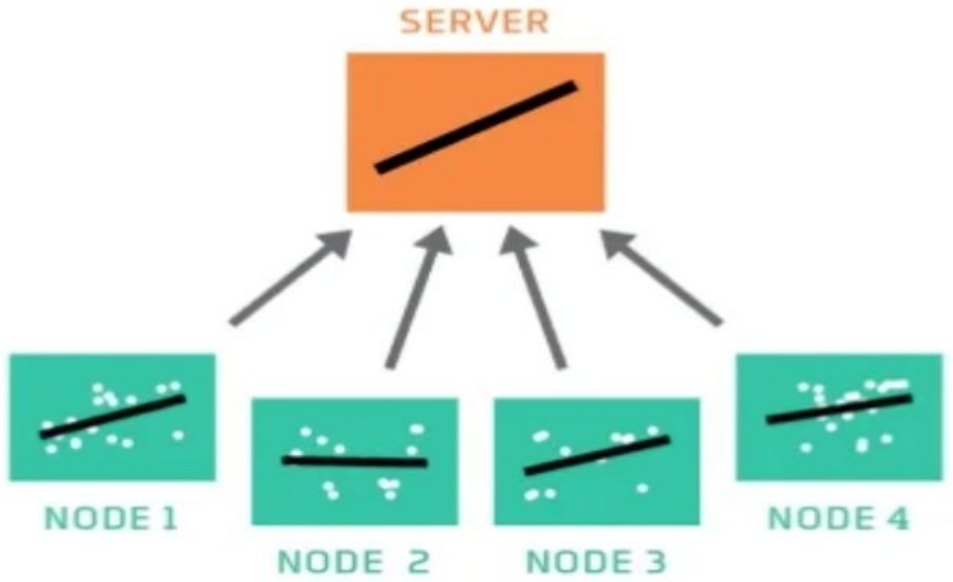
How Federated Learning works?

- **The Node trains and fits data .**



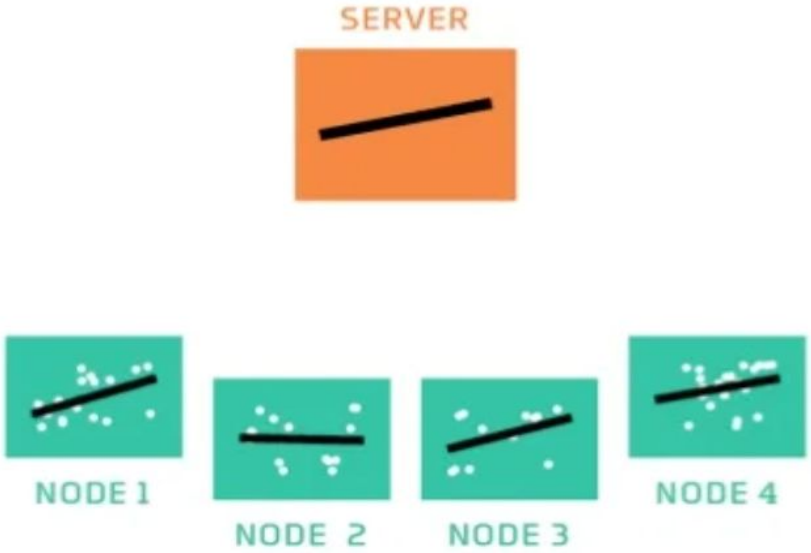
How Federated Learning works?

