Harikrishnan Gopal

✉ harikrishnangopal0411@gmail.com 📞 +91 9372296398 ⚙ GitHub in LinkedIn 🌐 harikrishnan.tech

Profile Summary

B.Tech Computer Engineering student specializing in data science and machine learning, with a portfolio of projects, internships, and hackathon experience. I work comfortably with cloud platforms and modern web development stacks, allowing me to design and deploy complete, end-to-end solutions.

Education

K J Somaiya School of Engineering   *2022 – 2026*
Bachelor of Technology in Computer Engineering
- CGPA: 9.26/10

Experience

AI/ML Intern   *Mumbai*
CDSL - Central Depository Services (India) Limited   *May 2025 – July 2025*
- Built an Azure-hosted Flask API based micro-services pipeline that ingests and cleans investor-grievance & help-desk records, standardising 40+ text fields with Python/SQL for downstream modelling
- Fine-tuned BERT, RoBERTa, and Sentence-Transformers classifiers to auto-tag complaint category and escalation probability, with F1-score of 0.87 and accelerating response SLA
- Integrated encoder-decoder (T5) and instruction-tuned LLMs (Mistral, LLaMA) from Hugging Face along with LangChain to orchestrate prompt workflows to generate suggested resolution actions.
- Risk-Score Model: merged four years of complaint frequency and turnaround metrics to predict a risk score for every DP/broker, powered by a matplotlib dashboard used by compliance teams for early-warning alerts

Software Engineering Intern   *Mumbai*
Software Development Centre (KJSCE-SDC)   *May 2024 – Dec 2024*
- Developed two MERN stack web applications to streamline faculty management: one for faculty appointments (examiner assignments, paper setting, evaluations) and another for faculty reimbursements (submission/approval workflows)
- Focused on seamless integration, high performance, and a responsive, user-friendly experience to optimize operational efficiency

Projects

SafeView   *GitHub* ↗
- Designed and fine-tuned a multimodal NSFW detection pipeline using Tensorflow on 500k labeled images, automating data cleaning and augmentation in Python. Used YOLOv5 for region detection and MobileNetV3/ResNet classifiers for refined classification, achieving a 92% F1-score.
- Engineered a JavaScript-based real-time chrome extension using Tesnorflow.js to stream and process video frames, classifying NSFW content with ¡50ms latency. Integrated asynchronous buffering for auto-blur flagged content, enforce whitelisting, and prevent visibility of NSFW frames

UniqScan   *GitHub* ↗
- Developed a web platform (FastAPI + MERN) that processes scanned or uploaded documents using OpenCV for image preprocessing (denoising, binarization, deskewing), followed by OCR with Tesseract and text extraction using NLTK and docling
- Encoded the extracted text using a fine-tuned RoBERTa and T5 models to detect plagiarised or LLM-generated content with 93% F1-score, and maintained detailed plagiarism records in a structured MySQL table
- Provided a virtual-classroom dashboard and on-demand visualization reports using Matplotlib, enabling educators to review submissions and uphold academic integrity

FinCredible

- Designed an ETL pipeline to integrate real-time market ticks and 10K+ daily news articles into MongoDB, powering user-specific feeds and portfolio analytics
- Built gradient-boosted & LSTM predictive models that improved 1-day return-prediction, exposing results via REST APIs through a FastAPI micro-service
- Visualized risk metrics and price-impact signals with Plotly dashboards, and deployed the solution on GCP for seamless scalability and faster investor action

StockSage

- Built a Flask micro-service that streams live OHLC data from Alpha Vantage, feeds it to an LSTM (TensorFlow), and serves rolling price forecasts via REST API
- Visualized actual vs predicted prices and confidence bands with Matplotlib on a responsive UI, giving traders an at-a-glance decision aid
- Deployed model via REST API using FastAPI/Flask, containerized with Docker and hosted on Azure for real-time inference

CodeSync

- Built a real-time collaborative coding platform enabling simultaneous multi-user code editing and execution using Socket.IO for live sync and Firestore for session management
- Integrated a VS Code-style editor with session sharing and remote Python code execution via Flask APIs, containerized in Docker for secure, isolated backend runtime
- Implemented a CI/CD pipeline using GitHub and GCP Cloud Build, enabling automated deployments and version-controlled infrastructure updates

Technical Skills

Programming Languages: Python, R, Java, C

Data Visualization: Matplotlib, Seaborn, Plotly, Tableau, Power BI

Data Analysis & Machine Learning: NumPy, Pandas, Scikit-learn, StatsModels, TensorFlow, Keras, PyTorch, OpenCV

Artificial Intelligence: Hugging Face Transformers, Sentence Transformer, LangChain, NLTK

Web Development & APIs: MERN Stack, Flask API, FastAPI, REST API, HTML, CSS, JavaScript

Databases & Querying: MySQL, PostgreSQL, SQL, NoSQL (MongoDB)

Big Data & Cloud Platforms: Apache Spark, Microsoft Azure, Amazon AWS, Google Cloud Platform

Tools & Technologies: Git, Jupyter Notebook, Docker, Postman

Key Achievements

Datathon (Data Science Hackathon) - Finalist

- Developed a machine learning-based system to predict customer churn in subscription-based services, leveraging predictive analytics to identify high-risk customers and recommend targeted retention strategies. Containerized the solution using Docker for consistent deployment and scalability

IIT Kharagpur Data Science Hackathon - Semi Finalist

- Built an AI pipeline that evaluates manuscript "publishability" by generating concise abstracts using BART-based summarization, extracting DocBERT embeddings, and classifying suitability using a BERT-based multi-label classifier combined with an XML-CNN head
- Engineered a real-time streaming workflow to ingest new papers, perform on-the-fly scoring, and deliver recommendations within seconds