Privacy Preservation

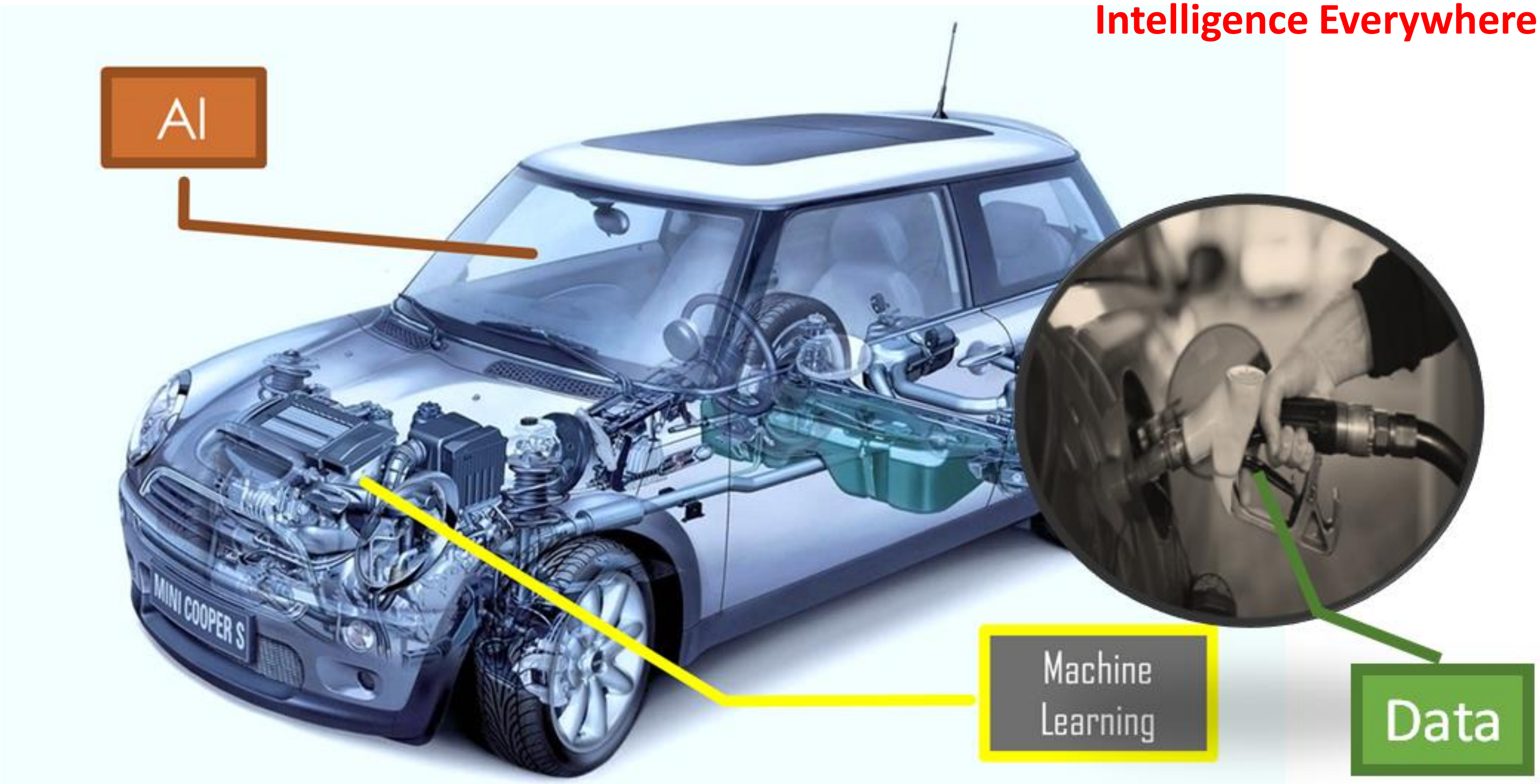
Yu Han

han.yu@ntu.edu.sg

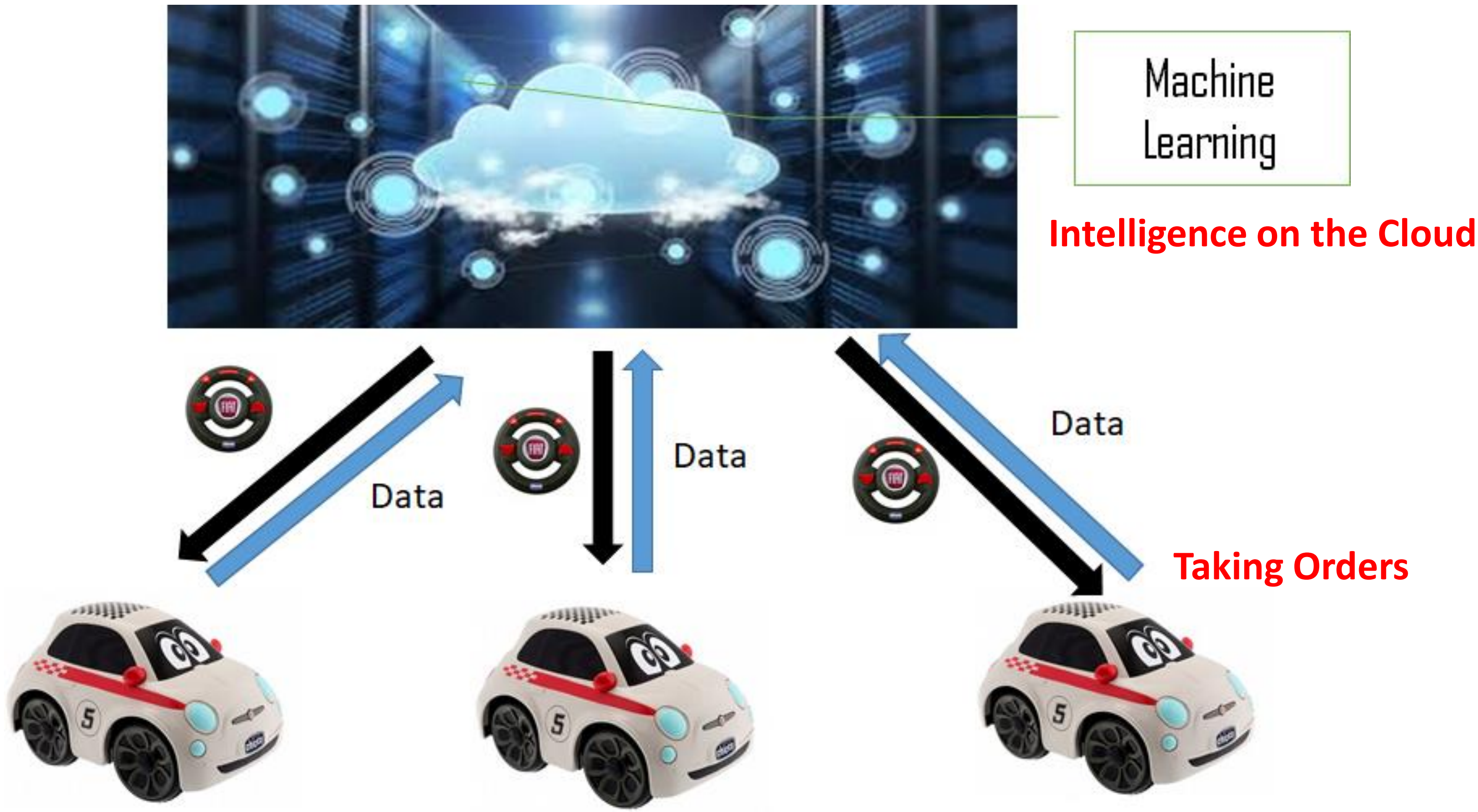
Nanyang Assistant Professor
School of Computer Science and Engineering
Nanyang Technological University

Data, ML & AI (<u>Ideally</u>)

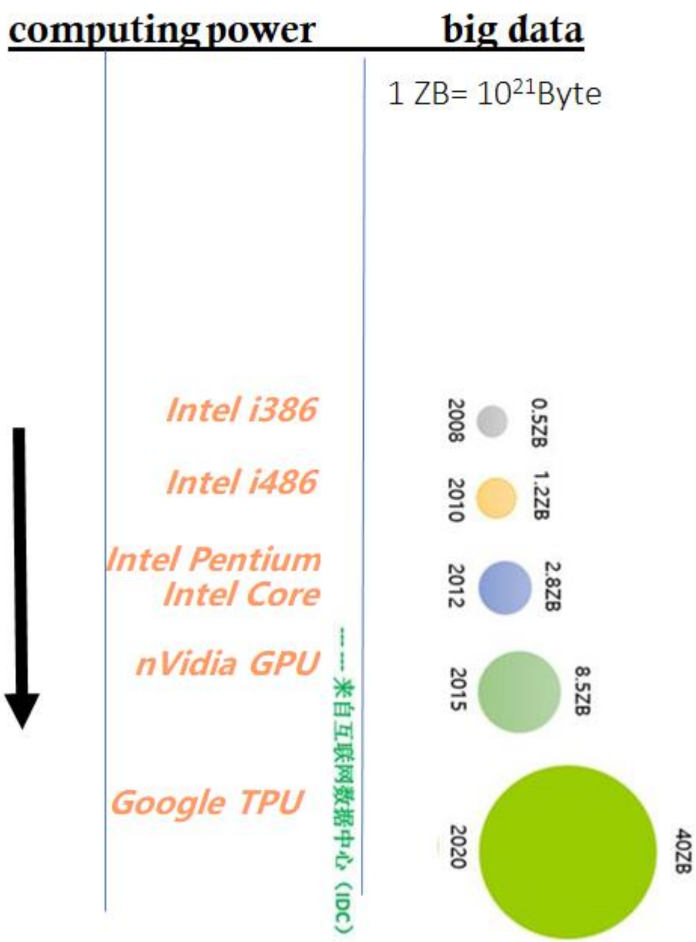
Intelligence Everywhere?



