Alex Chen

Brooklyn, NY | itsalexchen@gmail.com | LinkedIn | GitHub | Portfolio

EDUCATION

New York City College of Technology, CUNY **Brooklyn, NY**
Bachelor of Science in Data Science *06/24*

- **Relevant Coursework:** Data Structures, Database Fundamentals, Information and Data Management (Python), Artificial Intelligence, Machine Learning for Physics, Linear Algebra, Probability and Statistics

TECHNICAL SKILLS

Languages: Python, SQL, HTML, CSS, Sass, JavaScript, TypeScript, Java, MQL, Bash, R

Libraries/Frameworks: Pandas, Scikit-learn, Matplotlib, React, Jest, TensorFlow, NLTK, Node.js, Express

Technologies: Git, GitHub, Jupyter Notebooks, Postgres, MongoDB, Figma, Excel, Data Studio, Tableau

PROGRAMMING PROJECTS

Job Application Tracker Web App *Jan 2023 - Present*

- Develop a full-stack web app using TypeScript React and Express that allows users to create, view, update, and delete items in their job tracker
- Create and form a database using PostgreSQL to store information about the roles the user applied for
- Implemented RESTful API using Express.js to support web application

Emotion Detection Web App - GitHub *10/22 - 12/22*

- Constructed a web application that predicts emotions within a piece of text using natural language processing
- Implemented Multinomial Naive Bayes Classifier using Scikit-learn and NLTK with an accuracy of 88%

Image Classification Web App - GitHub *10/22*

- Developed a web application that classifies images of natural scenes around the world using machine learning
- Implemented Random Forest Classifier using Scikit-learn, Numpy, and Pandas with an accuracy of 53%

Reddit Client App - Project Page *08/22*

- Created an interactive Reddit client application that retrieves over 50 posts and comments from Reddit using HTML, CSS, JavaScript, and React.js to deepen understanding of React and the virtual DOM
- Constructed a wireframe and component tree for the application within Figma in preparation for development

Biodiversity in National Parks Analysis - GitHub *05/22*

- Analyzed 30,000 data points using python libraries, Pandas, SciPy, NumPy, Seaborn, and Matplolib, to answer questions about different species in 4 national parks and the 191 endangered species within the parks

RELEVANT EXPERIENCE

CUNY Tech Prep **New York, NY**
Data Science Fellow *07/22 - Present*

- Selected for a competitive data science fellowship with students from CUNY senior colleges where Fellows create technical projects using tools like Python 3, Jupyter Notebooks, Pandas, Numpy, Scikit-learn, and SQL
- Participate in weekly courses and learn industry best practices for exploratory data analysis (EDA), feature engineering, data collection and processing, statistical modeling, data visualization, machine learning techniques, data science process, and big data

The City University of New York, CUNY Conserves Team **New York, NY**
Energy Technology Intern *07/22 - 08/22*

- Tested and improved an import app, by reviewing and collecting feedback on 30+ test cases, with the lead software engineer resulting in faster speeds by over 20%.
- Developed a Python script to clean raw data and create a report that can be referenced, saving over 15% of time
- Implemented unit tests using Python to ensure that code was written correctly before deploying to production

New York City College of Technology **Brooklyn, NY**
Undergraduate Researcher *05/22 - 06/22*

- Created 25+ case studies working from an online resource containing datasets to assess how data science students compare and differ from the industry standard for interviews.
- Developed a pipeline using a Jupyter Notebook for searching details about interview questions for NYCCT CST faculty and students

The Urban Assembly **New York, NY**
Operations & Data Intern *03/22 - 05/22*

- Identified opportunities for improvement within the data science process over 12 different data sources
- Constructed data visualizations for easier and faster access to information for members across 3 different teams
- Analyzed over 10,000 data points across 3 departments to help better plan resources placement