

UTKARSH GUPTA

utkarsh.gupta@duke.edu • (919)282-2533 • [LinkedIn](#) • [Portfolio](#) • [GitHub](#)

EDUCATION

DUKE UNIVERSITY, THE FUQUA SCHOOL OF BUSINESS

Durham, NC

Master of Science in Quantitative Management: Business Analytics (STEM)

May 2025

Selected Coursework: Data Science for Business, Applied Statistics, Data Visualization, Data Infrastructure, Strategic Decision Analytics & Modelling, Empirical Economic Analysis, Enterprise Risk Management, Fraud Analytics, Cybersecurity, AI/ML, Deep Learning, Neural Networks, Strategic Management, People Analytics

SRM INSTITUTE OF SCIENCE AND TECHNOLOGY

Chennai, India

Bachelor of Technology: Computer Science and Engineering

Jun 2024

Research paper: [Psychological Analysis Using Social Media Tweets](#) at IEEE 2024 3rd International Conference on Applied Artificial Intelligence and Computing.

SKILLS

Tools & Tech: SQL, NoSQL, Python (NumPy, Pandas, Scikit-Learn, PyTorch, TensorFlow, Keras, XGBoost), R, Tableau, Power BI, Jupyter Notebook, Advanced Excel (Power Query, VBA), C++, Java, JavaScript, NodeJS, Jira, Looker Studio, API Development.

Modelling & Analytics: A/B Testing, prototyping, Feature Engineering, Regression, model programming, Time series, Predictive Modelling, Data Mining, Forecasting, Financial Modelling & Reporting, Portfolio Monitoring, Causal inference, NLP

Big Data & Cloud: Cloud Security (AWS Certified), Data Pipelines, HPC, GCP, Azure, Snowflake, DataBricks, Alteryx, Hive, Spark

EXPERIENCE

DUKE HEALTH (Capstone Project)

Durham, NC

Student Analyst & Consultant

Mar 2025 – May 2025

- Analyzed **500K+ multi-source patient records** (EMR, claims, public health) using SQL, Python & R; identified **\$3.2M in potential cost savings** and a **15% increase in early disease detection rates**, strengthening Duke Health's Medicaid strategy.
- Built Delivered **executive-ready Tableau dashboards** that revealed critical care gaps across 12 regions; insights directly shaped **3 statewide Medicaid outreach initiatives**, projected to impact **50K+ underserved patients annually**.

HCLTECH, Analytics Team

Noida, India

Business & Product Analyst Intern

Nov 2023 – Jun 2024

- Analyzed & Applied **ML and statistical analysis** on **100GB+ Microsoft Bing search interaction data**, uncovering product usage patterns & trends, driving a **20% increase in user engagement** and informing strategic product and marketing decisions.
- Built automated data pipelines and optimized SQL workflows to support A/B testing, cohort analysis, and real-time user tracking, **reducing data processing time by 40%** and **query execution time by 25%**.
- Developed dynamic **Power BI dashboards** to visualize product KPIs and feature adoption metrics, **boosting reporting efficiency by 30%** and enabling faster, insight-driven decision-making for cross-functional teams.

SIEMENS, IT

Gurugram, India

Data & Operations Analyst Intern

Jan 2023 – Jul 2023

- Conducted in-depth **System performance analysis** of baggage handling systems to identify product bottlenecks and efficiency gaps; delivered actionable insights that supported automation strategy and roadmap decisions.
- Evaluated and enhanced the **Basic Bag IQ product** for **microservices deployment**, contributing to a **14% increase in system scalability and accuracy**, aligning with product growth objectives.
- Developed and implemented Java-based automated error detection, optimized database queries, resulting in a **20% increase in baggage throughput** & **15% reduction in operational error rates**, enhancing real-time system reliability and efficiency.

SELECTED PROJECTS

Analyzing Psychological Traits Using Social Media Tweets (ML, Python). Designed and fine-tuned **LSTM** and **CNN** models on **50K+ tweets**, achieving **99.02% classification accuracy** in detecting depression-related content. Conducted sentiment and temporal analysis to uncover key social and linguistic factors linked to depressive behavior.

Causal Analysis of Customer Outreach on Bank Subscription Rates (R, Python). Analyzed **4500+ banking clients** using **PSM**, **DiD**, Causal **Random Forests** to quantify outreach effectiveness. Identified initial contact increased subscription rates by 11.1%. Repeated calls decreased by 2.6%. Delivered data-driven recommendations to improve marketing ROI, reduce outreach costs.

Semantic Analytics Engine for Vendor Ranking for Pyramid AI (Python, LightGBM, FastAPI). Built a **FastAPI-based analytics engine** for vendor qualification, leveraging **MiniLM embeddings** and **KeyBERT** to extract 20-keyword profiles from **1,000+ CRM tools**. Designed a **weighted scoring model** (60% semantic similarity, 10% token overlap, 30% normalized ratings) to rank vendors with high relevance and decision-making precision.