# Nitheesh Koushik Gattu

Boston, MA (857) 218-9878 gattu.n@northeastern.edu LinkedIn

## Education

**Northeastern University, Khoury College of Computer Science** Boston, MA
*Master of Science in Data Science, GPA: 3.43* December 2023
Relevant Courses: Programming for Data Science, Algorithms, Supervised Machine Learning, Data Management and Processing, Unsupervised Machine Learning, Database Management Systems, Natural Language Processing

**Jawaharlal Nehru Technological University, VNRVJIET** Hyderabad, IN
*Bachelor of Technology in Civil Engineering, GPA: 8.03* June 2021
Relevant Courses: C Programming, Data Structures, Ordinary Differential Equations and Laplace Transformations

## Technical Skills

| | |
|---|---|
| Programming Languages: | Python, R, SQL, JavaScript, MATLAB, C |
| Databases: | MySQL, SQLite |
| Data Analysis: | Pandas and Numpy |
| Data Visualization: | Tableau, ggplot2, Matplotlib, Seaborn, Folium |
| Machine Learning: | Regression, Tree-based modeling, Neural Nets, Clustering, Recommendation Systems |
| Web Scraping: | Selenium, BeautifulSoup |

## Work Experience

**Sales Analyst Intern** Spring 2023
*Mimecast*

- Collected data using web automation libraries like Selenium to increase an account's data quality.
- Collected about 150,000 data points that were directly imputed in the company's Salesforce and Demandbase database.
- Performed Chi$^2$ tests on the obtained data to make sure that the data collected is not different from manually collected.
- Used Google's Geocoding API to breakdown address information and to perform data validation.
- Used Tableau Server Client API and Simple Salesforce API for report generation to reduce manual errors and intervention.

## Academic Projects

**Aircraft Bird Strikes Database** Fall 2022
*Database Management Systems, Northeastern University, Boston MA*

- Created a MySQL database on the AWS Relational Database Service platform containing aircraft bird strike data.
- Performed ETL from the FAA published XML file and used to R to load the data to MySQL database.
- Analyzed data using SQL queries to find patterns and determine preventative measures to reduce bird strikes.

**Building Machine Learning Models** Summer 2022
*Unsupervised Machine Learning, Northeastern University, Boston MA*

- Built own clustering algorithms (kNN, kMeans, soft kMeans, DBSCAN), dimensionality reduction algorithms (PCA, t-SNE, Haar feature reduction), text summarization (KL Divergence).
- Tested these algorithms and compared with the inbuilt Python libraries for performance optimization.

**Boston Crime Data Analysis** Spring 2022
*Introduction to Data Management and Processing, Northeastern University, Boston MA*

- Visualized crimes in Boston to recognize geographical and seasonal patterns using Folium, GeoPy and Matplotlib.
- Used time-series analysis models such as ARMA, ARIMA, SARIMA and SARIMAX to predict the district- wise by month crime count and obtained a least Mean Absolute Percentage Error of 15.6%

**Image Classification** Spring 2022
*Supervised Machine Learning, Northeastern University, Boston MA*

- Solved an Image classification problem by building Linear Regression, SVM, CNN and Random Forest models after preprocessing the data using Principal Component Analysis
- Built an ensemble model using the above models and achieved an accuracy of about 83%

## Certifications

- Extract, Transform and Load *(CertNexus, Coursera)*
- Data Warehouse Concepts, Design, and Data Integration *(University of Colorado System, Coursera)*
- Distributed Computing with Spark SQL *(University of California, Davis, Coursera)*
- SQL for Data Science *(University of California, Davis, Coursera)*