

Anoop Krishna

Data Scientist

anoopkrishna.k9@gmail.com | linkedin.com/anoop | github.com/anoop2752

Professional Summary

Data Scientist with a strong foundation in Python, machine learning, and statistics, experienced in implementing GenAI-driven solutions to transform software development. Proficient in GPT, Claude, Copilot, and AI architecture, with expertise in automating workflows, minimizing errors, and accelerating delivery. Committed to driving next-gen innovations that boost efficiency and business impact.

Experience

Data Scientist (Rubix, Bengaluru)

Jan 2024-Feb 2025

- Designed AI-driven systems automating architecture generation, boosting software development speed and consistency.
- Leveraged LLMs such as GPT, Claude, and Mistral to solve complex business software challenges.
- Developed predictive coding models improving time-to-market and elevating customer satisfaction rates significantly.
- Led R&D initiatives to explore emerging AI trends and integrate them into core development pipelines.
- Conducted peer code reviews, ensuring adherence to best practices and optimizing code efficiency.

PROJECTS

AI-Based Face Recognition System

[github.com/repo](https://github.com/anoop2752/face-recognition)

- Designed an AI pipeline using LLaMA and Groq API for real-time face recognition and tagging.
- Automated identity verification by integrating LLM responses with facial feature embeddings and classification.
- Optimized API calls for low-latency face detection, enabling high-speed recognition across multiple input sources.
- Built a modular AI architecture to decouple image preprocessing, embedding generation, and LLM-based inference.
- Employed zero-shot learning capabilities of LLaMA to identify players without pre-labeled training data.

RAG – Powered Document Intelligent System

[github.com/repo](https://github.com/anoop2752/rag-system)

- Developed a Retrieval-Augmented Generation (RAG) pipeline to extract contextual answers from business PDFs.
- Used embeddings and vector databases to retrieve relevant content chunks for each user query.
- Integrated GPT-based LLMs for generating human-like answers from unstructured document inputs.
- Automated document processing with real-time question answering, reducing manual analysis efforts.
- Applied prompt engineering and context management to ensure high-quality, grounded LLM outputs.

Skills

- Programming & Tools:** Python (Pandas, NumPy, Scikit-learn, Matplotlib, Seaborn), SQL
- Machine Learning & AI:** Supervised Learning, Deep Learning, Generative AI, Computer Vision
- Large Language Models (LLMs):** GPT (OpenAI), Claude (Anthropic), LLaMA (Meta), Mistral
- GenAI Tools:** GitHub Copilot, Cursor.sh, v0.dev
- Frameworks & Libraries:** Transformers (HuggingFace), LangChain, FAISS, OpenCV, scikit-learn, NumPy, Pandas
- Cloud & Deployment:** Google Colab, Streamlit Sharing, API Integration
- Soft Skills:** Problem-Solving, Rapid Prototyping, Research-Oriented, Innovation-Driven, Continuous learning.

Education

University of Calicut

B.Sc., Computer Science

- Coursework:** Database Management Systems, Probability, Statistics, Computer Networks

Certification

- Certified Data Scientist- DataMites**
Completed real-world projects covering machine learning, deep learning, and data analysis. Gained practical skills in model deployment, data preprocessing, and performance tuning.
- Certified Data Scientist – IABAC (International Association for Business Analytics Certification)**
Globally recognized certification validating skills in data science, statistical analysis, and predictive modelling using industry-standard tools and techniques.