



# SIDDHARTH UMATHE

AI & DEEP LEARNING ENTHUSIAST | RESEARCH-ORIENTED



Passionate about building intelligent systems across multiple AI domains including speech, generative AI, computer vision, reinforcement learning, and data science. Strong foundation in deep learning, sequence modeling, and AI systems design with hands-on experience in end-to-end pipelines and multi-module AI applications. Eager to contribute to impactful research and develop scalable AI systems across diverse real-world applications.

## CONTACT

✉ 22f2001536@ds.study.iitm.ac.in

☎ 7020853476

📍 Nagpur, India

🐙 github.com/output9

🌐 linkedin.com/in/siddharth-umathe

## EDUCATION

Indian Institute of  
Technology Madras

B.S. in Data Science and  
Applications

2022 – Present

## TECHNICAL SUMMARY

- Built AI systems across Speech, NLP, Vision, RL, and Data Science
- Developed end-to-end pipelines (ASR, LLM-based systems, data workflows)
- Strong focus on representation learning, sequence modeling, and system design
- Experience with model training, evaluation, and inference optimization

## ACADEMIC HIGHLIGHTS

Best Software  
Engineering  
Project Award

Built scalable AI  
systems integrating  
multiple GenAI  
modules

Hands-on experience  
across Speech, NLP,  
Vision, RL, and  
Data Science

Strong focus on  
research-oriented  
problem solving

## PROJECTS

### AI-Powered Software Engineering System (GenAI System)

Best Project Award

- Developed and integrated 8 GenAI modules: AI chatbot, multi-level summarization (week + multi-week), AI coding assistant (error detection & explanations), AI-based problem generation, and more.
- Built backend using Flask and SQLAlchemy with modular, scalable architecture.
- Integrated LLM-based pipelines into production-ready workflows.
- Focused on system design, feature scalability, and real-world usability.

### Automatic Speech Recognition (ASR) – Kaggle

- Built end-to-end ASR pipeline using wav2vec2.0 and HuBERT representations with CTC-based alignment and decoding.
- Experimented with Whisper for robust speech-to-text performance.
- Implemented efficient preprocessing (lazy loading, batching, tokenization, and GPU-optimized training workflows).
- Focused on model stability, training efficiency, and inference performance.

### Business Data Analysis – Native Chefs

- Analyzed real-world dataset to identify revenue leakage (unpaid orders) and dish-level performance trends.
- Performed data cleaning, exploratory analysis, and built analytical dashboards.
- Delivered actionable insights to support data-driven business decisions.

## RESEARCH INTERESTS

Generative AI &  
LLM Systems

Retrieval-Augmented  
Generation (RAG)

Speech  
Representation  
Learning

Reinforcement  
Learning  
Systems

AI Systems  
Design &  
Applications

# SIDDHARTH UMATHE

AI & DEEP LEARNING ENTHUSIAST | RESEARCH-ORIENTED



## TECHNICAL SKILLS



### AI / ML & DEEP LEARNING

- Deep Learning (DL)
- Natural Language Processing (NLP)
- Computer Vision
- Reinforcement Learning (RL)
- Sequence Modeling
- Model Training, Fine-tuning & Evaluation
- Generative AI
- Retrieval-Augmented Generation (RAG)



### PROGRAMMING & TOOLS

- Python
- PyTorch
- Hugging Face (Transformers, Datasets, Accelerate)
- LangChain & LangGraph
- vLLM, LlamaIndex
- FAISS, Chroma, Weaviate
- Scikit-learn, NumPy, Pandas
- Matplotlib, Seaborn
- Weights & Biases, MLflow
- Docker, Git, Linux



### PLATFORMS & TECHNOLOGIES

- FastAPI
- REST APIs, JSON
- SQL, SQLite
- Flask
- Postman
- Google Colab, Kaggle
- Jupyter, VS Code
- CUDA, cuDNN
- Linux (Ubuntu)

## COURSEWORK



Deep Learning (DL)



Deep Learning Practice (DLP – includes ASR, NLP, Vision)



Introduction to Large Language Models



Mathematical Foundations of Generative AI



Reinforcement Learning (RL)



MLOps



Data Visualization Design (DVD)



Market Research (MR)



Deep Learning for Computer Vision (ongoing)



Speech Technology (ongoing)

## CONTACT

✉ 22f2001536@ds.study.iitm.ac.in

☎ 7020853476

📍 Nagpur, India

🐙 github.com/output9

🌐 linkedin.com/in/siddharth-umathe

## EDUCATION



Indian Institute of Technology Madras

B.S. in Data Science and Applications

2022 – Present

## TECHNICAL SUMMARY

- Speech AI, Generative AI, and Deep Learning systems
- Reinforcement Learning
- Data Science: analysis, preprocessing, and model building

## ADDITIONAL PROJECTS & KAGGLE COMPETITIONS



### NLP Competition – LLM Fine-tuning with Google Gemma

- Fine-tuned Google Gemma model for a specific NLP task using parameter-efficient techniques (LoRA).
- Focused on improving instruction-following, accuracy, and robustness on domain-specific data.
- Optimized training pipeline and inference for efficient deployment.



### Computer Vision Competition – 4 × Super Resolution

- Built and trained a model to perform 4× image super-resolution.
- Used advanced CNN architectures with perceptual and adversarial losses.
- Improved high-frequency detail recovery and achieved strong PSNR/SSIM scores.



### Generative AI Competition – Train GAN Style Architecture

- Trained a GAN-style architecture to generate realistic 32×32 images from base64-encoded input.
- Focused on minimizing Fréchet Inception Distance (FID) using stabilization techniques.
- Implemented custom data pipelines, augmentation, and evaluation workflows.

## RESEARCH INTERESTS



Speech Representation Learning



End-to-End ASR (CTC, Encoder-based)



Speech + LLM Integration (SpeechLLMs)



Generative AI & RAG Systems



Reinforcement Learning for Adaptive Systems