

# ANIKET DESHPANDE

Data Scientist | Machine Learning Engineer | NLP & Generative AI

+1 (480) 796-0191 | adeshp61@asu.edu | [linkedin.com/in/aniketdeshpande791/](https://www.linkedin.com/in/aniketdeshpande791/) | [github.com/Aniket19j8](https://github.com/Aniket19j8) | [aniketdeshpande.me](https://aniketdeshpande.me)

## PROFESSIONAL SUMMARY

Data Scientist with 2+ years of experience building production-grade machine learning and Generative AI systems, including predictive modeling, RAG pipelines, and real-time data workflows. Proven track record of deploying scalable ML solutions that improve business outcomes, reduce latency, and enable data-driven decision-making.

## TECHNICAL SKILLS

**Programming & Querying:** Python, SQL, R, PySpark

**Machine Learning:** Scikit-learn, XGBoost, TensorFlow, Keras, Optuna, Regression, Classification, Clustering, Time Series (ARIMA, Prophet)

**Generative AI & NLP:** LLMs, Transformers (BERT, T5), RAG, Semantic Search, Prompt Engineering, NER

**Data Engineering:** Apache Spark, Hive, Presto, Databricks, ETL Pipelines, Data Modeling

**MLOps & Deployment:** Docker, CI/CD, Model Monitoring, Drift Detection, Experiment Tracking

**Cloud:** AWS (S3, EMR, SageMaker), Azure ML

**Visualization & BI:** Tableau, Power BI, Matplotlib, Plotly, Excel

## EDUCATION

**Arizona State University** | **M.S. Data Science, Analytics & Engineering** | GPA: 3.8/4.0 Tempe, AZ | Aug 2024 – May 2026

Relevant Coursework: Machine Learning, Statistical Modeling, Big Data Analytics, Probability, Cloud Computing

**MIT World Peace University** | **PG Diploma, AI & Machine Learning** | GPA: 3.7/4.0 Pune, India

**MIT World Peace University** | **Bachelor of Computer Application** | GPA: 3.8/4.0 Pune, India

## PROFESSIONAL EXPERIENCE

**EdPlus at Arizona State University** | **Instructional Design Assistant – Analytics** Scottsdale, AZ | Aug 2024 – Present

- Built and deployed machine learning models (Python, scikit-learn) to identify at-risk students using behavioral and engagement data, improving early-risk detection accuracy by **30%** and enabling proactive academic interventions
- Engineered scalable ETL pipelines (Python, SQL, Hive) processing **10K+ daily events**, reducing data latency by **25%** and improving data availability for downstream analytics and ML workflows
- Conducted advanced EDA, statistical modeling, and feature engineering on large-scale datasets, uncovering key engagement drivers that informed course redesign and improved student retention outcomes
- Developed dashboards (Tableau, Power BI) tracking KPIs (engagement, completion, performance) and performed QA aligned with instructional design standards, improving data-driven decision-making and learning experience consistency

**Suzlon Energy Pvt. Ltd.** | **Data Analyst – Predictive Analytics** Pune, India | May 2021 – Aug 2021

- Developed predictive maintenance models (Random Forest, XGBoost) on **500GB+ SCADA time-series data**, achieving **95% failure detection accuracy** and reducing unexpected turbine downtime
- Built real-time anomaly detection pipelines using Python and SQL, decreasing manual diagnostic efforts by **25%** and improving operational monitoring efficiency
- Designed and deployed Tableau dashboards visualizing KPIs across **200+ wind turbines**, enabling engineering teams to optimize maintenance schedules and improve asset utilization

## PROJECTS

**Fraud Detection with Drift-Aware Adaptive ML - Fintech System** | Python · XGBoost · RF · LR · PCA · KS-Test · PSI · Scikit-learn

- Built an ensemble fraud detection system (XGBoost, Random Forest, Logistic Regression) on **280K+ transactions**, achieving **83% recall** while reducing false positives in high-risk scenarios
- Implemented concept drift detection (KS-test, PSI) with automated retraining and designed a production-ready ML pipeline with monitoring, evaluation, and scalable deployment architecture

**Stockio - Financial Forecasting & Market Anomaly Detection Platform** | Prophet · ARIMA · VADER · SHAP · Azure DevOps · HF

- Developed time-series forecasting (ARIMA, Prophet) and sentiment analysis (VADER) pipelines on multi-source financial data (market + news), achieving 93% anomaly detection precision and enabling proactive market signal identification
- Built explainable ML workflows (SHAP) and scalable data pipelines (API ingestion + preprocessing), reducing false alerts by 40% and latency by 35%, supporting real-time analytics for 1,000+ users

**Hybrid RAG Ticket Resolution System – Enterprise GenAI Capstone** | Py · ChromaDB · BM25 · MiniLM · Cross-Encoder · Ollama · Qwen3

- Built an end-to-end RAG pipeline (ChromaDB, BM25, cross-encoder reranking) combining dense + keyword retrieval to enable high-accuracy semantic search and significantly improve ticket resolution speed and relevance
- Leveraged LLMs with prompt engineering and context-aware retrieval to automate response generation, reducing manual support workload and enabling scalable enterprise-grade ticket resolution systems

## CERTIFICATIONS & ACHIEVEMENTS

- AWS Certified Machine Learning – Specialty | Google Data Analytics | IBM Data Science | HarvardX (PH525.4x) | Applied NLP (Infosys)
- Claude Hackathon @ ASU (2026) | Dean's List | Top 3% Academic Rank
- Research Assistant - Speech Emotion Recognition (NLP, Deep Learning)