

DIANA MININE GUDINHO

+1 716-939-5682 | dianagudinho16@gmail.com | [LinkedIn](#) | [GitHub](#) | New York

SUMMARY

Data Scientist with 4+ years of experience designing AI solutions that drive business impact across healthcare, supply chain, and retail. Skilled in deploying LLMs, RAG, and multi-agent systems to improve decision-making and automate knowledge-intensive workflows. Experienced in building end-to-end pipelines and cloud-native deployments that deliver scalable, production-ready AI while translating complex data into actionable strategies.

EDUCATION

University at Buffalo <i>Master of Science in Data Science (CGPA: 3.7/4.0)</i>	New York, USA <i>Aug 2023 – Feb 2025</i>
Visvesvaraya Technological University <i>Bachelor of Engineering in Information Science (CGPA: 3.5/4.0)</i>	Karnataka, IN <i>Aug 2016 – Aug 2020</i>

WORK EXPERIENCE

Research Assistant University at Buffalo	<i>Feb 2025 – Present</i>
---	---------------------------

- Improved diagnostic QA accuracy by **22%** on **10K+ MRI** scans by combining embedding-driven retrieval with **fine-tuned LLMs** to surface relevant prior cases for radiologists.
- Prototyped **multi-agent GenAI workflows** that coordinated retrieval, reasoning, and summarization tasks, reducing validation by **30%**.
- Optimized model training with distributed **PyTorch** and GPU scheduling, cutting training cycles by **25%**.

Graduate Student Assistant University at Buffalo	<i>Aug 2024 – Dec 2024</i>
---	----------------------------

- Developed a customer intelligence solution using **LLM-based text generation** and **prompt engineering** to analyze unstructured campaign data, driving an **18% increase** in marketing **ROI**.
- Designed a **RAG pipeline** to integrate structured and unstructured records, enabling **semantic search** and faster reporting by **45%**.

Data Analyst Tata Consultancy Services Ltd.	<i>Oct 2020 – June 2023</i>
--	-----------------------------

- Built **demand forecasting** models using **XGBoost** and **LSTM** on **50K+ SKUs**, improving stock accuracy, which cut procurement spend by **\$470K** annually.
- Designed **pricing and margin** models on transactional data, uncovering low-performing SKUs and guiding adjustments that lifted product **profitability** by **12%**.
- Engineered pipelines on **AWS Glue, Redshift, and S3**, processing **5M+** daily records to give executives near real-time visibility.
- Applied **clustering** and **anomaly detection** on **100K+** distribution routes to uncover inefficiencies, **reducing delays** by **17%** and improving fleet utilization.
- Conducted **regression** and **root cause analysis** on supplier and claims data, improving compliance report accuracy by **25%** while also lowering financial risk exposure.
- Built **Tableau** dashboards for fulfillment and cost **KPIs**; cut refresh time by **40%** and improved stakeholder adoption by **30%**.
- Automated retraining workflows and **CI/CD** pipelines, cutting release cycles by **30%** and improving production model reliability.
- Defined KPIs with **finance, logistics, and product** teams, embedding metrics like **SLA compliance** and **forecast accuracy** into daily decisions.

Full-Stack Software Engineering Intern QSpiders	<i>Jul 2019 – Aug 2019</i>
--	----------------------------

- Optimized **PostgreSQL** queries and indexing for **8K+** sessions, cutting response times by **45%** and stabilizing peak-hour performance.
- Built automated **Excel** templates for weekly summaries, streamlining issue tracking for managers.

SKILLS

- Programming:** Python (Pandas, NumPy, Scikit-learn, PyTorch, TensorFlow, Keras), SQL, R, Scala
- Machine Learning & GenAI:** Regression, Classification, Gradient Boosted Trees, Random Forests, Neural Networks, Clustering, Anomaly Detection, Recommender Systems, Forecasting, NLP, LLMs
- Statistical Analysis:** Hypothesis Testing, Causal Inference, Bayesian Methods, Time-Series Modeling
- Pipelines & Deployment:** PySpark, Airflow, dbt, Kafka, AWS (S3, Glue, Redshift, SageMaker), GCP BigQuery, Snowflake, Jenkins
- Business Intelligence & Visualization:** Tableau, Power BI (DAX), Looker, Matplotlib, Seaborn

PROJECTS

- AI-Powered Skill Gap Analyzer:** Deployed a containerized microservice with Ollama and NLP models to identify skill gaps and recommend personalized learning paths, driving a 35% increase in feature adoption.
- RAG Compliance Copilot:** Built a retrieval-augmented assistant to auto-summarize and flag risks in regulatory documents, cutting manual review time by 35% and improving issue detection accuracy.
- Customer Retention Insights:** Integrated classifiers with LLM-driven narrative generation to produce human-readable churn insights, improving adoption of retention strategies by 15%.
- E-Commerce Product Catalog Enrichment (LLM Ensemble):** Fine-tuned LLaMA + GPT ensemble for attribute extraction, boosting catalog completeness by 28%.