Joseph Chuang

jcc436@cornell.edu (607)-261-2174 jochuang.com github.com/josephch405 in/jochuang

Education

Cornell University – Ithaca, NY **Fall 2016 – Spring 2019**

- Bachelors of Science in Computer Science, College of Engineering - GPA: 3.999
- Relevant coursework: Algorithms, Machine Learning and Intelligent Systems, Computer Systems, Functional Programming, Advanced Topics in ML, Advanced ML Systems, Operating Systems, Stochastic Processes

Experience

Research Intern – Taiwan AILabs **Dec 2018 – Jan 2019**

- Adapted VoiceFilter technique for denoising task on audio streams with Pytorch CNN/LSTM operating on spectrograms; containerized + deployed with Docker for K8 deployment and integration with Chinese ASR pipeline

Software Engineering Intern – Asana, New York City **Sep – Nov 2018**

- Drove data model migration for core product by redesigning Typescript client datastore + rendering, then building distributed migration in Node.js, enabling project type conversions between Boards and Lists
- Worked on ETL migration from Redshift to Spark by working with Airflow pipelines and Scala data processors

Software Engineering Intern – Google, Sunnyvale, Google Cloud Shell Tutorials **May – Aug 2018**

- Implemented metrics pipeline for sending usage stats to external authors using JS and TypeScript
- Developed a Google Docs based WYSIWYG authoring workflow by constructing a Markdown importing tool with C++, a previewing tool for walkthroughs on Golang App Engine, and a Docs plugin in Apps Script hosting the walkthrough runtime; increased the count of official Cloud Walkthroughs by 83% with new workflow
- Designed and implemented tutorial recommendation service on internal Java serverless platform, first by processing corpus and keywords via TF-IDF and NMF in Python, then serving based on author-provided tags

Machine Learning Teaching Assistant **Jan – May 2018**

Algorithms Teaching Assistant **Aug – Dec 2017**

Kaggle Subteam Lead – Cornell Data Science Project Team (CDS) **Aug – Dec 2017**

- Directed team effort to compete in the Zillow Dataset challenge, using Pandas, Numpy, XGBoost, LGBM, Catboost and general ensembling methods to generate a top ~5% submission

Researcher – Team Ursa **Jul – Sep 2017**

- Researched solution for audio alignment tasks, using FFMPEG, Numpy and Scikit to downsample, preprocess, and predict offsets between audio recordings of the same source event, deployed as AWS Lambda serverless function

Projects

Quiklogs – Side Project **Jun – Current**

- Currently designing intelligent online journal using Next.js, Firebase and ElasticSearch

Paranet – Academic Research Project **Oct – Dec 2017**

- Designed and evaluated novel deep architecture, an extension from DenseNet with novel cascaded skip connections allowing early inference at varying stages for practicality
- Implemented and ran experiments in Pytorch involving logit matching, alternate parameter sharing architectures

Engag-Ed – Big Red Hacks Cornell - Microsoft Grand Prize + Best UI **Sep 2017**

- Directed 4-person team in building an intelligent classroom suite featuring automatic attendance tracking, classroom engagement analysis, and instant polling; implemented via RESTful Express server, Microsoft Cognitive Services Emotion/Face APIs, and backed by persistent MongoDB backend

Machine Learning Playground – ml-playground.com **Jun – Aug 2017**

- Designed a visual playground using React and HTML5 featured on front page of Product Hunt
- Highlighted differences between ML algorithms through end-device data generation, training and inference
- Implemented library of trainable ML algorithms (KNN, decision trees, neural networks etc) with JS, math.js

SkinnerDB – Academic Research Project **Jan – May 2018**

- Extended reinforcement-learning based database system by parallelizing query execution and join order exploration across 30 processors with Java stream-based lambda execution
- Provided up to 4x speed increases on select TPC-H benchmarks against original serial executor

Skills

- Data Science and Machine Learning: Python, Numpy, Pandas, Scikit-Learn, Pytorch, Tensorflow, Keras
- Software Development: Java, C, C++, Golang, Protobuf
- Web Development: JavaScript, TypeScript, Node.js, React, Redux, Firebase, Express.js, Google Cloud Platform