ABYLAY AMANBAYEV

+1 (209) 500-8452 ◇ Merced, CA
aabylay@gmail ◇ linkedin ◇ github ◇ amanbayev@ucmerced.edu

SUMMARY

Pursuing a PhD in Computer Science focusing on Agentic LLM/LMMs, RAG applications, and Vector Databases. Currently seeking 2025 Summer Internships.

SKILLS

Languages	`Python, SQL, C/C++, HTML/CSS,` LaTeX
DS/ML	`Pandas, NumPy, PyTorch, TensorFlow, Sklearn, Matplotlib, HuggingFace, LangChain`
DBMS	`PostgreSQL, MySQL, PG-vector, MilvusDB, Pinecone, PASE`
Tools	`SciPy, BS4, Docker, Git, Bash`

EDUCATION

University of California Merced Aug 2023 - Present
PhD in Computer Science
- Scientific Advisor: Prof. Florin Rusu
- Courses: DBMS Implementation, Deep Learning, Reinforcement Learning, Approximation Algorithms

National Research University Higher School of Economics Feb 2024
Master of Data Science
- Courses: Applied Machine Learning, SQL, Computational Complexity, Advanced Algorithms

EXPERIENCE

University of California Merced Aug 2023 - Present
Research/Teaching Assistant
- Developing query optimization strategies for vector databases
- Evaluation of indexing algorithms and designing benchmarking tools for RAG applications
- TA-ship for the course CSE 111-Database Systems (Fall 2023, Fall 2024)

Lawrence Livermore National Laboratory July-Aug 2024
Team Lead in Data Science Challenge
- Led a team of two undergraduates and one MS student in cardiac research utilizing machine learning methods
- Developed algorithms for cardiovascular diseases diagnostics with over 98% accuracy
- Trained neural networks to reconstruct cardiac activation and transmembrane potential maps from 12-lead ECG data

PROJECTS

Multi-Purpose RAG Application Aug 2024 - Present
- Developing multi-purpose agentic RAG application with foundational LLM/LMMs: `ChatGPT-4o, Llama-2/3`
- Fine-tuning LLM/LMMs using PEFT and few-shot learning
- Integration with Vector Databases and Embedding Models: `PG-Vector, Milvus, OpenAI, Mistral, BGE`

RESEARCH

1. A. Amanbayev, A. Datta, B. Tsan, F. Rusu. Query Optimization Strategies for Vector Databases *(in progress)*

2. A. Amanbayev, A. Datta, B. Tsan, F. Rusu. A comprehensive overview of Vector Databases and ANNS with Filters *(preparing for submission)*

3. Y. Izenov, A. Datta, B. Tsan, A. Amanbayev, F. Rusu. Spanning Tree-Based Query Plan Enumeration, arXiv: 2403.04026v1 [cs.DB] 6 Mar 2024