Data, ML & AI (Reality)



Data is the “New Oil”



The New Rich

Challenge: Data Privacy Protection

French regulator fines Google $57 million for GDPR violations

Share on Facebook Share on Twitter +

Google hasn't transparently implemented GDPR rules, French regulator claims.

- More than 50 million people involved
- UK fined Facebook for £500,000
- **The worst single-day market value drop for a publicly listed company in the US**, dropping $120 billion, or 19%

GDPR


GDPR
No Autonomous Modeling and Decision
Interpretability of Model Decisions
Users'Right for Data to be Forgotten
Data Privacy By Design
Explicit Consent for Data Usage

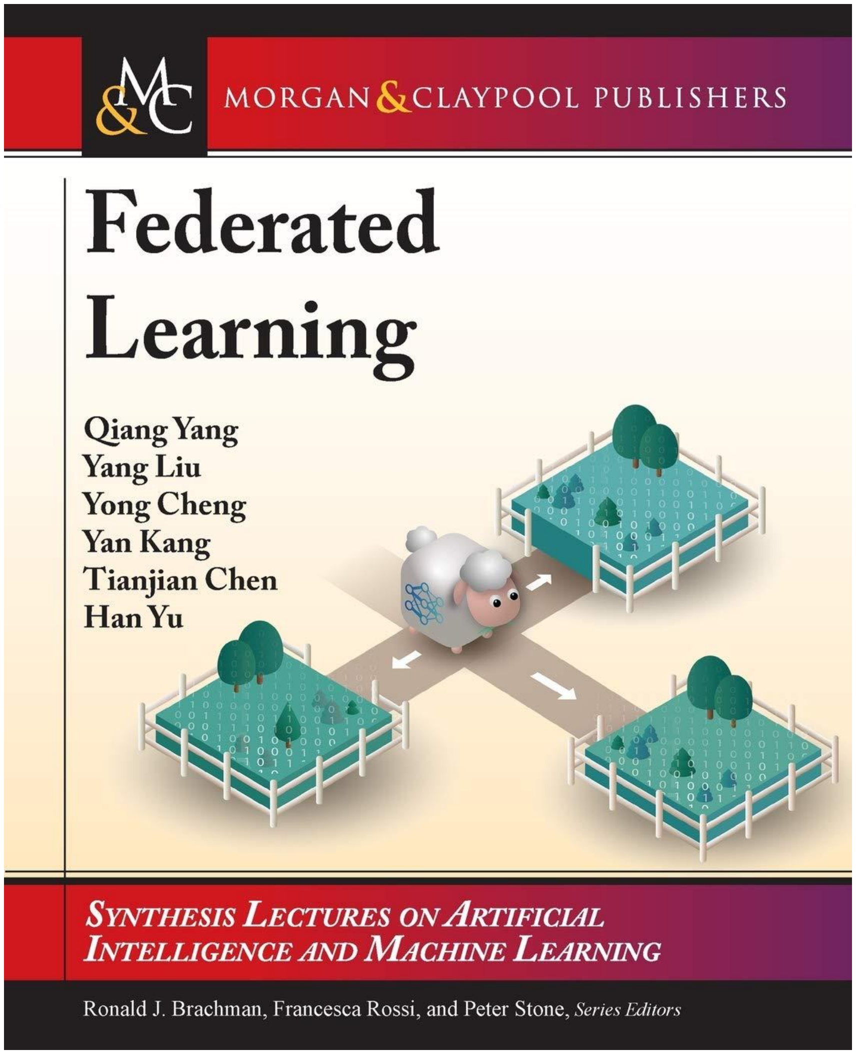
Why Federated Learning?

- Traditional machine learning methods need all data to be gathered in a central entity
- In many real-world applications data are isolated across different organizations and data privacy is being emphasized
- Federated learning (FL) is well suited for these scenarios due to its distributed and privacy-preserving nature

What is Federated Learning?

- A new approach for models trained from user interaction with distributed devices.
  - **distributes** the machine learning process over to the edge.
  - enables devices to **collaboratively learn** a shared model using the training data on the device and **keeping the data on device**
  - decouples the **need for doing machine learning** with the **need to store the data** in the cloud

Text Book

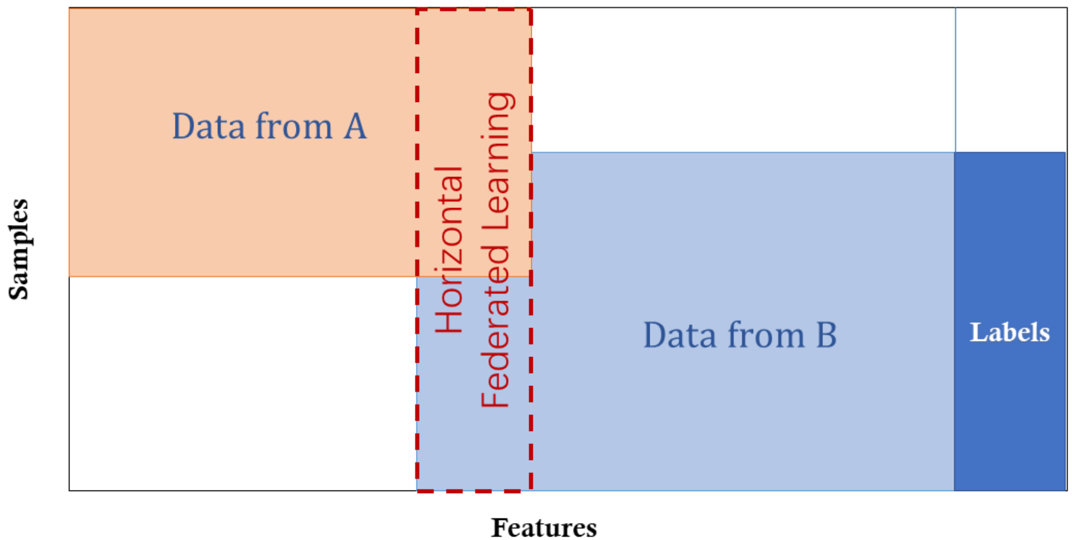


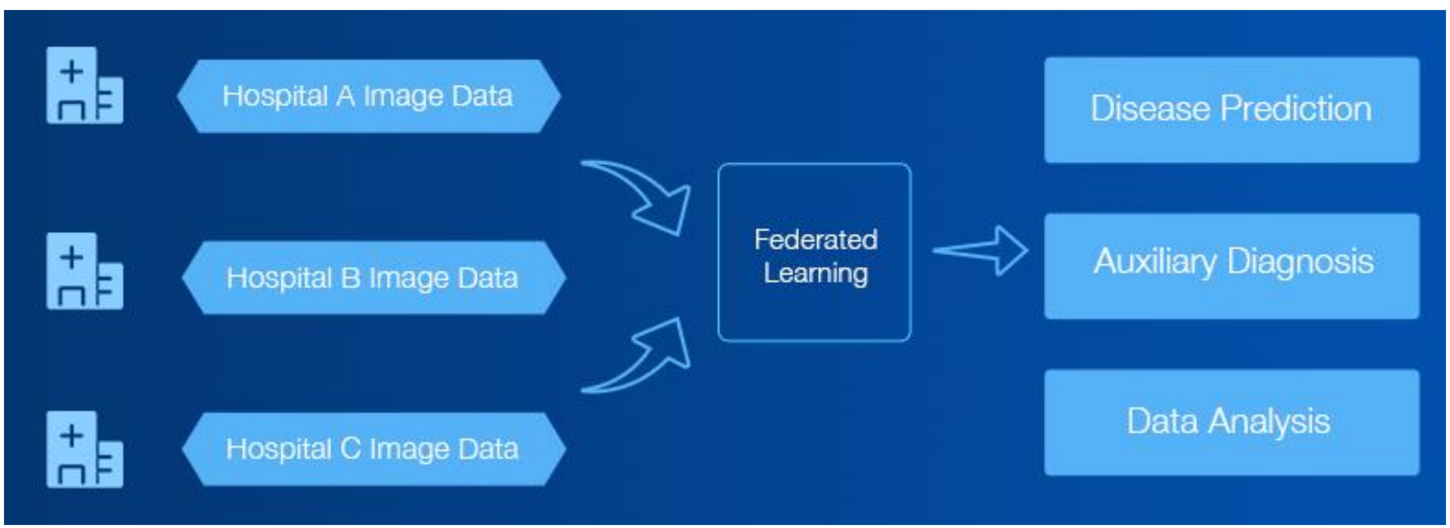
E-Book can be found from NTU Online Library:

https://ntu-sp.primo.exlibrisgroup.com/discovery/search?vid=65NTU_INST:65NTU_INST&lang=en

