

# Ricky Martin

[rickylenmartin@gmail.com](mailto:rickylenmartin@gmail.com) | (661) 808-3055 | New York, NY 10009 | [LinkedIn](#) | [GitHub](#)

## WORK EXPERIENCE

---

### MACHINE LEARNING ENGINEER

Jan 2022 - Present

Ardent MC

- Led the creation and deployment of a scanned document classifier using AWS Textract OCR tool and XGBoost classifier on heavily skewed data, achieving best F1 score of .92 after employing SMOTE techniques, built in AWS
- Built time-series forecasting model to query/monitor 5000+ login data points from Splunk, accessing security risk
- Overseeing the R&D of prediction based machine learning model to dynamically manage and allocate network resources to decrease various application processing times from an average 166 seconds to under 30 seconds
- Used SparkSQL and PySpark to explore patterns/trends of bias on 5 million+ application results within Databricks

### DATA SCIENTIST

Oct 2020 - Oct 2021

Redica Systems (Formerly Govzilla)

- Produced production-level Unsupervised DBSCAN algorithm in Python to cluster incoming FDA inspection reports, utilizing TF-IDF techniques in order to separate documents by similar subjects
- Created and presented weekly easy-to-digest customizable Machine Learning model visualizations in Tableau showcasing live data analytics and KPI visualization as requested per customer
- Improved prediction recall/precision of semi-supervised document classifier through feature engineering

### ROBOTICS RESEARCHER (Machine Learning Specialization)

May 2019 - Jul 2021

Bioinspired Robotics and Design Lab

Co-Author for Published Paper in 2020 IEEE-RAS International Conference on Soft Robotics

- Spearheaded the technical construction of 5 distinct iterations of affective touch-based artificial skins using Arduino Uno, force sensors manufactured in lab to mimic/recreate human mechanosensory neurons, and Python/C++ data pipeline to clean, transform, train, and test various Supervised Deep Learning models on live touch based categorization to achieve an average 93% accuracy across different users and skin mediums
- Devised a Random Forest parameter tuning pipeline that utilizes impactful features for 200% faster classification

### DATA SCIENCE INTERN

Oct 2019 - Oct 2020

GoSite, Inc.

- Pitched and developed production-level Python application that web-scraped new leads and scored them on likelihood of purchasing website development services using KNN model based on past conversion data

## PROJECTS

---

### DIY CHATBOT

- Created, dockerized, and deployed customizable CICD portfolio website with a chatbot built from the ground up using Python, NLTK, and Pytorch deployed through GitHub, Flask, and Heroku ([Website](#))

### ROBOT ORGANIZER

- Applied Q-Learning Reinforcement Learning techniques to allow a warehouse sorting robot arm to learn and find optimal path through an artificially created maze via UARM Python API (robot arm control API)

## TECHNICAL SKILLS

---

**Frameworks/Libraries:** NumPy, Pandas, Sklearn, Matplotlib, Seaborn, OpenCV, TensorFlow, Keras, NLTK

**Technologies:** Git, Jupyter, Docker, Databricks, Kubernetes, Spark, AWS, H2O, Tableau, Arduino, Linux-Ubuntu

## EDUCATION

---

University of California, San Diego (UCSD)

Graduated July 2021

B.S. in Cognitive Science with a specialization in Machine Learning