- **Each Node sends results back to the server.**



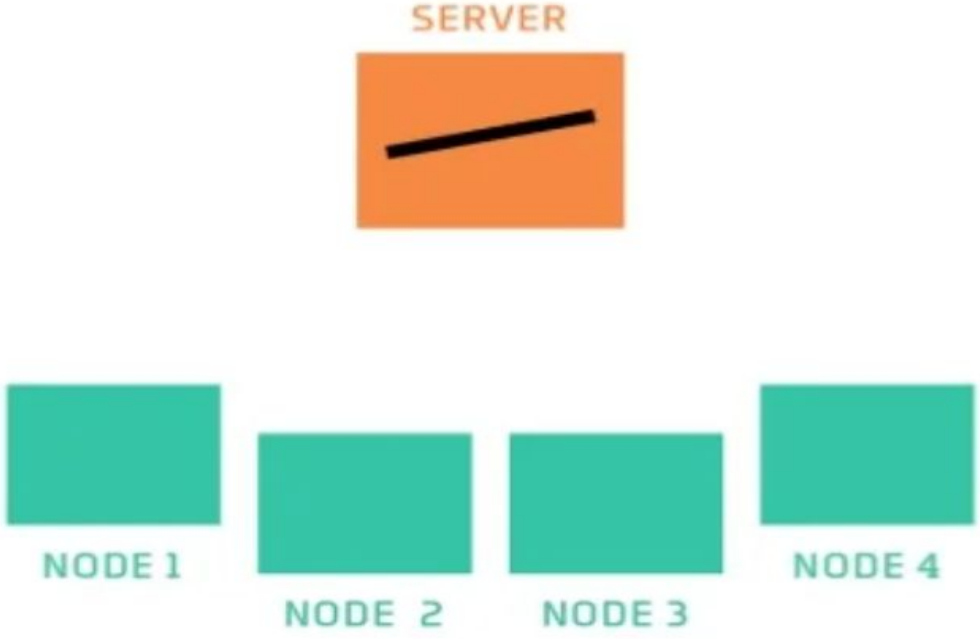
How Federated Learning works?

- **The server combines the models by taking an average.**



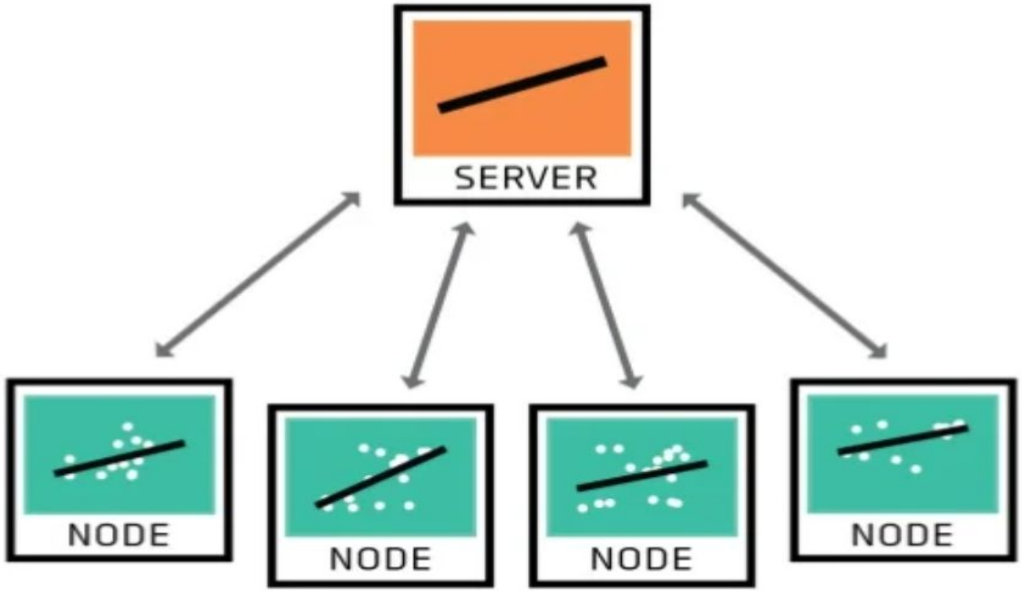
How Federated Learning works?

- The server now has a model that can recognize patterns and after each communication round the connection is invoked..



How Federated Learning works?

