VICTORIA SUAREZ

Data Scientist

CONTACT

✉ vasuarez1@gmail.com

📞 (305) 310 6991

🏠 San Francisco, CA

in linkedin.com/in/vasuarez

🐦 @Victoria_Data

🗣 English (fluent)
Spanish (fluent)

EDUCATION

MASTER OF SCIENCE IN DATA SCIENCE

July 2019
University of San Francisco
GPA 3.97/4.00

BACHELOR OF SCIENCE IN BUSINESS ADMINISTRATION, INFORMATION SYSTEMS

Minor Statistics
May 2018
University of Florida
GPA 3.83/4.00

COURSE WORK

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

SKILLS

Python	Docker
Spark	Flask
SQL	Concourse CI
MongoDB	R

WORK EXPERIENCE

DATA SCIENTIST II — *July 2020 - Present*

Ultimate Kronos Group (formerly Ultimate Software) | San Francisco, CA

- Lead research initiative of 4 data scientist to develop **multi-class sentiment classifier** using **ensemble techniques**, **transfer learning**, and **recurrent neural network** (**RNN**); successfully delivered the model into production with a performance of **0.85 AUC**, a 20% increase in performance from previous classifier
- Piloted first machine learning **fairness audit** at the company for 3 of the predictive models at highest risk

DATA SCIENTIST I — *July 2019 - July 2020*

Ultimate Software | San Francisco, CA

- Experimented with **neural networks** to model match score between candidates and job opportunities using **PyTorch**, **Siamese Networks**, and **GPUs**
- **Dockerized** and deployed **Flask** microservices on **Kubernetes** to serve 6+ machine learning models to product partners
- Mentored and collaborated with engineers, data scientists, and product managers on the data science team

DATA SCIENCE INTERN — *October 2018 - July 2019*

Ultimate Software | San Francisco, CA

- Researched and implemented **Natural Language Processing** (**NLP**) techniques such as **word embeddings** to detect gender bias in performance reviews using **Python**
- Built web scraper to acquire 1000+ resume and certification data using **Selenium**

DATA ANALYTICS INTERN — *May 2018 - June 2018*

Norwegian Cruise Line Holdings Ltd. | Miami, FL

- Designed and developed Twitter **sentiment analysis** program using **R** and **Hadoop**; created **Tableau** visualizations for analysis
- Built and automated **R** script to scrape cruise itineraries from websites for competitive analysis
- Developed **SQL Server** stored procedures for dynamic itinerary data manipulation

ACADEMIC PROJECTS & PUBLICATIONS

IMPACT OF BIKE-SHARING ON AIR QUALITY — *January 2019 - August 2019*

- Predicted Air Quality Index with **RMSE of 7.8** using **Random Forest**
- Assembled scalable data science framework using **AWS**, **MongoDB**, and **Spark**
- Published paper in 2019 IEEE SmartWorld conference (https://bit.ly/324qcjk)

PREDICTING IN-APP PURCHASES — *February 2019 - March 2019*

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