Additional Resources can be found at: http://federated-learning.org/

Horizontal Federated Learning (HFL)



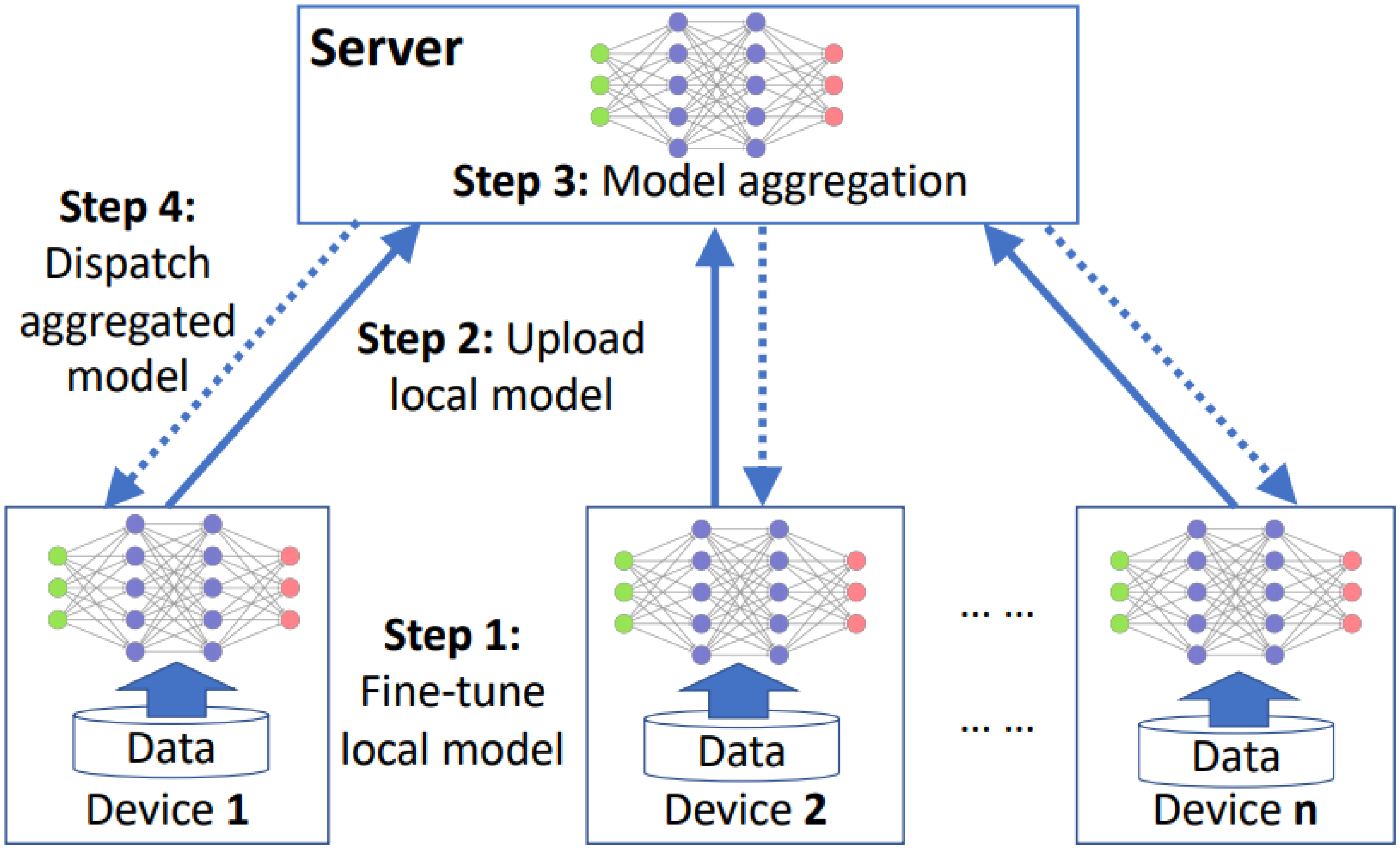


Horizontal Federated Learning (HFL)

- HFL assumes that datasets from different participants share the same feature space, but may not share the same sample ID space
- Existing FL approaches mostly focus on HFL

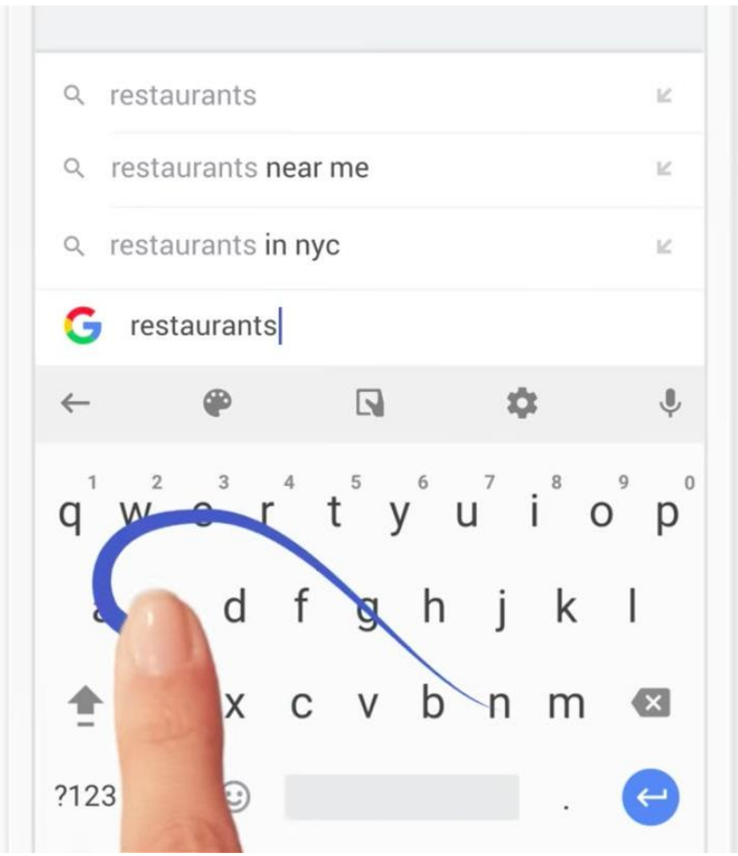
Yang, Q., Liu, Y., Cheng, Y., Kang, Y., Chen, T. & Yu, H. (2019) *Federated Learning*. Morgan & Claypool Publishers, San Rafael, CA, USA, p. 207.

