

Siddhant Gupta

+91 7836015759 | writetosiddhant@gmail.com | [LinkedIn](#) | [Github](#) | [Google Scholar](#) | [Personal Website](#)

EDUCATION

Indian Institute of Technology Roorkee

Bachelor of Technology

Roorkee, Uttarakhand

Oct 2022 – July 2026

EXPERIENCE

Observe.ai

AI platform for contact centers

June 2023 – Present

Machine Learning Intern

- Co-authored a research paper submitted to EMNLP 2025 on bias in large language models (LLMs), proposing a novel taxonomy of bias dimensions and a framework to evaluate and mitigate such biases systematically.
- Conducted in-depth research on autonomous agent systems, focusing on orchestration strategies and practical integrations for real-world enterprise NLP workflows.
- Enhanced the evaluation pipeline of the Automatic Quality Assurance (Auto-QA) system to improve robustness and performance analysis in production environments.

Cohere Labs

Research Lab and Open Science Community

June 2023 – Present

NLP & ML-Agents Lead

- Engaged in 50+ technical discussions and speaker workshops on topics such as NLP, multi-agent systems, contextual learning, synthetic data generation, and mechanistic interpretability, contributing to the community's NLP knowledge base [NLP]
- Led efforts to introduce a DAG-based open-source agentic framework under the ML-Agents program, collaborating globally to evaluate and benchmark emerging agentic systems [ML-Agents]
- Speaker at Cohere ML Summer School on *Introduction to Transformers and the Evolution of LLMs* [Talk].
- Developed a speech synthesis method using ASR data during the 8-week long hackathon Expedition Aya – 2024.
- Led 5 research projects for Expedition Aya – 2025, including work on reasoning evaluations, fine-grained machine-generated text detection, cultural bias in VLMs, multilingual LLM enhancements, and image-caption dataset development — with some accepted for publication.

Artificial Intelligence and Electronic Society (ArIES)

Indian Institute of Technology, Roorkee

May 2023 – Present

ML Executive

- Spearheaded teams in AI hackathons, providing mentorship in CV and NLP research alignment, leading to the successful implementation of 10+ innovative projects including participation in Inter-IIT competitions.
- Organized and conducted workshops and talks for 100+ participants, focusing on deep learning and image processing concepts such as edge detection, depth estimation, object detection, and character recognition, boosting technical proficiency across attendees

M2ai

AI Research Lab

February 2025 – Present

Research Head

- Led and co-authored multiple research projects resulting in publications at top-tier conferences, including key contributions to machine-generated text detection and the development of LLMs/VLMs tailored for South Asian languages.
- Actively working on improving Speech-to-Text (STT) and Text-to-Speech (TTS) models, designing efficient audio tokenizers, and building vertical agent tools and applications.
- Oversaw projects funded by grants totaling over \$20,000 from organizations such as OpenAI, Anthropic, and Cohere.

Traversaal.ai

AI Research Lab

April 2025 – Present

Research Fellow under Hamza Farooq

- Working on benchmarking data science agents, focusing on automated graph generation and ReAct-style agent workflow design and evaluation.

Computational Intelligence and Operations Lab (CIOL)

Shahjalal University of Science and Technology

September 2024 – November 2024

Research Collaborator

- Conducted research on hate speech detection across multilingual datasets using RoBERTa, addressing model bias and improving classification metrics and designed advanced solutions for Retrieval-Augmented Generation (RAG) tasks, enabling seamless integration of external knowledge retrieval improving F1k, MRR, precision and recall.

PROJECTS

SpeechAya | Multilingual LLM that can hear and speak

August 2024 – September 2024

- Engineered a novel multilingual LLM pipeline integrating speech and text modalities, processing over 1000 hours of audio data from LibriSpeech and Mozilla CommonVoice datasets across 5 languages.
- Implemented and optimized speech tokenization using state-of-the-art models (MMS, mHuBERT, XEUS), reducing processing time by 32% through efficient batching and parallel processing.
- Achieved a score of 112 in Word Error Rate (WER) on the PolyAI/minds14 benchmark dataset by fine-tuning a Qwen2-1.5b model architecture with custom speech embeddings.
- Developed a modular training pipeline supporting multiple speech tasks (ASR, TTS, voice cloning, translation) through a unified model architecture.

