Shreyas Darade

Pittsburgh, PA | sdarade@andrew.cmu.edu | +14127289474 | linkedin.com/in/shreyasdarade/ | **Portfolio**

EDUCATION

Carnegie Mellon University, Pittsburgh, PA, United States — August 2024 - December 2025
Master of Information Systems Management - Business Intelligence and Data Analytics **[Dean's List: All Semesters]** **GPA: 4.01**
Coursework: Machine Learning (10601), Deep Learning (11785), ML for Problem Solving, Data Science and Big Data, Econometrics

SKILLS

Data Science: Python, SQL, NumPy, Pandas, Scikit-Learn, PyTorch, A/B Testing, SpaCy, NLP, Topic Modeling, Deep Learning, TF-IDF, Supervised Learning, Unsupervised learning, Regression, Classification, Clustering, Optimization, Statistics
Analysis & Visualization: PowerBI, Tableau, Exploratory Data Analysis, Matplotlib, Seaborn, Plotly, Spark, MapReduce
Tools & Platforms: Git, GitHub, Figma, Latex, Jupyter, Google Colab, PowerPoint, n8n, Wandb, Snowflake

EXPERIENCE

[Capstone] Armada Partners, Pittsburgh, PA **- Data Scientist** — August 2025 – December 2025
Led a team of 4 graduate students to improve working capital utilization by doing **safety stock optimization** for the client. Proposed various models to accurately predict lead time variability and demand forecast to calculate safety stock at an item-location level. Defined key metrics aligning with the Data science lifecycle to measure business impact by conducting extensive stakeholder research.

Naptic, Pittsburgh, PA - **AI/ML Engineer Intern** — June 2025 - August 2025
Worked in a fast-paced startup to solve business problems for clients using interesting, efficient and secure AI agent workflows. Researched and implemented MCP and RAG agent strategies to make low-code AI agents faster, structured, reliable, and compliant to work with sensitive data.

GEP, Mumbai, India - **Senior Analyst** — September 2022 - August 2024

- [**Consulting**] Consulted for a Fortune 100 US client with a $4.5B indirect spend on a procurement transformation project including analytics insights and reporting 25+ metrics/KPIs to 220+ members for measuring the progress of the client program
- [**Analytics | Python | Excel**] Performed exploratory data analysis (EDA) on large scale procurement data to test hypothesis on cost drivers and savings opportunities; collaborated with data scientists and leaders to align models and priorities. Used contract forecasting to analyze data over 15+ months and help understand the need for new resources to reduce backlog
- [**Optimization | Power BI | Stakeholder Management**] Spearheaded redesign of 5 key client dashboards after understanding the stakeholder problems. Automated dashboard pipelines in PowerBI and Python, optimized dashboards achieving a 60% reduction in complexity and saving 4+ hours per week. Observed 100% improvement in customer engagement upon deployment
- [**Cross-functional collaboration**] Designed a modular BI workbench on the client analysis platform while collaborating with technology team and facility managers enabling data-driven decisions on spend, savings, contracts and suppliers
- [**Presentation | Data Visualization | Business Intelligence**] Created executive level monthly reports summarizing 15+ strategic KPIs, highlighting performance trends and actionable opportunities bridging raw data to business insights
- [**Product Metrics**] Defined and monitored different product engagement metrics to improve visibility and user retention over 25%

ACADEMIC RESEARCH EXPERIENCE AND PROJECTS

Personalized Carbon Emission Prediction and Recommendations:
[Regression | Python | XGBoost | Pandas | Data Cleaning | Neural Network | KNN-imputation] Developed a machine learning based personalized carbon emission predictor which gives an estimate of monthly emissions based on various key features.

- Set a clear baseline of the analysis using ridge regression and used different correlation analysis and imputation techniques to handle missing values and improve the validation data accuracy. Used grid search cv to choose the best hyperparameters.
- Trained 5+ models using various regression, Ensemble methods and achieved 98% accuracy on test data using XGBoost
- Delivered personalized insights on what better choices can the user make to decrease emissions using key contributing features

Airbnb New York City Market Analysis – Revenue Optimization:
[Exploratory Data Analysis | Data Cleaning | Visualization| Pandas] Conducted EDA on 30K+ Airbnb listings in combination with multiple data sources like listings reviews and NY neighborhoods to understand market trend and pricing based on some key features in the dataset to understand opportunities to increase revenue and create business impact.

Smart Traffic Management System:
[Object Detection | Python | Problem-Solving] Collaborated in a team of three to analyze, design and to solve a real-time problem Smart Traffic Management System for prioritizing emergency vehicles and efficiently controlling smooth traffic flow

- Trained object detection model, designed and integrated a traffic scheduling algorithm with 95%+ accuracy detection
- Efficiently managed the traffic congestion by developing a custom algorithm considering the vehicle waiting time, count, frequency

Research Paper: Dynamic Traffic Scheduling Using Emergency Vehicle Detection, INCOFT 2022, IEEE [***link***]