ANURAG SINGH

📞 7000859910 ✉ anurag756singh@gmail.com Github 🌐 Portfolio Website

Education

Indian Institute Of Technology Palakkad   Nov 2021 – July 2025
Bachelor of Technology in Civil Engineering   *Palakkad, Kerala*

Experience

1stop.ai   June 2024 – July 2024
Data Science Intern   *Remote*

- Applied advanced NLP techniques including tokenization, TF-IDF, and word embeddings to create a **hate speech detection model**, achieving **92% accuracy** and an **F1-score of 0.87** for real-time toxic content identification.
- Reduced manual review effort for moderation teams by **40%** and enhanced **community safety** by accelerating detection of abusive language.

Projects

Research Tool | Python, Langchain, UnstructuredURLLoader, FAISS | Link

- Developed a URL ingestion and semantic retrieval pipeline using LangChain, UnstructuredURLLoader, OpenAI embeddings, and FAISS, enabling seamless parsing of **95%+ URLs** and improving information discovery with **70% higher accuracy** and sub-second response times.
- Integrated a ChatGPT-based Q&A interface that delivered context-aware answers with source references, reducing manual workload by **80%** and research time by **60%**, significantly enhancing productivity for researchers, analysts, and journalists.

Chicken Disease Classification | Python, Scikit-Learn, Numpy, Pandas, CI/CD, Docker | Link

- Implemented a CNN-based chicken disease classifier with advanced augmentation and optimized loss, improving accuracy by **35%**; deployed via **Flask** with a user-friendly interface.
- Incorporated **DVC** for dataset/model versioning and a **CI/CD pipeline** for automated training and deployment, enhancing reproducibility and cutting deployment time by **70%**.

Network Security | Python, Scikit-Learn, MLflow, Docker, Github Actions | Link

- Built a scalable machine learning pipeline for a network security project, covering ingestion, validation, transformation, training, and evaluation—resulting in a robust threat detection system that improved detection efficiency by **40%** and reduced manual monitoring by **50%**.
- Blended **MLflow** for experiment tracking and model versioning across **50+ iterations**, reducing experimentation time by **30%** and accelerating reliable model deployment—enhancing threat response speed by **35%**.

AI-Powered LinkedIn Post Generator | Python, NLP, Groq LLM, Streamlit | Link

- Designed an AI-powered content generation workflow using **Groq LLM** and **Streamlit**, reducing creation time by **70%** while preserving the tone and style of LinkedIn influencers through few-shot learning.
- Built NLP modules to analyze 100+ past posts, extracting **topic**, **language**, and **length**, enabling personalized and stylistically aligned post generation across **English** and **Hinglish**, boosting engagement and ensuring brand consistency.

Technical Skills

Languages: Python, SQL, C++, HTML/CSS
Libraries: Pandas, Numpy, Matplotlib, Seaborn, Scikit-Learn, Tensorflow, Pytorch, NLTK, Regex, SpaCy
Devops and Frameworks: Docker, Git, Github Actions, Github, Flask, MLflow, DVC, Flask, Streamlit
GenAI: Langchain, HuggingFace, LLM, VectorDBs, RAG

Leadership

Deputy Technical Secretary   May 2023-March 2024

* Spearheaded the technical council, organizing high-impact activities and hands-on workshops that engaged **150+ students** in IT security and programming, driving a **30% increase in skill proficiency** and fostering a culture of technical excellence.