# Siddhant Gupta

sidworks01.github.io | writetosiddhant@gmail.com | +91-7836015759 | huggingface.co/minemaster01

## Education

**Bachelors in Industrial Engineering** | *Indian Institute of Technology (IIT) Roorkee* **Oct 2022 - July 2026**

## Work Experience

**Machine Learning Intern** | **Observe.ai** **May 2025 – Present**

Co-authored a research paper (EMNLP 2025) on bias in LLMs, proposing a taxonomy of bias dimensions and an evaluation framework. Conducted research on autonomous agent systems and improved Auto-QA evaluation pipelines for enterprise NLP workflows.

**NLP & ML-Agents Community Lead** | **Cohere Labs** **Jun 2023 – Present**

Led 5+ research projects under Expedition Aya (2024–25) on reasoning evals, cultural bias in VLMs, speech synthesis from ASR, machine-generated text detection using fine-grained methods, image caption datasets, and multilingual LLMs. Designed a DAG-based open-source agentic framework for Cohere Labs and collaborated globally on agent evaluations and A2A orchestration and memory sharing as part of ML-Agent Program. Spoke at Cohere ML Summer School on Transformers and LLM evolution; drove 50+ speaker and paper reading sessions on NLP, agents, interpretability and synthetic data as part of the NLP Program.

**Research Fellow** | **Traversaal.ai** **Jan 2025 – Apr 2025**

Benchmarking data-science agents - automated graph generation along with ReAct based agent workflow and design

**Software Head** | **Artificial Intelligence and Electronic Society (ArIES)** **May 2023 - Present**

Headed teams from my college society in AI hackathons, mentoring CV and NLP research alignment, leading to 10+ successful projects including Inter-IIT participation. Organized and conducted workshops for 100+ participants on deep learning and image processing (edge detection, depth estimation, object detection, OCR).

**Research Head** | **M2ai** **May 2024 – May 2025**

Led and co-authored multiple research projects resulting in top-conference publications, including work on machine-generated text detection and LLMs/VLMs for South Asian languages. Currently improving STT/TTS models and audio tokenizers, with projects funded by over $20,000 in grants from OpenAI, Anthropic, Cohere, and others.

## Publications

**Lexical Reranking of Semantic Retrieval (LeSeR) for Regulatory Question Answering**
*Jebish Purbey, Drishti Sharma, Siddhant Gupta, Khawaja Murad, Siddartha Pullakhandam, Ram Mohan Rao Kadiyala*
*arxiv.org/abs/2412.06009* - *Accepted at RegNLP @ COLING 2025 (3rd position)*

**SeQwen at the Financial Misinformation Detection Challenge**
*Jebish Purbey, Siddhant Gupta, Nikhil Manali, Siddartha Pullakhandam, Drishti Sharma, Ashay Srivastava, Ram Mohan Rao Kadiyala*
*arxiv.org/abs/2412.00549* - *Accepted at FinNLP-FNP-LLMFinLegal @ COLING 2025 (3rd position)*

**Multilingual Hate Speech Detection and Target Identification in Devanagari-Scripted Languages**
*Siddhant Gupta, Siddh Singhal, Azmine Toushik Wasi*
*arxiv.org/abs/2412.17947* - *Accepted at Chipsal @ COLING 2025*

**Robust and Fine-Grained Detection of AI Generated Texts**
*Ram Mohan Rao Kadiyala, Siddartha Pullakhandam, Siddhant Gupta, Jebish Purbey, Kanwal Mehreen, Drishti Sharma, et al.*
*arxiv.org/abs/2504.11952* - *Under review at EMNLP 2025*

**Uncovering Cultural Representation Disparities in Vision-Language Models**
*Siddhant Gupta, Ram Mohan Rao Kadiyala, Jebish Purbey, Srishti Yadav, Alejandro Salamanca, Desmond Elliott*
*arxiv.org/abs/2505.14729* - *Under review at EMNLP 2025*

