Yuchen (Romee) Zhou

yz7258@nyu.edu · Homepage · Yuchen Zhou · ROMEEZHOU

EDUCATION

New York University Shanghai | Shanghai, New York Sep. 2020 – May 2024

Bachelor of Science in **Computer Science** *, minor in* **Philosophy** *and* **Mathematics** *(minor posted in Spring 2024)*

GPA 3.91/4.0 **Computer Science GPA** 3.98/4.0

Core Courses Data Structures (A), Algorithms (A), Operating Systems (A), Introduction to Databases (A), Computer Networking (A), Open Source Software Development (A), Machine Learning (Data Science Graduate Course) (A-), Randomized Algorithms (A)

WORK EXPERIENCE

SAP China Intern | SAP Labs China | Shanghai, China Aug. 2023 – Present

- Led generative AI learning sessions for CRT (compliance, risk, and tax) team, covered fine-tuning, embedding, AI agent, etc.
- Developed generative-AI powered US internal tax configuration bot and deployed a demo in docker; reached 80%+ accuracy when the bot answers questions regarding 300+ pieces of tax policies.
- Fine-tuned StarCoder to generate ABAP (a programming language that SAP developed) programs with a global team.

Learning Assistant for Introduction to Computer and Data Science | NYU Shanghai | Shanghai, China Aug. 2023 – Present

- Held regular office hours and review sessions; tutored over 80 students.

AsiaInfo Algorithm Development Intern | AsiaInfo Telco AI Labs | Beijing, China May 2023 – Aug. 2023

- Authored file-parsing programs to parse 10k+ lines content in different file formats, such as pdf, xlsx, to natural language texts.
- Developed new LLM-based document question-answering MaaS service, which innovatively realized question-answering over pictures and tables in files with file-parsing scripts and served 50k+ users worldwide; submitted for patent publication
- Explored integration of Federated Learning in large-scale network log processing and co-authored a published patent CN116647471A.

RESEARCH EXPERIENCE

Production Suite for Large-scale Distributed System Failure Detectors | Advisor: Prof. Olivier Marin Sep. 2023 – Present

- Researched on failure detectors (FDs), large-scale failure detectors, and simulation environment for distributed systems.
- Improved an existing production environment for FDs to support multiple window operations on former heartbeats when FDs predict the expected arrival time of the next heartbeat, which increased the number of FDs supported by the environment by 50%+.
- Designed data structures to simulate interactions between multiple FDs and further simulate large-scale distributed system FDs.

Large Graph Visualization with Embedding & Dimension Reduction | Advisor: Prof. Jie Xue May 2022 – Sep. 2022

- Implemented graph embedding algorithms in Python, including DeepWalk, Node2Vec, and Laplacian Eigenmap; designed ShortestPath embedding and SPLEE (shortest-path Laplacian eigenmap embedding).
- Implemented t-SNE dimension reduction in Python and amplified t-SNE to t-SGNE by using nearest neighbor approximation.
- Designed metrics based on density to measure algorithms' quality of visualizing different clusters of nodes.
- Increased visualization quality by 5%-10% on graphs with 1k-7k nodes and 5k-27k edges; decreased the compution time to less than 5min on graphs with 300k nodes and 750k edges.

PROJECTS

Open Source Contribution to Python scikit-learn | Individual Contributor Mar. 2023 – Present

- Added automatic parameter validation for public functions to standardize API maintainance; influenced 600k+ users.
- Opened pull requests for adding verbose to guassian process regression and wrote comprehensive tests for this newly added feature.
- Achieved 9 merged pull requests among 11 pull requests; Python scikit-learn currently has 56.1k+ stars.

Open Source Firefox Web-extension | Leader of Group of 3 Feb. 2023 – Feb. 2023

- Developed back-end programs to change browser themes separately on different windows and to self-create window themes.
- Authored a front-end popup window in HTML and JavaScript to navigate the web-extension; project influenced 30+ users.

COMPETITION EXPERIENCE

L'Oreal Brandstorm Competition 2022, Top 200 in Mainland China Feb. 2022 – May 2022

- Designed an app that applies traditional machine learning, facial recognition and scene recognition to analyze user location and environment to provide personalized makeup tutorials; the app was predicted to have 30 million users by 2025.

PATENT & PUBLICATION

- Xinyu Li, Yao Xiao, and **Yuchen Zhou**. Efficiently Visualizing Large Graphs. *arXiv: 2310.11186*.
- Yong Song, Zhiyong Yuan, **Yuchen Zhou**, et al. Data Processing Method, Device, Electronic Equipment, and Computer Storage Medium. *CN116647471A*. Aug 25, 2023.

AWARDS AND SKILLS

- **Awards:** [1] Dean's Lists (every academic year) [2] 2020 Peng Yachao Elite Scholarship ($16,500)
- **Programming:** Python, C, C++, Java, HTML, JavaScript, CSS, R, PyTorch, TensorFlow
- **Other:** SQL, Git, Linux, Docker, Spring Boot, LangChain, RapidMiner, Flask, Wireshark, Cisco Packet Tracer