Leo Ding

NEW YORK CITY, NY | (626) 228-7378 | leo.ld.ding@gmail.com
[LinkedIn](#) | [GitHub](#) | [Personal Website](#)

Education
September 2020 - Present (Expected May 2023)
New York University – B.A. in Computer Science and Data Science (GPA: 3.90)
Relevant Coursework: Processing Big Data for Analytics Applications, Data Management and Analysis, Probability and Statistics, Basic Algorithms

Projects
- [Coffee Shop](#) (Docker, Golang, SQL, HTML, CSS, JavaScript)
  - **Golang** handled **password** and **session authentication**, HTTP request procensing, and **PostgreSQL** manipulation
  - **Docker** used to containerize components and allow for communication
  - Hosted on AWS (**Elastic Beanstalk** and **Relational Database Service**)
  - **JavaScript** used to send HTTP requests and update frontend elements
- [Coin Flip](#) (Docker, Golang, SQL)
  - Used **Docker Compose** to link frontend, backend, and database containers
  - **Golang** backend handled HTTP requests and **PostgreSQL** communication
- [Soccer Analytics](#) (Java, SQL, Scala)
  - Analyzed a decade worth of data from three top European leagues
  - Classified winners and losers using **Logistic Regression**
  - Grouped potential skill difference using **K-Means**
- [Spam Classifier](#) (Python)
  - Determined word importance in SMS messages using **TfidfVectorizer**
  - Classified spam messages using a **Logistic Regression** model
- [Movie Ratings Analysis](#) (Python)
  - Reduced features for analysis using **Principal Component Analysis**
  - Evaluated relationships between movies using **Multiple Linear Regression**

Technical Skills
Fluent: Python
Proficient: R, Java, PostgreSQL
Familiar: Golang, Docker, HDFS, MapReduce, Hive, Impala, Spark, HTML, CSS, JavaScript

Experience
Data Science Lead – NASA SPOCS
August 2020 - December 2020
- Collaborated with university students from USC, UCLA, and UC Berkeley
- Team proposed an automated experiment regarding virulence and antibiotic resistance in E. coli in microgravity environments
- Determined **data collection and analysis needs** for proposed project