Ethics of Data Science – Part III

Week 4: Federated Learning

Implementation

Dr. Chris Anagnostopoulos, Hon. Senior Lecturer

Federated learning

What?

- Federated analytics: compute simple statistics in a federated fashion
- Federated validation: build ML model centrally, validate against 3rd party data
- Federated learning: co-train an ML model on federated data

Why?

- Obfuscating or aggregating data as a way to de-anonymise is challenging for unstructured data
- Homomorphic encryption is computationally expensive
- “Data does not move” is simpler to communicate and to comply with

Federated learning

Why is it hard?

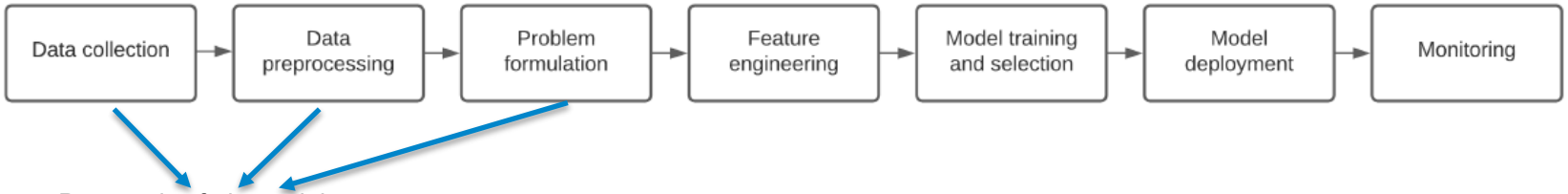
- Technologically, we need to distribute workloads given unreliable, heterogeneous client devices.
- Mathematically, we need to distribute computation.
- Statistically, local clients are operating in heterogeneous environments (dataset shift)
- Privacy-wise, we need to guarantee that data is not being leaked through model updates

Is it worth my time learning about this?

- Clear industry focus on federation (given privacy considerations and availability of edge computing)
- Forward-looking technological trend
- Technically very interesting

Federated learning

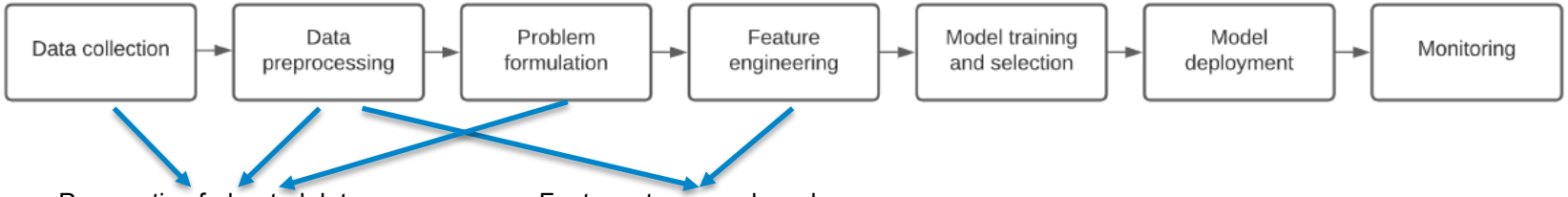
Federated learning



Prospective federated data collection requires strong strategic alignment

Retrospectively, federation requires conversion into a Common Data Model.

Federated learning



Prospective federated data collection requires strong strategic alignment

Retrospectively, federation requires conversion into a Common Data Model.

Feature stores or shared code on top of common data models

Visual checks for data quality decentralized

Federated learning

Prospective federated data collection requires strong strategic alignment

Retrospectively, federation requires conversion into a Common Data Model.

