

DATA SCIENTIST

Results-driven Data Scientist and MS Business Analytics candidate skilled in building, validating, and monitoring predictive models including regression, boosted models (XGBoost, LightGBM), decision trees, random forests, SVM, and neural networks, using Python, SQL, and cloud platforms (GCP, AWS). Proven track record deploying GenAI pipelines that generate measurable financial value, with the ability to communicate statistical findings and written recommendations clearly to both technical and non-technical audiences.

EDUCATION

Wake Forest School of Business, Winston-Salem, NC | Master of Science in Business Analytics | GPA: **3.96/4.0** | May 2026
Honors & Leadership: Beta Gamma Sigma, Wiseman Scholar, Graduate Student Ambassador, Center for Analytics Scholar

Plaksha University, Punjab, India | BTech – Data Science, Economics & Business | GPA: **3.69/4.0** | May 2025
Honors & Leadership: Full Ride Scholar, Student Ambassador, Pioneer Student Election Committee Member

SKILLS

Languages & Development: Python, Pandas, Statsmodels, Numpy, Scikit-learn, SQL, R, Excel, Stata, Git, MS SQL Server, ETL Pipeline
ML: Clustering, Regression, Deep Learning, XGBoost, LightGBM, Decision Trees, Random Forest, SVM, Experiment Design
Data Architecture & Cloud: GCP, Vertex AI, BigQuery, AWS S3, SageMaker, Microsoft 365, Azure Synapse Analytics, PostgreSQL
Advanced Analytics & AI: Tableau, Power BI, Neural Networks, Hypothesis & A/B Testing, LLM APIs, GenAI, Agentic AI, RAG, KPI
Strategy & Operations: Data Storytelling, Stakeholder Management, Agile Development, Supply Chain Analytics, Project Management

EXPERIENCE

Wake Forest University | *AI Engineer* | Winston-Salem, NC Mar 2026 – Present
▪ Developing Gemini-powered conversational chatbot using LLM APIs and prompt engineering to deliver personalized, interactive learning of the Ethical Decision-Action Making (EDAM) framework, serving 500+ students across the university.
▪ Building GenAI-driven personalization workflows optimizing admissions engagement for 2,000+ prospective students.

Castlery | *Graduate Consultant* | Winston-Salem, NC Oct 2025 – Present
▪ Architecting a data-driven recommerce decision tree integrating cost, vendor performance, and operational constraints to replace manual 2-vendor selection, reducing cost per returned product by \$500 through optimized routing logic.
▪ Developing robust predictive models to identify anomalies and patterns in operational data, reducing friction by 20% while translating models into actionable business strategies.

Benori Knowledge | *Data Scientist* | Gurgaon, India Jan 2025 – May 2025
▪ Headed a data science solution to deploy an automated pipeline for dataset curation, gathering and cleaning data from various sources to save 25 hours weekly and generate \$20K monthly.
▪ Applied statistical sampling and hypothesis testing to ensure unbiased model validation while enhancing feature engineering and ML pipeline logic, improving F1-score by 12% and reducing EOD by 18%.

Founder Forward | *Data Analyst* | New York, NY Jun 2024 – Aug 2024
▪ Built a Python-based ETL workflow to clean, transform, and load 5K+ records, improving data reliability by 45%.
▪ Developed 4 interactive dashboards in Tableau and Python for decision-makers, reducing time-to-insight by 40%.
▪ Partnered with operations teams to automate six database processes using JavaScript, enhancing system efficiency by 93%.

SBSF Agritech | *Data Scientist* | Berlin, Germany (Remote) Jul 2023 – Aug 2023
▪ Compiled and preprocessed 22K+ data rows to build a scalable data enrichment pipeline supporting growth curve modelling.
▪ Configured growth curves for 180 tree species in 40 agro-ecological zones, achieving a mean R-squared value of 0.86.
▪ Optimized ML models through Agile sprints, improving performance by 23%, strengthening species suitability assessments.

PROJECTS

DataCamp Data4Good 4th Annual Analytics Competition | *Team Lead* Oct 2025 – Jan 2026
Developed 12 NLP and generative AI models using embeddings, XGBoost, and LLM ensembles to classify answer factuality (0.97 macro-F1).

Wake Forest Data Analytics Challenge | *Tech Lead* Sep 2025 – Dec 2025
Engineered 10 fraud detection models using large datasets, achieving 84% capture rates and generating \$3.6M in annual fraud savings.

PovertyPred: Deep Learning Project | *Machine Learning Developer* Jan 2024 – May 2024
Constructed a hierarchical graph convolutional network (HGCN) to predict household poverty levels in Kenyan villages from 90K street images, achieving an accuracy of 91% (*Project nominated for paper publication as 1 of 7 top submissions*).

NOTABLE MENTIONS

- * 2026 **First Runner-Up** (East Coast Region) DataCamp’s 4th Data4Good Analytics Competition, competing against **120 teams**.
- * 2025 Recipient of the **Wiseman Full-Ride Scholarship (\$85, 000)**, Wake Forest University MSBA Program.
- * 2024 **Libra Internship Program** Select from a global pool with an acceptance rate of 2.8% for the summer internship, New York.
- * 2024 **Top 1% in NPTEL Database Management System** Course in a pool of **7134 students**.
- * 2024 **Caseify Winner:** Business Case Study Competition judged by consultants from BCG, India.