Jason Chen

718-866-7738 | jachen7@bu.edu |Boston, MA| https://github.com/jachen2956

Education:

Boston University	Boston, MA
Bachelor of Science in Data Science	Graduation Date: May 2026

- Honor: Dean's List (Fall 2023 – Present)

Skills:

Programming Languages: Python, SQL, R, Rust

Frameworks & Tools: NumPy, Pandas, Matplotlib, Seaborn, Scikit-learn, TensorFlow, Tableau, Power BI, Git, Azure, Keras, Statsmodels

Data Science Skills: Machine Learning (Supervised and Unsupervised), Neural Network, Time Series Analysis, Data Visualization, Statistical Analysis, Natural Language Processing, Feature Engineering

Professional Experience:

University of Massachusetts Amherst	Amherst, MA
CEET Research Assistant	June 2024 – Present

- Enhanced the model's accuracy and regional applicability by collecting and integrating data for **50+ regions** using custom Python scripts, resulting in improved **carbon intensity (CI) predictions**.
- Achieved optimal model selection for **first-tier forecasts** by analyzing forecast data using Pandas and Seaborn, resulting in an **15% increase** in forecasts accuracy.
- Conducted rigorous testing of **neural network models** to predict hourly CI for up to 96-hour forecasts, achieving a Mean Absolute Percentage Error (MAPE) of **less than 10%.**

Authentic Caribbean Foundation Inc	Boston, MA
Data Analyst Intern	November 2023 – February 2024

- Utilized **Excel** to analyze restaurant performance data, creating **pivot tables and charts** that lead to a **30% rise** in repeat business and online reservations, resulting in improved customer engagement.
- Identified trends among specific racial groups in Massachusetts regions by **leveraging census data**, leading to a **15% increase in patronage** through targeted marketing campaigns.

Project Experience:

Cyberbullying Classification (Python)

- Applied **NLP techniques**, including tokenization, stop word removal, part-of-speech tagging, lemmatization, and sentiment analysis, **to 47,000+ tweets**, enhancing text data processing.
- Transformed data into **binary classification** format and evaluated models using Support Vector Machine and Logistic Regression, achieving a **sensitivity of approximately 98%,** ensuring high detection rates.
- Achieved approximately **83% accuracy** by training and evaluating **multiclass classification** models, including Naïve Bayes, Random Forest, and Gradient Boosting, improving model performance.

Apple Stock Prediction (Python, SQL)

- Enhanced data insights by **preprocessing 12,000+** Apple stock data points and conducting feature analysis using pair plots, correlation analysis, and line plots.
- Achieved a Mean Squared Error (MSE) of 0.03 by developing a **Long Short-Term Memory** (LSTM) model to forecast future stock prices, resulting in **highly accurate predictions** .
- Ensured data accuracy and reliability **by executing SQL queries** on the Apple stock database to retrieve statistical information and validate predictions.