KAI WU

MS Computer Science, Expected Graduation: Dec 2020

@ kaiwu@ucdavis.edu ☎ (626) 475-1645 ◉ California, USA in https://www.linkedin.com/in/kai-wu/

EDUCATION

Northeastern University
MS Computer Science

📅 Sept 2018 – Dec 2020   ◉ San Jose, CA

GPA 3.76

Specialization: Data Science and Artificial Intelligence

University of California, Davis
BS Neurobiology, Physiology, Behavior

📅 Sept 2015 – Dec 2017   ◉ Davis, CA

GPA 3.14

EXPERIENCE

Informatics Data Specialist Co-op
Genentech

📅 Jan 2020 – June 2020   ◉ South San Francisco, CA

- Supported data management efforts in the Developmental Sciences Informatics department.
- Improved data conformance checking scripts (Python) for major vendors responsible for roughly 50% of monthly data transfers.
- Automated previously manual checks and processes involved with next-generation sequencing and pathology data transfers by integrating more functionality in scripts.
- Laid foundation for easily extending scripts to encompass data transfers from vendors that previously could not be checked by adding Google Sheets integration, accepting JSON file input, making existing scripts more modular, and providing clear documentation.

Graduate Teaching Assistant
Northeastern University

📅 May 2019 – Dec 2019   ◉ San Jose, CA

- Communicated course material in different perspectives to help students grasp abstract coding concepts during in-person office hours for Computer Systems and Algorithms courses.
- Returned personalized feedback to students by manually reviewing code for errors when test cases failed.
- Resolved academic dishonesty issues in a fair, but firm manner.

ER Medical Scribe
CEP America

📅 Mar 2017 – Aug 2017   ◉ Sacramento, CA

- Documented ER patient admissions and discharges using an electronic medical record in a fast-paced, dynamic environment.
- Tracked roughly 10-15 patient charts per shift ensuring patients could be discharged in timely fashion.
- Concisely summarized patient encounters in writing – ensuring pertinent information was captured and presented objectively.

SKILLS

Languages

Python, Java, C/C++, Bash, R

Tools

Scikit-learn/pandas, Jupyter, SQL, Git/GitHub

Knowledge

Data science, Machine learning, Database management

PROJECTS

Home Credit Default Risk Prediction
Northeastern University

📅 Oct 2019

- Based on 2018 competition from Kaggle to predict whether an applicant is capable of repaying a loan.
- Performed exploratory data analysis using pandas and numpy libraries in Python.
- Transformed data complex SQL queries and techniques such as one-hot, non-linear transformations, and normalization (z-score and L2).
- Trained a linear regression model using gradient descent in Ruby for predictions, coded without reliance on existing libraries.

Identifying Personal Attacks in Wikipedia Comments
Northeastern University

📅 Nov 2019

- Developed a text classification model using natural language processing that classified Wikipedia comments as either a personal attack or not.
- Performed text cleanup, data extraction, and feature selection using scikit-learn in a Jupyter notebook.

Order Manager Database
Northeastern University

📅 Jun 2019

- Designed an SQL database that could serve as the backend for an online store using Apache Derby in a team of two.
- Encouraged teamwork by making sure naming conventions and git commits were consistent for easier collaboration, and communicating often in-person.
- Developed the database's API in Java for interactions with the database that allowed looking up customer orders, checking if an item is in stock, etc.