

# Aditya Kumar Singh

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

## EDUCATION

### Vellore Institute of Technology, AP

B.Tech in Computer Science Engineering — 8.54/10

Sep. 2022 – Jun. 2026

## EXPERIENCE

### AI Research Intern

Josh Talks Pvt Ltd

Oct 2025 – Present

Gurugram, India

- Executed end-to-end production grade fine-tuning of Whisper, Wav2Vec2 and NVIDIA ASR models on massive audio datasets, significantly improving WER for Indic languages.
- Curated and preprocessed massive audio datasets for training and evaluating several closed-source and open-source ASR models on benchmarks
- Architected and managed an end-to-end cloud-integrated TTS evaluation platform, automating the pipeline from raw audio processing to real-time data collection. Transitioned manual research into a scalable operation that streamlined human onboarding and ensured data integrity for model performance analysis

### Machine Learning Intern

TechtoGreen Drone & Robotics Pvt Ltd

May 2025 – Sep. 2025

Vijayawada, India

- Developed a Pearson-correlation-driven LSTM-Attention model to predict 7-day dissolved oxygen in aquaculture, applying linear interpolation,  $3\sigma$  outlier removal, and min-max normalization for preprocessing, improving robustness and achieving 12% lower RMSE versus a baseline LSTM.
- 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.
- Explored multilingual ASR and speech-to-speech systems as part of Bhashini Challenge, aligning with real-world accessibility use cases.

## PROJECTS

### AI Workflow Builder | FastAPI, React, PostgreSQL, ChromaDB

Oct. 2025

- Built a full-stack visual workflow builder for custom LLM applications with Retrieval-Augmented Generation.
- Designed drag-and-drop UI components supporting real-time workflow execution and multi-model orchestration (OpenAI & Gemini).
- Implemented RESTful APIs for workflow management, file uploads, and context-aware AI responses.

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

Aug. 2025

- Fine-tuned BERT on Kaggle's Toxic Comment dataset for multi-label toxicity classification.
- 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