YIYANG LIU

2333 McClintock Ave, Los Angeles, CA 23333

233-233-233 | 233@usc.edu | linkedin.com/in/yiyangiliu | github.com/yiyangiliu

EDUCATION

University of Southern California   Los Angeles, CA
MS, Applied Data Science   *Aug. 2021 – May 2023*

Beijing University of Technology   Beijing, China
BE, Computer Science (with Hons.); GPA: 2.33/4, WAM: 23.3/100, Ranking: 23/233   *Sep. 2016 – July 2020*

PROFESSIONAL EXPERIENCE

ByteDance (TikTok's Parent Company)   Beijing, China
Software Engineer Intern   *Feb. 2021 – July 2021*

- **Action Center:** Imported the user action statistics from database to action center, and improved the recommender system's feed & search accuracy. Used stream processing framework to accelerate dumping speed.
- **User Event Storage:** Automatically parsed events from the user event log to generate rocksdb database file, and deployed this service by daily running docker on distributed instances.
- **A/B Test Toolkit:** Provided control tool for machine learning engineers to check different models' end-to-end effect on our recommender system's feed service. Reduced 23.3% of average delivery time of deep-learning models.
- **DevOps & QA:** Renovated service deployment scripts and transfered offline files from Alibaba Cloud to ByteDance's HDFS, saving $2333 per month. Wrote unit test, system test, end-to-end test code to ensure code quality.

FinTech Startup & Tsinghua University   Beijing, China
Machine Learning Engineer Intern   *Nov. 2019 – June 2020*

- **Sentiment Analysis**: Used Tensorflow & Keras training multiple depth-learning models on multiple GPUs, fine-tuning with multiple techniques such as pseudo-label, cross-verification and model fusion.
- **Paper Implementation (GitHub):** Implemented BERT model based on its paper, containing all components, such as multi-head self attention mechanism, transformer, embedding layer.
- **Financial News Dataset:** Wrote web crawler to make 23.3G financial news corpus dataset with millions of records by using Requests, Beautifulsoup, Selenium. Used pandas for cleaning data.
- **Kaggle Model Evaluation (GitHub):** Implemented best solution of kaggle house-prices forecasting competition and evaluated various machine learning methods' performance, such as XGBoost, random forest, tf-idf.
- **Bert4Keras (GitHub):** Submitted several merge requests to Bert4Keras, a prestigious 3.4k stars open source framework for training, fine-tuning deep-learning language models.

Carnegie Mellon University   Pittsburgh, PA
Research Assistant   *Sep. 2019 – Dec. 2019*

- **Cryptocurrency Transaction System (GitHub):** Mocked Bitcoin System to implement all crucial components such as block mining, PoW (proof of work), wallet, public & private key pairs, money transaction.
- **Technical Report:** Investigated Bitcoin system, and wrote out the whole operation process, including block mining, generation, proof, dissemination, transacting.

Ritatsu Software Tokyo   Tokyo, Japan
Software Engineer Intern   *Aug. 2019 – Sep. 2019*

- **Quality Assurance:** Designed and wrote unittests, function tests, end-to-end tests to guarantee the robustness of code delivered to some prestigious financial institutions like Nomura Securities.

SELECTED PROJECTS

RescueTime-Visualization (Demo, GitHub): Visualize RescueTime data and update it to GitHub to increase visibility.

Deep Learing Based Advertising System (GitHub): Use DL model to identify objects in video to increase Ad accuracy.

SELECTED SKILLS

Languages: Golang, Python, C++, Shell

Developer Tools: Git, Docker, Cloud Platform, Jetbrains Suite