- **A network of Nodes share the training results rather than actual data.. This is How Federated Learning preserves Privacy**



4

Federated Learning Algorithm(FedAvg)

#participants

#samples of participant k

central model parameter

$$w_{t+1} \leftarrow \sum_{k=1}^{K} \frac{n_k}{n} w_{t+1}^{k}$$

local model parameter of participant k

#samples of all participants

Federated avg learning Algorithm

Deep learning model training:

Traditional

- For a training dataset containing ***n*** samples **(x_i, y_i), 1 ≤ *i* ≤ *n*,** the training objective is:

$$\min_{w \in \Re^d} f(w) \quad \text{where,} \quad f(w) = \frac{1}{n} \sum_{i=1}^{n} f_i(w)$$

$f_i(w) = l(x_i, y_i, w)$ is the loss of the prediction on example *(x_i,y_i).*

- Deep learning optimization relies on *SGD* and its variants,

$$w_{t+1} \leftarrow w_t - \eta \nabla f(w_t; x_k, y_k)$$

Federated

- Suppose *n* training samples are distributed to *K* clients, where P_K is the set of indices of data points on client *k*, and $n_k = |P_k|$.

For training objective: $\min_{w \in \Re^d} f(w)$

$$f(w) = \sum_{k=1}^{K} \frac{n_k}{n} F_k(w)$$

where,

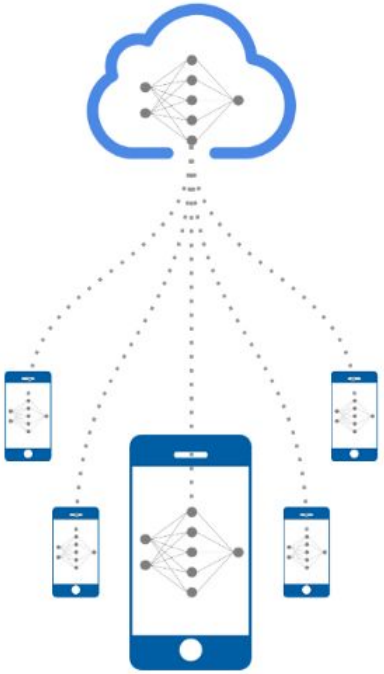
$$F_k(w) = \frac{1}{n_k} \sum_{i \in P_k} f_i(w)$$

5

So, what are its advantages?

