# Ricky Martin

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

## WORK EXPERIENCE

### MACHINE LEARNING ENGINEER                                             Jan 2022 - Present
Ardent MC
- Led the creation and deployment of a scanned document classifier using XGBoost classifier on heavily skewed data, achieving best PRAUC score of .91 after employing SMOTE techniques. Built in AWS using Sagemaker
- Oversaw R&D of time-series models to dynamically monitor network resource usage of 50+ microservices predicting anomalies prior to occurrence, resulting in 90% reduction in time taken to find and address issues
- Developed GraphQL and Python ETL Pipeline to gather weekly performance metrics of 3000+ internal repos
- Used SparkSQL and PySpark to report 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 to separate documents by learned similar subjects
- Collaborated with cross-functional teams to identify key business metrics and KPIs, leading to the creation of customizable Machine Learning model visualizations through Tableau Dashboards as requested per customer
- Improved precision/recall of semi-supervised document classifier by 10% through use of feature engineering

### ROBOTICS RESEARCHER (Machine Learning Specialization)               May 2019 - Jul 2021
UCSD 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 extract, clean, transform, train, and test Sequence-to-Vector Deep Learning models on live touch based classification 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.
- Built recommendation system that web-scraped leads and ranked them on predicted call time length using SVM and probability of purchasing website development services using KNN model based on past conversion data

## PROJECTS

### DAD JOKE CHATBOT USING FINE TUNED CHATGPT MODEL
- Created, dockerized, and deployed personalizable Fine Tuned ChatGPT Chatbot clone using OPENAI API, LLaMA, and LangChain to query user input and return dad jokes ([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 3D-printed maze via UARM Python API (robot arm control API)

## TECHNICAL SKILLS

**Frameworks:** NumPy, Pandas, Sklearn, Matplotlib, Seaborn, OpenCV, TensorFlow, Keras, NLTK
**Domain Knowledge:** Ensemble Learning**,** DNN, CNN, AutoEncoders, GAN, RNN, GRU, LSTM, Transformer, BERT, GPT, LM
**Technologies:** SQL, Git, Jupyter, Anaconda, Docker, Databricks, Kubernetes, Spark, AWS, 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