Tejashri Kelhe

+1(979) 344-9009 | tejashri.kelhe@gmail.com | LinkedIn | Github | Research & Innovation

EDUCATION

Master of Science, Data Science *Aug 2023 - May 2025*
Texas A&M University *College Station, Texas*
Distributed Computing, Optimization, Statistical Analysis, Scientific Computing, Computational Statistics, Advanced Analytics, Applied Math, Big Data Mining, Data Stream Algorithms, Machine Learning, Deep Learning, NLP, Reinforcement Learning

Bachelor of Technology, Electronics and Telecommunication Engineering *July 2017 - July 2021*
Cummins College of Engineering for Women *Pune, India*
C++ Programming, SQL, DBMS, Operating Systems, OOP, Data Structures, Algorithms, Signal Processing, Linear Algebra

WORK EXPERIENCE

Texas A&M AgriLife Research – Machine Learning Engineer *Remote, USA · Sept 2025 - Current*
• Built and deployed distributed Machine Learning models simulating spatiotemporal disease spread in livestock to guide biosecurity interventions; improved scenario coverage by 40% and robustness to stochastic variance by 25% (PyTorch, Spark).

Micron Technology – Machine Learning Engineer *India · July 2021 - Feb 2023*
• Developed a latency-critical defect detection model (Autoencoders, Computer Vision, Python) running under 30 ms on embedded devices, demonstrating expertise in developing end-to-end ML systems to meet autonomy-grade SLAs.
• Built a real-time recommendation system utilizing Graph ML & Bayesian optimization to dynamically model causal dependencies in semiconductor chip fabrication, leveraging exabyte-scale data to reduce defective chip manufacturing by 30%.
• Trained and optimized ML models for early yield prediction & fab parameter optimization, improving storage-related data processing throughput by 20% through distributed training and efficient cloud resource utilization.
• Integrated ML inference into factory control via REST APIs on a C++ backend, enabling secure & low-latency interoperability.
• Defined project deliverables and led code/design reviews for distributed ML pipelines across global teams.

AT&T – Data Scientist (Intern) *Dallas, USA · June - Aug 2024*
• Built a low-latency anomaly detection microservice (LSTM + XGBoost) with ranking and fail-safe fallbacks, incorporated customer feedback into evaluation loops, reducing late payments by 15%.
• Designed and deployed a real-time fraud risk scoring API (FastAPI, AWS Lambda, Bayesian Neural Network, Python) serving millions of requests per day with MLflow-driven A/B testing.
• Engineered scalable Spark SQL pipelines on Databricks ingesting 10M+ rows per day from Snowflake to Delta Lake; added drift and data quality checks for time-ordered telemetry.
• Automated CI/CD pipelines (MLflow, Docker, A/B testing), accelerating safe experiment-to-production deployment by 35%.
• Built real-time observability dashboards (Plotly, Tableau) to track system health, latency metrics, and anomalies.

Texas A&M University – Machine Learning Researcher *College Station, USA · Jan 2024 - May 2025*
• Researched Bayesian LSTM forecasting models for geospatial yield prediction, improving accuracy by 87% while providing robust uncertainty quantification for decision support.
• Benchmarked cross-domain audio/image ML models with MIT Lincoln Laboratory using CUDA and distributed GPU clusters.
• Mentored 150+ students in ML, Data Structures and Algorithms as a TA, fostering rigorous problem-solving skills.

PATENTS, RESEARCH PUBLICATIONS & JOURNAL PEER-REVIEW

U.S. Patent : "Attention-Based Feature Selection in DNNs" **(2024), Micron Technology**
Journal Peer-Reviewer: *Expert Systems with Applications (Elsevier)* - On Invitation, critically assessed a deep learning manuscript on Underwater Acoustic Target Recognition utilizing CNN, Attention architecture & Few-Shot learning **(2025)**
IEEE: "Cross-Domain Knowledge Transfer for Acoustic Classification" **(2025),** "Collision Avoidance in Self-Driving Vehicles Using Deep Learning and PSO" **(2021),** "A Smart Early Warning System for Prevention of Disease Outbreaks" **(2020)**

CORE SKILLS

Programming & Systems: Python, R, SQL, C++, C, Bash, Linux, Git, Spark, Databricks, Networking
ML/AI: Regression, Classification, Deep Learning (CNNs, LSTMs, Transformers), Graph ML, Bayesian ML, Recommender Systems, Causal Inference, Anomaly Detection, Explainable AI, Advanced Statistical Analysis, GANs/Diffusion
MLOps & Cloud: GCP, Azure ML, AWS, CI/CD, Docker, Kubernetes, BigQuery, MLflow, REST APIs, Profiling, Embedded ML, Snowflake, Hadoop

PROJECTS

LLM Bias Detection (Responsible AI): Developed interpretable NLP pipelines (BERT + SHAP + LIME) to surface hidden bias and unanticipated behaviors in large language models (F1 ≈ 0.93), aligning with principles of trustworthy and explainable AI.
GenAI Research: Researched and prototyped GAN-based Generative AI simulators for synthetic ECG generation (in collaboration with Baylor Scott & White Health), improving cardiac anomaly detection precision by 22%.
Health Accessibility Tech: Built an open-source audio-visual accessibility library for visually impaired users (Python, PyAudio).
Forgery Detection (Security & Trust): Built a CNN-based deepfake classifier for enhancing media authenticity (F1 ≈ 0.96).