Guangbo Yu

guangboyu@alumni.usc.edu · 213-422-4717 · GuangboYu · Los Angeles

Education

University of Southern California (USC), Los Angeles 2015 – 2017

M.S. in Computer Science (GPA: 3.2/4.0)
Relevant Coursework: Machine Learning, Data Mining, Analysis of Algorithms, Applied Probability.

University of Electronic Science and Technology of China, Chengdu, China 2011 – 2015

B.E. in Software Engeering (GPA: 3.7/4.0, Top 5%)

Projects

Byte Cup Machine Learning Competition. Aug. 2016 – Dec. 2016

- An international Machine Learning Competition of match community raised questions with domain experts held by IEEE-China and Toutiao
- Preprocessed data and generated features, implemented and applied **TF-IDF** on 100 GB data using **Pythıon** in **Google Cloud Linux Environment**
- Explored algorithms and implemented **Factorization Machine** and **Stacking** to increase the NDCG score by 9 %(**Linear Regression, SVM, Keras Neural Network, XGBoost, Collaborative Filtering**)

Tap News, Real-Time News Scraping and Recommendation System Apr. 2018 – May. 2018

- Implemented a data pipeline which monitors, scrapes and dedupes latest news (**MongoDB, Redis, RabbitMQ**)
- Built a single-page web application for users to browse news (**React, Node.js, RPC, SOA, JWT**)
- Implemented a click event log processor which collects users' click logs, then updates a news preference model for each user (**NLP, TF-IDF**)
- Designed and built an offline training pipeline for news topic modeling (**Tensorflow, DNN, NLP**)
- Deployed an online classifying service for news topic modeling using the trained model

NYC Building Energy Star Score Prediction Jan. 2018 – Feb. 2018

- Built a data pipeline to clean, analysis predict the energy star score via **Pandas**, **NumPy** and **Matplotlib**
- Built a **Gradient Boosting** model and tuned hyperparameters to increased the MAE by 35% compared to Linear Regression

Movie Recommender System Aug. 2017 – Sep. 2017

- Built a recommender system to recommend movies based on an adapted Netflix user dataset via **Hadoop**
- Computed top 5 recommendations for each user (**Item-based Collaborative Filtering, JAVA**)
- Processed 1GB data by Hadoop MapReduce jobs in the environment set up by **Docker**

Weenix OS Jan. 2016 – May. 2016

- Built a Mini Unix-like OS kernel written in **C** in Linux Environment
- Implemented key components including process management, drivers for terminals and hard disks, VFS, and page-based virtual memory
- Implemented system calls including fork, waitpid, execve, open, read/write, mmap, sbrk, etc

Skills

- Programming Languages: Python, Java, Javascript, C, Scala.
- Data Machine Learning: Hadoop, Spark, Data Mining, Machine Learning, Numpy, Pandas, Matplotlib.

EXPERIENCE

Java Software Developer Intern, Luzhou Hospital Jan. 2015 – May. 2015

Responsible for developing backend API of an E-commerce System via **SSM** framework