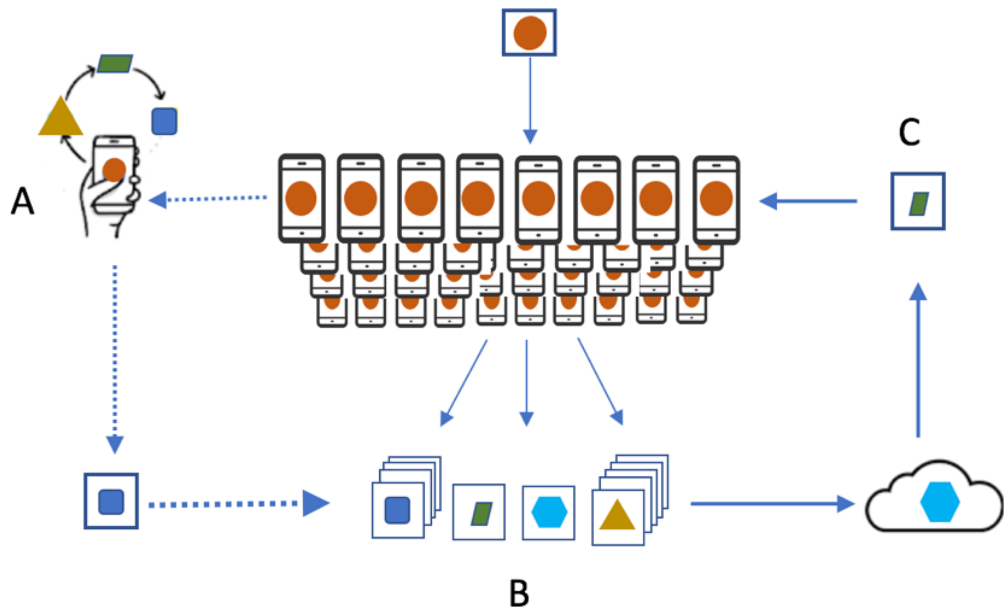
HFL Key Steps



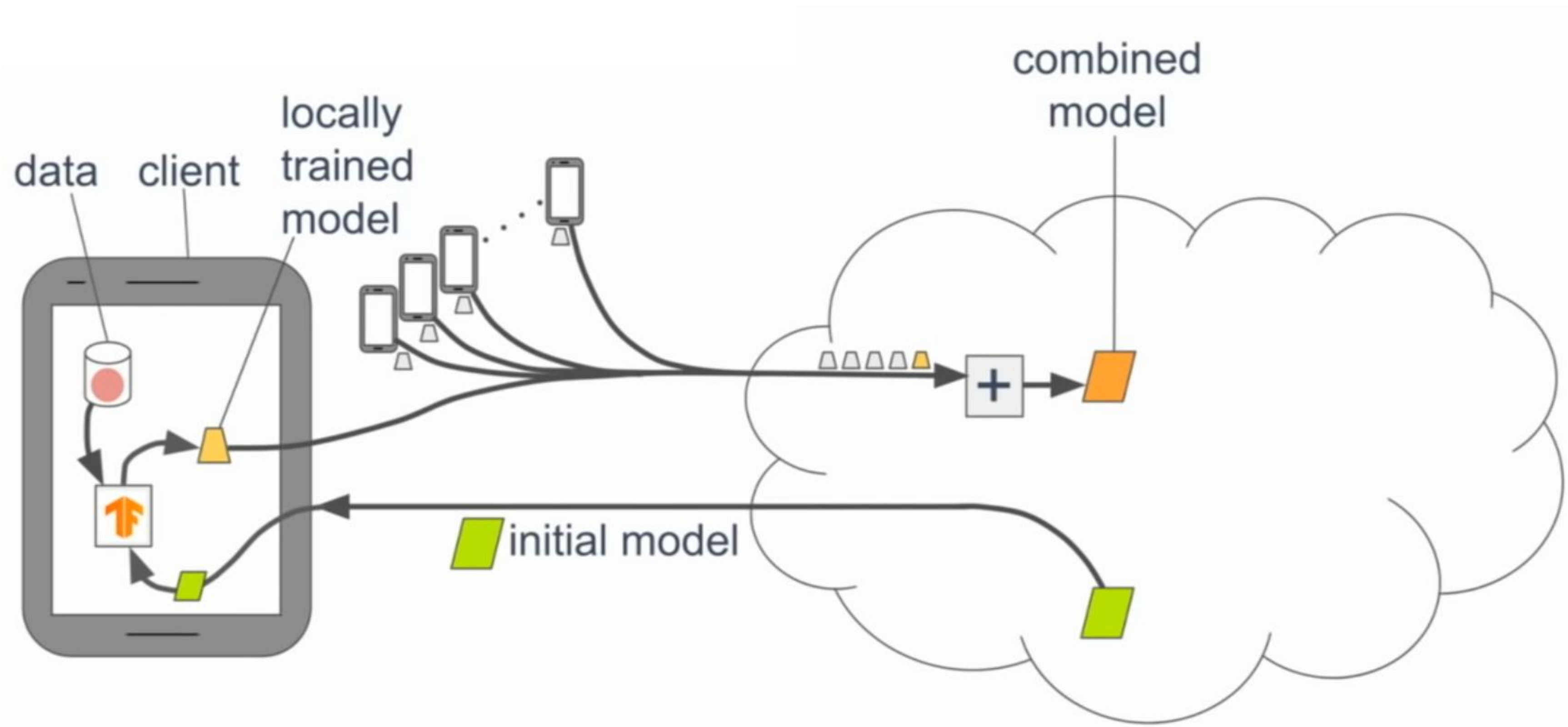
Source: https://arxiv.org/abs/2005.01026

Federated Learning (Google)

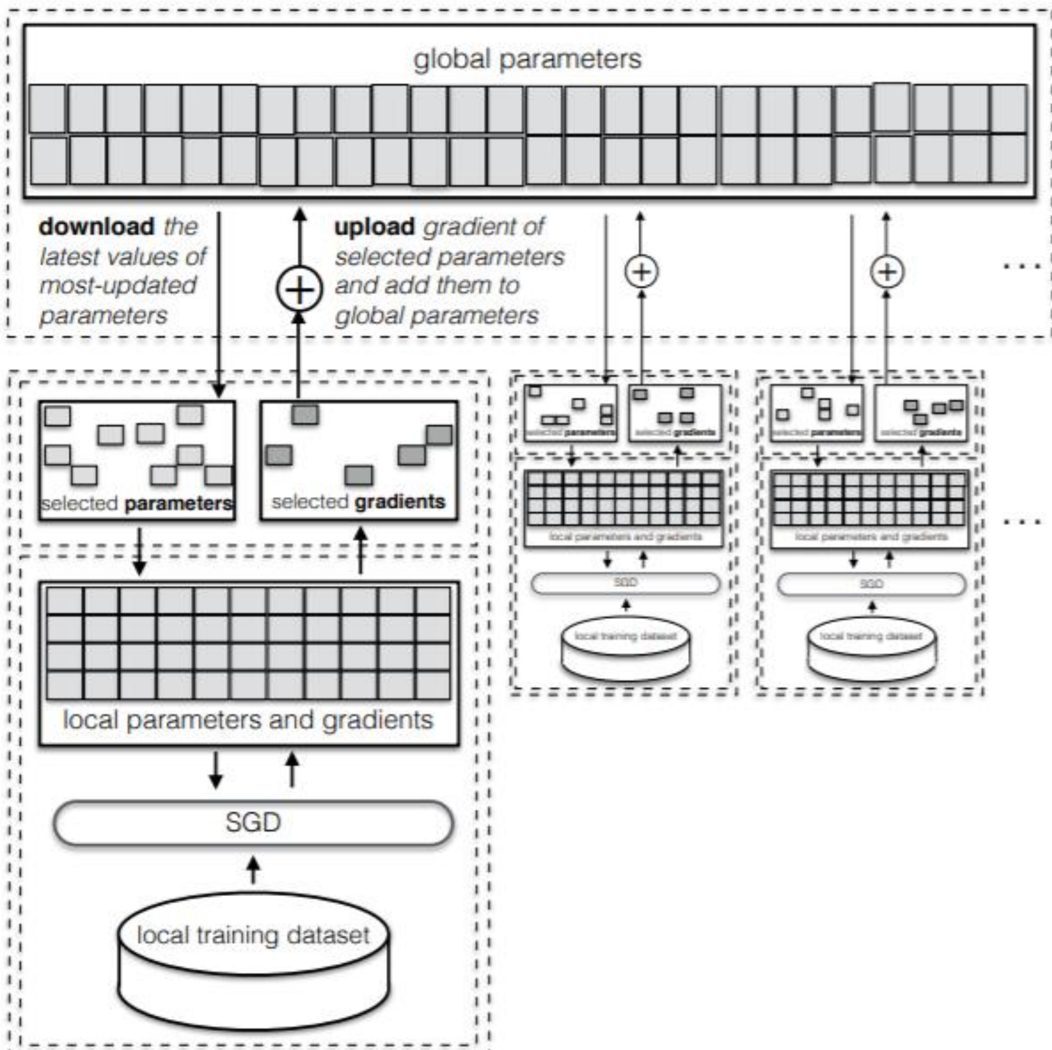




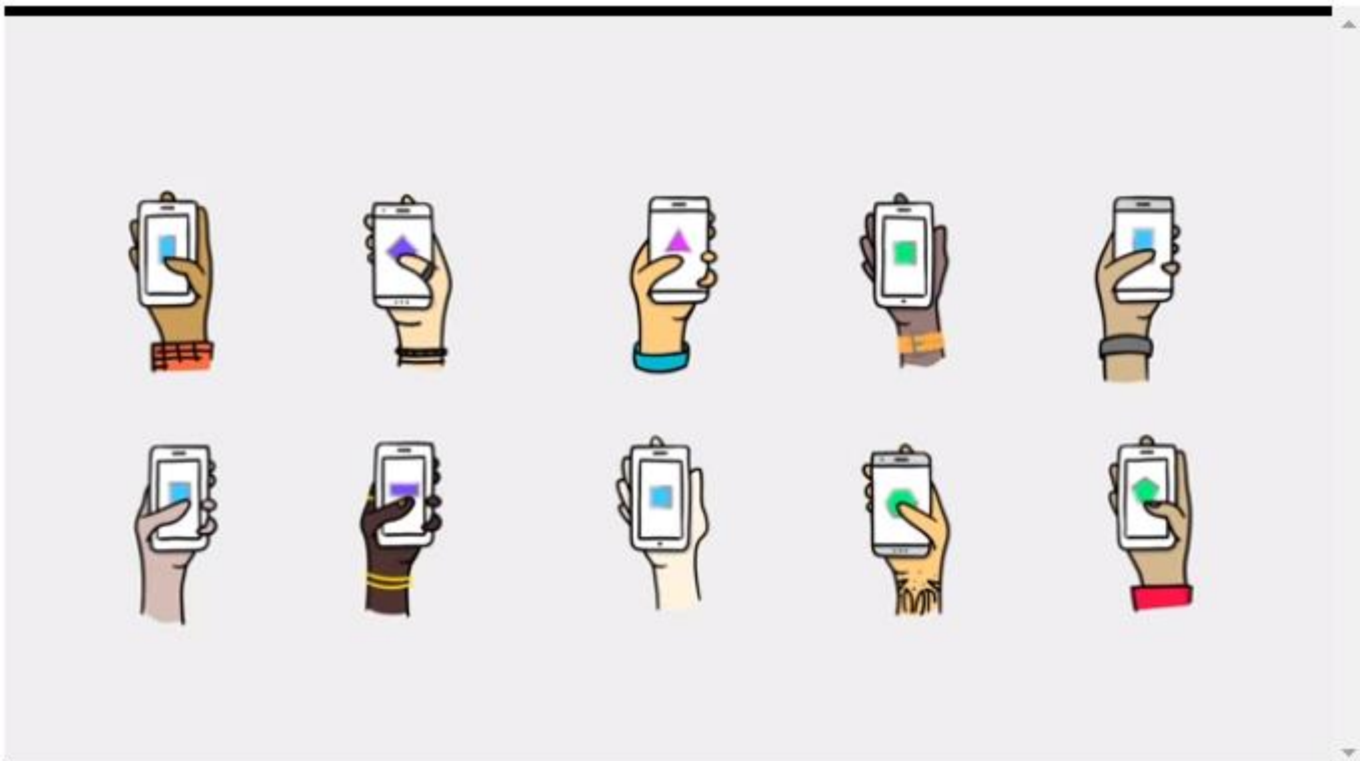
Federated Learning (Google)



