

Leo Ding

626-228-7378 | leo.ld.ding@gmail.com | leoding.com | linkedin.com/in/leo-ding | github.com/leoldding

EDUCATION

New York University

Sep 2020 – May 2023

B.A. in Computer Science and Data Science, GPA: 3.90

New York City, NY

- Honors: Magna Cum Laude, Phi Beta Kappa, University Honors Scholar
- Coursework: Deep Learning, Machine Learning, Parallel Computing, Data Structures and Algorithms, Processing Big Data for Analytics Applications, Data Management and Analysis, Probability and Statistics

TECHNICAL SKILLS

Languages: Go, Python, Java, SQL, JavaScript, HTML, CSS, C, R

Technologies: Docker, AWS, TensorFlow, PyTorch, PostgreSQL, OpenMP, MPI, React, Nginx
Tailwind, NumPy, Pandas, SciPy, matplotlib, scikit-learn, Apache Hadoop, Apache Hive,
Apache Impala, Apache Spark

PROJECTS

Coffee Shop Status V3 | *Go, PostgreSQL, Docker, AWS, React, Tailwind* | [Code](#) | [Website](#) Jul 2023 – Oct 2023

- Developed a status display web application with a Go backend and React frontend hosted on AWS
- Implemented user password authentication for admin privileges in order to change displayed status
- Designed to show acquaintances if I am at a coffee shop I frequent

Chat Room App V2 | *Go, Docker, AWS, React, Tailwind* | [Code](#) | [Website](#)

Jul 2023 – Oct 2023

- Created a chat room application with a Go backend and React frontend hosted on AWS
- Utilized WebSockets in order to enable real-time communication between the frontend and backend
- Incorporated a publisher-subscriber mechanism to separate different chat rooms and their respective text

News and Social Media Sentiment Analysis | *Python, Pandas, scikit-learn* | [Code](#)

Nov 2022 – Dec 2022

- Integrated a proportional frequency scoring system to determine sentiments associated with different words
- Tuned model hyperparameters with grid search cross validation with 10 folds using 1.6 million data points
- Yielded a 20% and 32% decrease in error for the news and social media data respectively compared to baseline

Soccer Analytics | *Java, SQL, Spark, Hadoop, Hive, Impala* | [Code](#)

April 2022 – May 2022

- Analyzed 250k data points related to the English Premier League, Spanish La Liga, and Italian Serie A
- Utilized shell scripting to automate each step of the data analysis process
- Trained Logistic Regression models to classify winners and losers to determine factors that contribute to winning
- Models performed with an average Area Under Receiver Operating Characteristic Curve score of 0.81

EXPERIENCE

Lead Data Scientist

Aug 2020 – Dec 2020

NASA SPOCS Initiative

Remote

- Collaborated with university students from USC, UCLA, and UC Berkeley
- Determined experimental data collection logistics and post-experiment data analysis processes
- Proposed experiment regarding virulence and antibiotic resistance in *E. coli* in microgravity environments

Web Development Intern

June 2019 – July 2019

California State University, Los Angeles

Los Angeles, CA

- Developed a professor-student communication platform using HTML, CSS, JavaScript, SQL, and PHP
- Utilized PHP authentication and parameterized SQL queries to increase website security

Human Resources Data Analyst Intern

July 2019

BMW Brilliance Automotive

Beijing, China

- Screened over 1000 applicants who were applying to three separate company positions
- Analysed data regarding accepted and rejected applicants to find trends in applications
- Designed a slide deck and presented data visualizations and statistics to the human resources manager