

ANURAG SINGH

📞 7000859910 ✉ anurag756singh@gmail.com
🌐 github.com/anurag-192

Education

Indian Institute Of Technology Palakkad

Bachelor of Technology in Civil Engineering

Nov 2021 – July 2025

Palakkad, Kerala

Experience

1stop.ai

Data Science Intern

June 2024 – July 2024

Remote

- Developed and optimized a machine learning model for hate speech detection, achieving an accuracy of 92% and an F1-score of 0.87, enhancing real-time harmful content identification.
- Led the collection and preprocessing of 100,000+ social media text samples, applying advanced feature engineering techniques such as sentiment analysis to improve model robustness.

Projects

Research Assistant Tool | *Langchain,FAISS,UnstructuredURLLoader,Python*

- Developed a URL ingestion pipeline using LangChain's **UnstructuredURLLoader**, automating content extraction for **researchers and analysts**, reducing manual copy-paste workload by **~80%** and enabling seamless parsing of **95%+** URLs.
- Designed a **semantic retrieval system** using **OpenAI embeddings** and **FAISS**, allowing users to locate relevant information with **70% higher accuracy** and **sub-second response times**, significantly improving content discovery.
- Integrated a **ChatGPT-based Q&A interface** that provided instant, **context-aware answers** with **source URLs**, cutting down user research time by **60%** and enabling faster, more reliable decision-making for **journalists, researchers, and students**.

Chicken Disease Classification | *Python,Scikit-Learn,Numpy,Pandas,CI-CD*

- Developed a deep learning-based chicken disease classification model using CNNs, improving model accuracy by **35%** through advanced data augmentation and optimized loss functions.
- Integrated **DVC** for version control of datasets, models, and pipelines, increasing experiment reproducibility and modularity by **60%**.
- Deployed the model using **Flask** with a user-friendly web interface, and implemented a **CI/CD pipeline** to automate the workflow, reducing deployment time by **70%**.

Network Security | *Python,Scikit-Learn,MLflow*

- Engineered a scalable **machine learning pipeline** for a **network security project**, covering critical stages such as **data ingestion, validation, transformation, model training, and evaluation**, resulting in a highly automated and robust **threat detection system** that improved overall detection efficiency by **40%**.
- Integrated **MLflow** for experiment tracking and **model version control**, enabling seamless comparison across **50+ model iterations** and reducing experimentation time by **30%**, leading to faster deployment of reliable models.
- Empowered **cybersecurity teams** with a streamlined and reproducible ML workflow, cutting down manual monitoring effort by **50%** and enhancing threat response speed by **35%**, thus strengthening organizational security posture.

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.