Vertically distributed learning for CVA

CORINNE ALLAART

ENABLING PERSONALIZED INTERVENTIONS

CONSORTIUM MEETING 22TH OF APRIL, 2021

Vertically distributed learning for CVA

- Vertically Distributed Deep Learning
  - Split Learning on vertically partitioned data
  - Evaluation on public medical datasets

- Use Case: Cerebrovascular Accident (CVA)
  - Prediction of Long Term Patient Outcomes
  - Antonius Ziekenhuis Nieuwegein and several rehabilitation clinics

Privacy Preserving Distributed Deep Learning
  - *Survey paper*
  - *Together with Saba Amiri (UvA)*

Vertically partitioned data

Vertically Partitioned Data

- Data of one patient in split up over multiple institutions

Distribution of features
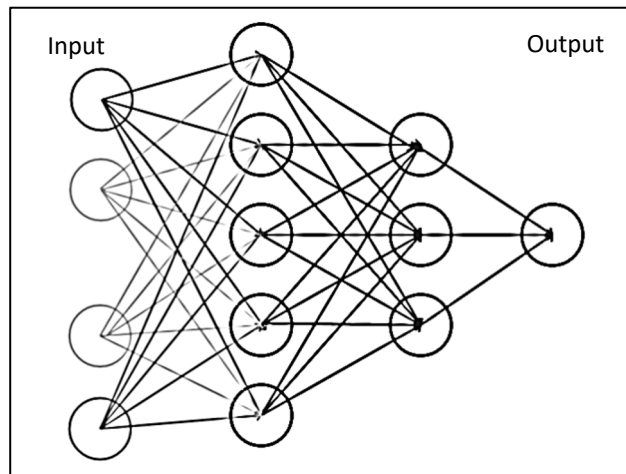
- Influence on feasibility of distributed learning

Vertical Split Learning

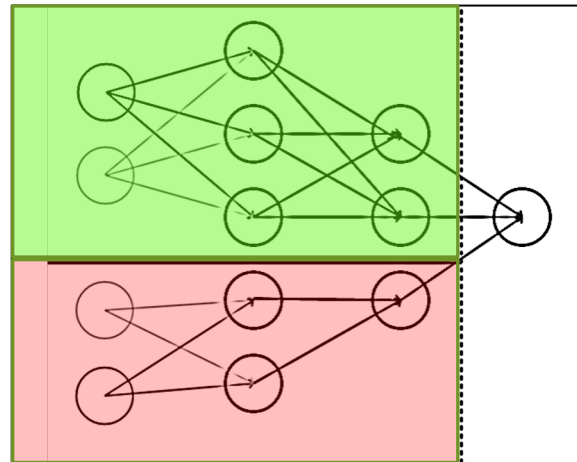Split Learning: neural network split over different locations
- Loss of interconnectedness
- Loss of predictive performance

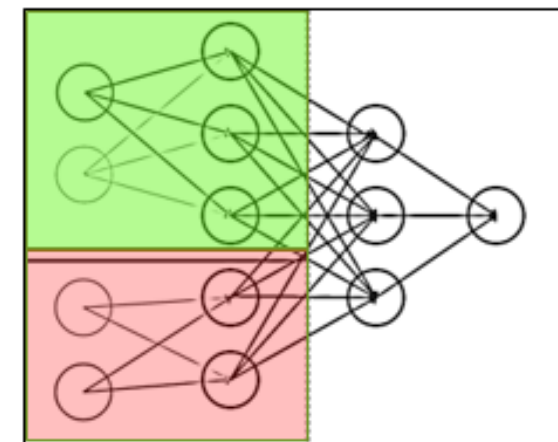Different possibilities for 'combining' outcomes of the split networks
- Payoff between privacy preservation and predictive performance

Centralized

Vertical Split

Vertical Split: situation 2

When is Vertical Split Learning beneficial?

