

RAUL G. MARTINEZ

@ gio.mtz3@gmail.com

+1 619-734-6231

San Diego, CA

in raul-giovanny-martinez-688b1bb4

raulgiovannymartinez

WORK EXPERIENCE

Data Engineer I, Illumina

July 2020 – Present

San Diego, CA

- Developed data integration strategies, ETL processes, and analytics for manufacturing of Next-Generation Sequencing (NGS) platform's test data.
- Configured SQL Server database and performed ad-hoc queries to gather instruments data for investigation or analysis.

Integration Research Associate I-II, Illumina

July 2018 – July 2020

San Diego, CA

- Executed integration and systems support for NGS platforms during different stages of product development, and functional areas, including optics, hardware, fluidics, software, and chemistry components.
- Created data pipelines (R, Python, MATLAB) for aggregation, analysis, and visualization (Dashboarding) of NGS instruments data including primary/secondary metrics, engineering tests, reports, and diagnostic logs.
- Developed algorithms for image processing, manipulation, and transformation using python libraries (i.e. NumPy, PIL).

Pharma Technical Development Intern, Genentech

February 2018 – July 2018

South San Francisco, CA

- Contributed with the virology group on process validation and viral clearance studies for the production of therapeutic monoclonal antibodies.
- Investigated virus titers by implementing DOEs, cell culture, infectivity assays, qPCR, and virus purification by following GLP and GDP standards.

Support Specialist, Dotmatics

December 2017 – February 2018

San Diego, CA

- Collaborated with Developers and Scientists to deliver product suite solutions by testing software UI and querying/updating relational databases.

Data Analyst, Retrovirox

May 2016 – July 2017

San Diego, CA

- Implemented fast turn-around data analysis and processing workflows for high-throughput screening (with FACS Flow Cytometry) of hundreds of novel small-molecules with antiviral activities.
- Facilitated data handling and documentation (i.e. aggregation, backup and management) for R&D experiments with script automation and by authoring Standard Operating Procedures (SOPs).

RESEARCH AND PROJECTS

Predicting Helpfulness of Amazon Reviews

April 2020 – Jun 2020

San Diego, CA

- Analyzed numerical and text data from 200,000 reviews in order to extract useful information with feature engineering and NLP techniques. Feature selection was done with an ablation study.
- Developed a Random Forest Regressor Model to predict the ratio between the number of helpful votes with total votes, performance metrics used include MAE, MSE, precision, and recall.

Identifying Oncogene-Specific Essential Genes

January 2016 – June 2016

San Diego, CA

- Evaluated a bioinformatics prototype using R language statistical software, based on distributional entropies and kernel-based density estimators to find drug sensitivity profiles that match patterns of gene expression with over 10,000 variables.

EDUCATION

Masters Degree in Data Science and Engineering, GPA 4.0/4.0

September 2019 – Present

University of California, San Diego
(June 2021)

Bachelors Degree in Bioengineering: Biosystems, GPA 3.4/4.0

October 2014 – June 2017

University of California, San Diego
Minor: Entrepreneurship and Innovation

PROFESSIONAL SKILLS

Domain Related

Data Visualization

Data Mining

Data Management

Data Integration

ETL Pipelines

Big Data

Neural Networks

Statistics

Machine Learning

Hadoop

Spark

Programming

Python

R

Matlab

C/C++

PyTorch

SQL

APIs

Bash Scripting

HPC Cluster

Object Oriented

Computer Software

Git

Tableau

Denodo

SAS JMP

Amazon S3

PostgreSQL

SQL Server

RELEV. COURSEWORK

Data Integration and ETL

Beyond Relational Data Models

Machine Learning

Data Analysis Using Hadoop and Spark

Probability & Statistics Using Python

Data Management Systems

Python for Data Analysis

ACHIEVEMENTS

- Chancellor's Associate Scholar - UCSD (2014-2017)

- Community College Scholarships - "ASO Campus Involvement" and "Neva Smith" (2014)