VANSHIKA RAM GURBANI

LinkedIn | GitHub

+1 (551) 332-0041 | vanshika.gurbani25@gmail.com | New Brunswick, NJ 08901

EDUCATION

Rutgers University, New Brunswick, NJ
Master of Science in Data Science; ***GPA: 3.45/4*** *2025*

Thadomal Shahani Engineering College
Bachelor of Engineering in Computer Engineering; ***GPA: 8.5/10*** *2023*

Relevant Coursework: Data Structure and Algorithm, DBMS, Data Warehouse and Mining, Data Wrangling, AI, ML, NLP, Quantitative Analysis, Big Data Analysis, Applied Data Science, Probability & Statistical Inference, Regression and Time Series, Statistical Learning, Algorithm Trading and Portfolio Management, Statistical Modeling and Computing

PROFESSIONAL EXPERIENCE

Rutgers University
Teaching Asst. - Part Time Lecturer & Grader *2023 - 2025*

- Instructed CS110/CS170 labs, covering foundational computing topics including **Scratch, HTML/CSS/JavaScript, Excel functions, and basic SQL queries**.
- Guided students through **hands-on assignments**; provided support during **exams** and **office hours** to reinforce conceptual understanding.
- Providing mentoring, office hours, and a technology-enhanced learning environment while **collaborating** with peers and enhancing **teaching skills**.

PROJECTS

Growth Mindset Causal Inference Analysis | Python, Causal Forests, IPW, Jupyter (Github) *2025*

- Analyzed synthetic observational data (10K+ student-level records across 76 schools) to estimate the **Average Treatment Effect (ATE)** of a growth mindset intervention on academic achievement.
- Applied multiple causal inference methods including T-Learner, S-Learner, Inverse Probability Weighting (IPW), and Causal Forests yielding consistent results with estimated **ATE ≈ 0.30** (95% CI: 0.24–0.36).
- Created publication-style plots for treatment effects, covariate overlap, and balance; delivered a **48"x36" academic poster** summarizing methods, assumptions, and limitations.

Biomedical Named Entity Recognition (BioNER) | Python**,** HuggingFace Transformers, PyTorch (Github) *2024*

- Developed a **multi-task learning model** using **BioBERT** to extract biomedical entities (e.g., genes, diseases, chemicals) across five benchmark datasets.
- **Fine-tuned pre-trained transformers** for token classification using a **shared encoder and task-specific heads**, boosting performance on low-resource datasets.
- Implemented evaluation using **IOBES tagging** and metrics like precision, recall, and F1-score, achieving **90%+ accuracy** across tasks.

Image Feature Detection & Captioning | Neural Networks, Object Detection, Streamlit (Github) *2023*

- Implemented **CNN** and **VGG-16** models for image feature extraction and **LSTM** (BLEU score: **0.65**)/ **Transformer** (BLEU score: **0.80**) models for caption generation.
- Created a user-friendly web interface using **Streamlit**, demonstrating **full-stack capabilities** in AI application development.

Enhancing Airbnb Revenue: A Multi-Modal Data Strategy for Dynamic Price Modeling | Python, R, SQL, Tableau *2023*

- Executed a **regression analysis** project, analyzing over **10,000 Airbnb listings**, resulting in a **25% increase** in prediction accuracy for revenue optimization.
- Employed **Python** for data wrangling, **R** for statistical analysis, and **SQL** for database management; utilized **Tableau** for visual insights, enhancing stakeholder presentations.

TECHNICAL AND OTHER SKILLS

- **Languages**: Python (NumPy, pandas, scikit-learn, PyTorch, Transformers), R, MATLAB, SQL, Java, Bash
- **Machine Learning**: Supervised & Unsupervised Methods, Transformers (BERT, GPT, T5), LLM Prompting, RAG, XGBoost, LightGBM
- **MLOps & Cloud**: AWS (EC2, S3), GCP (BigQuery, Cloud Functions), Docker, Kubernetes, Airflow, Jenkins
- **NLP**: NER, Text Classification, Document Extraction, Generative AI, Hugging Face, Spacy, LangChain, FAISS
- **Data & Distributed**: Apache Spark, DVC, Postgres, Hadoop, Ray
- **Visualization**: Tableau, Matplotlib, Plotly, Looker Studio, PowerBI
- **Other Tools**: Git, Jupyter Notebooks, Weights & Biases, Selenium

CERTIFICATIONS

Certified AWS AI Practitioner *2024*

Google Cloud Program *30 days of Google Cloud Program Certification* *2021*

Shape AI *Data Structures and Algorithm using C++ Certification* *2021*