Niketh Bayya Mahesh

(812)671-1867 nbayyam@iu.edu https://www.linkedin.com/in/niketh-bayya/ myPortfolio

Education

Indiana University Bloomington **Aug 2023 – May 2025**
Master of Science in Computer Science *Bloomington, IN, USA*

SRM Institute of Science and Technology **Jun 2019 – May 2023**
Bachelor of Technology in Computer Science and Engineering with specialization in IoT *Kattankulathur, India*

Experience

Indiana University Bloomington **Sep 2024 – Apr 2025**
AI Research Assistant *Bloomington, IN*

- Introduced an AI-driven content generation platform with GPT-based LLMs, cutting summarization time by 50%.
- Optimized embedding retrieval (Hugging Face Transformers), implementing vector search and fine-tuning embeddings, boosting search efficiency by 30%.
- Created a serverless AI pipeline (AWS Lambda, S3, API Gateway) for scalable, cost-efficient inference. Enhanced LLM responses via reinforcement learning and prompt engineering, increasing coherence by 35%.

Lennox International Inc **Jun 2024 – Aug 2024**
Software Developer Intern *Richardson, TX*

- Engineered a high-performance geospatial search engine using PostgreSQL, PostGIS, and spatial indexing, optimizing queries across a 193-million-record dataset and reducing execution time by 40%.
- Designed and architected a scalable .NET API in C# on Azure, implementing asynchronous processing and multi-threading, enabling seamless integration and efficient high-throughput spatial queries.
- Refined geolocation algorithms with Haversine, Vincenty's formulas, and quadtree partitioning, improving search accuracy by 25% and reducing computational overhead by 40%.
- Implemented search ranking with ML-driven dealer recommendations, leveraging clustering and anomaly detection on historical query patterns to improve result relevance and personalization.
- Boosted search logic migration to an Azure-hosted SQL database, utilizing parallel query execution and indexing strategies, enhancing system scalability and fault tolerance.

Projects

Intelligent QA System Using CRAG | *LangChain, LangGraph, OpenAI, Tavily, LanceDB, LLM* **Feb 2025**

- Created a pipeline using LangChain for modular orchestration, integrating OpenAI GPT-4 for query refinement and LanceDB for efficient vector-based document retrieval, achieving a 30% improvement in context relevance.
- Implemented ChatOpenAI-based document grading to assess the relevance of retrieved content, leveraging Pydantic for data validation and Tavily API for real-time web search integration to ensure high-quality answers.
- Defined end-to-end processing with LangGraph for flexible model chaining, reducing cycle time by 20%, and boosting semantic search speed by 40%, enabling scalable, efficient QA with enhanced response accuracy.

IAB Taxonomy Classification | *Python, Neo4j, PyTorch, Transformers, RAG* **Dec 2024**

- Led a team of 3 to design and implement a Neo4j-based RAG framework, developing graph-based semantic retrieval techniques. Achieved 70% Tier-1 and 43% Tier-2 accuracy in IAB Taxonomy classification.
- Improved precision in IAB taxonomy classification by 15% using LLM-driven synonym generation, advanced Natural Language Processing techniques, and weighting-based node refinement.
- Implemented a RESTful API for data ingestion and retrieval, enabling efficient querying of the knowledge graph and scalable deployment of the RAG framework.

Image and Audio Enhancement using Machine Learning | *Matplotlib, Librosa, STFT, ICA, SVM, DCT* **Mar 2024**

- Constructed an ML-based signal processing framework, enhancing feature extraction accuracy by 20% and improved overall signal clarity through optimized processing techniques.
- Applied spectral analysis and deep learning models, improving noise suppression by 35% across diverse datasets.
- Engineered a modular Python package, optimizing reusability and integration into larger ML pipelines, reducing code redundancy by 40%. And, leveraged advanced algorithms for real-time noise suppression.

Technical Skills

Languages/Developer Tools: Python, C/C++, C#, JavaScript, HTML/CSS, SQL, Git, Docker, VS Code, Visual Studio
Frameworks/Libraries: React.js, Node.js, Next.js, Spring Boot Flask, Fast API, .NET Core, Neo4j (Graph Data Modeling), NLTK, Hugging Face Transformers, Pandas, NumPy, Matplotlib, Scikit-Learn, TensorFlow, PyTorch, CUDA
Databases/Cloud: MySQL, PostgreSQL, PostGIS, MongoDB, Neo4j, Microsoft Azure, AWS
Machine Learning/AI: RAG, Linear/Logistic Regression, Clustering, Classification, SVM, Reinforcement Learning, Neural Networks, RNNs, LSTMs, Transformers (BERT, GPT-4o), LLMs, CRAG
Certification: AWS Certified Machine Learning Engineer - Associate