

Muhammed Hisham C

Machine Learning Engineer

[LinkedIn](#) | [GitHub](#) | [Kaggle](#) | [+91 9778477369](#) | [izamdeveloper1@gmail.com](#) | Kochi, Kerala

SUMMARY

As a Machine Learning Engineer, Having developed many software products utilizing different ML and non-ML tools and APIs. My curiosity is mainly in generative AI. Projects include ML methods for NLP, Computer vision, and GenAI problems. Experience in LLM libraries such as Langchain and Llama-index along with LLMs, Embedding models, agents, Vector Databases, etc.

PROJECTS

Ragrank - [GitHub](#)

- An Opensource evaluation library of Large Language Models (LLM) and Retrieval Augmented Generation (RAG) models.
- Developed in Python with strict Pydantic validation and strict linting with ruff on top of flake-8. Tested 96% of code with Pytest with GitHub actions.
- Have 4 pre-defined metrics for response and context in the RAG. Can create custom metrics for user's preference. All of the metrics are tested and benchmarked with a hugging face RAG dataset.
- Ingest data with multiple sources powered by pandas and can customize the internal LLM with Ragrank Langchain and llama index LLM wrapper.
- Detailed and user-friendly [documentation](#) and [API references](#) are available.

GemInsights - [GitHub](#)

- Objective: Build a tool that gives general insights from a CSV file to help the data analysts understand the data with a few clicks using LLM.
- Visualized the data and gave that to the multimodal Large Language Model Gemini-1.5 in vertex ai. (developed this 1 week after the Gemini release)
- Prompt engineered for getting better results concerning the data and evaluated the response with Trulens Eval. Improved the response by 20% by using 5-shot prompting.
- deployed the project using Streamlit and hosted it in Hugging Face.

TECHNICAL SKILLS

Libraries/Frameworks

- **Machine Learning:** Scikit-learn, XGBoost, LightGBM, CatBoost, Tensorflow, Keras, PyTorch
- **Natural Language Processing:** SapCy, NLTK, Gensim, Transformers by Hugging Face, TextBlob
- **Data Visualization:** Matplotlib, Seaborn, Plotly
- **Computer Vision:** OpenCV, Pillow
- **LLM:** Langchain, Huggingface Transformers, Llama index
- **MLOps:** MLflow, Zenml
- **Version Control:** DVC, Git
- **Database:** PostgreSQL, MongoDB

Familiar with

- Statistics and Mathematics (for ML)
- EDA and Preprocessing
- Machine Learning algorithms
- Data Structures
- Deep Learning Architectures (Neural Networks, CNN, RNN, GAN)
- Natural Language Processing (NLP)
- Computer Vision (CV)
- Large Language Models (LLM) and GAN
- Generative AI.
- MLOps
- ([refer to this repo](#) for mini projects)

EDUCATION

Machine Learning

Brototype, Kochi 2023 - Present