# Leo Ding

NEW YORK CITY, NY | (626) 228-7378 | ld2425@nyu.edu
**Personal Website** | **LinkedIn** | **GitHub**

## Education
September  2020 - Present (Expected May 2023)

**New York University** – Joint Computer Science and Data Science Degree
GPA: 3.90
Relevant Coursework: Processing Big Data for Analytics Applications, Data Management and Analysis, Probability and Statistics

## Projects
- Soccer Analytics
    - sought to find key metrics to predict winning in soccer
    - analyzed a decade worth of data from three top European leagues
    - used **Logistic Regression** to classify winners and losers
    - ran **K-Means** to determine potential difference in skill
    - found, unsurprisingly, that scoring goals is the main factor for winning
    - findings also reinforce 'home team advantage' idea
- Hidden Markov Model Part Of Speech Tagger
    - trained algorithm to assign parts of speech to words
    - made decisions based on prior and word **transition probabilities**
- Spam Classifier
    - used **TfidfVectorizer** to determine word importance in SMS messages
    - created a **Logistic Regression** model to classify spam messages
- Movie Ratings Analysis
    - applied **Principal Component Analysis** to effectively reduce features for analysis
    - used **Multiple Linear Regressions** to evaluate **relationships** between movies

## Technical Skills
- **Python**, Java, **PostgreSQL**, HDFS, MapReduce, Hive, Impala, Spark, **GitHub**, HTML, CSS, JavaScript

## Experience
Data Science Lead – NASA SPOCS
- automated experiment regarding virulence and antibiotic resistance in E. coli in microgravity environments
- **decided data collection** and analysis needs for proposed project
- collaborated with university students from USC, UCLA, and UC Berkeley

## Activities
I enjoy solving Rubik's Cubes and am a fan of Formula 1.