Room513, Pengyuan 20#
Tongji University
Shanghai, China
greilfang@gmail.com
Webpage : https://greilfang.github.io/
0086-18916093316

Pei Fang

Senior Undergraduate, Computer Science

Education

Tongji University, Shanghai, China
Bachelor of Engineering in Mechanical Engineering *Jul' 2016 - Jul' 2017*
Bachelor of Engineering in Software Engineering *Jul' 2017 - Jul' 2021 (Expected)*
GPA: 89.8/100 (∼16%) (Overall)**; 92/100** (Major)
Core courses:
Math: Advanced Mathematics (4), Linear Algebra (5), Probability and Statistics (5), Discrete Mathematics (5), Algorithm (5), Data Structure (5).
Computer Science: Data Mining (5), Database Principle (5), Operating System (4), Computer Network (5), Computer System Structure(5), Cloud Computing (5).

Publications

Manuscripts (TKDE Submitted)
Pei Fang, Zhendong Cai, Qingjiang Shi, "FLFE: A Communication-efficient and Privacy-preserving Federated Feature Engineering Framework"
See: https://arxiv.org/abs/2009.02557

Research Experiences

Federated Sports Prediction on Wearables | HKUST SymLab
Supervisor : Prof. Pan Hui *Sep '2020 - Present*

- Used Federated Learning techniques to solve privacy issues. Applied a model-centric system design: A parameter server holds a global model and each client trains the model with its local data.
- Split the training into many Cycles, where the phones send their updated model to the server, and the server aggregates them. The training data is collected by sensors in user's smart watches.
- Responsible for the performance improvement of the model, and the federated model implementation both in server and Android platform with PySyft and KotlinSyft
- The system predicts 6 different sports with 3 sensors at an accuracy of 95%, where users needn't upload their local data. Currently we are solving Malicious Client problem.

Federated Automated Feature Engineering | Tongji University
Supervisor : Prof. QingJiang Shi *Sep '2019 - May '2020*

- Implemented Learning Feature Engineering (LFE) (F Nargesian et al. IJCAI, 2017), whereupon we collected 120 datasets on OpenML, and generated quantile sketch arrays (QSA) to train a set of MLP classifiers. Performed automated feature engineering on a real-world insurance dataset.
- Applied the transformation of which the confidence score returned by the classifier is greater than 0.9 and compared the result with auto-encoder. LFE reduced the dimensions from 330 to 130 and improved the f1-score from 85.6% to 87.5%.
- Studied the potential of LFE in multi-party feature engineering problem. Proposed FLFE, a framework that performs automated multi-party feature engineering. Designed a feature exchange mechanism to preserve privacy and reduce communication overhead.

See:Federated Forest, Automated Feature Engineering, and Federated Learning Feature Engineering

Time Series Forecasting Models on Commodity Cluster | Tongji University
Supervisors : Prof. WeiXiong Rao *Apr '2020 - May '2020*

- Implemented FTC-Tree and Time Series Forecasting Models proposed in papers.
- Optimized the distance function of FTC-Tree and integrated it with a K-Mediod algorithm. The Silhouette Coefficient reached up to 0.29 (better than 0.26 in the paper).
- Applied SVM, Random Forest, and MLP to predict commodity sales. Applied PCA to enhance robustness. Reached a comparable Residual Standard Error (RSE) to the result in the paper.

See: https://github.com/Greilfang/Data-Mining

Course Projects

A simulative Red-Alert game
C++ Programming Language | Supervisor : Prof. Qingpei Zhao *Mar '2018 - Apr '2018*

Used Cocos-2d as the game engine and was responsible for Network Server programming and basic manipulations. Completed the LAN multi-player game by using the Asio.boost net module.
See: https://github.com/Greilfang/Our-Red-Alert

A simulative distributed file system
Cloud Computing | Supervisor : Prof. Jiangfeng Li *Dec '2019*

Implemented a simulative distributed system with Python, which includes mechanisms such as heart-jump, redundancy, and distributed IO. Referred to the structure of UFS and Google File System.
See: https://github.com/Greilfang/hive-hdfs-practise

An OLAP movie anaysis application
Data Warehouse | Supervisor : Prof. Hongming Zhu *Nov '2019 - Dec '2019*

Crawled 187,881 movies and more than 7,910,000 movie reviews. Analyzed the reviews' emotional inclinations. Tried MySQL (relational database), MongoDB (non-relational database) and Neo4j (graph database) to speed up the query.
See: https://github.com/Greilfang/Amzaon-movie-analysis

Work Experience

Lightelligence, Inc
Position : Algorithm Engineer Intern *Dec '2020 - Present*

- Read and reproduce the latest papers on Computer Vision, Natural Language Processing, and Datacenter Network, looking for the potential of optic AI chips on these fields.
- Develop products based on optic AI chips with C and Assembly.

Shenzhen Research Institution of Big Data
Position : Research Intern *Jul '2020 - Sep '2020*

- Participated in the development of the large-scale Linear-Programming-Solver with C++.
- Developed the solver based on Primal Simplex, Dual Simplex, and Interior Point Method.

Leadership & Activities

Tongji Open-Source Student Club | Core Member *Oct '2017 - Apr '2020*
Gave lessons on: a. Python data analysis with NumPy and Pandas. b. React.js, basically the DOM and the transaction mechanism.
Awarded **'Best Lecturer'** by **Tongji Google Camp Club**.

Skills & Others

Languages: C, C++, C#, Python, JavaScript, LaTeX
Research Tools: PyTorch, scikit-learn, PySyft, MPI, etc.
English: Toefl: 101 (Reading: 30, Listening: 30, Speaking: 22, Writing: 19).