

Amitesh Tripathi

Boston, MA | +1 857-376-1991 | amitesh.tripathi.neu@gmail.com | [linkedin.com/in/theamiteshtripathi](https://www.linkedin.com/in/theamiteshtripathi) | [Github](#) | [Portfolio](#)

SUMMARY

As a Data Analytics postgraduate with experience at Capgemini and Aasmo Digital, I've honed my machine learning, ETL optimization, and Data Visualization skills. Proficient in Python, R, Tableau, and SQL, I'm now seeking full-time/ internship roles to leverage my data-centric skills further

EDUCATION

Northeastern University, Boston, USA

Expected May 2024

Master of Data Analytics Engineering

Courses: Data Mining, Machine Learning, Data Management for Analytics, Data Computation and Visualization, Foundations of Data Analytics

Rajiv Gandhi Technical University, Bhopal, India

August 2017 - May 2021

Bachelor of Computer Science Engineering

Courses: Database Management Design, Data Visualization, Data Warehouse and Mining, Machine Learning, Artificial Intelligence

EXPERIENCE

Capgemini Technology Services India Ltd

Chennai, India

Data Engineer

July 2021 - August 2022

- Coordinated feature engineering, schema updates & dataset generation in financial data warehouse using **NoSQL** and **Hive**, ensuring data quality
- Orchestrated ETL workflow automation using **Apache Airflow**, reducing manual effort by 20% and improving turnaround time
- Developed 10+ **Python** scripts for analytics and reporting, providing accurate, real-time insights to the business intelligence team
- Utilized **Power BI** to create 5+ comprehensive financial dashboards, enabling effective decision-making by highlighting key metrics
- Streamlined financial decision-making, resulting in a 15% improvement in operational efficiency

Aasmo Digital Pvt. Ltd

Indore, India

Associate Data Analyst

May 2020 - June 2021

- Developed robust database schema and architecture using **Oracle SQL** for a large E-commerce site serving 100K+ users
- Streamlined and optimized MySQL databases using **Tableau**, reducing query time by 25% and enhancing server performance
- Revamped ETL process for customer database using **Alteryx**, reducing data loading intervals to near real-time
- Leveraged **K-means** clustering machine learning technique to categorize 5K+ products by customer behavior, boosting user personalization
- Increased conversion rates by 20% by integrating a collaborative filtering recommendation system via **R** programming

TECHNICAL TOOLKIT

- **Languages:** SQL, Python, R, Java, Bash, JavaScript, HTML-CSS, C++, React JS
- **Analytics:** Predictive Modeling, Machine Learning, Statistical Modeling, Feature Selection, Regression - Linear & Logistic, Hypothesis Testing, TF-IDF, A/B Testing, Principal Components Analysis (PCA), Clustering, Classification, Neural Networks, SVM, Decision Trees, Naïve Bayes, Regularization, Ensemble Models, Time Series
- **Tools and Frameworks:** Power BI, Tableau, MS Excel, MS PowerPoint, Google Analytics, Apache Airflow, Streamlit, Docker, PyCharm, VSCode, MS Word, SAS, Flourish, DataWrapper, JIRA, GitHub
- **Databases:** MySQL, Oracle Database, SQL Server, MongoDB, PostgreSQL, MapReduce
- **Python Libraries/Tools:** TensorFlow, Scikit-Learn, NumPy, Pandas, Keras, PyTorch, SciPy, matplotlib, Seaborn, SNS
- **Certifications:** Google Data Analytics Certificate
- **Competencies:** Deep Learning, Software Development, Cloud Programming, CI-CD

PROJECTS

Text Summarizer (Python, NLP, Flask, Docker): [GitHub](#)

- Engineered a Python-based NLP model to condense long texts into short summaries, enhancing readability and information accessibility
- Deployed the model via a Flask app in a Docker container, boosting user engagement and ensuring easy distribution and scalability

Stock Sentiment Analysis (NLP, ML): [GitHub](#)

- Developed a Python-based NLP and ML model, achieving 82% accuracy in predicting stock market trends from news headlines
- Transformed news headlines into actionable investment insights, enhancing the model's predictive power and translating text data into insights

IPL Score Prediction (Python, Linear Regression, Flask): [GitHub](#)

- Built an ML model in Python to predict IPL scores, providing valuable insights for sports analysts and IPL fans
- Deployed the model via a Flask app, enhancing user interaction with real-time score predictions and increasing the tool's accessibility

Ad Click Prediction (Python, Data Analysis, ML): [GitHub](#)

- Implemented a predictive model in Python to forecast ad clicks, assisting businesses in optimizing ad placements and maximizing ROI
- Conducted data analysis and feature engineering in Python, enhancing the ad click prediction model's accuracy and robustness

Heart Diseases Tableau Project (R, Data Visualization, RStudio, RMarkdown): [GitHub](#)

- Analyzed heart disease data using R, uncovering key patterns and factors associated with heart diseases, providing deep insights
- Generated comprehensive visualizations using R, effectively communicating complex data patterns in an easily interpretable form

Adult Income Dataset Visualization (Tableau, Data Visualization, Data Analysis): [GitHub](#)

- Explored adult income dataset using Tableau, revealing key trends in income distribution among different demographic groups
- Crafted an interactive Tableau workbook, enabling stakeholders to dynamically explore the data and gain insights into income patterns

CERTIFICATIONS

Google Data Analytics Certificate | (Issued By Google) [Credential](#)

GPT-4 Foundations: Building AI-Powered Apps | (Issued By LinkedIn) [Credential](#)

Data Science For Engineers | (Proctored Exam Silver Medal, Issued By NPTEL) [Credential](#)