**Improving Multilingual Capabilities in LLMs with Cultural and Local Knowledge**
*Ram Mohan Rao Kadiyala, Siddartha Pullakhandam, Siddhant Gupta, et al.*
*arxiv.org/abs/2504.09753* - *Under review at EMNLP 2025*

**Evaluating Generalization Capabilities of LLM-Based Agents in Mixed-Motive Scenarios Using Concordia**
{*20+ authors*}, *Siddhant Gupta*
*arxiv.org/abs/TBD* - *Under review at NeurIPS 2025*

**Spot the BlindSpots: Quantifying Fine-Grained LLM Biases in Contact Center Summaries**
*Kawin Mayilvaghanan, Siddhant Gupta, Ayush Kumar*
*arxiv.org/abs/TBD* - *Under review at EMNLP Industry Track 2025*

**DSBC : Data Science task Benchmarking with Context engineering**
*Ram Mohan Rao Kadiyala, Siddhant Gupta, Jebish Purbey, Giulio Martini, Suman Debnath, Hamza Farooq*
*arxiv.org/abs/TBD* - *Under review at EMNLP 2025*

## PROJECTS

**SpeechAya: Multilingual Speech Synthesis via Unified LLM Architecture**
*Open-Source Project — Cohere Labs Hackathon*
Engineered a multilingual LLM pipeline integrating speech and text modalities using 1000+ hours of audio (LibriSpeech, CommonVoice).
Optimized tokenization using MMS, mHuBERT, XEUS; achieved 0.112 WER on PolyAI/minds14 with Qwen2-1.5b model and custom embeddings.
Developed a modular pipeline for ASR, TTS, voice cloning, and speech translation tasks.

**Advanced Attribute Extraction & Classification Pipeline**
*Amazon ML Hackathon 2024*
Applied OCR on 400K+ product images with 88% recognition accuracy. Fine-tuned DistilBERT & LLaMA 3.2 for entity extraction.
Optimized LayoutLM for attribute classification (weight, height, width), reducing misclassification by 10–15%.

**DocAI: Smart Lab Report Generator**
*Hackathon Project*
Built full-stack Django app for doctors/patients; integrated NLP-based report suggestions reducing manual input by 40%.
Embedded chatbot and Plotly-based dashboard for real-time analytics and UX improvements.

**Carbon Footprint Detector**
*3-Day Hackathon Project — Feb 2023*
Created Chrome extension analyzing 200+ websites' carbon emissions. PostgreSQL + Node.js backend for real-time tracking.
Built CI/CD pipeline and Dockerized deployment; achieved 30% faster dev cycles and higher user engagement.

**Music Genre Classifier using Audio Signals**
*Apr 2023 – May 2023*
Achieved 91.2% accuracy on 6-genre classification with CNN + Librosa pipeline.
Reduced overfitting by 38% with dropout, weight decay, and batch norm; used ensemble methods and GridCVSearch.

**Deep Space Image Classifier**
*Mar 2023*
Classified celestial objects using multiclass image classifier with 78% accuracy.
Compared KNN vs. mean imputation; tested SVM, RF, ANN, LightGBM, XGBoost, CatBoost.

**Design Automation for Mechanical Workflows**
*B.Tech Thesis — Jul 2025 – Present*
Developing software to automate design/planning workflows for mechanical systems.
Built custom workflow design modules integrated into engineering pipelines.

## TECHNICAL SKILLS

**Languages:** Python, C++, Julia, JavaScript
**Libraries and Frameworks:** PyTorch, Django, TensorFlow, Sklearn, Librosa, NLTK, TRL, Transformers, LoRA, OpenCV, Numpy, Pandas, Matplotlib, Gradio, BitsandBytes
**Databases:** PostgreSQL, SQL, SQLite

## REFERENCES

**Suman Debnath** ..... *Principal Developer Advocate, Amazon, USA*
suman.san14@yahoo.in

**Ram Mohan Rao Kadiyala** ..... *NLP Research Head, Traversaal.ai*
https://www.rkadiyala.com/