Avinash Madhav Bhurke

avinashbhurke8@gmail.com | +91 9075807884
GitHub: https://github.com/avin-1 | LinkedIn: www.linkedin.com/in/avinash-bhurke

Skills

Languages: Python, C++, C, Java, JavaScript, TypeScript

AI/ML Frameworks: TensorFlow, PyTorch, scikit-learn, HuggingFace Transformers, LangChain

Technologies & Tools: Machine Learning, Deep Learning, Natural Language Processing (NLP), Computer Vision, Retrieval-Augmented Generation (RAG), Generative AI, LLMs, ChromaDB, FAISS, NodeJS, React, MongoDB, MySQL, SQLite3, Git, Docker, OpenCV, Tesseract

Project Work

- **Semantic Document Intelligence** (Link): Designed an offline semantic analysis pipeline for structured PDF extraction and contextual relevance ranking in the domain of Document AI and Information Retrieval. Implemented heading-based section parsing with semantic embeddings (all-mpnet-base-v2), cosine similarity ranking, and TF-IDF keyword extraction for refined summaries. Optimized for CPU-only environments, achieving sub-60s processing of multiple documents while ensuring complete data privacy. **Tech Stack:** PyMuPDF, sentence-transformers, scikit-learn, NumPy, Docker.

- **RAG Insurance Query Processor**: Developed a full-stack Retrieval-Augmented Generation (RAG) system to answer natural-language insurance queries against uploaded policy documents. Built multilingual NLP and Question Answering capabilities with document ingestion (OCR fallback), FAISS-based vector search, and Gemini LLM for structured intent parsing. Integrated JSON-based decision generation with justification and source clauses. Built React frontend for file uploads, query submission, and results visualization. **Tech Stack:** FastAPI, LangChain, FAISS, HuggingFace embeddings, React, OpenCV, Tesseract, Poppler, Gemini API, Docker.

- **Voice Over Generation Bot:** Commentary Bot with API Integration – Developed a Python-based tool to generate voiceover scripts from video frames using OpenAI and Gemini APIs. Applied Computer Vision and Video Processing (OpenCV) with text-to-speech (gTTS), automating speech synthesis and commentary generation for 100+ frames.

Work Experience

DSC | AIML-Team Member Nov 2024 – Present

- Conducted research on integrating Convolutional Neural Networks (CNNs) with hybrid neural networks, contributing to team whitepapers.
- Explored applications in Deep Learning, Computer Vision, and AI-driven automation.
- Organized speaker sessions, boosting attendance by 15% through effective publicity efforts.

EPEC | Outreach and Publicity Volunteer Aug 2023 – Jun 2024

- Supported publicity for college Techfest, increasing attendance by 20% via social media campaigns.
- Recruited 50+ participants for the tech festival, increasing the visibility of the event by 15%.

Education

Vishwakarma Institute of Technology, Pune, India | B.E. in Computer Engineering 2023 – 2027 (Expected)
CGPA: 8.63/10

Blue Bells Jr. College, India | Class XII 2022 – 2023
Percentage: 63.4%

Universal English Medium School, India | Class X 2020 – 2021
Percentage: 93.20%

Certificates and Achievements

- **Introduction to DevOps Certificate by IBM** (Link)
- **2x Scopus Indexed Research Papers Published**
- Published *"Smart Automated Table Cleaning"* research paper in a Scopus-indexed journal (Link)
- Published research paper titled *"Morse Code Based Two-Way Communication Device for Visually Impaired and Mute"* at IEEE International Conference (DICCT-2025), Graphic Era University, Dehradun (Link)
- Ranked **482 out of 5,000 students** in GeeksforGeeks Institute Coding Leaderboard