Federated Learning (Google)



Federated Learning (Google)



Video Demo: https://youtu.be/gbRJPa9d-VU

How to Send Gradients to Server?

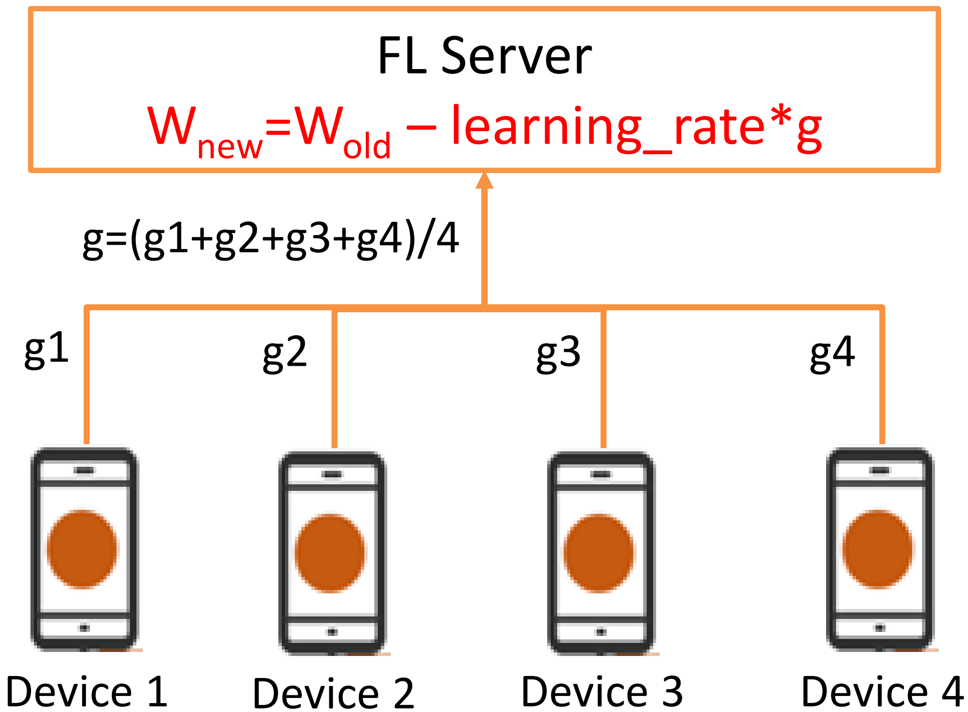
- Federated Stochastic Gradient Descent (FedSGD)

- Federated Averaging (FedAvg)

FedSGD

- Devices send gradients/parameters to server
- Server averages these gradients/parameters to obtain a new model
- Server sends the new model back to devices
- High communication overhead

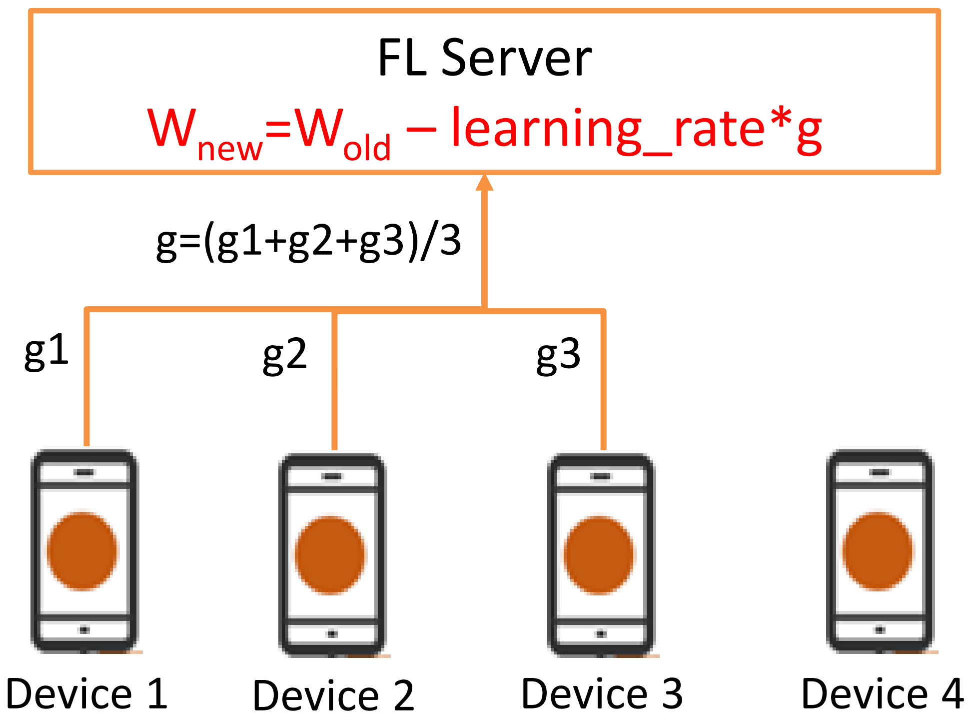
FedSGD, C=1



Version 1:

- Sending gradients
- The gradient descent operation happens on the FL server
- We set **C=1**, meaning 100% of the devices participate in FedSGD

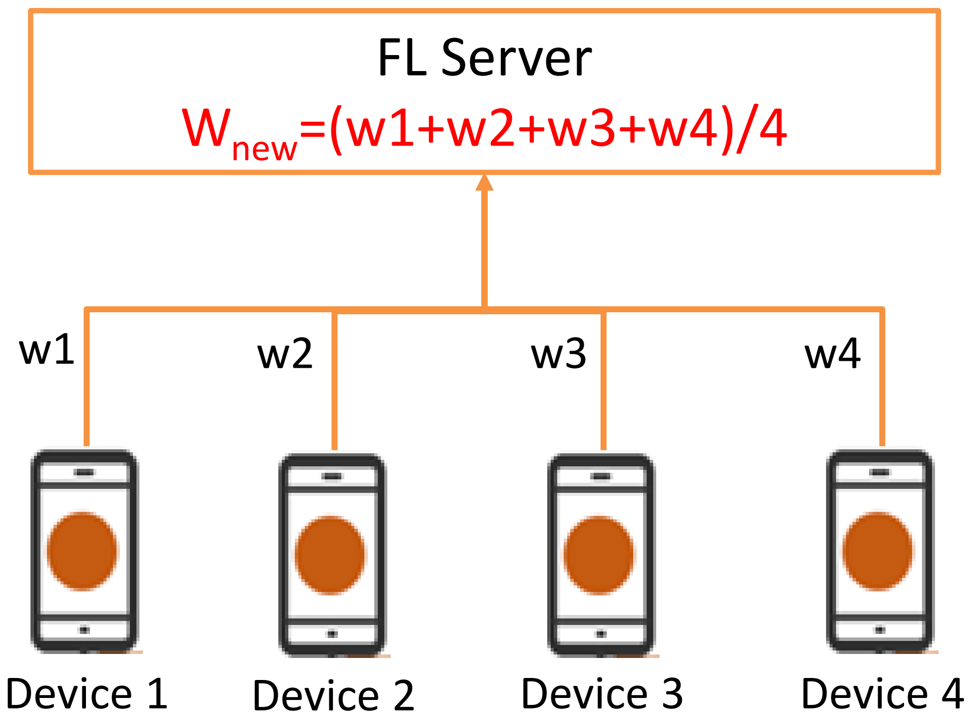
FedSGD, C=0.75



Version 1:

