

# YEABSIRA GIRMA

AI / ML Engineer | NLP & Generative AI | LLMs | Prompt Engineering | Data Engineering

 Email: [yeabsiragirma5@gmail.com](mailto:yeabsiragirma5@gmail.com)

 LinkedIn: [linkedin.com/in/yeabsira-girma-hailegiorgis](https://linkedin.com/in/yeabsira-girma-hailegiorgis)

 Portfolio: <https://portfolio-website.com>

 GitHub: <https://github.com/YEAB-G>

---

## PROFESSIONAL SUMMARY

AI/ML Engineer specializing in LLM application development, NLP pipelines, prompt engineering, and data engineering. Experienced in building end-to-end AI systems including RAG chatbots, generative applications, optimization toolkits, and analytical pipelines. Strong foundation in Python, ML workflows, data processing, and deploying production-ready AI tools. Passionate about developing real-world AI solutions with measurable impact.

---

## PROFESSIONAL EXPERIENCE

### DeepAxis Computing

AI Engineer & LLM Application Developer  
Accra, Ghana | 2021 – Present

- Designed and built LLM-powered applications including a RAG PDF chatbot, AI email styler, and a prompt optimization toolkit using Groq Llama models.
- Implemented backend LLM logic, structured system prompts, few-shot examples, relevance ranking, and reusable prompt templates.

- Developed RAG-style retrieval pipelines using embeddings, vector search, and context injection for grounded question answering.
- Created a Prompt Optimization Playground to benchmark prompt performance using latency, token usage, and output consistency.
- Built Streamlit frontends with custom theming, animations, and responsive layouts for web and mobile.
- Deployed applications to Streamlit Cloud, configuring secrets, runtime environments, and monitoring workflows.

**Skills:** LLMs (Llama 3), Groq API, Prompt Engineering, RAG, NLP, Streamlit, Python, Model Evaluation, REST Integration, Git/GitHub, UI/UX for AI tools

---

## AlgoMine Research Group

**Machine Learning Engineer & Automation Specialist**  
Addis Ababa, Ethiopia | 2018 – 2021

- Developed text preprocessing pipelines for NLP tasks including tokenization, normalization, lemmatization, and feature extraction.
- Built sentiment analysis and review analytics workflows for fintech application data, identifying customer themes and actionable insights.
- Prepared clean, labeled datasets for experimentation, classification, and downstream ML models.
- Explored embeddings, similarity metrics, and vector search as foundations for modern retrieval-augmented systems.
- Automated ML workflows and preprocessing steps using Python and Bash, improving reproducibility and consistency.
- Documented preprocessing assumptions, metadata, and dataset versions to maintain auditability and experiment tracking.

**Skills:** NLP, Text Preprocessing, Sentiment Analysis, Python, Pandas, Scikit-learn, Embeddings, Vector Stores, Bash Automation, Data Labeling

---

## DataWave Software Solutions

**Data Engineering & Backend Assistant**  
Addis Ababa, Ethiopia

- Contributed to ETL pipelines processing multi-format datasets (CSV, JSON, logs, sensor streams).
- Wrote Python scripts for data cleaning, validation, transformation, and quality checks.
- Performed exploratory data analysis on time-series and tabular datasets to identify trends and anomalies.
- Assisted in configuring cloud storage, indexing structures, and dataset access layers for analytics teams.
- Standardized schemas, naming conventions, documentation, and data lifecycle workflows.

**Skills:** Data Engineering, ETL, Python, Pandas, Data Cleaning, EDA, Time-Series Analysis, Cloud Storage, Documentation

---

## PROJECTS

### RAG PDF Chatbot

- Built a document-aware chatbot using embeddings + vector search + Groq Llama models.
  - Implemented retrieval logic, context injection, and PDF parsing for accurate responses.
- Tech: RAG, LLMs, Streamlit, Python, Groq API

## AI Email Styler

- Developed a generative AI tool that converts rough messages into polished emails across tones and styles.  
Tech: LLMs, Prompt Engineering, Streamlit

## Prompt Optimization Playground

- Built a benchmarking tool to compare prompt quality, latency, and token efficiency across models.  
Tech: Streamlit, Groq API, LLM Evaluation, UI/UX

## Solar Energy Analytics

- Performed EDA on solar production data, visualized trends, detected anomalies, and created analysis dashboards.  
Tech: Python, Pandas, Data Cleaning, Visualization
- 

## TECHNICAL SKILLS

AI & ML: LLMs (Llama 3.x), NLP, RAG, Generative AI, Prompt Engineering

Languages: Python, SQL, Bash

Frameworks: Streamlit, FastAPI (basic), Scikit-learn

Data Engineering: ETL, Pandas, Numpy, Data Cleaning, EDA

Tools: Git, GitHub, VS Code, Postman

Deployment: Streamlit Cloud, API Secrets Management

Other: Vector Search, Embeddings, Token Estimation, UI/UX for AI Apps

---

## EDUCATION

### BSc in Computer Science and Engineering

#### AI/ML Engineering Program

Specializations: NLP, LLMs, RAG, Data Engineering, Deployment, Prompt Optimization

---

## CERTIFICATIONS

- Microsoft Azure Fundamentals (AZ-900 or equivalent)
  - AWS Certified Cloud Practitioner
  - 10 Academy Applied AI & ML Sprint Training
- 

## CONTACTS

 Email: [yeabsiragirma5@gmail.com](mailto:yeabsiragirma5@gmail.com)

 LinkedIn: [linkedin.com/in/yeabsira-girma-hailegiorgis](https://linkedin.com/in/yeabsira-girma-hailegiorgis)

 GitHub: [github.com/YEAB-G](https://github.com/YEAB-G)

 Portfolio: *Add your link when deployed*