

Education

- **University of California at Los Angeles (UCLA)** Los Angeles, United State
M.S in Computer Science. 2017 - 2019 (expected)
 - Courses: Web Information Management, Network Protocol and Systems Software Design for Wireless.
- **Shanghai Jiao Tong University (SJTU)** Shanghai, China
B.S in Computer Science and Technology, GPA: 3.7/4.0, Outstanding Graduate of SJTU. 2013 - 2017
 - Key Courses: Massive Data Processing, Algorithm and Complexity, Software Engineering, The Principle of Database System, Computer Network, Operating System, Internet-based Information Extraction Technologies.

Projects

- **Hybrid-Based Academic Recommender System (Java, Python, SQL)** Aug. 2016 - Jun. 2017
 - Designed hybrid-based recommendation model for suggesting papers to users through a large scale academic articles.
 - Trained LSI model for 6 million research literatures to generate similar papers based on abstract in Python, and combined paper citation network factor using collaborative filtering on the Hadoop platform in Java.
 - Worked as the team leader of four and displayed final recommendation results on website *acemap.sjtu.edu.cn*.
- **Sentimental Analysis Based on Naïve Bayes Model (Python)** Nov. 2016 - Dec. 2016
 - Developed an enhanced Naïve Bayes Model including Laplacian smoothing, negation handling, feature pruning for classifying positive and negative attitude of writers using Python.
 - Trained this model on 25,000 IMDb reviews set with the precision of 85.3% on test set (1,000 reviews).
- **Gaussian Process-Based Radio Map Reconstruction (Matlab)** Mar. 2016 - Aug. 2016
 - Proposed a novel model analyzing the spatial and temporal characteristic of radio signal.
 - Applied machine learning technology, Gaussian Process to regress the spatio-temporal model.
 - Solved the device variance problem by adding dynamic calibrate phase between offline training phase and online localizing phase. Developed transfer function using Gaussian Process Regression in Matlab.
- **Large Scale Patents Classification (Python)** Apr. 2016 - May. 2016
 - Classified large-scale patents with over 110,000 sample sizes using supervised learning model.
 - Utilized different methods to classify, including SVM and Max-Min Algorithm (both Random-based and Prior Probability-based), and obtained the highest accuracy of 96.81%.
 - Utilized MPI (Message Passing Interface) to achieve parallelism and better performance.
- **Efficient Parallel String Searching Algorithm (Python, C, CUDA)** Mar. 2016 - Apr. 2016
 - Implemented the multiprocessing programs for string searching algorithm using OpenMP and CUDA via GPU, based-on Naïve algorithm, BMH (Boyer-Moore-Horspool) algorithm and KMP (Knuth-Morris-Pratt) algorithm.
 - Proposed a novel method based-on BMH algorithm for string searching which is more suitable for parallel.
- **Tiger Compiler Implementation in Java (25k lines)** Oct. 2015 - Nov. 2015
 - Implemented a Tiger Compiler, including lexical and syntactics specification, semantic analysis, control flow and data flow analysis.
- **Indoor Localization Android App Development (Java)** Apr. 2015 - Jun. 2015
 - Proposed an indoor localization algorithm combing the Bluetooth-based method and pedometer.
 - Developed an usable indoor localization Android App, including Bluetooth data transfer, steps analyzing function and real-time localization display.

Internship

- **Firstbrave Information Technologies Inc.** Shanghai, China
Data Engineer Intern Mar. 2017 - Jun. 2017
 - Predicted whether the coming movies are worth purchase based on a data regression model using information such as film stars and directors' previous works crawling from IMDb and Douban Movie (Chinese IMDb) in Python.
 - Crawled the Weibo (Chinese Twitter) information of film stars using the random walk approach based on their social network through "follow" information in Python.

Skills

- **Languages:** Java, C/C++, Python, SQL, HTML, CSS, JavaScript, PHP.
- **Tools:** Hadoop, Android Studio, Git, MATLAB, ElasticSearch, Mathematica, L^AT_EX.