- Sending gradients
- The gradient descent operation happens on the FL server
- We set **C=0.75**, meaning 75% of the devices participate in FedSGD

FedSGD, C=1



Version 2:

- Sending parameters (i.e. weights)
- The gradient descent operation happens on the devices
- We set **C=1**, meaning 100% of the devices participate in FedSGD

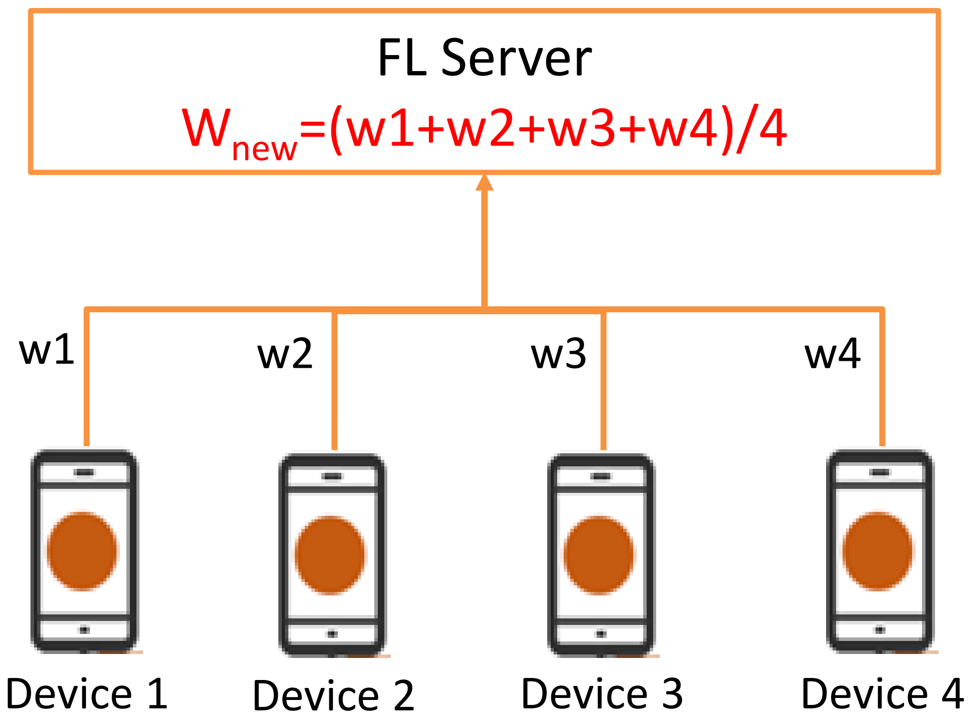
FedAvg

- Devices perform mini-batch training locally, and update their local parameters using gradient descent

- Devices send parameters to server
- Server averages these parameters to obtain a new model
- Server sends the new model back to devices

- Less communication than FedSGD

FedAvg, C=1, E=1, B=∞



- We set **C=1**, meaning 100% of the devices participate in FedAvg
- **E=1**, meaning the local SGD epoch=1
- **B=∞**, meaning all local data are used for training. Setting it to a smaller means we have mini-batch training locally.

Under this setting, FedAvg = FedSGD

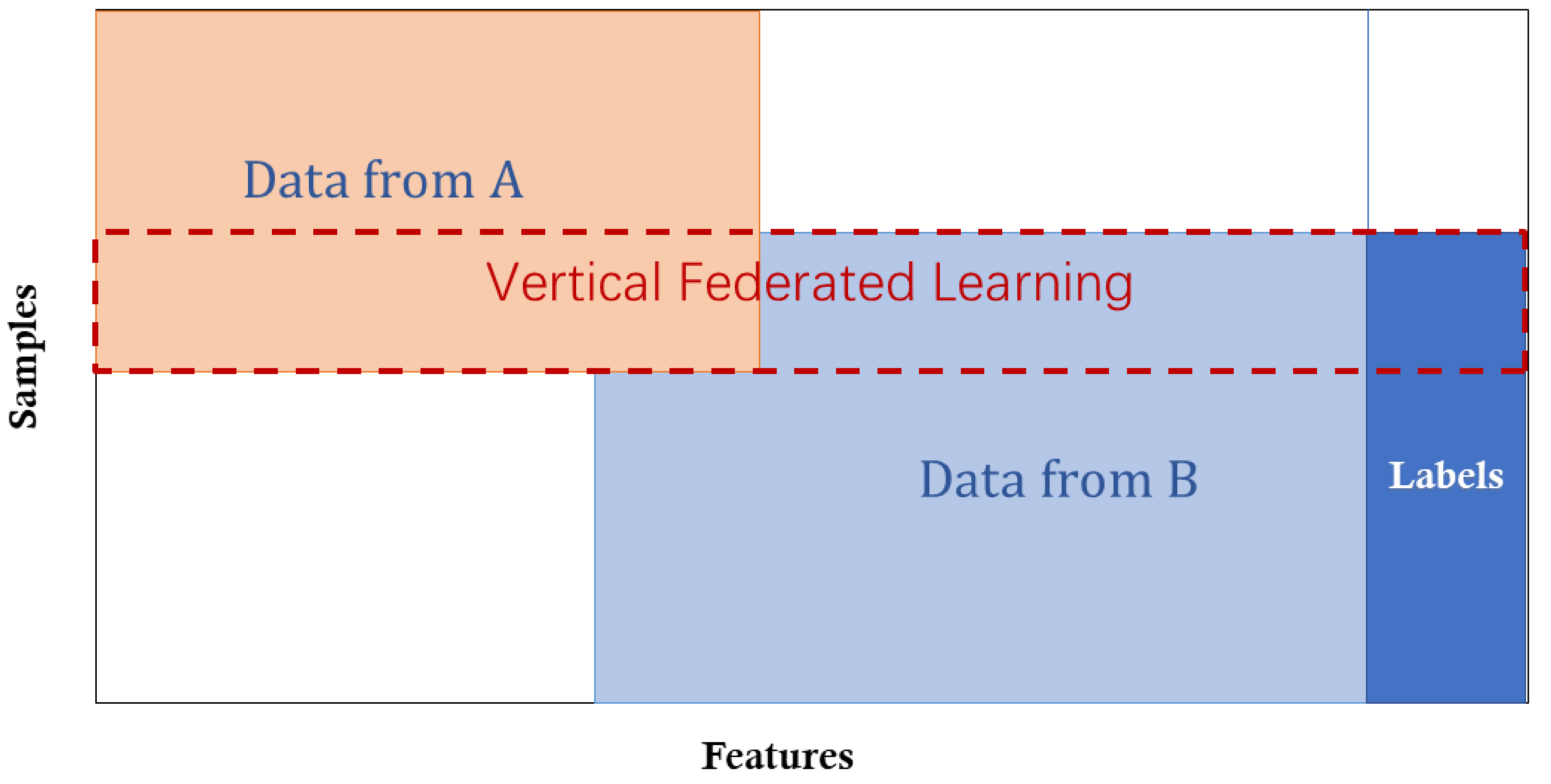
FedAvg

- You can increase **E** and reduce **B** to make more use of local device computing power to train the model and reduce communication overhead.

