Zhoujie Ding

Berkeley, CA | +1 (510)717-1017 | zhoujie.ding@berkeley.edu | https://jasonding0401.github.io

EDUCATION

University of California, Berkeley **Expected Graduation: 12/2023**
B.A. in Computer Science and Applied Mathematics **GPA: 3.96/4.00**
Honors: Phi Beta Kappa (top 10% L&S students); Upsilon Pi Epsilon (top one third CS students); Dean's List
Core Coursework: Machine Learning, Deep Learning, Computer Vision, Reinforcement Learning, Optimization Models, Database Systems, Operating Systems, Theoretical Statistics, Real Analysis, Complex Analysis

WORK EXPERIENCE

Research Assistant **05/2022 – 08/2022**
Advisors: Joseph Gonzalez, Ion Stoica, at UC Berkeley SkyLab
- Researched on Skyplane project, a tool for fast inter-cloud data transfers. Gained over 5000 PyPI downloads.
- Benchmarked data transfer speed and cost against AWS DataSync: up to 100x faster and 85% lower cost.
- Implemented user usage metrics collection to improve core APIs and prioritize bug fixes.
- Familiarized with AWS, GCP, Azure CLIs, Python APIs, object storage, and object transfer services.

Software Developer Intern **06/2021 – 08/2021**
Manager: Huahua Chen, at Hundsun Technologies Inc., Hangzhou, China
- Devised Word and pdf content information-extraction, and tested over 200 fund contract documents.
- Resolved ~3000 conflict data by Python program for NLP model training, saving 30% time in hand labeling.
- Deployed a Python API with tornado for document info-extraction and tested 100 HTTP requests with Postman.

RESEARCH EXPERIENCE

Undergraduate Researcher **02/2022 – Present**
Advisors: Yizheng Chen, David Wagner, at Berkeley Security Group
- Researched deep learning for vulnerable program detection.
- Curated ~10,000 hand-labeled commits from the top 100 GitHub C/C++ repos; used them for model pretraining.
- Benchmarked vulnerability detection on code changes, commits, and issues among 3 SOTA models.
- Now focus on using Snorkel for hand labeling and fine-tuning language models for classifying issue messages.

Undergraduate Researcher **03/2021 – 12/2021**
Advisor: Joseph Gonzalez, at UC Berkeley RISELab
- Researched on Kernel-as-a-Service project, which manages shareable GPU memory and schedules user kernels across this pool of available GPUs.
- Implemented KaaS by Ray, and integrated Python API-based interfaces for users to easily deploy it.
- Benchmarked on online and offline workloads against Ray actors: 12x lower tail latency, 6x higher endurance.

PROJECTS

Facial Keypoint Detection **10/2021 – 12/2021**
Class Individual Project
- Modified ResNet-50 to train on iBug dataset of 6666 images of varying image sizes with data augmentation.
- Turned the ground truth keypoint coordinates into pixel-aligned heatmaps and modified FCN-ResNet to train.
- Reached top 10 in class Kaggle competition among over 150 students.

Pintos Operating System **01/2021 – 05/2021**
Class Group Project
- Designed an efficient alarm clock and strict priority scheduler to support multithreaded user programs.
- Added a buffer cache for Pintos file system, support for extending files, and syscalls for directory manipulations.
- Wrote design documents in four parts – data structures, algorithms, synchronization, and rationale.

SKILLS

Programming Languages: Python, C**,** Java, SQL**,** RISC-V, Matlab;
Libraries: PyTorch, NumPy, Skimage, Sklearn, Ray, Matplotlib; **Toolkits:** AWS, Docker, Postman, Git, Grafana