Emily M. Chen
Data Scientist

chen.emi@northeastern.edu
Boston, MA
(516) 350-7889
linkedin.com/in/emily-chen-817ab6177/
Availability: January 2022 – June 2022

EDUCATION

NORTHEASTERN UNIVERSITY, Boston MA **September 2019 – Present**
Khoury College of Computer Sciences **Expected Graduation May 2023**
Candidate for Bachelor of Science in Data Science and Biochemistry, Minor in Mathematics
GPA: 3.7/ 4.0, Dean's List
Relevant Courses: Object Oriented Design | Information Presentation and Visualization | Database Design | Probability & Statistics | Linear Algebra | Differential Equations | Organic Chemistry | Biochemistry | Genetics & Molecular Biology
Activities: NU Women in Technology, Data Club, NUSCI, Club Volleyball, Big Sister, Chinese Student Association

SKILLS

Languages: Python, Java, HTML/CSS, JavaScript, TypeScript, Tableau, Scala, Visual Basic
General Office Tools: Microsoft Office (Word, Excel, Powerpoint)
Software: AWS, Spark, Eclipse, Excel, Git, Vim, Jupyter, Angular, React

WORK EXPERIENCE

GENONE TECHNOLOGIES, Cambridge, MA **July 2021 – Present**
Software Developer Co-op

- Develop an Angular based NativeScript app that uses an iPhone sensor to sense the position of two custom printed toys using machine learning methods (SVM and SVR).
  - Test and collect data from an iPhone sensor using HTTP requests.
  - Perform data analysis using Python (matplotlib and plotly) to evaluate the functionality of the app.
- Present project findings to lead software developer and founding partners during weekly standup.

NORTHEASTERN UNIVERSITY, Boston, MA **January 2020 – Present**
Research Assistant

- Collaborate with a PhD candidate to utilize DNA sequencing and bioinformatics approaches to identify the metabolic potential of microbial communities.

SIGNIFY HEALTH, New York, NY **June 2020 – August 2020**
Data Science Intern

- Formulated a novel model to represent heterogenous data in electronic health records using a 3D tensor.
  - Implemented model in a machine learning algorithm to predict potential surgeries of clients in advance.
- Developed perplexity-based probability models using Python and Scala to evaluate over 10,000 patients.
- Aggregated patient data from healthcare claims using Spark and AWS.

PERSONAL AND ACADEMIC PROJECTS

Northeastern Research Database:

- Created a Flask web application using PyMySQL for students to navigate and apply to research labs for all Northeastern's colleges using real-world data stored in a MySQL database schema.

Visualizing Boston Tap Water:
Project URL: https://neu-ds-4200-f20.github.io/project-group-07-boston-tap-water/

- Designed a website encoding a stacked bar chart, table and heatmap using JavaScript (d3.js), HTML and CSS.

Predicting Coronary Heart Disease:

- Explored machine learning algorithms (Logistic Regression, Support Vector Classifier, K Nearest neighbor, Naïve Bayes, Decision Tree Classifier) to predict if a given patient will develop heart disease using Python.
  - Determined that Logistic Regression and Support Vector Classifier worked the best with an accuracy of ~ 85%.

LEADERSHIP

Data Club, ***Communications Director*** **November 2020 – Present**

- Deliver and respond to information via emails and Slack and maintain a calendar with events and workshops.
- Facilitated a Tableau workshop to members (https://www.youtube.com/watch?v=sHlZz0FM4BU&t=1809s).

NU Women in Technology, ***Vice President*** **April 2020 – Present**

- Coordinate all club events and communications with general members.
- Manage relationships with corporations interested in partnering with NU Women in Technology.

NUSCI Science Magazine Club, ***Writer*** **September 2019 – Present**

- Published work in the fields of ethics, sustainable energy, political argumentation and quantum computing.

References available upon request.