

Bitan Majumder

Kolkata, WB, India
Email: bitanmajumder2002@gmail.com
Mobile: +91 6290600025

ABOUT

Ph.D. Computer Science student with research experience in Natural Language Processing and Machine Learning. Focused on Graph Machine Learning and NLP for Computational Social Science, with published work on code-mixed text and active research on large-scale social networks.

EDUCATION

Doctor of Philosophy <i>Ashoka University</i>	Aug 2026 – Present <i>Sonipat, Haryana</i>
Master of Science in Computer Science CGPA: 8.99/10 <i>Pondicherry Central University</i>	Jul 2024 – May 2026 <i>Kalapet, Pondicherry</i>
Bachelor of Science (Honours) in Computer Science CGPA: 8.8/10 <i>Ramakrishna Mission Vivekananda Centenary College</i>	Sept 2021 – May 2024 <i>Kolkata, WB</i>
WBBSE: 88% WBCHE: 89% <i>Ramakrishna Mission Vidyapith</i>	2013 – 2021 <i>Purulia, WB</i>

SKILLS

- Programming:** Python, C, Java, SQL
- Deep Learning:** PyTorch, TensorFlow, Keras
- Graph & Language Modeling:** NLP, Graph Neural Networks, PyTorch Geometric, NetworkX
- Web & Markup Technologies:** HTML, LaTeX
- Version Control:** Git, GitHub
- Languages:** Bengali (Native), English (Fluent), German (Elementary)

EXPERIENCE

Research Assistant <i>Ashoka University</i>	Jan 2026 – Apr 2026 <i>Sonipat, Haryana</i>
- Performs large-scale social network analysis on Twitter datasets and creates graph representations of the user behavior patterns. - Applying Graph Machine Learning to extract structural insights from large-scale network data from diverse real-world sources. Our research focuses on deep learning models of Graph Neural Networks (GNNs). - To add more contextual information, we use Natural Language Processing (NLP) together with network analysis and the existing social media text information. - We build scalable data pipelines for extracting and transforming large amounts of Twitter data (i.e., tweets, their authors, and their corresponding networks) required for graph construction and feature engineering.	
Intern <i>IEEE Computer Society, Kharagpur Section</i>	June 2024 – Aug 2024 <i>Remote</i>
- Conducted in-depth research on time series forecasting, with a focus on evaluating Large Language Models (LLMs) as zero-shot predictors. - Reproduced and extended methodologies from the paper “Large Language Models Are Zero-Shot Time Series Forecasters” to design experimental forecasting pipelines. - Implemented forecasting workflows using Gemini and OpenAI APIs, adapting prompt-based inference for structured temporal data. - Engineered and preprocessed custom time series datasets in Python to systematically assess model performance. - Analyzed forecasting accuracy and limitations of foundation models in temporal prediction tasks, gaining practical insight into LLMs for non-text modalities.	

Research Intern

Indian Institute of Engineering Science and Technology (IIST)

June 2023 – Dec 2023

Shibpur, WB

- Designed and developed an automated exam answer script evaluation system using Natural Language Processing (NLP) and unsupervised machine learning techniques.
- Implemented text preprocessing, feature extraction, and similarity-based scoring methods to predict marks for descriptive answers.
- Conducted experiments to compare model-predicted scores with human-evaluated marks, achieving competitive and consistent performance.
- Gained hands-on experience in applying machine learning to real-world educational assessment problems, focusing on fairness and scalability.

Intern

West Bengal State University

Aug 2022 – Feb 2023

Kolkata, WB

- Contributed to the development of an Optical Character Recognition (OCR) system for the Bengali language, addressing challenges in low-resource scripts.
- Designed and implemented character segmentation pipelines using contour detection, smoothing, and noise removal techniques.
- Applied image processing algorithms to segment individual Bengali characters from scanned document images.
- Utilized OpenCV and Python to preprocess datasets and improve segmentation robustness.
- Gained strong foundations in digital image processing and practical experience in building OCR systems for Indic languages.

PUBLICATIONS

1. **Bitan Majumder** and Anirban Sen. "Small Wins Big: Comparing Large Language Models and Domain Fine-Tuned Models for Sarcasm Detection in Code-Mixed Hinglish Text." arXiv preprint, arXiv:2502.21933 [cs.CL], 2025. (Under review at ACL Student Research Workshop 2026)

DATASETS

1. **Bitan Majumder** and Jishnu Bandyopadhyay. Broad-Answer Type Q&A Dataset for Automated Exam Paper Evaluation. IIST Shibpur, 2024. Available: https://github.com/TheRealJishnu/Automated_Exam_Paper_Evaluation

CERTIFICATIONS

- Deep Learning Specialization - Deeplearning.ai
- IELTS (Band: 8) - IDP

RELEVANT LINKS

LinkedIn: <https://www.linkedin.com/in/bitam-majumder-672801237>

Portfolio: <https://bitanm.github.io/>

GitHub: <https://github.com/bitamM>

Kaggle: <https://www.kaggle.com/bm2002>

HackerRank: <https://www.hackerrank.com/profile/bitamajumder201>