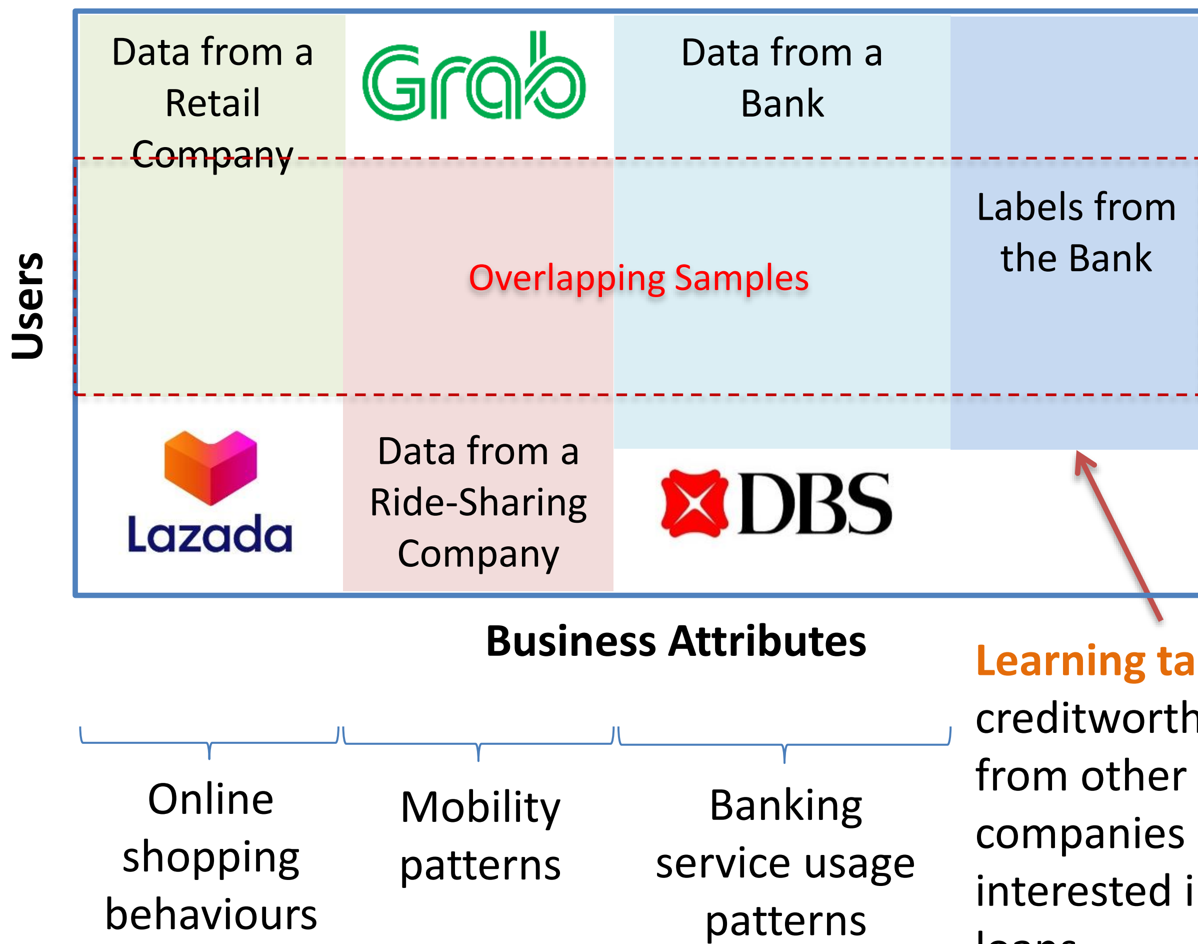
- FedAvg provides you with more flexibility to adjust local computing power utilization and communication overhead during FL model training compared to FedSGD.

H. Brendan McMahan, Eider Moore, Daniel Ramage, Seth Hampson, Blaise Agüera y Arcas. Communication-Efficient Learning of Deep Networks from Decentralized Data. *CoRR*, arXiv:1602.05629, 2016.

Vertical Federated Learning

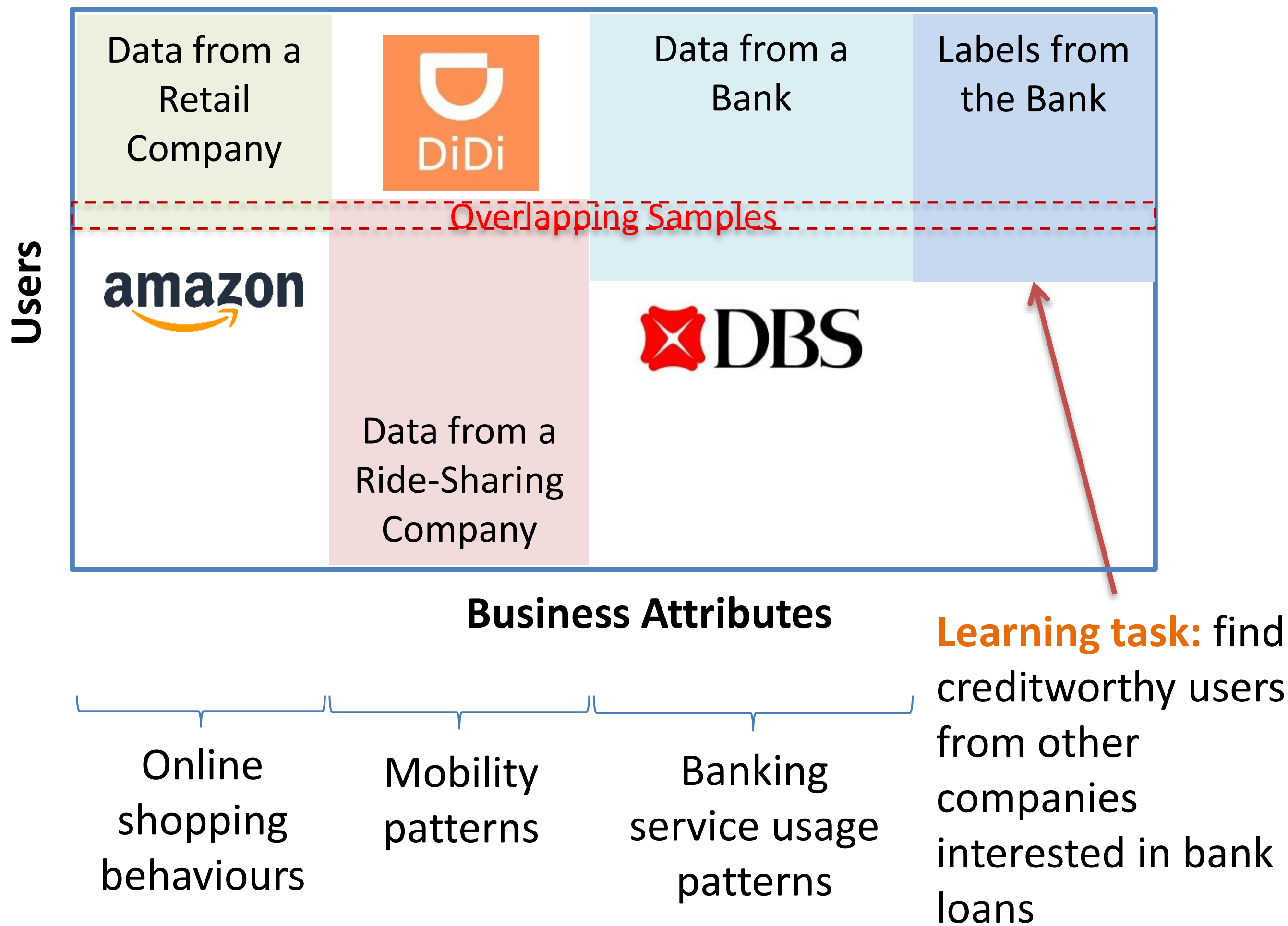


Vertical Federated Learning (VFL)

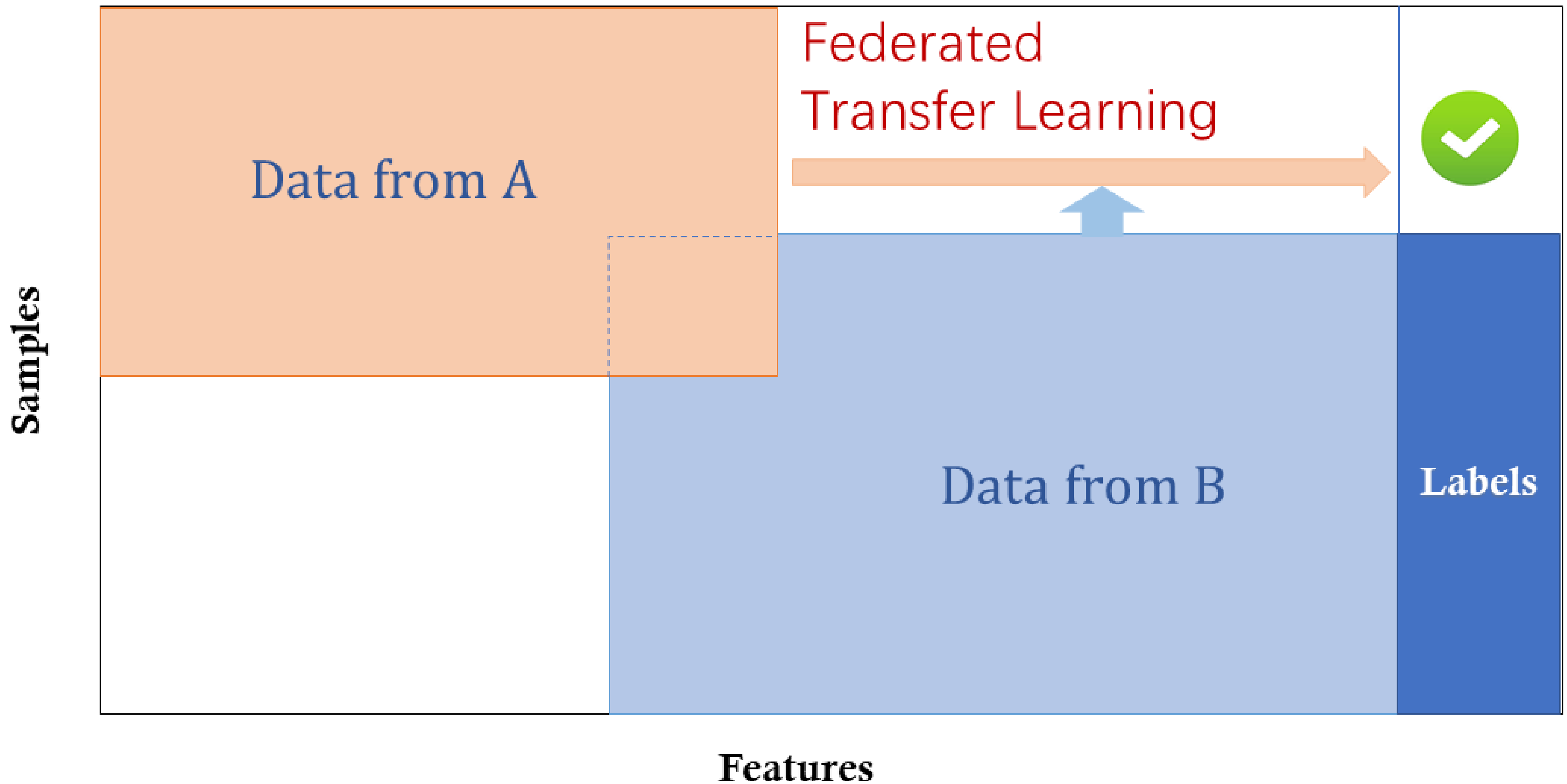


Learning task: find creditworthy users from other companies interested in bank loans

