TEMPLATE FOR A PROJECT REPORT

Author Name, Affiliation

Report for a project within the course CS-E4740 - Federated Learning.

ABSTRACT

Instructions (Remove before submission). Provide a concise summary of the project, including the FL application, how you modelled it as a FL network, the GTV minimization on top of it and the FL algorithms used.

***Index Terms*—** Federated Learning, Networks, Personalized Machine Learning, Trustworthy AI

1. INTRODUCTION

Instructions (Remove before submission) This section introduces the background and motivation of your FL project. It should include:

- **A real-life scenario motivating your FL application, consisting of networked devices that train personalized models (see Lecture "FL Networks").**
- **A summary of state-of-the-art methods relevant to your project.**
- **A brief outline of the structure of your report.**

2. PROBLEM FORMULATION

Instructions (Remove before submission). A core task for the project is the modelling of your FL application as an federated learning (FL) network (see Lecture "FL Networks" and [2, Ch. 3]). This section must clearly state and explain:

- **Nodes: How many and what real-world devices (smartphones, smart light-bulbs, smart meters, a fitness tracker or a desktop computer) do they represent.**
- **Local Models: Describe the models used at each node. You can also state several different candidate models for a node.**
- **Loss Functions: For each node, specify the local loss functions for training the corresponding local model.**
- **Edges: How do you choose the edges and their weights? (see [2, Sec. 7])**

3. METHODS

Instructions (Remove before submission). Design FL algorithms for your application by formulating FL as GTVMin which is solved using distributed optimization methods. Clearly state and explain:

- **Your choice for the variation measure. For example, when using parametric local models you can use $\phi\big(\mathbf{w}^{(i)} - \mathbf{w}^{(i')}\big)$ to measure the variation across edge $\{i, i'\}$ (see Lecture "FL Design Principle" and [2, Ch. 3]).**
- **Your choice for FL algorithm, which must be implemented as message passing across the edges of the FL network (see Lecture "FL Algorithms" and [2, Ch. 5]).**

4. NUMERICAL EXPERIMENTS

Instructions (Remove before submission). Provide details on the implementation of the FL algorithms and analyze the results. Clearly state and explain

- **the used data sources. One example for a data source is the open data interface of the Finnish Meteorological institute [3].**
- **Your approach to model validation, selection and diagnosis (see [1, Sec. 6.6]). In particular,**
  - **how did you implement model validation?**
  - **If you have considered different options for the local model used by some node i, how did you choose between those options?**
  - **how happy are you with the performance of the resulting trained local models? can you think of a baseline to compare against?**
- **for each trained local model, state and compare the average loss on the training set and the validation set.**
- **how you obtained the final trained local models. This includes**

  - **the choice of FL network (including choice for local model and edge set)**
  - **the choice for the variation measure (penalty function for parametric models).**
  - **the FL algorithm (which must be implemented as message passing)**

- **For the finally obtained trained local models, report the average loss on a test set (which has neither been used for training or validation of local models).**

Important: Your report must be accompanied by a Python notebook that allows reproduce the numerical results. Try to use only basic libraries such that the notebook can be executed without hassle on `https://jupyter.cs.aalto.fi/`.

5. CONCLUSION

Instructions (Remove before submission):

- **Discuss whether the obtained numerical results suggest that the problem is solved satisfactorily or if there is room for improvement.**
- **Identify limitations of the methods used and suggest potential improvements.**

6. REFERENCES

[1] A. Jung, *Machine Learning: The Basics*, Springer Singapore, 1st edition, Feb. 2022. working draft: https://github.com/alexjungaalto/MachineLearningTheBasics/blob/master/MLBasicsBook.pdf

[2] A. Jung, "Federated Learning: From Theory to Practice," Aalto, 2025. https://github.com/alexjungaalto/FederatedLearning/blob/main/material/FLBook.pdf

[3] Finnish Meteorological Institute, "Open Data," 2025, https://en.ilmatieteenlaitos.fi/open-data, Accessed: 2025-02-27.