WANSHIKA PATRO

+1-716-390-2392 | wanshika85@gmail.com | LinkedIn | Github

SUMMARY

Data Science graduate with a strong foundation in large-scale machine learning, agentic workflow development, and scientific modeling. Experienced in building AI pipelines, optimization algorithms, and LLM-integrated systems. Skilled at developing predictive models, statistical simulators, and production-grade analytics for payment, fraud detection, and compliance domains. Seeking to contribute to Amazon Payments' mission through high-impact, scalable AI research and deployment.

EDUCATION

- **University at Buffalo, SUNY** | Buffalo, NY
  Master of Science in Data Science | *Jun 2025*
  Relevant Coursework: Machine Learning, Distributed Computing, Database Management Systems, Deep Learning Bayesian Networks, Scalable Pipelines

- **KIIT University** | Bhubaneswar, India
  Bachelor of Technology in Computer Science | *Aug 2021*
  Undergraduate Projects: Credit Card Fraud Detection, Object Recognition using Neural Networks

TECHNICAL SKILLS

- **Programming:** Python, Java, SQL, C++, R, Hive, REST APIs
- **ML/AI Frameworks:** Scikit-learn, TensorFlow, PyTorch, Hugging Face Transformers, XGBoost, Autoencoders, RNNs, LLMs
- **Optimization & Modeling:** Time Series Forecasting, Anomaly Detection, Reinforcement Learning, Statistical Simulation, SHAP
- **Tools & Platforms:** Docker, Flask, Streamlit, GitHub Actions, CI/CD, Tableau, Power BI
- **Data Infrastructure:** Oracle SQL, PostgreSQL, BigQuery, ServiceNow, PySpark

CORE COMPETENCIES

Agentic AI Workflows | Large-Scale Machine Learning | Deep Learning Models | Optimization Algorithms | Generative AI (LLMs) | Scientific Modeling | Risk Simulation | Real-Time Prediction Systems | Time Series Forecasting | Statistical Inference | Reinforcement Learning | SQL & Data Warehousing | Production-Grade ML Deployment | Distributed Computing | Anomaly Detection |Cross-Functional Collaboration

PROFESSIONAL EXPERIENCE

Application Development Analyst
ACCENTURE - Pune, India| *Jul 2021 - May 2022*

- Automated 100K+ ServiceNow API-based logs, improving SLA compliance by 40% across transaction pipelines.
- Designed 3+ backend analytics dashboards monitoring 10K+ records/month for real-time SLA visibility and availability.
- Conducted root-cause simulations on 3K+ recurring incidents using SQL + Python, cutting repeat failure patterns by 22%.
- Partnered with 6+ global teams across 3 time zones to align platform risk flows, improving data observability by 30%.

PROJECTS

Stock Volatility Risk Simulation | *Python, Docker, SHAP, Streamlit* | Jan 2024 – May 2024

- Designed agentic volatility workflows simulating synthetic trade risks across 100K+ pricing scenarios.
- Benchmarked real-time vs batch inference pipelines, improving model response latency by 38%.
- Strengthened containment strategies, reducing simulated systemic risk by 40%.

Revenue Forecasting & Regulatory Insights | *XGBoost, Streamlit, Tableau* | Feb 2025 – Present

- Modeled 145K+ payment records to forecast revenue across business flows with 92% accuracy.
- Delivered 4+ regulatory dashboards with dynamic filters and alerts, improving decision latency by 20%.
- Enabled stakeholders to detect anomalies 30% faster with interpretable ML outputs.

Spelling Error Clustering for Transaction Cleaning | *BERT, Autoencoder, LLMs* | Jan 2024 – May 2024

- Clustered 50K+ transaction log anomalies using unsupervised BERT embeddings to improve data consistency.
- Embedded feedback-based audit loop using Streamlit, reducing manual correction effort by 35%.
- Improved downstream data pipeline quality by over 30% through pre-cleaning validation.

Toxic Comment Moderation Pipeline (LLM Workflow) | *Hugging Face, Transformers, Flask* | Jan 2025 – May 2025

- Built moderation engine with switchable BERT layers, classifying 5 classes with 88% F1 score.
- Developed real-time web app UI reducing moderation cycle time by 30%.
- Integrated policy violation flags, increasing content coverage across 3 moderation tiers by 20%.

Crime Trend Monitoring & Risk Alerts | *PySpark, MLlib, Pandas* | Aug 2024 – Dec 2024

- Processed 1M+ regional event records to identify geographic and behavioral clusters.
- Created 5+ heatmaps and dashboards to visualize high-risk areas, reducing risk planning delays by 25%.
- Built anomaly detection logic that flagged 12% more real-time incident deviations compared to static rules.

CERTIFICATION & LEADERSHIP

- **Supervised Machine Learning –** Stanford/DeepLearning.AI *(In Progress)*
- **Career Skills in Data Analytics –** LinkedIn Learning
- **Meta Kaggle Hackathon –** Identifying predictive trading signal patterns *(Jun 2025 – Present)*
- **Digital Strategy Lead – MUN Leadership Team | Volunteer Analyst – Animal Rescue NGO**