Zehao (Zach) Guan

zehaog@cs.cmu.edu | (412) 641-0276
linkedin.com/in/zach-guan | github.com/zach96guan

Education

Carnegie Mellon University Pittsburgh, PA
M.S. in Artificial Intelligence and Innovation, School of Computer Science, GPA: 3.83/4.0 *May 2020*

Zhejiang University Hangzhou, China
B.E. in Information Engineering, School of Information and Electronic Engineering, GPA: 3.76/4.0 *Jun. 2018*

Skills

Coursework: Distributed Systems, Practical Data Science, Scalable Machine Learning, Natural Language Processing(TA), Deep Learning, Machine Learning, Text Mining, AI Engineering, Intro to Computer Systems, Database Systems.
Languages: Python(various data science libraries), Java, C/C++, JavaScript, SQL, Scala, Assembly.
Tools: Linux, Git, AWS, K8s, MySQL, MongoDB, Hive, MapReduce, HDFS, Spark, TensorFlow, PyTorch, MATLAB.

Experience

Walmart Labs Sunnyvale, CA
Machine Learning Engineer Intern, Personalization Team *May 2019 – Aug. 2019*

- Designed a personalized recommender for implicit data, and developed latent factor model based collaborative filtering, including pointwise/pairwise non-negative matrix factorization and factorization machine methods.
- Aggregated million-scale Walmart.com transaction data by Hive, extracted and preprocessed raw data with PySpark.
- Visualized insights from sampled data by Tableau and utilized t-SNE and K-means to analyze latent representations.
- Applied ranking-based information retrieval metrics such as Hit Ratio, Mean Average Precision and Normalized Discounted Cumulative Gain in Python for graded relevance evaluation based on the customers preference.

University of Michigan Ann Arbor, MI
Research Assistant, Electronic Engineering and Computer Science Dept. *Jul. 2017 – Oct. 2017*

- Developed a prototype system integrated with wearable sensors and remote monitor display for U-M athletic swimmers.
- Assembled varied sensors, Bluetooth module and micro-controller on Arduino by C++ to collect real-time detected biometric data, and utilized Eagle to redesign schematics and circuit layouts to improve communication efficiency.
- Designed Android App to receive transmitted data and multi-thread programs to establish synchronous connections.

Nokia Siemens Networks Co., Ltd. Hangzhou, China
Software Engineer Intern *Aug. 2016 – Nov. 2016*

- Implemented image processing, segmentation and feature extraction by OpenCV in Python, and applied CNNs based model to build a human face recognition system in TensorFlow.
- Embedded the framework in Raspberry Pi with JSON-based front-end display interface to complete a human-machine interactive demo, which is capable of pushing customized information of recognized targets.
- Created an open source platform for mobile development on IoT and succeeded in application for Utility Model Patent.

Selected Projects

Speech-to-Text Transcription System *Feb. 2019 – Apr. 2019*

- Developed Attention-based deep neural networks in PyTorch, with a combination of RNNs, CNNs and Dense Nets to design an End-to-End system for speech utterance to corresponding text transcription.
- Evaluated on WSJ dataset with the Levenshtein distance and perplexity, and ranked top 5% on Kaggle leaderboard.

Question Generation & Answering System *Sept. 2018 – Nov. 2018*

- Designed a QA system to generate/answer questions from the source of Wikipedia articles based on linguistic features.
- Applied sentence segmentation, Porter stemming and tokenization by spaCy, parsed POS taggers and NER recognition by dependency tree, transformed declarative sentences to ranked binary/WH questions.
- Utilized scikit-learn to preprocess and vectorize input sentences, identified question types and coreference resolution, calculated Jaccard similarity and TF-IDF relevance to select top candidates for answering.

Intro to Computer Systems Labs *May 2018 – Aug. 2018*

- Wrote a general-purpose cache simulator, and optimized a small matrix transpose kernel to minimize number of misses.
- Realized a simple Unix shell that supports basic commands of job control and implemented I/O redirection.
- Implemented a dynamic memory allocator by designing segregated free list and manipulating bits in header/footer with LIFO and best-fit policy to improve both space utilization and throughput.
- Built a concurrent HTTP proxy that caches recently accessed web objects to handle multiple client requests in parallel.