VICTORIA SUAREZ

Data Scientist

CONTACT

✉ vasuarez1@gmail.com

📞 (305) 310 6991

🏠 San Francisco, CA

in linkedin.com/in/vasuarez

🐦 @Victoria_Data

EDUCATION

MASTER OF SCIENCE IN DATA SCIENCE

July 2019
University of San Francisco

BACHELOR OF SCIENCE IN BUSINESS ADMINISTRATION, INFORMATION SYSTEMS

Minor Statistics
May 2018
University of Florida

COURSE WORK

Machine Learning
Deep Learning
Distributed Computing
A/B Testing
Regression
Time Series

SKILLS

Python	Docker
Spark	Flask
SQL	Concourse CI
MongoDB	R

EXPERIENCE

DATA SCIENTIST — July 2019 - Present

Ultimate Software | San Francisco, CA

- Experimented with **Neural Networks** to model match score between candidates and job opportunities using **Pytorch**
- **Dockerized** and deployed **Flask** microservices on **PCF** to serve models to product partners
- **Mentored** and **collaborated** with engineers, data scientists, and product managers on the data science team to deliver 6+ models

DATA SCIENCE INTERN — October 2018 - July 2019

Ultimate Software | San Francisco, CA

- Researched and implemented **NLP** techniques such as **word embeddings** to detect gender bias in performance reviews using **Python**
- Built web scraper to acquire resume data using **Selenium**

DATA ANALYTICS INTERN — May 2018 - June 2018

Norwegian Cruise Line Holdings Ltd. | Miami, FL

- Developed sentiment analysis model using **R**, **Hadoop**, and **Hive**; created **Tableau** visualizations for analysis
- Coded and automated **R** script to scrape cruise itineraries which helped teams with competitive analysis
- Programmed **SQL** Server stored procedures for dynamic itinerary data manipulation

PROJECTS

PREDICTING IN-APP PURCHASES — February 2019 - March 2019

University of San Francisco, Group Project

- Built gradient boosting classifier using **LightGBM** to predict customer purchases; model performed at 97% AUC
- Placed first in class Kaggle competition from 25 groups

IMPACT OF BIKE-SHARING ON AIR QUALITY — January 2019 - February 2019

University of San Francisco, Group Project

- Predicted Air Quality Index with RMSE of 7.8 using **Random Forest**
- Assembled scalable data science framework using **AWS**, **MongoDB**, and **Spark**
- Presented research paper at Creative Activity & Research Day 2019