CONG PENG

740 Weyburn Terrace, APT #64, LA, CA, 90024 | (424)325-8496 | pengcong@ucla.edu

SKILLS:

Programming: Java, Python, JavaScript (React Native), C++, PHP, HTML/CSS, MATLAB, SQL(Mongo/MySQL)
Language: English, Chinese(native)

EDUCATION:

University of California, Los Angeles (UCLA) Expected 2017.12
- Electrical Engineering, Signals & Systems (current GPA: 3.6/4)

Beijing University of Posts and Telecommunications (BUPT) 2012.08 - 2016.07
- Communication Engineering (major GPA : 88/100)

Instituto Superior Técnico, Lisboa, Portugal 2015.02 - 2015.06
- Electrical and Computer Engineering

EXPERIENCE:

[INTERNSHIP] Data Mining Intern, PingAn Tech, Shanghai (Python, Mongodb, Selenium) 2017.07 - present
- Designed a web crawler based on Scrapy, Python based framework for web mining. The spider can currently crawl information from Chinese main social network: (1) Weibo (2) Baidu Headlines (3) WeChat Headlines (4) Baidu Index.
- The spider dealt with requests from html content by BeautifulSoup and Xpath and smoothly handled 302 redirect issues, including logging captcha verification, gesture verification with Selenium.
- The result of (1) is fed as the input of the recommendation system for further interest classification. The results of (2), (3), (4) are part of framework of Big Data platform.

[COM SCI 130] Software Engineering, iOS React-Native Mobile Application Development – Front End (JavaScript, React Native) *UCLA* 2017.04 - 2017.06
- Modified on the existing user interfaces and designed some styled and fitted interfaces for the Mobile App.
- Interacted with back end to fetch the data from the database and display the information requested by the users.
- Implemented several useful components (buttons, bars, links) to let users better interact with the App.

[COM SCI 230] Software Engineering, Dynamic Universal Software System Visualization Kit – DUCK (C++, Python, & JavaScript) *UCLA* 2017.04 - 2017.06
- Static analyzer: implemented a parser to analyze the codebase into intermediate representation module (JSON format). (C++)
- Intermediate Representation Module: designed this module to take the input from the parser, upload the JSON format result to Django server and feed it to the Attribute Extractor. PageRank is also applied to provide information based on usage dependency. (Python)
- Attribute Extractor: built this query-able module to extract the desired visualization metrics (total lines of code, number of methods and number of fields) to the plotter module. This module provides search option to quickly locate target function.
- Plotter: Implemented 5 different plots (Hierarchy Tree, Sunburst, Edge Bundling, Circle Packing and PageRank) to visualize the software infrastructure with D3.js. (JavaScript)
- Survey: Designed a survey to investigate the influence on software developers to better understanding the codebase with the help of DUCK.

[EL ENGR 219] Large-Scale Data Mining: Models and Algorithms (Python & MATLAB) 2017.01 - 2017.03
- Regression Analysis: Basic implementation of Regression Models (Linear Regression, Ridge Regression, Logistic Regression, Polynomial Regression, etc.) on network backup and housing dataset, along with basic techniques to handle over-fitting by cross validation and different regularization methods.
- Classification & Clustering Analysis: Implemented text data modeling strategies and different feature extraction strategies (TFxIDF, LSI, PCA and etc.) on '20 *Newsgroups'* dataset with various learning algorithms (K-means, soft margin SVM, Naïve Bayes, logistic regression and etc.). (Both two-class and multiclass Classification)
- Collaborative Filtering: Implemented collaborative filtering with "Alternating Least Squares" to build a recommendation system on the *'MovieLens'* dataset with a volume of 100,000 movie rating data. (Matrix Factorization Toolbox in MATLAB & NMF in Python)

[EL ENGR 210A] Adaptation & Learning (MATLAB) *UCLA* 2017.01 - 2017.03
- Designed a generalized CNN structure, implemented feed forward and back propagation algorithms for entire CNN structure, also dealt with problems concerning padding, partitioning, pooling, and permutation.
- Proposed a strategy on memory saving and computation efficiency during the training phase.
- Trained CNN with MNIST dataset and ImageNet database with 60000 28*28 digit images in ten classes and 1000 256*256 animal images in four classes respectively. The test results on accuracy end at 94.11% and 45.24% for each test cases.

[Undergraduate Dissertation] Advanced Algorithm Analysis on Wi-Fi Locating Based on Automatic Path Tracking (Java, Python & SQL) *Tsinghua University* 2015.11 - 2016.05

- Implemented K-means clustering algorithm to learn features of actual human trajectories.
- Summarized 3 important patterns along the human trajectories, set a criteria for further evaluation in localization and revised the localization trajectory with given pattern rules.
- Improved localization accuracy according to a well-defined localization algorithm by overall 10%.

[COM SCI 143] Relational Database Management Design (C++, SQL, PHP & HTML) *UCLA* 2016.9 - 2016.12

- Design several required SQL queries on given Movie Database with desire speed and efficiency.
- Built an open-ended Movie Database system allowing the user to search for the information of Movies and Actors through a Web interface. (SQL, Web application)
- Bruinbase Design, implemented B+ tree index on Bruinbase which can efficiently retrieve the information from the database. (System Design, RDBM)

Learning Bayesian Networks (Java) *Instituto Superior Tecnico* 2015.02 - 2015.06

- Implemented Greedy Hill Climbing algorithm and random restarts in Java to learn the experimental data from the training set.
- Discovered the best structures supported by 2 different scoring algorithms, log-likelihood and minimum description length.
- Computed the parameters of the structure and predicted exact value successfully.

Weather parameters monitoring and sharing based on WeChat Common Platform (PHP) *BUPT* 2014.03 - 2014.06

- Obtained data from the database collected by the sensors.
- Built up one WeChat platform which can interact with the subscribers by informing weather parameters collected by sensors, and displayed requested information on the screen in user-friendly way.