VISHAL KRISHNA KUMAR

Seattle, WA | (206) 843–4386 | vishal-krishna.me

vishalkrishnakkr@gmail.com | github.com/Vishal-Krishna-Kumar | linkedin.com/in/vishal-krishna-kumar

EDUCATION

City University **Expected Graduation: June 2027**

Master of Science, Computer Science (Data Science) – Dean's Scholar **Seattle, WA**

- **Relevant Coursework:** AI Cloud Computing, NLP, Large Language Models, Secure Systems, Networking, VR/AR, HCI

EXPERIENCE

Research Assistant – Applied Machine Learning & Intelligent Systems

Graduate Researcher **August 2025 – Present**

- Built modular Retrieval-Augmented Generation (RAG) pipelines with hybrid dense-sparse retrieval, reducing LLM hallucinations by 30% and improving factual grounding by 35% via structured chunking and graph-based reasoning
- Developed deep learning models for CV and NLP in PyTorch, improving task performance by 25% through architecture iteration, hyperparameter tuning, and efficient training pipelines
- Authored reproducible experiment workflows and research-grade documentation, reducing debug and rerun time by 40%

Walmart – Software Engineering Virtual Experience

Software Engineer Intern **Remote – January 2026**

- Implemented a custom heap-based priority queue in Java for shipping logistics scheduling, reducing average operation latency by 30% with O(log n) inserts/extracts in high-volume workflows
- Produced UML-based data processing architecture enabling modular execution modes and database-backed services for multi-workflow retail simulations
- Designed a normalized relational schema with ER modeling for a retail pet department, improving query efficiency by 25% and ensuring consistency across 10+ entities

BCG – Data Science Virtual Experience

Data Science Intern **Remote – August 2026**

- Built an end-to-end churn prediction pipeline in Python (Pandas/NumPy), performing EDA and feature engineering on 10K+ customer records to extract high-signal behavioral features
- Trained a Random Forest model for imbalanced data, increasing recall from 32% to 50% via threshold tuning, class-weight calibration, and error analysis
- Presented retention insights by identifying top churn drivers using feature importance and model evaluation metrics

PROJECTS

GraphAugmented Intelligence – *1st Place, SCE Hacks* **Python, GPT-4, LLaMA, Neo4j | September 2025**

- Designed prompt-aware RAG combining Knowledge Graph reasoning with LLMs, reducing hallucinations by 40% and improving factual grounding over 27M nodes and 53M edges
- Implemented graph-traversal retrieval to extract minimal context-relevant subgraphs (vs. static chunks), enabling explainable biomedical QA
- Engineered end-to-end pipeline with vector DBs, multi-LLM orchestration, and CLI execution; evaluated on BiomixQA benchmarks with improved reasoning accuracy and consistency

3D Attention UNet++ Brain Tumor Segmentation – *2nd Place, Hack for Research* **PyTorch, BraTS | April 2025**

- Built a 3D Attention UNet++ for multimodal MRI segmentation using 3D CNNs, skip connections, and attention gates to learn volumetric spatial features
- Achieved Dice score >0.97 across tumor subregions on BraTS via optimized preprocessing, augmentation, and tuning, improving segmentation accuracy by 20%

Deepfake Detection System – *Published in TDEMAS* **TensorFlow, XceptionNet, OpenCV | December 2024**

- Developed a deepfake detection framework using XceptionNet, improving detection accuracy by 35% on highly manipulated video datasets via optimized TensorFlow training and feature extraction
- Published results in TDEMAS, demonstrating impact on media authenticity, ethical AI practices, and AI-driven forensics

NutriChat – Local RAG Pipeline – *Cup Winner* **LLaMA, FAISS, NVIDIA GPU | August 2024**

- Built a local RAG pipeline for privacy-preserving document QA with FAISS vector search and PDF ingestion, enabling textbook querying on local GPU infrastructure
- Improved response accuracy and reduced hallucinations by grounding generations with similarity-retrieved factual passages during inference

SKILLS

Languages: Python, C++, Java, SQL, JavaScript

ML/AI: Machine Learning, Deep Learning, NLP, Computer Vision, Reinforcement Learning, RAG, GraphRAG

Frameworks: PyTorch, TensorFlow, Scikit-learn, OpenCV, Hugging Face, LangChain

Tools & Platforms: AWS, Docker, Git, Linux, PostgreSQL, MongoDB, Vector Databases, Neo4j, Jupyter