Setting up split learning is time consuming: can we estimate when this will lead to a worthwhile increase in predictive performance?
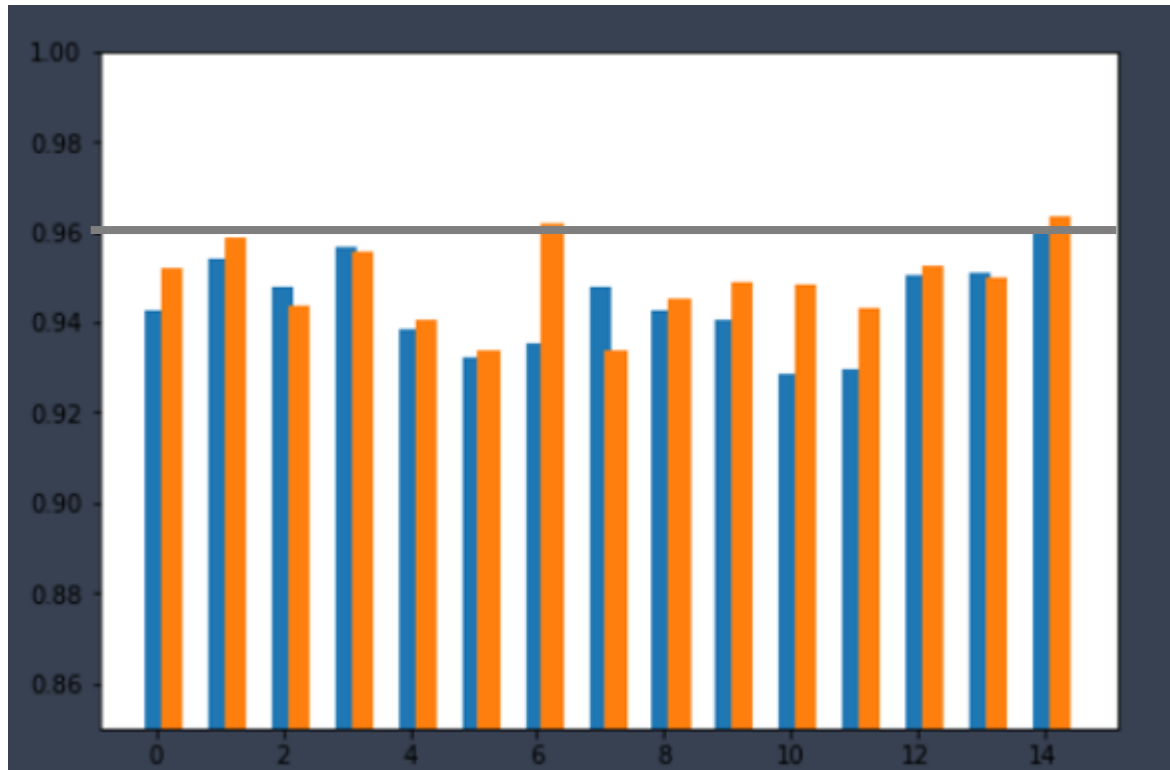
Expectation for
- loss of accuracy compared to centralized learning, and
- gain of accuracy compared to no data sharing

Best set-up for vertical split learning

Evaluating on 4 public medical datasets
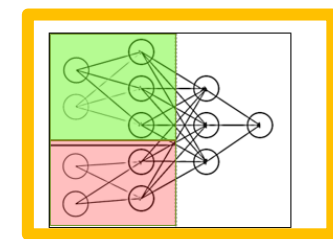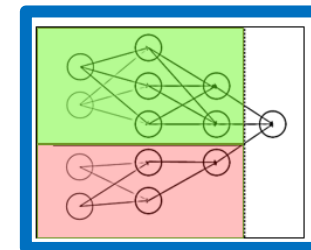- Expanding to financial domain (fraud) with a bachelor student
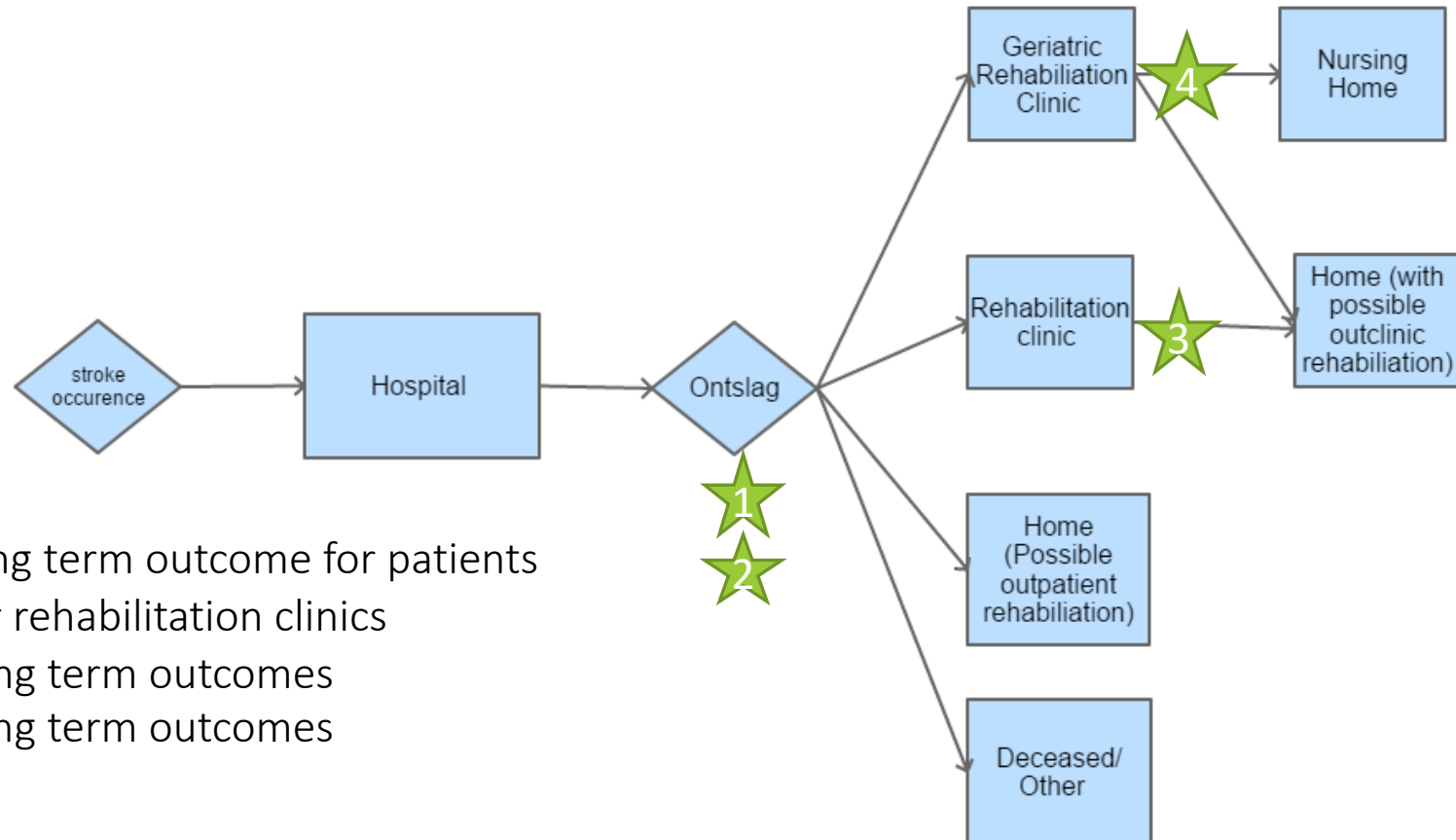
Preliminary results

