Muhammad Umar Khan

Email: m.umark2002@gmail.com | (226) 506-9244 | LinkedIn| GitHub | Portfolio

EDUCATION

Wayne State University | Detroit, MI, USA **Jan. 2025 – Aug. 2026**
MS in Computer Science GPA: 3.93

University of Windsor | Windsor, ON, Canada **Sep. 2020 – Apr. 2024**
BS Honours Computer Science with Software Engineering Specialization

TECHNICAL SKILLS

Machine Learning & AI: Deep Learning (CNN, GCN), Computer Vision, Image Segmentation, Clustering (K-Means), Feature Engineering, RAG, Prompt Engineering, Model Evaluation, TF-IDF
Frameworks & Libraries: PyTorch, TensorFlow, Scikit-learn, NumPy, Pandas, Matplotlib, FastAPI, Flask
Languages: Python, Java, SQL
Databases & Storage: PostgreSQL, AWS RDS, AWS S3
DevOps, Cloud & Tools: Docker, Git, GitHub Actions, CI/CD, AWS CloudFront, REST APIs
AI Platforms & Models: Ollama, LLaMA 3.2, Gemini API, Vertex AI (in progress)

EXPERIENCE

University of Windsor| Windsor, ON, Canada **Apr. 2023 – Nov. 2024**
Research Assistant

- Built an image segmentation prototype in Python (NumPy, SciPy, scikit-learn) that partitions images into regions using centroid-based assignment and K-Means color clustering with Euclidean distance matrices, plus visualizations of regions and centroids.
- Studied and presented a published CNN+GNN superpixel classification paper (SLIC radius graphs, GAT, hybrid loss) and reviewed the authors' open-source implementation.
- Implemented a Keras CNN on MNIST and a 2-layer GCN in PyTorch Geometric on Cora (~81% test accuracy), with t-SNE embedding visualizations; documented both in technical presentations.

University of Windsor| Windsor, ON, Canada **Sep. 2022 – Apr. 2024**
Teaching Assistant

- Assisted 100+ students through code reviews and debugging sessions across Data Structures, OOP, and Database Management Systems.
- Led weekly lab sessions for 30+ students on Java, data structures, and algorithms, while grading 25+ weekly assignments and exams.

PROJECTS

MindForge AI Suite (StudyAI + ResearchAI) | Python, Flask, HTML, Ollama, LLaMA 3.2, RAG, TF-IDF, ThreadPoolExecutor

- Architected a fully local LLM platform (Flask, LLaMA 3.2 via Ollama) analyzing 6 document formats (PDF, DOCX, PPT/PPTX, TXT, MD) with zero API cost and no data leaving the machine.
- Engineered a RAG pipeline with TF-IDF retrieval over 400-word overlapping chunks, passing only top-ranked passages to the model to reduce prompt context across multi-document research sessions.
- Parallelized per-document summarization across 4 workers (ThreadPoolExecutor) with real-time SSE streaming; integrated DuckDuckGo web retrieval and automated APA citation generation.

K-Means Market Basket Analysis | Java, Apriori, K-Means

- Built Apriori (candidate generation, pruning, support counting) and supporting algorithms from scratch in Java, as part of a 2-person team developing a K-Means pre-filter that clusters items on revenue and co-purchased revenue.
- Benchmarked against Apriori-only and a published clustering method on 130K+ records: ~10x faster at the largest input (~7 hrs → ~45 min) with comparable top-100 rule value.

Transfer Learning & Domain Generalization | Python, PyTorch, ResNet-50, EfficientNet-B3

- Ran 90 controlled fine-tuning experiments (ResNet-50 and EfficientNet-B3 × CIFAR-10/100 and brain tumor MRI × 3 data sizes × 5 freeze depths); built the PyTorch pipeline and layer-freezing logic.
- Found freezing half the backbone costs <1 pt (CIFAR-10: 96.7% → 96.0%) while cutting training time ~29%; reached 95.9% on brain tumor MRI.