# 北 京 邮 电 大 学

## 本科毕业设计（论文）任务书

| | | | | | |
|---|---|---|---|---|---|
| 学院 | 网络空间安全学院 | 专业 | 网络空间安全（实验班） | 班级 | 2019211806 |
| 学生姓名 | 卢亭松 | 学号 | 2019212443 | 班内序号 | 11 |
| 指导教师姓名 | 陆月明 | 所在单位 | 网络空间安全学院 | 职称 | 教授 |
| 设计(论文)题目 | （中文）一种基于 FATE 框架的联邦学习方法设计与实现 | | | | |
| | （英文）Design and Implementation of a Federated-Learning Method Based on FATE Framework | | | | |
| 题目分类 | 工程实践类☐　　研究设计类☑　　理论分析类☐ | | | | |
| 题目来源 | 题目是否来源于科研项目　　是☑　　否☐ | | | | |
| | 科研项目名称： | | | | |
| | 科研项目负责人： | | | | |

主要内容：

由于隐私保护要求，使得某些行业领域的数据共享面临困难，数据效用无法充分利用，形成了所谓的数据孤岛问题，联邦学习通过中央服务器在保护隐私的同时从本地数据中学习，为跨设备、跨孤岛机器学习问题提供了解决方案。毕业设计将基于开源项目 FATE (Federated Al Technology Enabler)，学习机器学习与隐私保护相关知识，搭建 FATE 联邦学习集群，对机器学习隐私保护，联邦学习进行调研，基于 FATE 框架设计实验，在联邦场景实现两种以上主流机器学习场景（计算机视觉、自然语言处理等）的样例算法，完成联邦学习机器学习算法的训练和测试，并与传统机器学习进行对比，从准确率、安全性等层面进行分析，完成毕业设计论文。

主要（技术）要求：

1. 学习机器学习与隐私保护相关知识，学习一种传统的机器学习框架（pytorch、tensorflow），学习并搭建 FATE 联邦学习集群。

2. 对机器学习隐私保护，联邦学习进行调研，分析当前机器学习隐私保护现状和联邦学习的优势，对联邦学习的场景，特点，实现原理进行调研，撰写文献综述。

3. 基于 FATE 框架设计实验，在联邦场景实现两种以上主流机器学习场景（计算机视觉、自然语言处理等）的样例算法。

4. 完成联邦学习机器学习算法的训练和测试，并与传统机器学习进行对比，从准确率、安全性等层面进行分析。

主要参考文献：

[1]谭作文, 张连福. 机器学习隐私保护研究综述. 软件学报, 2020, 31(7): 2127-2156.

[2] Abreha HG, Hayajneh M, Serhani MA. Federated Learning in Edge Computing: A Systematic Survey. Sensors. 2022; 22(2):450. https://doi.org/10.3390/s22020450

[3] K. M. Ahmed, A. Imteaj and M. H. Amini, “Federated Deep Learning for Heterogeneous Edge Computing,“ 2021 20th IEEE International Conference on Machine Learning and Applications (ICMLA), 2021, pp. 1146-1152, doi: 10.1109/ICMLA52953.2021.00187.

[4] Y. Liu, J. J. Q. Yu, J. Kang, D. Niyato and S. Zhang, “Privacy-Preserving Traffic Flow Prediction: A Federated Learning Approach,“ in IEEE Internet of Things Journal, vol. 7, no. 8, pp. 7751-7763, Aug. 2020, doi: 10.1109/JIOT.2020.2991401.

[5] Liu JC, Goetz J, Sen S, Tewari A. Learning From Others Without Sacrificing Privacy: Simulation Comparing Centralized and Federated Machine Learning on Mobile Health Data. JMIR Mhealth Uhealth, 2021;9(3):e23728

[6] 胡健龙. 联邦学习在车联网数据共享与保护技术中的研究[D]. 电子科技大学,2022.DOI:10.27005/d.cnki.gdzku.2022.004716.

[7] M. Nasr, R. Shokri and A. Houmansadr, Comprehensive Privacy Analysis of Deep Learning: Passive and Active White-box Inference Attacks against Centralized and Federated Learning. In Proceedings of 2019 IEEE Symposium on Security and Privacy (SP), 2019, pp. 739-753, doi: 10.1109/SP.2019.00065.

[8] 王慧超. 机器学习中的数据隐私保护研究[D]. 中国科学技术大学,2021.DOI:10.27517/d.cnki.gzkju.2021.001722.

[9] Wang, X.; Wang, J.; Ma, X.; Wen, C. A Differential Privacy Strategy Based on Local Features of Non-Gaussian Noise in Federated Learning. Sensors 2022,22(2424). https://doi.org/10.3390/s22072424

[10] 师兆森. 联邦学习中的隐私保护技术研究. 电子科技大学,2022. DOI:10.27005/d.cnki.gdzku.2022.000987.

进度安排：

1、学习机器学习、隐私保护相关的知识，对研究背景进行调研。查找并阅读联邦学习、机器学习隐私保护的相关论文，为后续的实验以及论文撰写打下基础。秋季学期 17-18 周

2、开始 FATE 集群的搭建，完成开题报告的撰写。春季学期 1-2 周

3、构建论文框架，完成 FATE 集群搭建。春季学期 3-4 周

4、完成论文前两章的撰写，确定两种以上的机器学习场景及代表性问题，设计相应的联邦机器学习算法。春季学期 5-6 周

5、总结上一阶段的工作，完成中期检查。春季学期 7 周

6、完成两种以上的联邦机器学习算法的传统实现与联邦环境实现，能够在 FATE 集群上进行训练。春季学期 8-9 周

7、在 FATE 平台上完成预测与评估，对比分析算法在集中式环境与联邦环境的差异，从准确率，安全性，通信效率等进行分析。春季学期 10-11 周

8、根据上述的实验结果，完成论文的撰写。春季学期 12-13 周

9、完成论文，整理本毕设课题的全部成果。春季学期 14 周

| 指导教师签字 | | 日期 | 年　月　日 |
|---|---|---|---|