Pranathi Bombay

Boston, MA| (857) 384-9844 | bombay.p@northeastern.edu | LinkedIn| GitHub

EDUCATION

Northeastern University, Boston, MA **December 2026**
Master of Science in Data Analytics Engineering
Related Courses: Data Management, Foundation in Data Analytics, Data Mining, Computation and Visualization, Deterministic Operations Research, Natural language Processing, Machine Learning Operations

National Institute of Technology Tiruchirappalli, India **May 2021**
Bachelor of Technology - Electronics and communication Engineering
Related Courses: Programming Languages, Data Structures and Algorithms, Object Oriented Concepts, Operating Systems, Database Management System, Computer Architecture, Data Analytics and Data mining

SKILLS

Programming: Python, R, SQL, C, C++, Java, HTML, CSS, React
Libraries & Frameworks: TensorFlow, Scikit-learn, Pandas, NumPy
Databases & Tools: MySQL, PostgreSQL, Oracle, MongoDB (NoSQL), AWS, Docker, Git/GitHub
Data Visualization: Tableau, Power BI, Matplotlib, Seaborn, Excel (PivotTables, Charts, Dashboards)
Data Engineering: ETL Process, Data Cleaning, Data Wrangling, Exploratory Data Analysis (EDA)
Productivity Tools: Advanced Excel (Macros, VLOOKUP), Microsoft Office Suite
Certifications: Google Data Analytics, Microsoft Power BI
Communication: Strong verbal & written communication, Stakeholder collaboration, Presentation delivery

WORK EXPERIENCE

E-MECH Solutions Pvt Ltd, Bengaluru **July 2021 - October 2024**
Data Analyst:

- **Built** and **optimized SQL**- and **Python**-based **ETL pipelines**, automating data ingestion and cleansing, which eliminated **50% of manual effort** and reduced processing time by **15%**, accelerating enterprise reporting.
- **Developed** automated validation scripts in SQL & Python (null checks, range checks, duplicate detection), which decreased reporting errors by **30%** and ensured **high data quality for dashboards and analytics**.
- **Designed and deployed 10+ interactive dashboards** in Tableau & Power BI to track KPIs across sales, operations, and data quality, cutting **ad-hoc reporting turnaround by 40%** and improving executive visibility.
- **Automated recurring business analyses with Python (Pandas, NumPy, Matplotlib)**, improving insight delivery speed by **35%** and enabling **faster decision-making across departments**.
- **Collaborated** with 10+ **cross-functional teams** (engineers, analysts, data scientists) to define enterprise-wide **data quality standards**, improving **consistency, scalability, and trust in analytics outputs**.

PROJECTS

Blood And Organ Donation Network **January 2025**

- **Constructed** a MySQL & NoSQL database to organize **1,000+ simulated donor, patient, and hospital records**, showcasing scalability and efficiency in blood and organ allocation.
- **Formulated** SQL queries (joins, nested, aggregates) and NoSQL pipelines to achieve **real-time donor–recipient matching** and generate hospital reporting outputs.
- **Applied** data preprocessing techniques and **produced** Python visualizations (bar, pie, histogram), enabling **25% faster trend analysis** on donation datasets.
- **Orchestrated** ETL workflows to unify data from blood banks, hospitals, and transplant centers, reducing **manual record-keeping by 40%** in the project dataset.

Healthcare Patient Clustering & Risk Stratification **January 2025**

- **Processed** and consolidated 58K+ patient records (demographics, lab tests, vitals) into a **30K+ clean dataset** using imputation, normalization, and feature selection techniques.
- **Applied** K-Means and Hierarchical Clustering, and **validated** results with Elbow Method, Silhouette Score, and PCA to uncover patterns in large-scale healthcare data.
- **Identified** 4 patient clusters associated with diabetes, kidney disease, hypertension, and immune disorders, supporting **targeted risk assessment** and clinical decision-making.
- **Performed** exploratory data analysis (EDA) on demographics, lab tests, and vitals, detecting anomalies and reducing data noise by **15%**, which improved clustering accuracy.

Facial Emotion Recognition in Cartoons **January 2021 - May 2021**

- Built a CNN–RNN model in TensorFlow/Keras, tuning hyperparameters and processing **20K+ video frames** (Python, OpenCV) to boost accuracy by **18%**.
- Implemented a real-time feedback loop with computer vision analytics, improving **engagement tracking accuracy by 30%** and demonstrating practical application of deep learning in user interaction analysis.