Introduction:

Increasingly, personal computers and laptops’ performance has gained huge progress, even smart phones have powerful computing power, as they could deal with unprecedented amount of data now, running machine learning programs becomes possible. Also, mobile devices like smart phones and video cameras could collect data all over the world, and for the lack of communications and the constraints of privacy protection, collecting all data from mobile devices and processed with a centralized computing could be infeasible, which means the traditional machine learning strategies could not be effectively performed.

A work from McMahan et. had proposed a decentralized approaches named federated learning, starts a on-device training trend, both of academic and industry have done a lot of work on it. McMahan’s work advanced a algorithm called FedAvg, which lead to distributed training on local devices, devices collect amount of data and train them to get the local model, devices upload the model and the central server fuse the model to get a global one and distribute it to the devices and loop this program.

The key challenge in Federated learning is the data distribution, the local devices aggregates data from different context, which leads to a tremendous difference on the distribution of data. For example, the pictures captured by people living in country could be more possible containing a live chicken, in contrast, the people living in the city are more likely to take photos of vehicles. This lead to the non-IID data,