

Aditya Kumar Singh

+91-9785750064 | adityaintoi@gmail.com | portfolio | linkedin | github

EDUCATION

Vellore Institute of Technology, AP

B.Tech in Computer Science Engineering — 8.54/10

Text

Sep. 2022 – Jun. 2026

EXPERIENCE

AI Research Intern

Oct 2025 – Present

Josh Talks Pvt Ltd

Gurugram, India

- Developed production-grade fine-tuning pipelines for Whisper, Wav2Vec2, and NVIDIA ASR(Automatic Speech Recognition) models, achieving significant Word Error Rate (WER) reductions for Indic languages through rigorous model benchmarking and validation.
- Architected a TTS(Text-to-Speech) models evaluation platform to automate the validation of generative speech models, integrating MOS, CMOS, and MUSHRA frameworks with rigorous statistical hypothesis testing (p-values, confidence intervals) and A/B testing to quantify "humanness" and prosody against production baselines.
- Developed and deployed GenAI-powered internal tools by orchestrating LLM APIs (OpenAI, Gemini, Groq) to automate batch processing of audio transcripts and metadata extraction, significantly reducing manual research overhead.
- Optimized research operations by transitioning manual experiments into a scalable, automated framework, streamlining human-in-the-loop onboarding and performance analysis

Machine Learning Intern

May 2025 – Sep. 2025

TechtoGreen Drone & Robotics Pvt Ltd

Vijayawada, India

- Engineered a Pearson-correlation-driven LSTM-Attention architecture for environmental forecasting, implementing a robust data pipeline featuring 3σ outlier removal and min-max normalization that achieved a 12% reduction in RMSE over baseline models
- Led the finalist team of *Bhashini Domain Innovation Challenge*, driving end-to-end development of an AI-powered solution for multilingual accessibility, and demonstrated innovation and leadership among 200+ competing teams.
- Conducted validation experiments and exploratory data analysis (EDA) to align deep learning research with production-grade accessibility use cases in Agritech industry improving crop yields ensuring reduced environmental impacts.

PROJECTS

AI Workflow Builder | FastAPI, React, PostgreSQL, ChromaDB

Oct. 2025

- Architected a full-stack orchestration platform for custom LLM applications, featuring a drag-and-drop UI for multi-model chaining (OpenAI, Gemini) and scalable RESTful APIs for real-time execution monitoring.
- Implemented a Retrieval-Augmented Generation (RAG) pipeline using ChromaDB for vector embeddings, enabling context-aware AI responses through optimized document indexing and retrieval.

Hate Speech Detection with Fine-tuned BERT | Hugging Face, Flask

Aug. 2025

- Fine-tuned a BERT-based transformer model for multi-label toxicity classification, achieving high precision across 6 distinct categories of toxic speech
- Deployed it as an interactive Flask app providing real-time classification with a clean frontend interface.

Deepfake Detection Using EfficientNet-B0 | Pytorch, OpenCV, Torchvision

Feb. 2025

- Developed a CNN-based model achieving 90.36% accuracy and 0.9748 ROC-AUC for deepfake image detection.
- Applied interpretability techniques (confusion matrix, precision-recall, Grad-CAM) to validate and explain predictions.

SKILLS

Languages: Python, Java, Swift, C, HTML-CSS, JavaScript

Frameworks: FastAPI, Django, Streamlit, Flask, React, Hugging Face Transformers

Libraries & Toolkits: PyTorch, TensorFlow, Scikit-learn, OpenCV, Matplotlib, Seaborn, SQLAlchemy, Pandas

Speech & Audio: Whisper, wav2vec2, HuBERT, IndicWav2Vec, Librosa, torchaudio, WER/CER evaluation

Databases: MySQL, PostgreSQL, Firebase, Amazon S3

Developer Tools: AWS, Linux, Git, Docker, Postman, Google Colab, VS Code, Power BI