Advantage of Federated Learning

- Smarter models
- Less power consumption
- Ensuring privacy



Hyper-Personalized

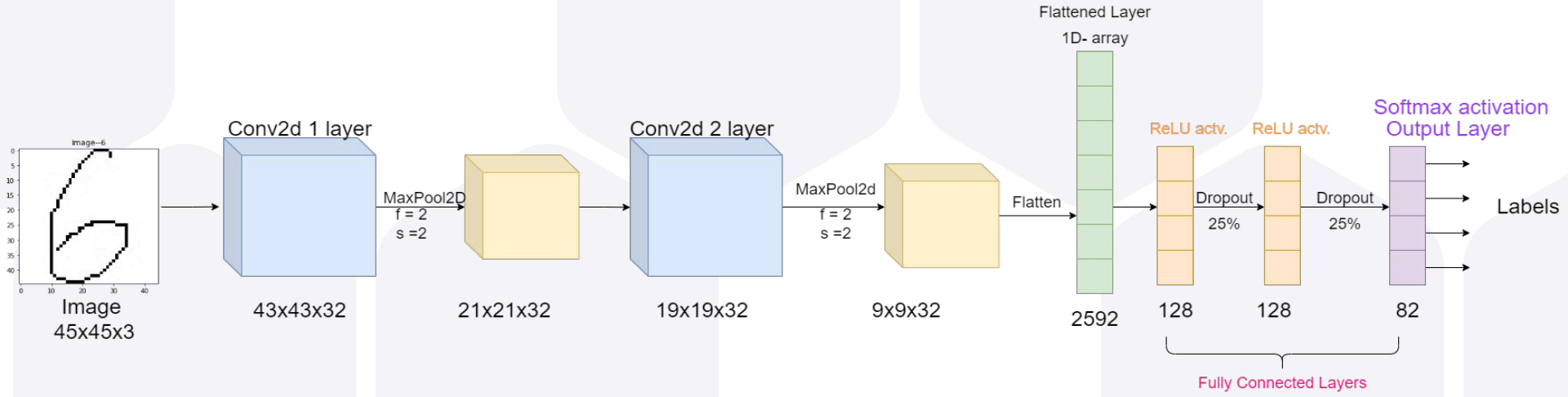
Low Cloud Infra Overheads

Minimum Latencies

Privacy Preserving

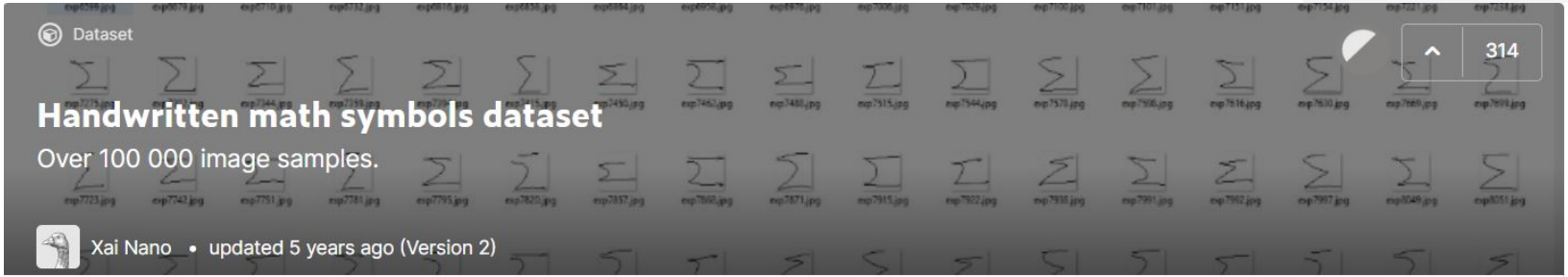
6

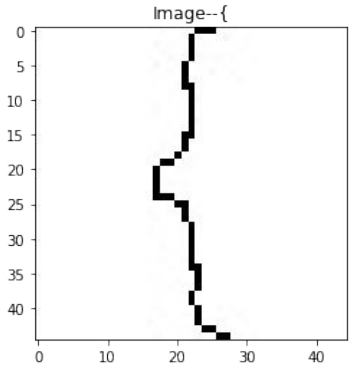
Our Implementation: (Handwritten math equation recognizer)

