# Bairi Raju

+918179443848 rajubairi02@gmail.com LinkedIn Github

## Education

**Indian Institute of Information Technology, Kottayam** **December 2021 – April 2025**
*Bachelor of Technology in Computer Science and Engineering* ***CGPA: 8.00***

## Technical Skills

**Languages**: C++, Python, JavaScript, Java
**Frameworks**: LangChain, TensorFlow, Pytorch, Django, FastAPI
**Libraries**: OpenCV, NLTK
**Databases** :MySQL, PostgreSQL

## Experience

**Data Science Intern** **July 1, 2024 - June 30, 2025**
*Faceintel Technologies*

- As a Data Science Intern at FaceIntel Technologies, I Worked on developing a Face Anti-Spoofing model to detect fraudulent facial authentication attempts.
- Worked on real-time video streaming pipelines using Apache Kafka for multi-camera setups, enabling smooth and scalable face detection, head pose estimation, anti-spoofing, and face recognition processing.
- Worked on Milvus vector database for efficient storage and retrieval of high-dimensional face embeddings, accelerating fast similarity search in face recognition models.
- Optimized and evaluated model deployment by converting models to ONNX and TensorRT, and applying post-training quantization (FP32 → FP16/INT8), resulting in faster, more efficient real-time inference.

**Machine Learning Intern** **Nov 2025**
*Hyperworks Imaging*

- Evaluated and integrated open-source and commercial STT systems (AssemblyAI, Deepgram, Kyutai), and built a high-reliability WebSocket pipeline for real-time, low-latency audio streaming and transcription.
- Built an MCP-driven browser automation framework, orchestrated via Claude Desktop Sonnet, with UGround-72B powering vision-based DOM interpretation and GPT-120B enabling precise structured data extraction.
- Applied Vision-Language Models (VLMs) for video analytics, designing inference pipelines for automated scene understanding and event detection.

## Projects

**Model Merging for Automatic Speech Recognition** **August 2024 – April 2025**

- Developed and evaluated advanced model merging techniques on the FLEURS multilingual dataset covering 15 languages, to create high-performing joint ASR models without additional training data.
- Achieved significant reductions in Character Error Rate (CER) for merged models, with EvoMerge method delivering CER as low as 18.5 for related language pairs like Kannada-Telugu and 20.1 for Tamil-Malayalam.
- Applied CMA-ES evolutionary optimization to effectively discover optimal model merging parameters, enhancing transcription accuracy and advancing scalable multilingual ASR model integration.

**Face Anti Spoofing** | *Mobilenet* **July 2024 – September 2024**

- Developed and trained a Face Anti-Spoofing model using real-time data and the CelebA-Spoof dataset.
- Utilized MobileNet architecture for lightweight yet efficient feature extraction.
- Built the model to work frame-by-frame, enabling detection of both print and replay attacks in real-time.
- Achieved 99% accuracy, 0.99 AUC and 0.5 ACER demonstrating high effectiveness in detecting spoofing attacks.

**RAG chatbot for Medicos** | *LangChain, Faiss, Streamlit* **July 2025 – August 2025**

- Developed a Retrieval-Augmented Generation (RAG) system integrating Tesseract OCR for extracting text from scanned medical textbooks to generate high-quality embeddings.
- Leveraged OpenAI API (GPT-3.5-turbo) to generate contextually accurate, medically precise answers based on relevant retrieved content, improving chatbot response relevance for medical queries.
- Integrated a reranker module to re-score and prioritize retrieved text passages, enhancing the precision of context selection and improving the accuracy of OpenAI API generated answers.

**Image Captioning** | *vgg-16, LSTM* **March 2024 – May 2024**

- Developed an innovative image captioning model using a combination of Long Short-Term Memory (LSTM) networks and Convolutional Neural Network (CNN) architecture, specifically VGG16.
- Employed VGG16 to process images and obtain high-level visual feature representations.
- Utilized Long Short-Term Memory (LSTM) networks to train and predict captions based on extracted image features
- Integrated convolutional neural networks with recurrent neural networks to develop a model that links visual content with natural language which achieved a BLEU score of 0.52