FARGO: Federated leARninG for human moBility

Radu-Ioan Ciobanu, **Ciprian Dobre**

ciprian.dobre@upb.ro

Objectives

- Study how the different dimensions of **federated learning** interact and interfere with each other, when devising a generic distributed software platform supporting a wide range of federated learning solutions in the context of **mobility analysis, modelling and prediction**
- Build open-source **software artefacts** supporting federated learning for mobility prediction
- Make the federated learning approach **intelligible** for an average developer wanting to use it

Federated learning (FL)

- Uses the computational power of the edge devices in order to reduce information traffic
- The model is downloaded on the device and updated using the locally stored data
- These models are then sent back to the central server and averaged, creating an aggregated global model
- The end goal is to achieve a minimized global loss

FARGO

- The problem addressed in is intrinsically **multidimensional**, comprising aspects related to:
    - human mobility
    - distributed algorithms
    - software architecture
- Providing reusable support for federated learning comes with several other generic **dimensions**, such as the considered data and algorithmic partition model, security and privacy constraints, scalability requirements, etc.

FARGO (2)

- The overall objective is thus to investigate to what extent these dimensions can be made **generic** and how they **interact**, and possibly **interfere** with each other
- We aim to offer various kinds of federated learning solutions for human mobility data in a **modular** and **flexible** platform
- This should be done in a **generic** fashion that will allow the developers to easily add and express their **custom** solutions
- For this reason, we proposed a model for federated learning for mobility, that will be used for describing the desired solutions

FARGO (3)

- We envision a federated learning solution as the set of unique answers to **ten key questions** that characterize it, representing the generic dimensions in the space of all potential FL solutions
- We ensure the genericity of our model by expressing these dimensions as parameters which must be set by the developers at the beginning of a new project in the FARGO platform
- There are **three categories** that the ten questions are split into:
  - machine learning algorithm
  - distributed system model
  - sharing and privacy models

Machine learning algorithm questions

- Question 1: What is the learning problem?
  - the main goal of the developer
  - various types of data: indoor, outdoor, vehicular
- Question 2: What is the learning model?
  - neural networks, support vector machines, linear regression, decision trees, reinforcement learning, etc.
- Question 3: What is the learning method?
  - regression, classification, clustering, dimensionality reduction, ensemble methods, neural networks and deep learning, transfer learning, • reinforcement learning, NLP, etc.
- Question 4: What is the partitionability of the problem?

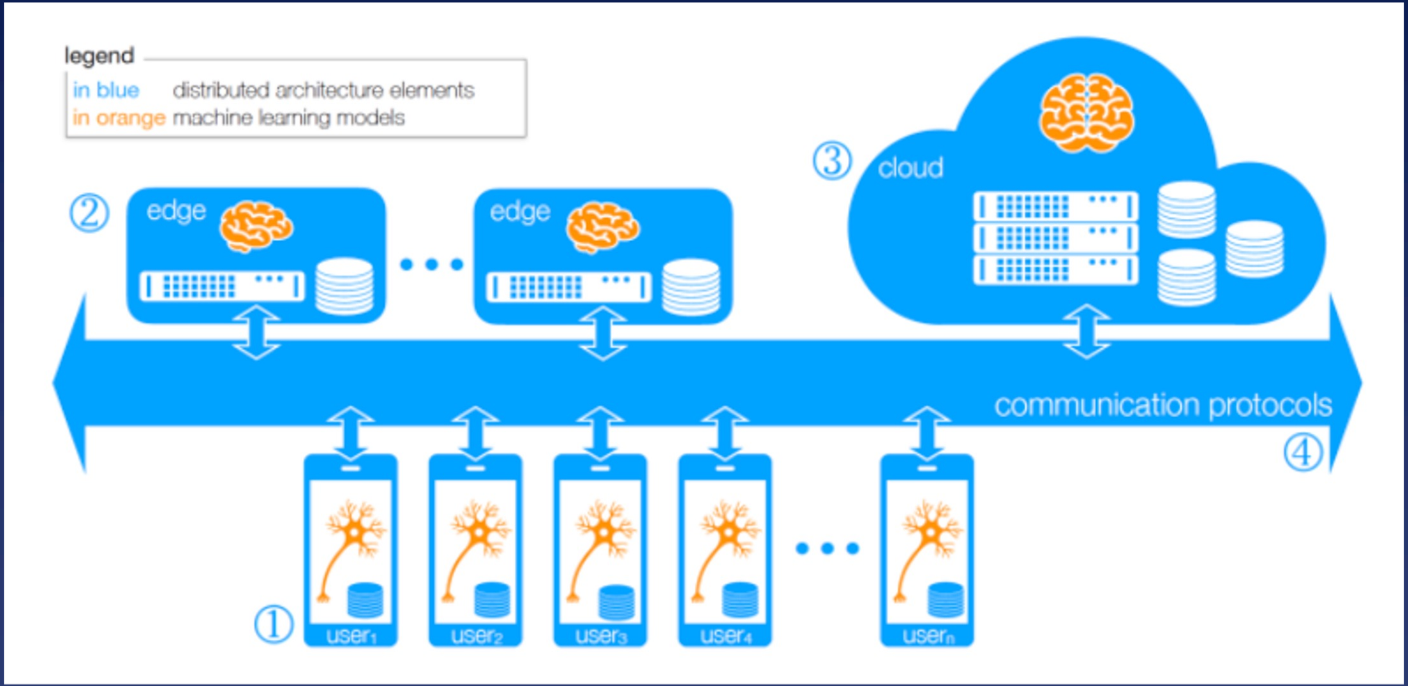
Distributed system model questions

- Question 5: What is the distribution model?
  - client-server, peer-to-peer
- Question 6: What communication protocol should be used?
  - Wi-Fi, Bluetooth, 4G, 5G, etc.
- Question 7: What failure model should be employed?
  - processing failure models (benign failures, crash failures, byzantine failures), message failure models (delay failures, duplication failures, disorder failures, drop failures)
- Question 8: What should the consistency model be?
  - strict, sequential, causal, processor, cache, slow, general, local, eventual, etc.

Sharing and privacy models questions

- Question 9: What sharing model does the problem require?
  - refers to the way data will be shared from the user's device towards the processing backend and is based on the fact that federated learning takes this into account inherently
- Question 10: What privacy model is needed?
  - inherent to federated learning

Architecture



Architecture (2)

- User nodes (1)
  - typically mobile devices associated with end users
  - the main acquisition sources of the location data used to train the FL models
  - modest computing capacity and a rather limited storage space
- Edge nodes (2)
  - computing devices located close to user nodes, at most one or two hops away
  - while more powerful than user nodes, edge nodes are also resource-constrained compared to cloud nodes
- Cloud nodes (3)
  - typically resides in a datacentre and benefits from a large computing capacity and a virtually unlimited storage space
  - they might however be located rather far from user nodes and edge nodes and might potentially become bottlenecks
- Communications protocol stack (4)

Thank you!

Questions?