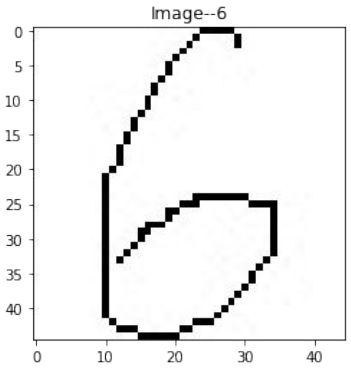


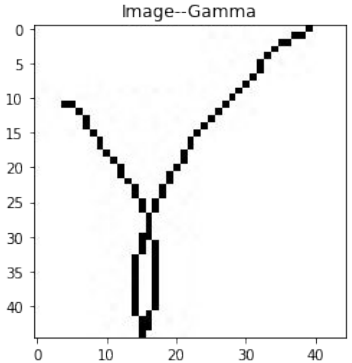
Server and Local CNN Architecture

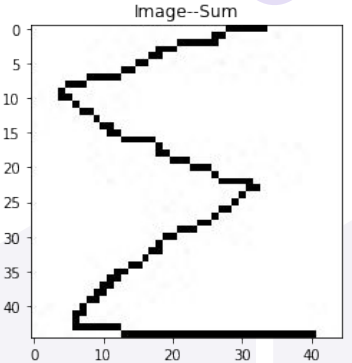
Image Dataset Used

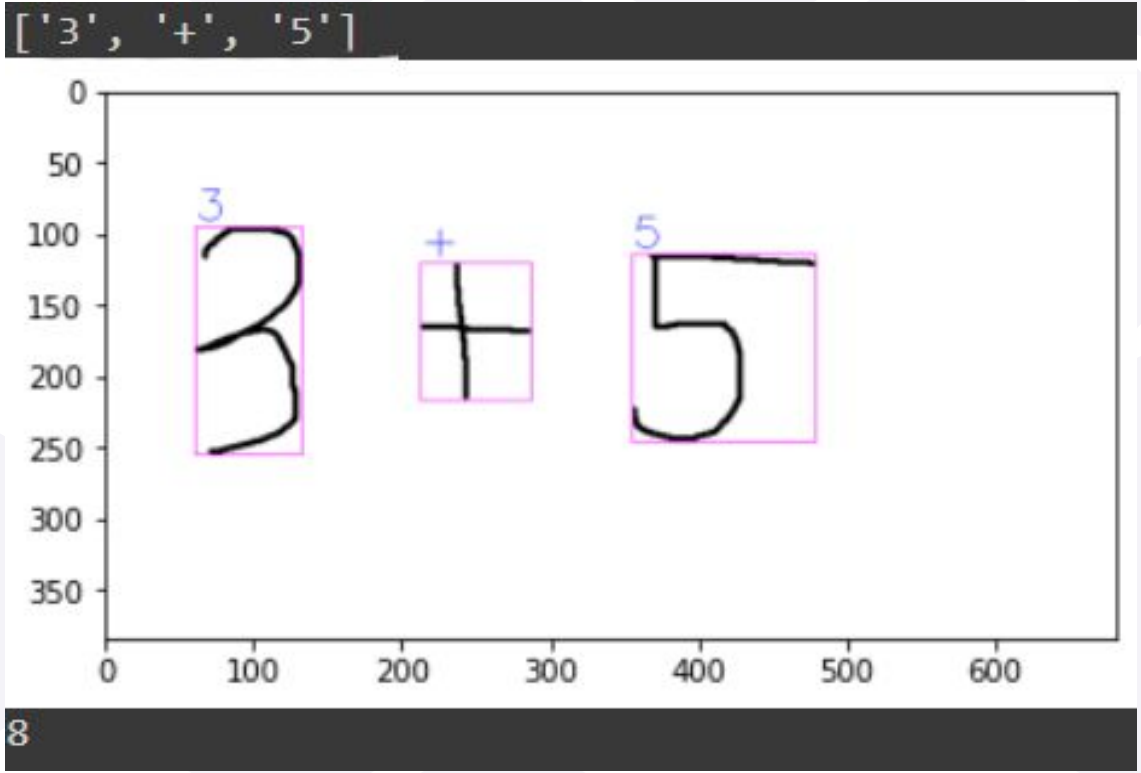




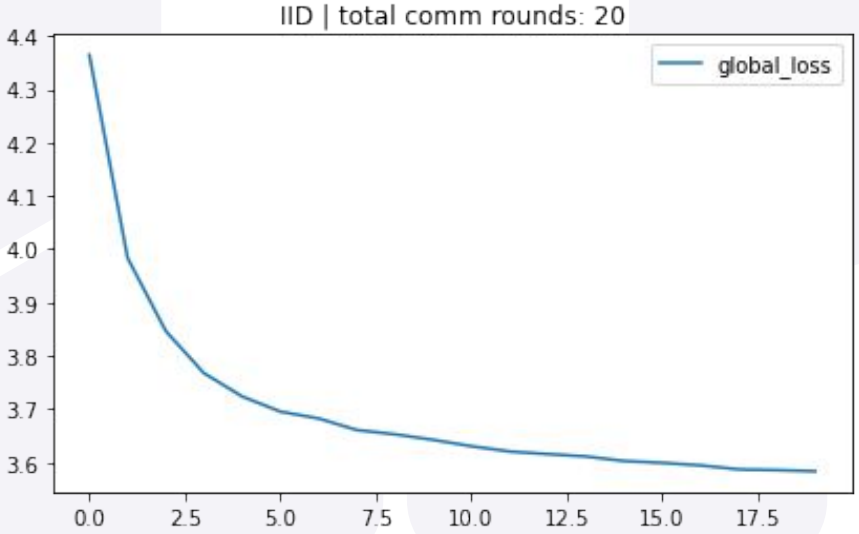


