Yi ZHOU

yizhou3@andrew.cmu.edu
(412) 773-3378

Education

Carnegie Mellon University (CMU) - School of Computer Science Pittsburgh, PA
Master of Science in Artificial Intelligence and Innovation (Mid-Semester GPA: 4.22/4.30) May 2020
Courses: Machine Learning (in progress), Natural Language Processing (in progress), Computer Systems

The Hong Kong Polytechnic University (PolyU) Hong Kong
Bachelor of Science in Computing, Minor - Applied Mathematics (GPA: 3.85/4.00) Jun. 2018
Courses: Machine Learning, Database Systems, Data Structures, Software Engineering, OOP

Work Experience

XtalPi Inc. Beijing, China
Algorithm Engineer Intern May. 2018 – Aug. 2018

- Improved a molecule vectorization algorithm to be three times faster in Numpy by balancing multiprocessing load with MPI4py and reducing unnecessary computation and visualized the vectorization results using Python with Plotly.

PolyU, Internet Infrastructure and Security Research Laboratory Hong Kong
Research Assistant Jun. 2017 – Aug. 2017

- Collected wireless data using MQTT protocol with Raspberry Pi's deployed across the campus.
- Developed an anomaly detector that raises alerts based on round-trip time and packet loss rate with Python.
- Used Python with Matplotlib and Pandas to preprocess and analyze the measurement results across the campus.
- Determined the major factors that correlate with the performance of the wireless network.

Selected Projects

Question Generation and Answering System, Course Group Project, CMU Sep. 2018 – Dec. 2018

- Designed and implemented a system to ask and answer intelligent questions based on Wikipedia articles.
- Used Stanford Parser and Named Entity Recognizer, NLTK Tokenizer, and WordNet Lemmatizer to retrieve information, analyze syntactic structure and semantics, and manipulate English text.

Implementation of Machine Learning Algorithms, Course Assignments, CMU Sep. 2018 – Dec. 2018

- Implemented and conducted experiments on multiple machine learning models including Decision Trees, Logistic Regression, Neural Networks, Hidden Markov Models, Reinforcement Learning (Q-Learning), and K-Means with Numpy.

System Programming (C, Linux), Course Assignments, CMU Jun. 2018 – Aug. 2018

- Implemented `malloc`, `free`, and `realloc` in C by using segregated lists and clustering tiny memory blocks and ranked 4^{th} in memory utilization and 10^{th} in throughput in the class of 230.
- Implemented a concurrent caching Web proxy by multi-threading with semaphores in C and tested on real web pages.
- Wrote a cache simulator and optimized a matrix transpose kernel to minimize the number of misses on a simulated cache.

Data Analysis on Campus Wi-Fi, Final Year Project, PolyU Sep. 2017 – May. 2018

- Clustered Wi-Fi latency by drawing Q-Q plots and conducting kernel density estimation in Python (Jupyter).
- Trained a logistic regression model with cross-validation to predict high latency with 69% precision and 70% recall.
- Retrieved data from Cassandra using Spark with Scala and from MongoDB with Python.
- Presented and explained analysis results to the staff of Information Technology Services Office (ITS) at PolyU.

Image Classification using Brainwave Data, Course Group Project, PolyU Oct. 2016 – Dec. 2016

- Designed an SVM majority vote algorithm to differentiate brainwaves of people viewing faces and landscapes.
- Implemented the algorithm using Matlab and achieved 1st place in differentiation rate in the class.

Skills

Programming Languages: Python (proficient), Java, C, Scala, Bash, SQL, Javascript, PHP, C++, Matlab
Technologies: Numpy, sklearn, Pandas, Matplotlib, MongoDB, MySQL, Linux, Git, Spark, Docker