ASHRITHA GUGIRE

gugireashritha3@gmail.com • LinkedIn • +1 (949) 204-6952 • Virginia (Ready to Relocate)

SUMMARY

Data Scientist with 4+ years of experience building scalable Machine Learning solutions across pharmaceutical, financial, and commercial domains. Expert in Python, SQL, and cloud platforms with proven ability to deploy production models and collaborate with engineering teams. Leading IARPA-sponsored research on anomaly detection using ensemble models and MLOps. Skilled in translating business requirements into data-driven solutions, A/B testing frameworks, and communicating insights to technical and non-technical stakeholders.

TECHNICAL SKILLS

- **Technologies:** Python (SciKit-learn, Matplotlib, TensorFlow), R, SQL, Excel (VBA, pivot tables, array functions, power pivots), PowerBI, Tableau, Looker, Hadoop, Hive, Git, S3, Databricks, Jupyter Notebook, PyCharm, Visual Studio Code, Jira
- **Statistical & ML Expertise:** Advanced Statistical Modeling, Causal Inference, Time Series Analysis, Ensemble Learning, Anomaly Detection, Predictive Analytics, A/B Testing, Experimental Design, Hypothesis Testing.
- **Production Deployment:** Model Deployment, MLflow, Version Control (Git), Agile/Scrum Methodologies, Software Development Lifecycle (SDLC), Automated Testing
- **Data Management**: SQL Server, MySQL, PostgreSQL, AWS (S3, SageMaker, EC2), GCP, Hadoop, Hive, PySpark, Streaming Data

EXPERIENCE

Deloitte | Data Analyst | Virginia **January 2024 - Present**

- Designed and deployed scalable ML pipelines using Python (Pandas, NumPy, Scikit-learn) to analyze 500K+ behavioral datasets with automated model monitoring and performance optimization.
- Built and trained ensemble ML models (Random Forest, Isolation Forest) reducing fraud detection time by 40% and enabling real-time anomaly identification across 16K+ agents
- Collaborated with engineering teams to deploy production models using Git, Docker, and Agile methodologies throughout the software development lifecycle
- Designed A/B testing frameworks with statistical validation (t-tests, chi-square) improving user engagement by 20% and informing $2M+ product roadmap decisions
- Engineered feature extraction techniques from multi-source datasets, improving model performance by 23% across business units
- Communicated findings through interactive Tableau visualizations and executive presentations.
- Automated reporting workflows with data governance and lineage tracking, improving operational efficiency by 20%

Infosys | Data Analyst | India **July 2020 - December 2022**

- Developed predictive models using Python (Pandas, Scikit-learn) and SQL for supply chain optimization, boosting stock availability by 25% and reducing inventory costs
- Built data infrastructure processing 500M+ records with Spark Streaming, reducing batch processing time by 60% for real-time commercial operations
- Applied machine learning algorithms (K-Means, DBSCAN, Random Forest, XGBoost) for customer segmentation, improving identification accuracy by 10%
- Created executive Tableau dashboards visualizing 50+ KPIs across revenue and operations, accelerating timelines by 25%
- Conducted consumer research through 50 interviews and 400 surveys, uncovering commercial opportunities worth $250K annually
- Engineered ETL pipelines with data quality checks and version control practices ensuring data integrity across business units
- Documented processes and methodologies following software development best practices for team collaboration

ACADEMIC PROJECTS

January 2024 - December2024

Urban Mobility Analysis| LLM's, Gradient Boosting, Random Forest, Logistic Regression

- Delivered insights on road hierarchy, time-of-day trends, and route efficiency using transformer models and prompt engineering.
- Established MLOps workflows with version control, MLflow tracking, and automated retraining to maintain model performance.
- Collaborated with technical teams to implement semantic network graphs using graph-based ML techniques for route classification and behavioral analysis, enabling advanced anomaly detection capabilities

Medical Image Classification| ResNet, Pytorch, TensorFlow

- Developed and evaluated deep learning models using PyTorch and TensorFlow to fine-tune ResNet-X101 architecture for skin cancer classification, achieving 80%+ accuracy through transfer learning and data augmentation
- Implemented model optimization **techniques** to address class imbalance challenges and enhance generalizability on unseen datasets, following software development lifecycle best practices
- Enhanced model generalizability on unseen samples through robust augmentation and transfer learning techniques.

EDUCATION

George Mason University| Master of Science, Data Analytics and Computational Science

Coursework: Advanced Machine Learning, Decision Analytics, Marketing Analytics, Computing, Optimization, Demand/Pricing Analytics

Vignana Bharathi Institute of Technology | Bachelors in Computer Science and Engineering