

# Sajeeb Das

+880 1537224180 | hi5.sajeeb@gmail.com | hi5sajeeb.pages.dev | github.com/sajeeb007 | linkedin.com/in/hi5sajeeb

Computer Science graduate with expertise in deep learning, LLMs, and multimodal AI, aiming to advance research in efficient, interpretable, and hallucination-resistant models that drive trustworthy decision-making and impactful societal applications.

## EDUCATION

---

**National Institute of Technology, Rourkela, B.Tech in Computer Science and Engineering** | Odisha, India 2024  
GPA: 8.1 / 10 | Gradecard

Winner: IIT Bombay Techfest & NIT Rourkela Innovation E-cell Hackathon

Funding: NIT Warangal Firefly Startup Incubation

Courses: DSA | Database Engineering | OS | Computer Networks | Machine Learning | Deep Learning | Distributed System  
Quantum Computing | Computer Security | Intrusion Detection System

**Government Hazi Mohammad Mohsin College, Chittagong, HSC** | Chittagong Board, Bangladesh 2019  
GPA: 5.0 / 5.0

**Chittagong Government High School, Chittagong, SSC** | Chittagong Board, Bangladesh 2017  
GPA: 5.0 / 5.0

## Test Scores

---

GRE: Quant: 163, Verbal: 151, AWA: 4.0

IELTS: Listening: 8.5, Reading: 8.5, Writing: 8.0, Speaking: 7.5 - Band Score: 8.0

## PUBLICATIONS

---

**Sajeeb Das<sup>†</sup>, Srijit Paul<sup>†</sup>, Ucchas Muhury, Akib Jayed Islam, Dhruva Jyoti Barua, Sultanus Salehin, Prasun Datta.** “**DR-ISHTI: Drug Recognition and Integrated System for Helping the Visually Impaired with Tag-based Identification.**” Proceedings of Recent Advances in Natural Language Processing (RANLP), Varna, Bulgaria, 2025. [\[PDF\]](#)

Srijit Paul<sup>†</sup>, **Sajeeb Das<sup>†</sup>**, Akib Jayed Islam, Sultanus Salehin, Prasun Datta. “**Federated Learning for Privacy-Preserving Gait Recognition on Edge Devices Using IMU Data.**” Proceedings of the 16th International Conference on Contemporary Computing and Informatics (ICCCNT), 2025. [\[In Press\]](#)

**Sajeeb Das<sup>†</sup>, Srijit Paul<sup>†</sup>, Akib Jayed Islam, Sultanus Salehin, Prasun Datta.** “**Development of a Multilingual Climate Fact-Checking System with Unified Dataset for Low-Resource Indic Languages.**” Proceedings of the 16th International Conference on Contemporary Computing and Informatics (ICCCNT), 2025. [\[In Press\]](#)

Prasun Datta, Sultanus Salehin, Akib Jayed Islam, Srijit Paul, **Sajeeb Das.** “**AGFormer-Net: A Transformer-Guided Attention-Gated Framework for 3D Brain Tumor Segmentation.**” Proceedings of the 16th International Conference on Contemporary Computing and Informatics (ICCCNT), 2025. [\[In Press\]](#)

Prasun Datta, Sultanus Salehin, Akib Jayed Islam, Srijit Paul, **Sajeeb Das.** “**Learning to Localize: A Weakly Supervised Deep Learning Framework with PCAM and Spatial Priors for Chest X-rays.**” Proceedings of the 16th International Conference on Contemporary Computing and Informatics (ICCCNT), 2025. [\[In Press\]](#)

## EXPERIENCE

---

**Commure, Operations Analyst (RCM Rules Engine)** | Onsite (Dhaka, Bangladesh) Sept 2025 - Present

- Designed SQL-based validation pipelines for the RCM Rules Engine integrating LLM endpoints with rule JSON payloads via chain-of-thought prompting, enabling context-aware claim adjudication across large-scale EHR datasets.
- Engineered an LLM-driven rule audit system using LangChain + GPT-4/Gemini with context caching over historical denial records, where the model surfaces rule fix suggestions from exception traces, cutting manual review by 40%.
- Built an internal RAG-powered AI agent using ChromaDB + pgvector embeddings over the backend codebase for context-aware retrieval of rule conditions and EHR schema, significantly reducing rule debugging time.
- Fine-tuned an in-house open-source LLM using QLoRA and a custom PyTorch training loop on sensitive claims denial payloads, improving model precision on payer-specific adjudication edge cases over the base model.
- Built iterative analytics dashboards surfacing LLM-generated diagnostic insights on rule coverage, denial rate trends, and exception clustering across production claim workflows.

