

Suvasish Das

Research Scholar at Indian Institute of Technology, Gandhinagar

 (+91) 62917-30738  suvasish@iitgn.ac.in  github.com/suvasish114  suvasish114.github.io

Education

- **Indian Institute of Technology, Gandhinagar** Gujarat, India
PhD in Artificial Intelligence - Ongoing Jul, 2026 - Jul, 2031
- **Indian Institute of Technology, Bombay** Mumbai, India
PhD in Intelligent Systems (IS1) - Dropped due to research misalignment Jan, 2025 - Jul, 2026
- **West Bengal State University** Kolkata, India
M.Sc (Hons) in Computer Science (CGPA: 9.78/10) Jul 2022 - Jul 2024
Secure 1st position in the department and its affiliated colleges - [Certificate]
- **Barrackpore Rastraguru Surendranath College** Kolkata, India
B.Sc (Hons) in Computer Science (CGPA: 9.51/10) Jul 2019 - Jul 2022

Recent Works

- Least Unsquared Deviation for Cryo-EM 3D Reconstruction** [Files] [Report] Jul, 2026
- Implementation of Algorithms and Comparative Study of Least Unsquared Deviation.
 - Used diffusion for denoising 2D projections, which significantly reduces noise in the input images.
- Document Image Machine Translation (ENG to CHI)** [Presentation] Nov, 2025
- Developed an end-to-end document image translation pipeline that converts English document images into structure-preserving Chinese Markdown, and obtained LLM finetuned structural accuracy of 88.47%
- Text-to-text Machine Transliteration (Hindi to Devangari)** [Presentation] [Code] Sep, 2025
- A transliteration system to convert Hindi (Roman) to Devanagari while preserving pronunciation accuracy.
 - Experimented with LSTM, Transformer, and SoTA LLM-based architectures and compared them.
- Part of Speech Tagging for English** [Presentation] Aug, 2025
- Developed a POS tagger for the English language using different architectures such as HMM, Bi-RNN, SoTA LLM, etc and compared them.
- Diverse and Improved Sparse Retrieval** [Report1] [Report2] Jul, 2025
- Developed a retrieval system to preserve diversity in the generated outputs on the WikiDump corpus.
 - Improved the system by adding a learned sparse matrix in the latent space, with added validation such as vocabulary mismatch, synonymy, or cross-lingual variation, etc.
- Two Stage Ranking System** [Report] Sep, 2025
- Stage-1 Ranker module generates relevant documents using a contrastive learning strategy.
 - Stage-2 Reranker module leverages Markov random fields to model term dependencies within a window.
- Reproducing Conditional Generative Adversarial Networks** [Report] [Code] Nov, 2025
- Reproduce a 50% scaled-down version of the Conditional-GAN paper by [M Mirza and S Osindero].
- Graph Neural Network based Question Answering system** [Report] [Code] Apr, 2025
- Developed a Graph Neural Network-based end-to-end system for information retrieval.
 - Used the FLAN T5 language model for in-context text generation for user-friendly responses.
- Information Bottleneck and Classification using VAE** [Presentation] Mar, 2025

Jailbreaking and Red-teaming DeepSeek Models [Presentation] [Files]	Nov, 2025
Studied adversarial attacks based on prompt injection to break the safeguard of DeepSeek models and compared our implementations with the current available benchmarks.	
Bengali Spelling Correction using Noisy Channel Approach [Code] [Report]	Jun 2024
- A web-based application to rectify Bengali spelling errors by using Bayesian Inference. - Implemented beam-search on top of the noisy-channel to improve overall language model accuracy. - Collected diverse handwritten texts from college students for real-time evaluation.	
Facility Detection and Popularity Assessment [Code] [Slides]	Mar 2023
- Facility detection from large-scale text classification of social media and crowdsourced data. - Achieve accuracy of 91% using Support Vector Machines and 85% using Naïve Bayes. - Scrape Twitter for a particular timeline with target facility type, used to determine the truthiness of the tweet.	
Non-adaptive Image Encryption based on Chaos [Code] [Report]	Jul 2022
- Implemented an image encryption method based on non-adaptiveness chaos. - Reduced noise up to 80% created by public channels and retrieved encrypted images at the receivers end.	
A full featured Technical Blog Website using Python-Flask [Code]	2021
Android based Covid-19 Vaccine Notifier Application [Code]	2020

Coursework / Certifications

Courseworks: Information Retrieval, Natural Language Processing, Foundations of Machine Learning, Advanced Image Processing (Inverse Problem), Human-centered AI, etc.

Online Certifications: Neural Networks and Deep Learning ([Coursera](#)), Neural Network Bootcamp ([Udemy](#))

Skills

Languages: Python (8000+ lines), C/C++, HTML/CSS, Java, JavaScript, $\text{\LaTeX}$, SQL

Library & Tools: Pytorch, Tensorflow, NLTK, Pandas, Numpy, Keras, Flask, Git, Jekyll

Experience

Teaching Assistant in CS101 (Fall-25), CS271 (Spring-26) IIT Bombay	Mumbai, India
<i>Assisted UG students by arranging class quizzes, assignments, and semester exams, etc.</i>	<i>Jul, 2025 – Jul, 2026</i>
Tech Content Reviewer/ Creator in CS101-outreach program by BodhiTree, IIT Bombay	Mumbai, India
<i>Content creator and reviewer for BodhiTree platform, sponsored by IBM and TowerResearch</i>	<i>Dec, 2025 – Jan, 2026</i>

Fellowships / Achievements

Received Honorarium of Rs. 1.2 Lakh from BodhiTree, sponsored by IBM and TowerResearch	<i>Feb, 2026</i>
INSPIRE Fellowship (round-1) awarded by Department of Science and Technology, Govt. of India	<i>Jun, 2025</i>
GATE Qualified with a score of 469 and 94.28% percentile (paper: CS), 76.56% percentile (paper: DA)	<i>Mar, 2024</i>
UGC-NET Qualified $\times 2$ for assistant professor and admission to PhD	<i>Dec, 2022 & Jun, 2024</i>