Thura Win Kyaw

Yuseong-gu Gung-dong, Daejeon,KR

thurawk@gmail.com | github.com/thurawk

Profile

Aspiring AI/ML engineer with hands-on experience in building RAG-powered assistants, developing full-stack AI applications, and designing LLM-based conversational agents. Skilled in transforming research ideas into practical systems.

Research Interests

I am passionate about artificial intelligence, machine learning, and conversational AI, with a focus on Retrieval-Augmented Generation (RAG), chat-bot development, and the integration of vector databases to enhance user interaction and contextual understanding. Committed to developing intelligent systems that deliver accurate, context-aware, and user-centric solutions for real-world applications.

Education

Chungnam National University, B.E. in Electrical Engineering — Sept 2022 – Present

- **Coursework:** Computer Programming (C), Computer Scientific Thinking, Linear Algebra, AI and Future Society, Modern Control Theory and Design, Sensor and Measurement Engineering
- **Certifications:** HarvardX: CS50's Introduction to Programming with Python, CS50: Introduction to Computer Science, CS50's Introduction to Artificial Intelligence with Python

Experience

AI, Software, Prompt Engineer, AIdenBio – Daejeon, KR — Dec 2024 – Present

- Designed and maintained backend systems with a focus on AI model integration, API development, and database management
- Engineered and optimized prompts for Large Language Models (LLMs) to ensure personalized, context-aware responses across various use cases
- Developed and deployed Retrieval-Augmented Generation (RAG) pipelines to enhance the accuracy and relevance of AI-generated outputs
- Contributed to building intelligent, production-grade systems for life sciences and biomedical applications

AI Software Engineer Intern, GRINDA AI – Daejeon, KR — Oct 2025 – Nov 2025

- Built an AI-powered Slack bot that converts issue reports into structured GitHub issues using Claude AI and FastAPI, with auto-labeling, translation, and monitoring

Projects

ChatPDB – Protein Structure Visualization Assistant (LLM + RAG + PyMOL)

- Built an assistant that interprets natural language queries to visualize protein structures with LLM + RAG and PyMOL rendering.
- Implemented a retriever pipeline (BM25 + FAISS + Cross-Encoder) using UniProt annotations and enabled session-based workflows.
- Integrated FastAPI services and a CLI using local LLMs (Ollama) for interactive usage.

Iris Flower Classification (SVM)

- Built an SVM classifier with StandardScaler preprocessing, GridSearchCV tuning, and evaluation via accuracy, confusion matrix, and classification report.
- Visualized decision boundaries with PCA and saved the best model pipeline for reuse.

Customer Segmentation (K-means)

- Applied K-means clustering on synthetic customer data (age, income, spending) to identify actionable segments.

- Used StandardScaler preprocessing, elbow and silhouette analysis for k selection, and cluster visualizations; saved the trained model and scaler.

Text Classification

- Built a supervised text classifier with TF-IDF features and multiple models (Naive Bayes, Logistic Regression, SVM) with GridSearchCV tuning.
- Evaluated with accuracy and precision/recall/F1, visualized results, and saved the model, vectorizer, and label encoder for deployment.

Quote Sentiment Analysis

- Built a sentiment analysis pipeline for short quotes by scraping data with requests and BeautifulSoup, labeling with VADER, and training Naive Bayes and Logistic Regression models.
- Evaluated with confusion matrices and visualizations like sentiment distributions and word clouds.

Skills

Programming Languages: Python, C, MATLAB.

AI/ML:

- Retrieval-Augmented Generation (RAG)
- LangChain
- Prompt Engineering
- Vector Databases
- LLM-Powered Chatbots
- Local LLM Deployment (Ollama)

Tools & Platforms:

- Development: VS Code, Cursor AI, Jupyter Notebook, PyMOL
- Version Control: Git, GitHub
- Productivity: Notion, Google Workspace, Microsoft Office
- Model Hosting & Inference: Ollama, Hugging Face, Transformers Inference API

Operating Systems: Windows, Linux, Mac.

Languages: Burmese (Native), English (Fluent), Korean (TOPIK 5).

References

Available upon request.