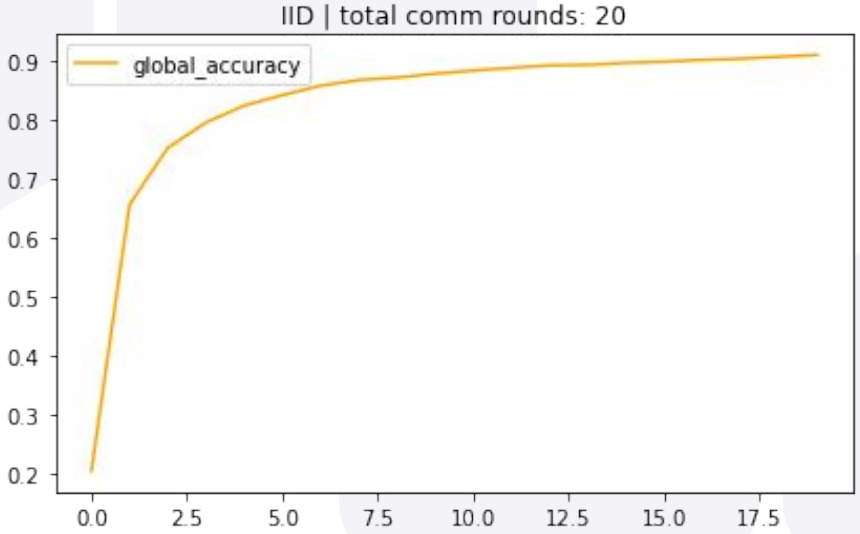






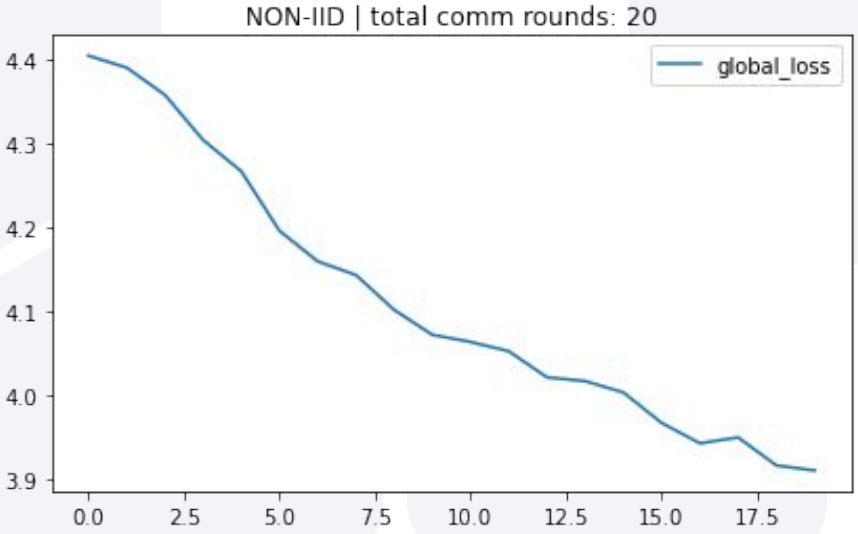
Predicted Output

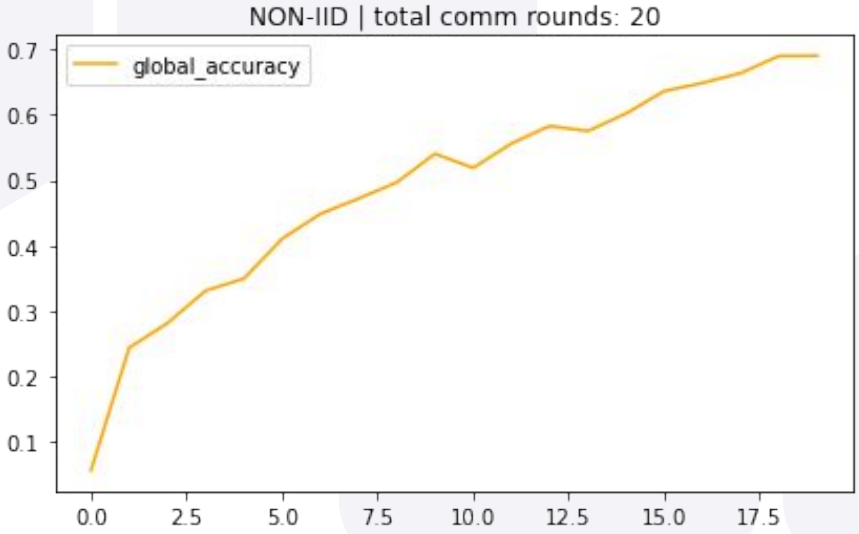




Global loss & accuracy (IID) , Accuracy peaks to 90%

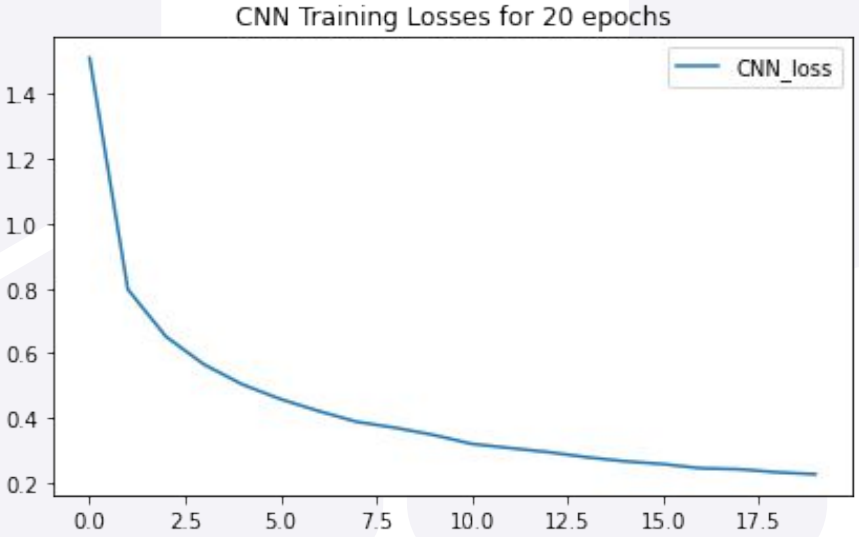
