CS-E4740 Federated Learning

"FL Project"

Dipl.-Ing. Dr.techn. Alexander Jung

FL Project

pick some FL application and model it as GTVMin

$$\left\{\widehat{\mathbf{w}}^{(i)}\right\}_{i=1}^{n} \in \underset{\text{stack}\left\{\mathbf{w}^{(i)}\right\}_{i=1}^{n}}{\text{argmin}} \sum_{i \in \mathcal{V}} L_i\left(\mathbf{w}^{(i)}\right) + \alpha \sum_{\{i,i'\} \in \mathcal{E}} A_{i,i'} \left\|\mathbf{w}^{(i)} - \mathbf{w}^{(i')}\right\|_2^2$$

you must:

- choose local datasets (train/val/test sets)
- choose local models
- choose loss functions
- choose edges and their weights

Local Datasets

- each local dataset $\mathcal{D}^{(i)}$

```
G.nodes[node_i]["ytrain"]
G.nodes[node_i]["Xtrain"]
```

consists of a

- train set: used to define local loss in GTVMin
- val set: used to select local models and/or edges
- test set: used for final performance assessment

Local Models

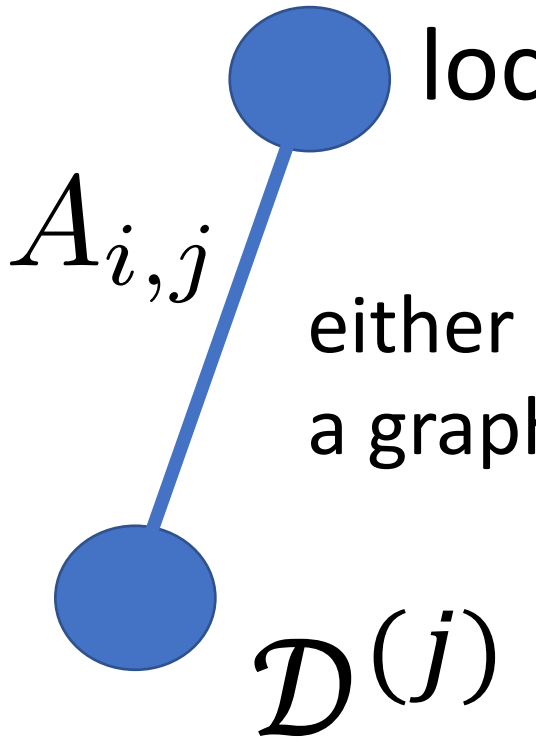
- local dataset $\mathcal{D}^{(i)}$

  local model can be anything listed in MyCourses Section “FL Project”

```
Gin.nodes[node_i]["model"] = DecisionTreeRegressor(max_depth=4)
```

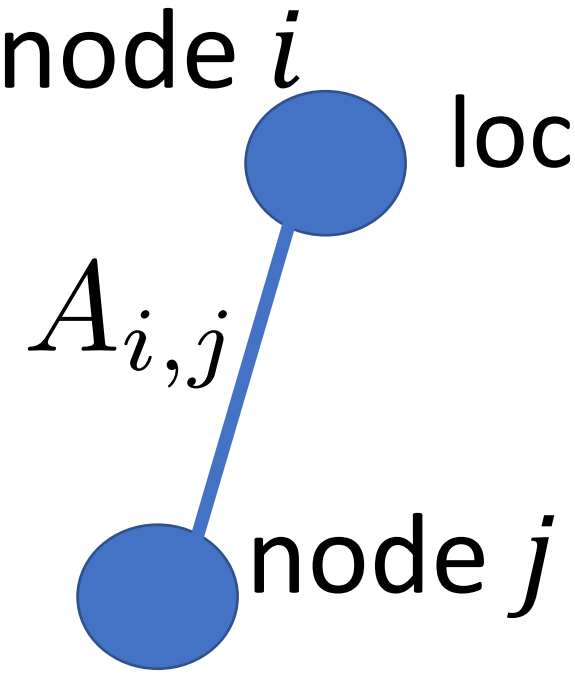
where did we choose loss function here?

Choose Edges and Weights



either choose manually (using your domain expertise) or a graph learning method (see Sec. 7.5 of lecture notes)

Variation of Parametric Models



local model param. vec. $\mathbf{w}^{(i)}$

$$\left\|\mathbf{w}^{(i)} - \mathbf{w}^{(j)}\right\|^2$$

local model parameters $\mathbf{w}^{(j)}$

Putting Together the Pieces

$$\left\{\widehat{\mathbf{w}}^{(i)}\right\}_{i=1}^{n} \in \underset{\mathrm{stack}\left\{\mathbf{w}^{(i)}\right\}_{i=1}^{n}}{\mathrm{argmin}} \sum_{i \in \mathcal{V}} L_i\left(\mathbf{w}^{(i)}\right) + \alpha \sum_{\{i,i'\} \in \mathcal{E}} A_{i,i'} \left\|\mathbf{w}^{(i)} - \mathbf{w}^{(i')}\right\|_2^2$$

FL Algos obtained by applying optimization methods (GD, SGD) to solve GTVMin [1, Sec. 5]

FL Project Credits and Grades

- extends course from 5 to 10 credits
- project points P2 (max. 80 report/max 20 review)
- basic variant points P1 (max 100)
- grade for 10 credit variant determined from

½ (P1 + P2)

1: 50-59; 2: 60-69; 3: 70-79; 4: 80-89; 5: 90-

FL Project Deliverables

- submit project report + notebook by 31.5.2024
- peer review during 31.5 – 30.06.2024
- By 30.09.2024: submit revised project report/ notebook AND response to peer reviews

FL Project Schedule

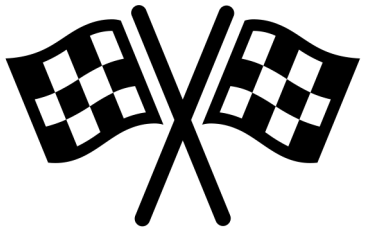
now

prepare project report +
notebook with experiments

31.05.2024

peer review

30.06.2024



revise project report +
notebook with experiments
prepare response letter

1. Introduction
2. Problem Formulation
3. Methods
4. Results
5. Conclusion

1. Introduction

- describe application domain
- summarize existing work
- outline of paper

2. Problem Formulation

provide qualitative description of local datasets:

- what are datapoints? [2, Ch. 2]
- what is the quantity of interest (label)? [2,Ch. 2]
- what is the data source ?
- intrinsic similarity betw. local datas.? [1, Sec. 6]

3. Methods

- which local models/local loss and why those?
- which FL algorithm used ? [1, Sec. 5]
- how did you validate trained models? [2, Ch. 6]

4. Results

- report (local) train/val errors for each model
- diagnose FL algorithm [2, Ch. 6.6.]
- final choice for edge weights, GTVMin param. λ
- final chosen local models and their test errors?

5. Conclusion

- recap your findings during the project work
- ponder about limitations and possible improvements

References

[1] AJ, “Lecture Notes CS-E4740”, link, 2024

[2] AJ, “Machine Learning: The Basics,” link, 2022

Peer Review Questions

draft for you to comment: click here

revise report and notebook based on peer review

Final Submission (30.09.2024):

- revised report (pdf)
- revised notebook with experiments (ipynb)
- response letter that explains how you used the peer review
- sample1, sample2, sample3

Thank you for your attention!