**Qualcomm, Software Engineer Intern (QDSP Team)** | Onsite (Hyderabad, India) 2023

- Enhanced power measurement accuracy for SNPE and QNN SDK, leading to a 30% reduction in energy consumption for ML models by providing the team with actionable insights for optimization.
- Built framework to measure and visualize inference-per-watt, aiding model selection for Qualcomm chipsets.
- Instrumented pipelines to log layer-level latency, memory, and power usage, identifying bottlenecks.
- Explored INT8/FP16 quantization in SNPE/QNN, enabling low-power inference with minimal accuracy loss.
- Analyzed throughput-efficiency trade-offs, proposing deployment strategies for edge hardware.
- Optimized inference pipelines via profiling and debugging, improving reliability and runtime stability.
- Authored internal documentation and guides, streamlining integration and reducing onboarding effort.
- Developed automated test suites and benchmarking playlists, cutting manual testing overhead.
- Collaborated with San Diego team to integrate bare-metal power module into local SDK builds.

- Engineered CNN and Vision Transformer architectures for MRI/CT segmentation, achieving >95% Dice scores.
- Developed LLM-augmented vision pipelines with CLIP embeddings to auto-generate accessible radiology reports.
- Designed an RFID-based medication management system with MQTT IoT integration for real-time monitoring.
- Implemented federated learning on edge clusters with secure aggregation for privacy-preserving patient analytics.
- Mentored undergrads on AI-healthcare prototypes and co-authored publications.

**National Institute of Technology, Rourkela, *Research Assistant* | GitHub**

2023–2024

- Fused POS, NER, and lemmas into Transformer encoders, boosting BLEU by 35% on low-resource Indic MT.
- Fine-tuned OpenNMT-py seq2seq pipelines on English–Indic Samanantar corpus, tuning LR, dropout, and warmup to achieve SOTA BLEU across 8 language pairs.
- Performed ablations with sacreBLEU, TER, and chrF; automated error analysis.
- Injected visual frame and audio spectrogram embeddings via cross-modal attention for semantic disambiguation.
- Built federated fine-tuning pipelines with secure aggregation and differential privacy for decentralized updates.
- Engineered beam-search back-translation and MLM paraphrasing pipelines, scaling under-represented data 4×.
- Implemented RAG using BM25 retriever with seq2seq decoder for open-domain QA.

**Instructor, Chittagong Junior Coders | Online (Bangladesh) | Youtube**

2021–2023

- Taught C programming to students from grades 2–10 through the “Chittagong Junior Coders” Facebook community.
- Volunteered to provide accessible coding education in underserved areas, fostering early interest in technology.
- Developed students’ problem-solving skills and built foundational knowledge in computer programming.

**SKILLS**


---

Languages	Python, C++, SQL, Git, Bash, LaTeX
Machine Learning	Pandas, Numpy, Spacy, Matplotlib, Scikit-Learn
Development	MongoDB, Express, ReactJS, Node, REST API
Retrieval-Augmented Generation	LangChain; LlamaIndex; RAG (FAISS, Chroma, Pinecone); RETRO
OS	Linux, Windows
ML Frameworks	Pytorch, Opennmt-Py, FairSeq, Flower

**PROJECTS****Automated Bottle Defect Detection System**

- Developed a real-time bottle defect detection system using YOLOv8n, identifying 8 distinct defect classes with 85–95% mAP@0.5.
- Built an inspection setup with a conveyor belt and relay-controlled arm, managed by an ESP32 microcontroller.
- Deployed the inference engine on an edge device, with the ESP32 acting as control node for coordinating movements.
- Enabled synchronized defect sorting by interfacing the control logic and detection results via serial communication.
- Integrated a Flask-based web interface for live system monitoring, analytics, and operational control.

**License Plate Recognition System**

- Built a license plate detection system using a classical vision pipeline (binarization, CCA, morphological operations) to isolate plates from vehicle images.
- Segmented characters on identified plates and trained an SVC for high-accuracy recognition.
- Validated the final model using 4-fold cross-validation to ensure robust performance on unseen data.

**Traffic Signs Classification with a Convolutional Network**

- Preprocessed and augmented the German Traffic Sign Dataset (to 63k+ images), utilizing localized histogram equalization to enhance features.
- Designed and trained a custom Convolutional Neural Network (CNN) with multi-scale features, dropout, and L2 regularization to prevent overfitting.
- Implemented a two-stage training and fine-tuning strategy, achieving an outstanding test set accuracy of 99.33%.

**Custom Image Processing Filters**

- Engineered a suite of image filters (e.g., Blur, Grayscale, Edge Detection) for BITMAP images from scratch in C.
- Implemented multithreading to significantly accelerate filter processing speeds and enhance computational efficiency.
- Developed comprehensive test cases to validate filter accuracy and ensure reliable, high-quality output.

**Achievements**

- Recipient of the **Bangladesh Government Merit Scholarship** for outstanding performance in SSC exams, recognizing academic excellence and dedication.
- Awarded the **Study in India Government Scholarship** by the Ministry of External Affairs, providing full financial support for the entire duration of Bachelor’s studies at NIT Rourkela (NITR).
- Secured funding from the **NIT Warangal Startup Incubation Center** for the development of our startup, focused on an innovative **EV Vehicle Ecosystem** app, supporting its growth and implementation.