

Ziheng Ni

☎ (412)577-8143 • ✉ ziheng.ni@gmail.com • 🌐 www.linkedin.com/in/zihengni

Experience

Amazon, Alexa Communications

Sunnyvale, CA

Software Development Engineer II

March 2020 - Present

- Built features for multiple calling stacks across the entire line of Echo & Alexa devices on both device and cloud sides
 - Designed a middleware service in C++ to process calling-related messages from Alexa Voice Service on devices
 - Architected a generic set of destination-independent libraries in Java to record and publish calling metrics on devices
 - Implemented an Android configuration management service fetching calling-related data from Alexa Cloud to devices
- Constructed services to enable Zoom Meetings across multiple Echo & Alexa devices on both device and cloud sides
 - Delivered a multi-threading Android service with coroutines in Kotlin to empower meeting sessions on devices
 - Implemented a serverless service using AWS Lambda to handle requests and manage sessions for meetings at runtime
 - Introduced a shortcut in messaging service to enable direct access to Zoom Meetings on the Alexa comms landing page
- Created features for Timers, Alarms, and Reminders (TAR) on Linux-based Echo & Alexa devices

Qualcomm, Research & Development

Beijing, China

Interim Engineering Intern

June 2019 - August 2019

- Utilized deep learning methods to recognize human activities with data collected by sensors in mobile devices
 - Designed and tuned deep neural networks (DNN) using TensorFlow based on accelerometer and gyroscope features
 - Reproduced a differentiable neural architecture search (NAS) method on the model by sampling the search space
 - Optimized the relaxed loss function by stochastic gradient descent and compared it with the hand-tuned one
 - Converted the models using TensorFlow Lite converter and deployed on-device with TensorFlow Lite interpreter

Projects

ETL Pipeline

- Implemented a Hadoop Streaming Job Flow analyzing Wikipedia Dataset with Amazon Elastic MapReduce (EMR)
- Implemented the PageRank algorithm based on Twitter Social Graph Dataset with Spark
- Utilized Samza APIs to set up a driver-matching service with streams of NYC taxi GPS data generated by Kafka

Text-based Multiple-model Search Engine

- Implemented a text-based large-scale search engine with Lucene API on a corpus of 500,000+ documents
 - Incorporated multiple retrieval algorithms including Unranked and Ranked Boolean, Okapi BM25, and Indri
 - Accomplished pseudo-relevance feedback using original queries expansion to improve the retrieval performance
 - Developed a learning to rank model by training an SVM classifier from human selected relevant features
 - Evaluated models by adjusting parameter values, analyzing trends, and testing performance on trec_eval service

Yixiaoban Website

- Developed an interactive webpage incorporating functions of making friends, hosting parties, and conducting events
 - Visualized nearby posts in map by integrating GeoLocation API and Google Maps API
 - Deployed website's back-end with Spring Boot as the server framework providing RESTful API
 - Built relational MySQL databases hosted on Amazon AWS to save restaurants and events data

Education

Carnegie Mellon University - School of Computer Science

Pittsburgh, PA

M.S. in Information Technology | Cumulative GPA: 3.62/4.00

September 2018 - December 2019

Southern University of Science and Technology

Shenzhen, China

B.E. in Computer Science and Engineering | Cumulative GPA: 3.63/4.00

September 2014 - June 2018

University of California, Berkeley

Berkeley, CA

Summer Exchange

June 2017 - August 2017

Skills

Programming Languages: Java, Kotlin, Python, C++, JavaScript

Platforms & Frameworks: Android, AWS, Hadoop, Spark, Kafka, Spring, TensorFlow