Federated Transfer Learning (FTL)



Federated Transfer Learning (FTL)

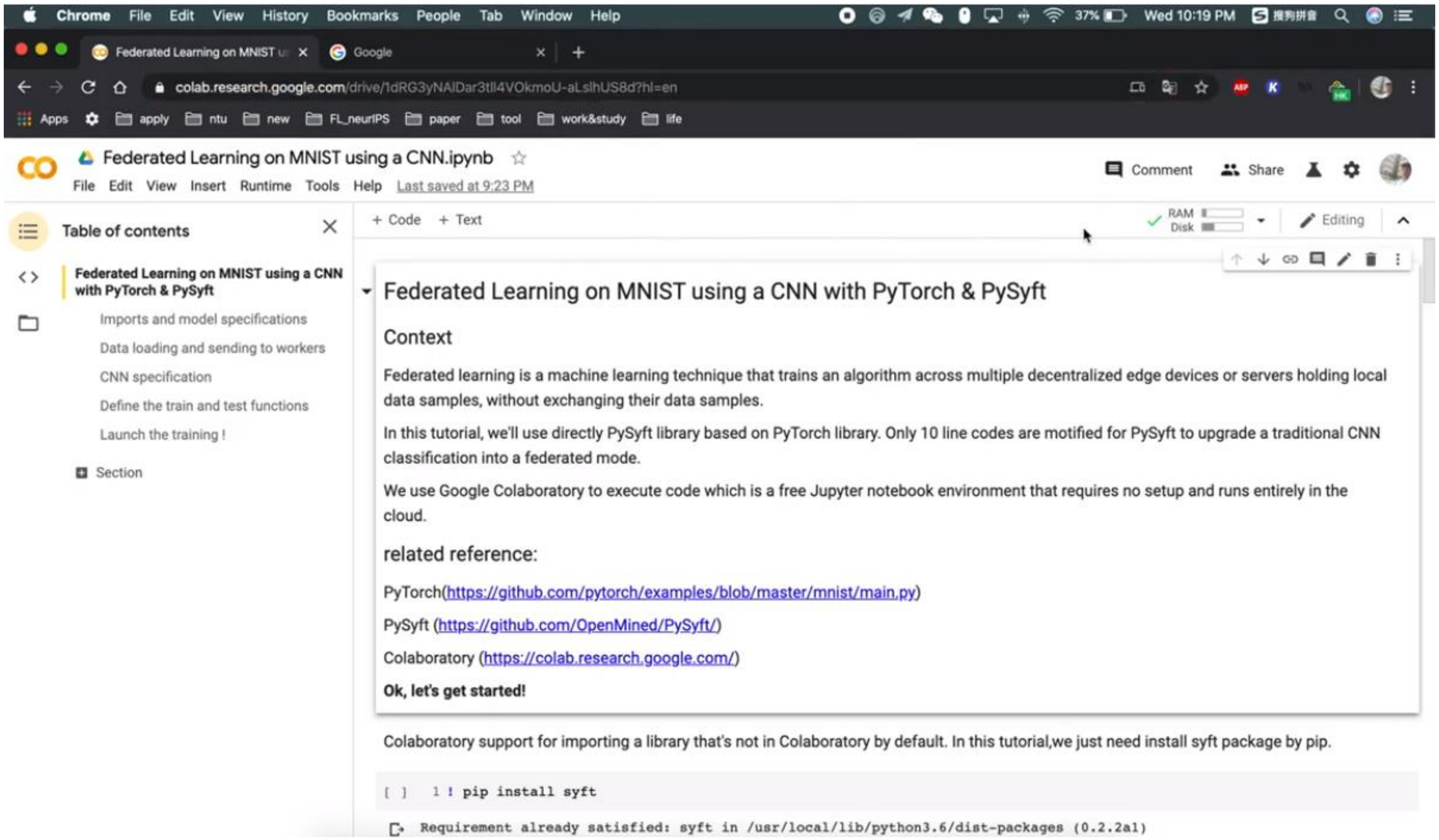


Video Explanation

https://www.youtube.com/watch?v=NPGf_OJrzOg&feature=youtu.be

Hands-on Practice

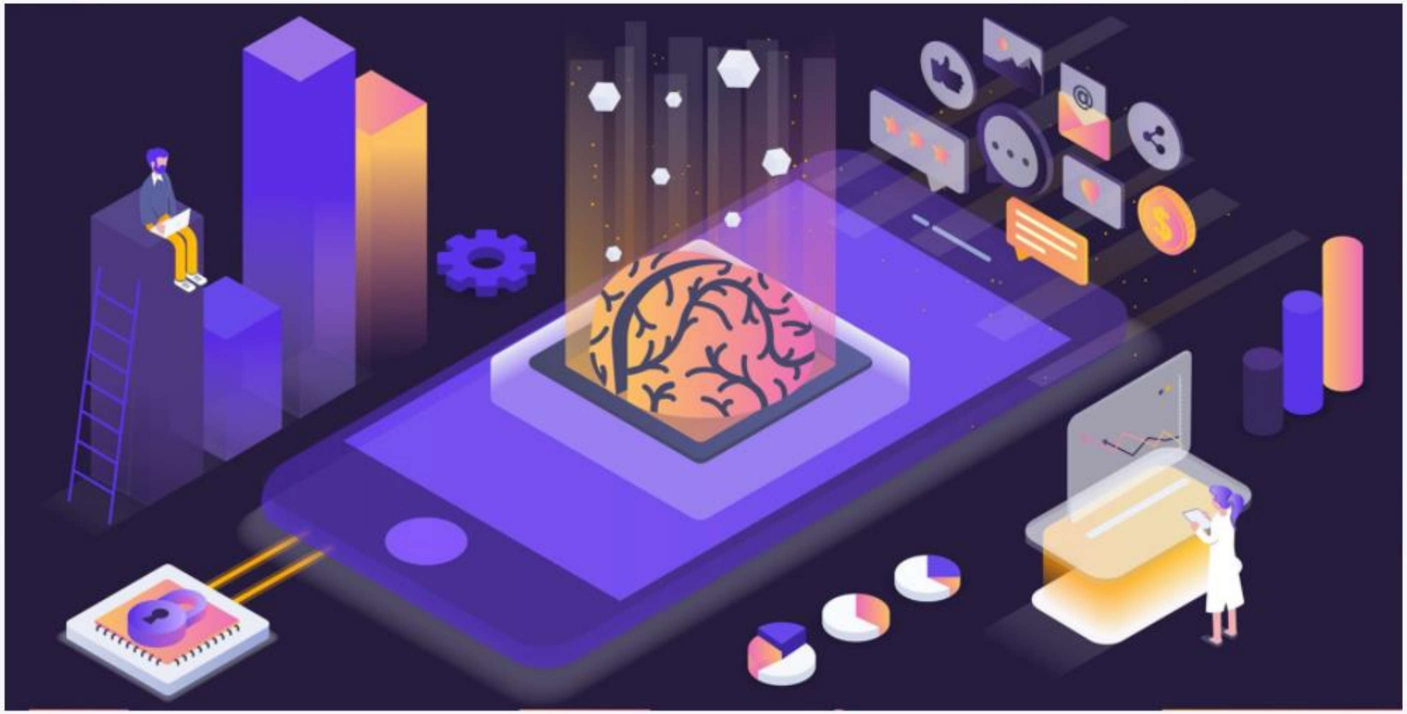
https://colab.research.google.com/drive/1dRG3yNAlDar3tll4VOkmoU-aLslhUS8d



Video Guide: https://www.youtube.com/watch?v=NPGf_OJrzOg&feature=youtu.be

Federated Learning Portal

http://federated-learning.org/



The Federated Learning Portal

In this webportal, we keep track of books, workshops, conference special tracks, journal special issues, standardization effort and other notable events related to the field of Federated Learning (FL).

Privacy Preservation

Yu Han

han.yu@ntu.edu.sg

Nanyang Assistant Professor
School of Computer Science and Engineering
Nanyang Technological University