

Debjyoti Ghosh

Kolkata, India · +91 70740 19679 · debjyoti1ghosh@gmail.com

debjyoti-ghosh.in · github.com/debjyoti71 · linkedin.com/in/debjyoti-ghosh-508bba281

Undergraduate researcher and AI/ML engineer with research and engineering experience spanning AI Security, Trustworthy AI, Multimodal Learning, and Intelligent Systems. My work focuses on deepfake detection, context-aware natural language processing, biometric authentication, Bayesian reasoning systems, and scalable AI infrastructure for real-world applications.

Research Interests

Trustworthy AI · Multimodal Learning · Representation Learning · AI Security · Deepfake Detection · Media Authenticity · Natural Language Processing · Computer Vision · Adversarial Machine Learning

Education

University of Engineering & Management, Kolkata	2023 – 2027
Bachelor of Technology in Computer Science and Information Technology	CGPA: 8.7 / 10

Experience

Backend Engineer (Python Developer)	Nov 2025 – Present
<i>ThreatPurge Private Limited</i>	

- Engineered scalable backend services for a multi-tenant cybersecurity platform using **Python**, **FastAPI**, and **PostgreSQL**, successfully handling over 50,000 daily security events with strict data isolation.
- Architected asynchronous **ETL pipelines** for phishing detection and brand abuse monitoring, accelerating the processing of large-scale threat intelligence data by 40%.
- Optimized **PostgreSQL** query execution and indexing strategies, reducing database latency by 35% and improving overall API responsiveness for production workloads.
- Built production-ready **REST APIs**, background workers, and scheduled jobs to support enterprise cybersecurity automation workflows, achieving 99.9% system uptime and eliminating manual threat triaging tasks.

Full-Stack AI Engineer (Freelance)	Mar 2026 – Present
<i>Elaina Legal AI</i>	

- Architected a dual-facing **Agentic AI** legal assistant using **Python** and **FastAPI**, conducting autonomous client intake and extracting structured facts for attorneys.
- Engineered an AI tool-calling framework empowering attorneys to schedule meetings, generate invoices, and draft letters directly via natural language chat commands.
- Built automated **OCR** evidence pipelines to process client uploads in real-time, proactively detect missing case information, and maintain a dynamic legal timeline.
- Formulated machine learning infrastructure utilizing a **Bayesian Neural Network (BNN)** for legal outcome prediction, providing confidence-scored case evaluations.
- Constructed embedding-based **semantic retrieval** systems to maintain long-term conversational memory, enabling context-aware contradiction detection.

Research Projects

Multimodal Deepfake Detection via Audio-Visual Fusion (IEEE INDICON 2025)

- Architected a scalable multimodal deepfake detection pipeline utilizing a **Transformer-based** attention mechanism to fuse spatial, temporal, and acoustic embeddings, achieving 95% accuracy and 0.97 AUC.
- Integrated a **ResNet-50** visual encoder, **BiLSTM** temporal encoder, and **1D-CNN** audio encoder processing **Mel-spectrograms** to capture complex cross-modal inconsistencies.
- Implemented a custom synchronization loss function to enforce cross-modal verification, effectively detecting audio dubbing by evaluating lip-speech (phoneme-viseme) alignment.
- Evaluated the framework on the FakeAVCeleb dataset, demonstrating superior robustness against modality-specific ambiguities compared to standard unimodal detection methods.

Context-Aware Hate Speech Detection via Post-Comment Joint Modeling

- Created a context-aware moderation framework replacing isolated text analysis with post-comment joint modeling, achieving 95% accuracy and an ROC-AUC of 0.9911.
- Structured a four-class reaction taxonomy mapping discourse to explicit moderation actions, reducing false positives caused by benign, high-arousal disagreement.
- Extracted and fused parallel features including **SBERT** semantic similarity, emotion distributions, and **DistilBERT** hate probabilities to accurately distinguish contextual intent.
- Validated framework robustness on a curated interaction dataset, proving the necessity of parent-post context for scalable, safe, and proactive online harm mitigation.

Edge-Based Voice Biometric Verification and Cloud-Assisted Anti-Spoofing

- Constructed a speaker recognition system combining edge-based audio acquisition with a centralized cloud backend, achieving 97.5% authentication accuracy in controlled environments.
- Formulated an embedding-based verification pipeline utilizing **MFCCs** and spectral features processed through a **Gaussian Naive Bayes** classifier for closed-set identification.
- Engineered **FastAPI** backend and web interfaces facilitating secure user enrollment, dataset management, and parallelized API-based inference.
- Evaluated the system against environmental noise and spoofing vulnerabilities, improving authentication resilience compared to classical isolated models.

Publications

Multimodal Deepfake Detection (22nd IEEE INDICON 2025, IIIT Bangalore)
DOI: 10.1109/INDICON68490.2025.11392906

Open Source & Software Projects

AntiGravity (Dev Timekeeper) – VS Code Extension [Open VSX Registry]

- Created and published a developer productivity and time-tracking extension for the VS Code ecosystem, currently amassing over **1,300 downloads** on the Open VSX Registry.
- Built the extension to integrate seamlessly with the editor's lifecycle, providing unobtrusive background tracking and workflow analytics to optimize coding efficiency.

Technical Skills

Languages:

Python, C, SQL, Bash

Machine Learning & AI Frameworks:

PyTorch, TensorFlow, Hugging Face Transformers, Scikit-learn, Bayesian Neural Networks (BNN), SBERT, DistilBERT

Backend Engineering & Databases:

FastAPI, Flask, PostgreSQL, RESTful APIs, Asynchronous Python (asyncio), ETL Pipelines, Microservices Architecture

Data Processing & Audio-Visual:

Pandas, NumPy, OpenCV, Librosa, Audio Processing (MFCCs, Mel-spectrograms), Semantic/Vector Embeddings

DevOps & Infrastructure:

Git, Docker, Background Workers, AWS

Achievements

- Presented research at the **22nd IEEE INDICON 2025**, hosted by IIIT Bangalore.
- Selected for the **Amazon ML Summer School 2025**.
- Achieved multiple podium finishes in hackathons focused on machine learning and intelligent systems.
- Achieved multiple podium finishes in robotics competitions involving autonomous systems and embedded AI.