

# Deven Shah

New York | 716-573-4760 | deven.rshaha@gmail.com | LinkedIn | Portfolio | GitHub

## Education

### University at Buffalo

Master of Science in Data Science (GPA: 3.4/4.0)

Buffalo, NY

August 2024 – December 2025

### MIT World Peace University

Bachelor of Technology in Computer Science and Engineering (GPA: 9.01/10.0)

Pune, IN

November 2020 – July 2024

## Experience

### Data Scientist / Researcher

March 2025 – Present

University at Buffalo

Buffalo, NY

- Developed an end-to-end admissions ML pipeline processing 3,000 applications per cycle, reducing approximately 300–400 manual review hours and admissions processing time by around 10% for a pilot CSE department using AHP-based interpretable feature engineering.
- Standardized and clustered 100+ academic major titles, reducing preprocessing effort by 20% for admissions data preparation workflows still in testing using TF-IDF, K-Means, and fuzzy matching pipelines.
- Trained Random Forest and GPU-accelerated PyTorch models on approximately 20,000 applications, improving minority-class recall and F1-score by 15% during model validation using Focal Loss to address class imbalance.
- Achieved 85.1% accuracy, 95.9% recall, and 90.7% F1-score across roughly 20,000 applications, automating approximately 20% of rejection reviews while improving fairness via feature removal using LIME-based model explanations.

### Python Developer

January 2024 – June 2024

Markytics

Pune, IN

- Optimized Django-PostgreSQL backend by refactoring complex ORM queries, improving response times by 30% and cutting query execution cost by 25% through indexing and caching strategies.
- Integrated REST APIs for internal and external systems, reducing data latency by around 10% and improving load performance for large datasets in production applications using standardized API-based communication mechanisms.
- Implemented code review checklists, improving review throughput by 50% and reducing post-deployment issues across a five-developer engineering team by enforcing clean-code and review standards.

## Projects

### AI-Powered SQL Analytics Agent

December 2025 – Present

- Engineered a Text-to-SQL Agent achieving 93.8% accuracy on 18-query benchmarks, reducing query latency to 3.03s and boosting interpretability through hybrid BM25 + semantic retrieval with rule-based clarification.
- Developed a “What-If” Simulator enabling multi-variable pricing and discount analysis with real-time sensitivity modeling, empowering decision-makers to test 4+ scenario tiers safely via non-destructive SQL CTE execution and historical comparison.
- Implemented a Chain-of-Thought SQL Engine solving 75% of multi-hop reasoning tasks, advancing autonomous query reasoning through recursive CTE generation and window functions (LAG) for analytical tasks spanning 4+ relational joins.
- Containerized the full analytics stack achieving reproducible deployment and CI/CD automation, improving scalability and monitoring transparency through Docker, GitHub Actions, and self-initializing TPC-H + ChromaDB environments.

## Certifications & Publications

- **AWS Certified Machine Learning Engineer Associate** “(Amazon Web Services) – Issued January 2025.
- **Shah, D.** et al. “Deep Learning for Exoplanet Exploration.” – ICDAI Springer Nature 2025.

## Technical Skills

**ML/AI:** Regression, Classification, Ensemble Models, LSTM, ANN, CNN, LIME/SHAP

**Deep Learning & NLP:** PyTorch (CUDA), TensorFlow/Keras, Sequence Modeling, Sentiment Analysis, TF-IDF

**Data Engineering & MLOps:** PySpark, ETL, SQL Optimization, MLflow, Airflow, Git, CI/CD

**Cloud/DevOps & Databases:** AWS, GCP, Docker, Kubernetes, PostgreSQL, MySQL, MongoDB

**Programming:** Python, SQL, Java, JavaScript