IID*(Identical and Independent Distribution)*

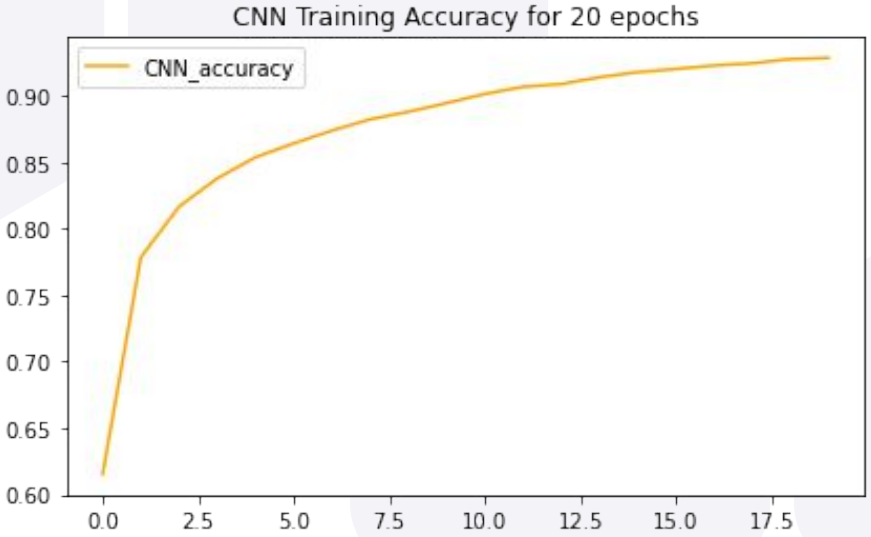




Global loss & accuracy (Non-IID) , Accuracy is about 70%

Non IID*(Non-Identical and Independent Distribution)*





While Training only on CNN for 20 epochs ,accuracy peaked to 92.85%

Thank You