

Anoop Krishna

anoopkrishna.k9@gmail.com | linkedin.com/anoop | github.com/anoop2752 | +91 9633442752

About Me

Data Scientist with expertise in machine learning, statistical modeling, and data-driven decision-making. Proven ability to build predictive models and optimize data pipelines using Python, Pandas, SQL, and Power BI. Experienced in collaborating with diverse teams to automate data processes and extract actionable insights. Committed to leveraging AI and analytical techniques to solve complex business challenges.

Education

University of Calicut

B.Sc., Computer Science

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

Experience

Rubixe

Jan 2024 - Oct 2024

Data Scientist

Bengaluru

- Developed and optimized machine learning models for predictive analytics, achieving a 25% improvement in accuracy while applying statistical techniques and hypothesis testing.
- Maintained and enhanced data pipelines using Python and SQL to automate data processing and ensure seamless model deployment scalability.
- Conducted A/B testing and exploratory analysis to validate business impact, collaborating with cross-functional teams to drive data-driven decision-making.

Projects

Spotify Recommender System

github.com/repo

- Designed a system that suggests songs based on user preferences, helping them discover music that matches their taste.
- Built a solution that enhances the listening experience by providing personalized recommendations, making music exploration easier and more enjoyable.
- Tools Used: Scikit-Learn, NLTK, Spotify (Spotify API).

RAG-Based PDF QA System

github.com/repo

- Developed a Retrieval-Augmented Generation (RAG) system that enhances question-answering by retrieving relevant information from documents before generating responses.
- Implemented an efficient document processing and indexing pipeline to ensure accurate and context-aware answers.
- Tools Used: RAG, DeepSeek, LangChain, LLMs, FAISS.

FACE-RECOGNITION SYSTEM

github.com/repo

- Integrated **LLaMA-4** with the **Groq API** to generate natural language descriptions of identified players, including bio and career highlights.
- Achieved real-time face matching accuracy by leveraging face_recognition and OpenCV for image preprocessing and encoding.
- Tools Used: LLaMA-4 (via API), Groq API, Docker, OpenCV, Git, GitHub.

Skills

- **Programming:** Python (pandas, NumPy, scikit-learn), SQL
- **Machine Learning:** Regression, Classification, Clustering, Time-Series Forecasting, Deep Learning
- **Data Visualization:** Power BI, Tableau, Matplotlib
- **MLOps & DevOps:** CI/CD, Docker, Kubernetes
- **Generative AI:** Transformer Models (PyTorch, TensorFlow)
- **Cloud Platforms:** AWS, GCP, Azure

Certification

- **DataMites , Certified Data Scientist:** Gained hands-on experience in machine learning, deep learning, and data analysis through real-world projects.
- Worked on end-to-end model deployment, focusing on optimizing and scaling machine learning applications.
- Developed practical skills in handling real-world datasets, model tuning, and performance optimization.