

# FARHEEN AKHTER

[linkedin.com/in/farheen-akhter](https://www.linkedin.com/in/farheen-akhter)

feenu.akhter@gmail.com

510-377-2570

<https://farheen-akhter-23.github.io/my-portfolio/projects>

## EDUCATION

MS Computer Science and Engineering | California State University | 2023 - 2025 | GPA: 3.9

## PROFESSIONAL EXPERIENCE

### CALIFORNIA STATE UNIVERSITY (AI & ML Lab)

San Bernardino, CA

#### Research Assistant

2024 – PRESENT

CSUSB's CSE Department serves 1,200+ students, specializing in applied AI, cybersecurity, and data science research. Contributed to Responsible AI initiatives by reducing LLM hallucinations, streamlining fine-tuning, and co-authoring a CAI 2025 paper on healthcare bias mitigation.

- Reduced hallucinations in academic QA by 40% by building a domain-aligned RAG pipeline with Milvus and Lang Chain, improving response quality in early deployments to 3 internal departments.
- Accelerated LLM fine-tuning workflows by 60% on GCP/HPC clusters by automating preprocessing, orchestrating training pipelines, and embedding SHAP for downstream explainability audits.
- Built pet re-identification system using real-time facial recognition, achieving 95.3% accuracy and reducing false positive rate by 25% through iterative model refinement and CV tuning.
- Led hands-on mentoring of 5 junior contributors to build reproducible ML pipelines, run bias attribution tests, and open-source explainability tools aligned to Responsible AI goals.
- First-author on accepted CAI 2025 paper surveying bias calibration frameworks for instruction-tuned LLMs, with a focus on NLP in healthcare and bridging fairness metrics with domain-specific utility.

#### Teaching Associate

2024 – PRESENT

- Designed and taught labs on ETL pipeline engineering, LLM prompt tuning, and vector DB integration (Milvus, FAISS) to equip students for real-world GenAI infrastructure roles.
- Led machine learning instruction using project-based modules on regression, classification, and transformer-based NLP, improving student model deployment rates by 40%.
- Delivered live coding walkthroughs and TA support in CS fundamentals, boosting hands-on engagement through interactive demos, Git-based assignments, and Slack peer support threads.
- Facilitated advanced algorithm workshops in Python and C++, covering recursion, graph search, and DP strategies tied to industry coding interviews and production-level challenges.

### MOROCCAN DATA SCIENTIST

Remote

#### Data Scientist Intern

Feb 2009 – May 2011

Moroccan Data Scientist is an AI-driven nonprofit focused on sustainable development, with initiatives in agriculture, healthcare, and education across North Africa. Worked with data scientists to build a crop disease detection model, real-time GCP pipelines, and fine-tuned LLMs for agronomic advisory tools.

- Built crop disease detection model using ResNet on field images, achieving 92% accuracy and enabling up to 25% reduction in preventable crop loss across pilot farms.
- Processed 100K+ annotated crop images using scalable GCP pipelines, delivering real-time inference and reducing image latency for field diagnostics by 40%.
- Engineered ETL pipelines to transform 1TB+ of soil and crop telemetry data, cutting training runtimes by 35% and enabling faster model iteration cycles.
- Fine-tuned open-source LLMs on domain-specific agronomy corpora, improving generation accuracy and retrieval relevance by 30% for farmer advisory content.

- Deployed AI-powered forecasting models into BI dashboards, raising crop yield prediction accuracy by 15%, and supporting early resource allocation decisions for agribusiness stakeholders.

## X CUBE LABS

Hyderabad, India

### Senior Software Engineer

July 2021 – Aug 2023

*xCubeLabs is a 500+ employee digital product engineering firm with \$30M+ revenue, delivering solutions across healthcare, real estate, agritech, and e-commerce. Led cross-functional teams on full-stack builds with Azure CI/CD, real-time dashboards, and intelligent automation for scalable platform delivery.*

- Built Power BI dashboards for agronomy ops, reducing manual reporting by 80% and enabling field managers to act on real-time crop KPIs and environmental signals.
- Led 12-person team to ship live crop monitoring platform, integrating IoT feeds, task scheduling, and yield tracking; boosted operational efficiency across farms by 50%.
- Developed a high-traffic real estate portal with map search, dynamic filters, and rec systems, increasing session duration and property inquiries by 25%.
- Architected OAuth2-based role access and JWT-based session auth for multi-tenant apps, ensuring 100% compliance with internal and external security standards.
- Delivered modular e-commerce engine serving 10K+ MAUs with sub-200ms latency on cart and checkout flows improved user retention by 30% over two quarters.

## VSN Technologies

Hyderabad, India

### Associate Software Engineer

Nov 2015 – Jun 2020

*VSN Technologies is a niche software firm with under 10 employees, focused on digital solutions for urban planning and insurance. Grew from intern to full-stack developer, building high-load approval systems and APIs using Java and ReactJS.*

- Developed building plan approval system for insurers and city agencies, reducing manual reviews by 60% through digital submissions, status tracking, and automated rule enforcement.
- Built secure, policy-compliant APIs enabling cross-agency data exchange between municipal systems, insurance providers, and planning authorities with zero compliance breaches.
- Designed stakeholder dashboards for architects, inspectors, and admins using Angular, reducing onboarding time by 40% and increasing self-service task completion.
- Refactored backend job queues and DB access patterns to support 3x peak load concurrency, boosting system throughput by 45% without additional compute resources.

## TECHNICAL SKILLS

**Languages & Programming:** Python, C++, Java, JavaScript (React.js, Node.js), GraphQL

**AI/ML Frameworks:** PyTorch, TensorFlow, Keras, scikit-learn

**LLMs & NLP:** LLM Fine-Tuning, RAG Pipelines, Prompt Engineering, Transformers (Hugging Face)

**MLOps & DevOps:** MLflow, Docker, Azure DevOps (CI/CD), GCP

**Backend & APIs:** Flask, Django, Spring Boot, REST, FastAPI

**Databases & Data Systems:** MongoDB, PostgreSQL, Milvus, FAISS

**Cloud & Infra:** Google Cloud, Azure, Vertex AI, HPC Clusters

**Tools & Platforms:** Power BI, Git, Jupyter, LangChain

## RESEARCH & PUBLICATION

“Bias Mitigation in Deep Learning: A Survey of Modern Techniques” – Conference on Artificial Intelligence (CAI 2025), Accepted for presentation in May 2025. Led a comprehensive survey of modern bias mitigation strategies across model training, data preprocessing, and post-hoc explainability in deep learning and LLMs.