Different distributions of features

Preliminary results
(Cervical Cancer UCI, 5-fold cross validation)
- Distribution of features is very influential on predictive performance
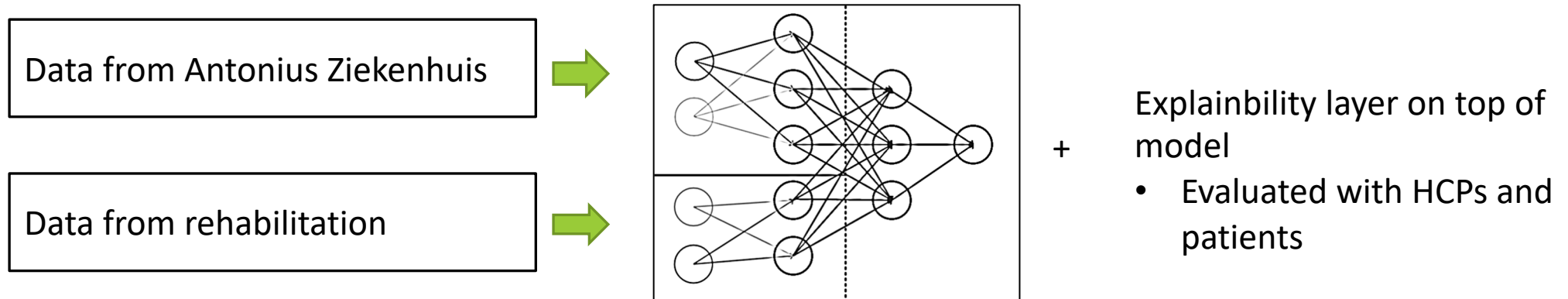- Vertical Split situation 2 seems to give higher performance

Use Case CVA: St. Antonius Ziekenhuis

Prediction moments: ⭐
1. Ontslaggesprek: Expected long term outcome for patients
2. Expected 'burden of care' for rehabilitation clinics
3. **Ontslaggesprek:** Expected long term outcomes
4. **Ontslaggesprek:** Expected long term outcomes

CVA use case: prediction model

Data from Antonius Ziekenhuis

Data from rehabilitation

+ Explainbility layer on top of model
- Evaluated with HCPs and patients

Patient 1:

Revalidatiekliniek ligduur :
10-14 dagen

Na 3 maanden
Functionele status : 1-2

Leeftijd <70
Geen complica
Variabel 3
Variabel 4
Variabel 5
Variabel 6
Variabel 7

Leeftijd <70
Geen complic
Variabel 3
Variabel 4
Variabel 5
Variabel 6
Variabel 7

Future work: Use Case CVA

Data collection and model development
- ◦ Complete data collection in rehabilitation clinics
- ◦ Apply Vertical Split Learning
- ◦ Focus on the other prediction moments

Development of infrastructure for data and model sharing between institutions
- ◦ Consortium effort

Evaluation of models and visualizations
- ◦ Focus groups with patients