José San Martin

919-753-6779, hello@josesmo.com
LinkedIn Personal Website

EDUCATION

Duke University, Trinity College of Arts and Sciences (Durham, NC), May, 2020

- Junior at Duke University with intended majors in *Computer Science* and *Statistical Science*
- **Cumulative GPA:** 3.6/4.0
- **Relevant Courses:** Regression Analysis(A), Statistical Computing(A-), Computer Architecture(A), Data Structures & Algorithms(A), Elements of Machine Learning (A-), Mathematical Statistics, Bayesian Statistics,

RELEVANT EXPERIENCE

DATA SCIENCE AND OPERATIONS INTERN IBM Global Chief Data Office (New York, NY), Summer 2019

- Worked primarily under Satyajeet Raje on many aspects of his Automatic Metadata Generation (AMG) project. These aspects included handling unstructured data and parallelizing machine learning models as a resource for a service
- Developed a back-end architecture to provide Machine Learning Models as a service toclients. The architecture made use of the Bazel, Docker, and Kubernetes software.
- Other tasks included maintenance of IBM's BIGSQL database via leveraging their catalog API and handling the unstructured data that is generated from IBM's Risk Insights service

DATA SCIENCE INTERN Duke University Department of Computer Science (Durham, NC), Summer 2018

- Research under Dr. Jeffrey Forbes and Dr. Kristen Stephens-Martinez, Where I prepared a new Data Science course at Duke based on UC Berkley's DATA8. The new course acts as an introduction to the Discipline of Data Science
- Worked on a Retention Study, which consisted of an in-depth analysis of past Duke graduates to create a model of predicting a student's chance of retention and success in Computer Science at Duke
- Database design and population using SQLite3

UNDERGRADUATE TEACHER ASSISTANT: Duke University Department of Statistics (Durham, NC), 2017 – Present

- Teaching Assistant for STA 210: Regression Analysis / STA 101: Data Analysis and Statistical Inference / STA 199: Intro to Data Science
- Instruct a weekly lab section, where I assist students in learning statistical concepts, data analysis using the R software, and additional help on the course homework
- Hold dedicated office hours to assist students in coursework material and coding

SUCCESS PARTICIPANT SEO Career (New York, NY) 2019-Present

- Receive 50+ hours of individualized coaching and online instruction to achieve targeted professional developmental goals mainly around technical and soft skills training
- Mastery of fundamentals specific to Data Science in order to maximize the likelihood of a return offer at the end of the summer

SKILLS & INTEREST

PROGRAMMING & SOFTWARE: R*****, JAVA*****, Python*****, JavaScript***, C***, Node.js***, React.js***

DATA ANALYSIS: Through my coursework in my statistics and math classes and my data science internships, my experiences working with real data, and my proficiency in the programming languages R and Python, I have obtained a passion as well as valuable skills in the data science and machine learning fields. In my website you can see some of the projects I have worked on during my coursework and my internships

WEB DEVELOPMENT: Completed the Udemy Web Development Bootcamp course, and created a website for a machine learning lab to display their SVM algorithm, as well as my own personal website

SOFTWARE ENGINEERING: Through my computer science coursework and software engineering courses, I have learned a deep understanding of software design, as well the best methods to create clean, flexible, and efficient software. In my website are examples of some of the projects I have worked on.

CAMPUS & COMMUNITY INVOLVEMENT

DIRECTOR OF OPERATIONS Duke University Undergraduate Machine Learning (Durham, NC) 2017 – 2018

PEER PRECEPTOR: David M. Rubenstein Scholarship, Duke University (Durham, NC), 2017 – 2018

MATH HELP ROOM TA FOR LINEAR ALGEBRA: Duke University (Durham, NC), 2017 – 2018