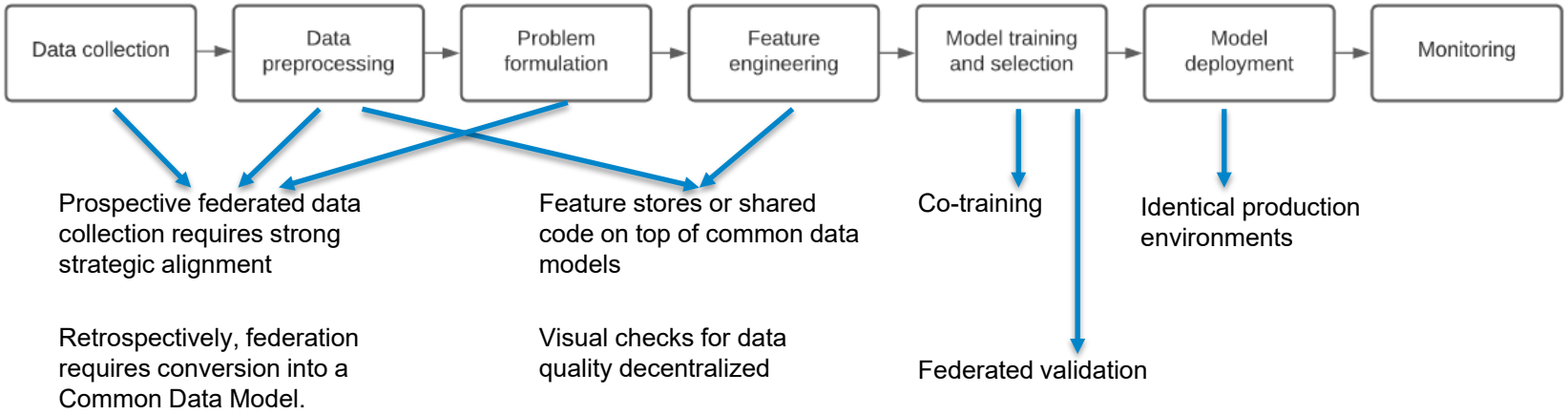
Feature stores or shared code on top of common data models

Visual checks for data quality decentralized

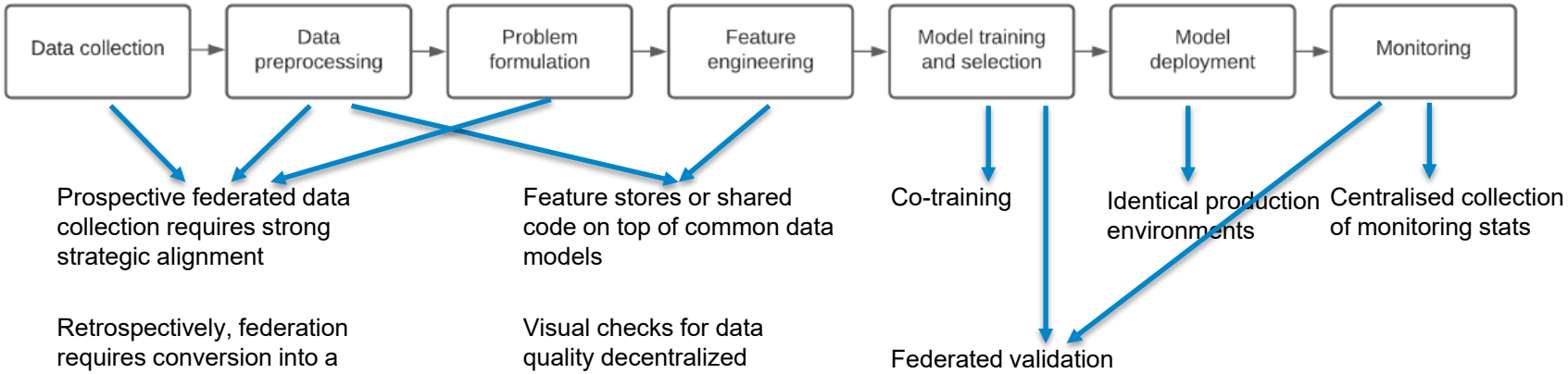
Co-training

Federated validation

Federated learning



Federated learning



Federated learning

Example 1: mobile phones

Example 2: hospitals