

Gowtham G

Trivandrum | imgowthamg@gmail.com | 0963 335 72 18 | linkedin.com/in/imgowthamg
github.com/imgowthamg

Education

University of Kerala , M.Sc in Data Science	Aug 2022 – Oct 2024
• GPA: 8.49/10.0	
University Institute of Technology , B.Sc in Computer Science	July 2019 – March 2022
• GPA: 7.51 / 10.0	

Experience

Project Associate , ICFOSS – Thiruvananthapuram , Kerala	March 2025 - Present
• Developed a Malayalam-to-Indic translation system.	
• Applied machine learning and deep learning for better language understanding.	
• Enhanced computational processing of regional languages.	
Language Computing/ML Intern , ICFOSS – Thiruvananthapuram , Kerala	Sep 2024 – March 2025
• Focused on NLP challenges such as tokenization, transliteration, and embeddings tailored to complex linguistic structures.	
• Applied machine learning and deep learning techniques to enhance natural language understanding.	
Data Scientist , Upwork Freelancer , Remote	June 2023 – Sep 2024
• Specialized in data analysis, web scraping, and machine learning model development to deliver data-driven solutions.	
• Built and deployed ML models for diverse client needs, improving decision-making and automation.	

Projects

Malayalam-to-Dravidian NMT Development	github.com/repo
• Collected and created a 3 crore sentence corpus using Azure API and Bing Translator for training a Transformer-based NMT with OpenNMT and PyTorch.	
• Designed data pipelines, optimized SentencePiece tokenization , and improved vocabulary handling	
• Conducted multi-model research and internally deployed for evaluation.	
RAG-Based Chatbot with PDF Knowledge Retrieval	github.com/repo
• Developed a RAG chatbot with LangChain, FAISS, and Hugging Face for PDF-based retrieval.	
• Designed a CPU-optimized Streamlit UI with real-time PDF uploads and vector search.	
• Integrated Ollama's Gemma3:1B LLM for efficient document-based question answering.	
Twitter Hate Analyzer	
• Built a dual-stage explainable AI pipeline using BERT for hate speech detection and CNN-LSTM for category classification.	
• Enhanced model transparency with LIME for responsive, interpretable insights.	

Technologies

Languages: Python, C++, C, SQL, HTML, CSS, Shell scripting

Frameworks and Libraries: NLTK, Spacy, ONMT, PyTorch, TensorFlow, Hugging Face, Looker Studio, Numpy, Pandas, Matplotlib, Sklearn, FAISS, Langchain, FastAPI

Databases: MongoDB, MySQL

Tools and Software: Postman, GitHub, Git, Docker

Interested Areas: LLM, NLP, Generative AI, Quantum Computing, MLOps, Data Analysis