

Muskan Jain

jain.musk@northeastern.edu | (857) 264 - 6541 | [LINKEDIN](#) | Boston, MA

EDUCATION

Northeastern University, Boston

May 2025

Master of Science in Data Analytics Engineering

CGPA 3.82/4

Coursework: Database Management Systems, Computation and Visualization for Analytics, Fundamentals of Data Analysis, Deep Learning and Neural Networks

Dwarkadas J. Sanghvi College of Engineering, Mumbai

May 2023

Bachelor of Engineering in Electronics & Telecommunication Engineering

CGPA 3.5/4

SKILLS

Technical Skills: Tableau, Power BI, Python, SQL, Excel, Hadoop, Snowflake, Alteryx, Pentaho, Airflow, Databricks
Libraries: Pandas, NumPy, Matplotlib, SciPy, Seaborn, Plotly, Ggplot2
Databases and Tools: MySQL, Oracle, NoSQL, SQL Server, Tableau, Power BI, Flourish, Looker, Data Wrapper, Excel
Machine Learning: Supervised Learning (Regression and Classification), Unsupervised Learning (Clustering), Predictive modelling
Technologies: Natural Language Processing, Exploratory Data Analysis, Regression Analysis, Amazon Web Services, Git, Cloud Computing, Advanced Statistical Analysis, Big Data Processing, Data Mining, Computer Science, Statistics

EXPERIENCE

Teaching Assistant- Marketing Analytics – Boston, MA

Jan 2024 -Current

- Contributed to the academic environment at Northeastern University's Business School by actively participating in the delivery of marketing analytics courses
- Graded assignments, provided constructive feedback, and held office hours to address 50 students and their queries
- Developed and executed interactive workshops on data visualization using tools such as Tableau, leading to a 20% improvement in student proficiency
- Collaborated with instructor to revamp curriculum by incorporating 10+ real-world case studies, empowering students to apply marketing analytics principles and achieve measurable results, resulting in a 40% increase in student engagement and application of concepts

Data Science Analyst - Intern, Play Time Media Solutions, India

Feb 2023- Aug 2023

- Collaborated with a diverse team of 5 computer vision engineers to design and fine-tune custom OCR models using Paddle OCR, resulting in a 15% accuracy improvement
- Arranged weekly meetings with solution leads and managers to address potential issues and ensure project progress
- Evaluated and presented pros and cons of deep learning models, leading to a 20% reduction in inference time while maintaining 95% accuracy, enhancing overall model performance

Business Intelligence Analyst Intern – Pravanzaa, Remote- Germany

Jan 2022- Aug 2022

- Supervised and mentored a team of 10 interns, designed ETL pipelines for the system, overseeing inventory management, customer assistance, and data-driven course selection
- Implemented monthly sales initiative that delivered a notable 60% increase in revenue compared to the previous year and satisfied their business objectives and managed each project successfully

PROJECTS

Feature Importance of Real-Time Movie Data Survey

Sep 2023-Dec 2023

- Developed Random Forest regression models to predict a movie's meta score and gross USA ticket sales, achieving an R^2 score of 0.999 and 0.994, respectively, indicating highly accurate models
- Utilized Python libraries such as pandas for data manipulation, scikit-learn for model building and evaluation, and matplotlib and seaborn for visualizing feature importance
- Achieved significant model performance improvements over baseline models, with the Random Forest model for meta score prediction showing a root mean square error (RMSE) of 0.522 and for gross USA ticket sales prediction an RMSE of 5,267,724.

Pharmacy Subscription and Management System

Sep 2023-Dec 2023

- Created and implemented a pharmacy and subscription management system database using both SQL and MongoDB, including 10 tables to store customer information, prescription information, subscription information, and more
- Inserted over 300 records of data into the database to ensure it was fully functional
- Developed a Python application to connect to the databases, retrieve records, and visualize data using Python libraries
- Delivered a comprehensive database solution and a Python application with visualizations that conveyed insights to the client

Los Angeles Crime Analysis

Sep 2023-Dec 2023

- Analyzed 2022 crime data for Los Angeles, revealing a 12% increase in the city's overall crime rate. Proficient in identifying crime trends and influencing factors
- Conducted research on crime incidents, highlighting Fridays as the day with the highest crime incidents of 982385
- Investigated 2022 gender-specific crime victim data, uncovering a significant increase in female victims