# Dana Rocha

858-353-4055 dbridgette.rocha@gmail.com linkedin.com/in/rochadana github.com/dana-rocha

## EDUCATION

**Northeastern University** **Boston, MA**
*Master of Science in Bioinformatics and Graduate Certificate in Data Analytics (3.66/4.0)* *Dec 2020*

**Northeastern University** **Boston, MA**
*Bachelor of Science in Biochemistry* *Aug 2018*

## TECHNICAL SKILLS

**Languages:** Python, R, SQL, HTML, CSS
**Tools:** PyCharm, Git, RStudio, Spotfire, Pipeline Pilot, DBeaver

## PROFESSIONAL EXPERIENCE

**Merck & Co.** **Boston, MA**
*Informatics Co-op – MRL Computational and Structural Chemistry* *Jan 2020 – Aug 2020*

- Trained machine learning models to identify relationships between chemical compounds, genes, and disease
- Queried chemical data from public and internal databases using Pipeline Pilot, Pilot Script, and SQL
- Delivered compound sets to three drug discovery research programs in Chemical Biology for initial testing
- Created interactive Spotfire data visualization dashboards for over 140,000 compounds and genes
- Presented data insights to chemists and data scientists within Chemical Biology and Informatics
- Led weekly meetings with co-workers to promote an inclusive virtual community during global pandemic

**Orig3n Inc.** **Boston, MA**
*Research Assistant Co-op* *Jul 2017 – Dec 2017*

- Maintained stem cell cultures for internal assays and animal studies
- Characterized stem cell cultures with fluorescence microscopy and flow cytometry
- Represented Orig3n as a brand ambassador at events in New York City, Baltimore, and Lake Tahoe

## VOLUNTEER EXPERIENCE

**Northeastern University** **Boston, MA**
*Student Facilitator - Graduate Leadership Institute* *Feb 2021 – Apr 2021*

- Mentored 40 graduate students on developing leadership competencies outlined by learning goals
- Led weekly online workshops and discussions with students
- Oversaw logistics for online workshops and discussions using Zoom and FlipGrid
- Hosted weekly virtual meetings with co-facilitators to plan workshops, activities, and presentations

## PROJECTS

**Heart Failure Prediction** *Dec 2020*

- Created classification models to predict survival of heart failure patients using R with over 81% accuracy
- Trained, tested, tuned classification models using k-NN, Support Vector Machines, and ANN algorithms
- Built a bootstrapped ensemble model with 86.66% prediction accuracy

**Mapping Manhattan** *Jun 2020*

- Created Python script to compute distance to nearest subway stations using Pandas, Scikit-learn, and Numpy
- Visualized data with an interactive choropleth map using the Plotly library in Python
- Processed data from NYC MapPLUTO dataset and New York City Capital Planning Platform