

# Leo Ding

626-228-7378 | [leo.ld.ding@gmail.com](mailto:leo.ld.ding@gmail.com) | [leoding.com](https://leoding.com) | [linkedin.com/in/leo-ding](https://linkedin.com/in/leo-ding) | [github.com/leoldding](https://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: Processing Big Data for Analytics Applications, Data Management and Analysis, Data Structures and Algorithms, Deep Learning, Machine Learning, Parallel Computing, Probability and Statistics

## TECHNICAL SKILLS AND CERTIFICATIONS

**Linux Foundation**, Certified Kubernetes Administrator

Jul 2024

**AWS**, Certified Developer - Associate

Apr 2024

**AWS**, Certified Cloud Practitioner

Feb 2024

**Languages:** Go (Golang), Python, Java, JavaScript, TypeScript, SQL, HTML, CSS

**Technologies:** AWS, Docker, PostgreSQL, Spring Boot, ReactJS, TailwindCSS, Nginx, NumPy, Pandas, SciPy, matplotlib, scikit-learn, TensorFlow, PyTorch

## PROJECTS

**Application Tracker CLI Tool** | *Go* | [Code](#)

Mar 2024

- Developed a command-line interface (CLI) tool in Go to streamline job application management
- Automated scraping and parsing for job description data sourced from LinkedIn, Greenhouse, and Lever web pages
- Integrated Notion's API to facilitate seamless transmission of scraped information into our database system

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

Jul 2023 – Oct 2023

- Developed a chat room application featuring a Go backend and React frontend, hosted on AWS infrastructure
- Created a CI/CD pipeline using Github Actions / Workflows and AWS Elastic Beanstalk with a modified AWS CloudFormation stack to automatically update and deploy AWS EC2 instances
- Implemented WebSockets to enable real-time communication between the frontend and backend components
- Incorporated a publisher-subscriber mechanism to deliver relevant texts to their respective chat rooms

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

Nov 2022 – Dec 2022

- Designed a proportional frequency scoring system to assess sentiments linked to various words
- Tuned model hyperparameters using grid search cross-validation leveraging 1.6 million data points
- Yielded a 20% reduction in error for news data and a 32% decrease for social media data compared to the baseline

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

Apr 2022 – May 2022

- Examined 250,000 data points associated with the English Premier League, Spanish La Liga, and Italian Serie A
- Automated the entire data analysis process using shell scripting techniques to improve development efficiency
- Trained Logistic Regression models for classifying winners and losers to identify contributing factors to winning
- Achieved noteworthy model performance with an average Area Under Receiver Operating Characteristic Curve (AUC-ROC) score of 0.81

## EXPERIENCE

### Data Scientist

Aug 2020 – Dec 2020

*NASA SPOCS Initiative*

*Remote*

- Formulated strategies for experimental data collection and devised procedures for post-experiment data analysis
- Collaborated closely with the engineering team to strategize and refine logistics for data collection processes
- Proposed an experiment exploring virulence and antibiotic resistance in E. coli within microgravity environments

### Web Development Intern

Jun 2019 – Jul 2019

*California State University, Los Angeles*

*Los Angeles, CA*

- Created a communication platform for professors and students using HTML, CSS, JavaScript, SQL, and PHP
- Strengthened website security by implementing PHP authentication and parameterized SQL queries