Advanced Attribute Extraction and Classification Pipeline

July 2024 – August 2024

- Applied advanced OCR techniques with pre-trained models to extract text from over 400,000 product images, achieving an 88% text recognition accuracy and significantly enhancing data extraction efficiency.
- Fine-tuned DistilBERT and LLaMA 3.2 for Named Entity Recognition (NER) tasks, using proper metrics for optimization, which resulted in an improvement in entity extraction precision and recall.
- Optimized LayoutLM for attribute classification tasks, such as identifying product dimensions (e.g., weight, height, width), reducing misclassification rates (False Positives) by 10-15% and streamlining attribute extraction workflows.

PUBLICATIONS

- Lexical Reranking of Semantic Retrieval (LeSeR) for Regulatory Question Answering**
Jebish Purbey, Drishti Sharma, [Siddhant Gupta](#), Khawaja Murad, Siddhartha Pullakhandam, Ram Mohan Rao Kadiyala
Accepted at RegNLP @ COLING 2025 — 4th position in workshop
- SeQwen at the Financial Misinformation Detection Challenge Task: Sequential Learning for Claim Verification and Explanation Generation in Financial Domains**
Jebish Purbey, [Siddhant Gupta](#), Nikhil Manali, Siddhartha Pullakhandam, Drishti Sharma, Ashay Srivastava, Ram Mohan Rao Kadiyala
Accepted at FinNLP-FNP-LLMFinLegal @ COLING 2025 — 3rd position in workshop
- Multilingual Hate Speech Detection and Target Identification in Devanagari-Scripted Languages**
[Siddhant Gupta](#), Siddh Singh, Azmine Tushik Wasi
Accepted at Chipsal @ COLING 2025
- Robust and Fine-Grained Detection of AI Generated Texts**
Ram Mohan Rao Kadiyala, Siddhartha Pullakhandam, Kanwal Mehreen, Drishti Sharma, [Siddhant Gupta](#), Jebish Purbey, Ashay Srivastava, Subhasya TippaReddy, Arvind Reddy Bobbili, Suraj Telugara Chandrashekhar, Modabbir Adeeb, Srinadh Vura, Hamza Farooq
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
Under Review at EMNLP 2025
- Improving Multilingual Capabilities with Cultural and Local Knowledge in Large Language Models While Enhancing Native Performance**
Ram Mohan Rao Kadiyala, Siddhartha Pullakhandam, [Siddhant Gupta](#), Drishti Sharma, Jebish Purbey, Kanwal Mehreen, Muhammad Arham, Hamza Farooq
Under Review at EMNLP 2025
- Evaluating Generalization Capabilities of LLM-Based Agents in Mixed-Motive Scenarios Using Concordia**
(Authors list omitted for brevity)
Under Review at NeurIPS 2025
- Spot the BlindSpots: Systematic Identification and Quantification of Fine-Grained LLM Biases in Contact Center Summaries**
Kawin Mayilvaghanan, Siddhant Gupta, Ayush Kumar
Under Review at EMNLP Industry Track 2025

- [9] **NepX-Hate: A Nepali Hate Speech Corpus with Fine-Grained Sociocultural Annotations**
Jebish Purbey, Sanjeeb Prasad Panday, Sudhan Shrestha, Rajan Kusi, Manoj Rokaya, Bikash Balami, Sajan Maharjan, Prabhat Bhatta, Ayush Purbey, Siddhant Gupta, Vatsal, Ram Mohan Rao Kadiyala
Under Review at EMNLP 2025

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: PostgresSQL, SQL , SQLite

REFERENCES

Suman Debnath *Principal Developer Advocate, Amazon, USA*
suman.san14@yahoo.in — +91-96201-02221 — LinkedIn

Ram Mohan Rao Kadiyala *MS in Machine Learning, University of Maryland, College Park*
rkadiyala.com