S M Jishanul Islam

Software Engineer, Apurba Technologies Ltd., Dhaka, Bangladesh
jishanlion@gmail.com — +880 1759338652 — Linkedin — Website — Google Scholar — GitHub

RESEARCH INTERESTS

Multimodal Learning - Multimodal Representation - Multimodal Understanding - Multimodal AI Applications.

EDUCATION

United International University, Dhaka, Bangladesh — 2020 — 2024
Bachelor of Science in Computer Science and Engineering (**Major**: Data Science) — CGPA: 3.77/4.00
Supervisor: Prof. Dr. Swakkhar Shatabda — Major CGPA: 4.00/4.00

EXPERIENCE

Apurba Technologies Ltd. — Dhaka, Bangladesh
Software Engineer — December 2024 — Present

- Worked on the core ML team of a government-adopted national Bengali OCR system https://ocr.bangla.gov.bd/.
- Built scalable and optimized RESTful APIs, improving performance by 10%.
- Co-created an improved word-level model for OCR, recording an accuracy of 95%+.
- Collaborated with stakeholders to implement, research, and complete project deliverables 2-3 weeks before the deadline.

United International University — Dhaka, Bangladesh
Lecturer (contractual) — June 2024 — October 2024

- Taught core CS concepts to students.
- Took exams to monitor students' performance to provide informative feedback, improving average performance by 7%.
- **Courses conducted:** Object-Oriented Programming Lab, Data Structures and Algorithms 1 Lab, and Database Management Systems Theory.

United International University — Dhaka, Bangladesh
Undergraduate Assistant — September 2023 — May 2024

- Managed to facilitate course content and materials with the lab faculty, checked and evaluated assessments, and judged lab projects.
- Created the first course materials for the country's first-ever undergraduate degree in Data Science.
- **Courses supervised:** Programming for Data Science, and Object-Oriented Programming for Data Science.

PUBLICATIONS

Journals

- **S M Jishanul Islam**, S. H. Mustakim, M. Hossain, M. Mashira, N. I. Shourav, M. R. Ahmed, S. Islam, S. Shatadba, and A. K. M. M. Islam. "An audio video-based multi-modal fusion approach for speech emotion recognition.", **Knowledge-Based Systems (Elsevier)**, I.F 7.6, under 2nd revision, 2024.

Conference and Workshop papers

- S. H. Mustakim, **S M Jishanul Islam**, U. M. Muna, M. Chowdhury, M. Jawwadul Islam, S. Ahmmed, T. Reeti, S. T. A. Dhrubo, and S. Shatabda, "Watch, Listen, Understand, Mislead: Tri-modal Adversarial Attacks on Short Videos for Content Appropriateness Evaluation", Long-paper (proceedings) track in **SVU Workshop at the IEEE/CVF International Conference on Computer Vision (ICCV)**, 2025. arxiv.2507.11968
- **S M Jishanul Islam**, S. H. Mustakim, S. Ahmmed, M. F. A. Sayeedi, S. Khandoker, S. T. A. Dhrubo, and N. Hossain, "MIMIC: Multimodal Islamophobic Meme Identification and Classification," 3rd **Muslims in ML Workshop – NeurIPS**, 2024. arXiv:2412.00681
- S. Ahmmed, T. Rahman, **S M Jishanul Islam**, A. Reyad, S. Dey, J. A. Purification, and M. D. Farid, "E-MedViTR: Enhanced Vision Transformers with Registers for Biomedical Image Classification," in Proceedings of the 6th **International Conference on Electrical Engineering and Information & Communication Technology (ICEEICT)**, 2024. 10.1109/ICEEICT62016.2024.10534573

NOTABLE RESEARCH PROJECTS

An audio video-based multi-modal fusion approach for speech emotion recognition.

- **Co-supervised by**: Md. Rayhan Ahmed, Dr. Salekul Islam, Dr. Swakkhar Shatabda, and Dr. A.K.M Muzahidul Islam.
- Under Review at Knowledge-Based Systems (Elsevier) [codebase].
- We present an approach to classify human emotions by fusing audio and visual inputs. Our approach sets new state-of-the-art results while keeping the architecture very simple. Additionally, we present a new frame filtering strategy to overcome the problem of spatiotemporal redundancy.

ChimeraBreak: a novel coordinated tri-modal attack on MLLMs

- **Supervised by**: Dr. Swakkhar Shatabda.
- Long paper (Proceedings), SVU Workshop @ ICCV 2025. [paper(arxiv)] [code] [dataset].

- We introduce SVMA, an adversarial dataset for content moderation in short-form videos, and ChimeraBreak. This coordinated strategy exposes systemic safety flaws in leading MLLMs for content appropriateness evaluation. We experiment with open and closed source MLLMs and utilize LLM-as-a-Judge to evaluate ethical reasoning and confidence.

MIMIC: Multimodal Islamophobic Meme Identification and Classification

- **Supervised by**: Nahid Hossain.
- Accepted at MusIML Workshop - NeurIPS 2024. Ongoing intl. collaboration. [arXiv version] [OpenReview] [codebase].
- First work to combat Islamophobic hate shown through memes by creating a novel dataset out of Islamophobic memes. We also propose a classifier based on the Vision-and-Language Transformer (ViLT) specifically tailored to identify anti-Muslim hate within memes by integrating both visual and textual representations.

NOTABLE PROJECTS

NurtureAid (github.com/sadia-ahmmed/AgeWell-Frontend)
Built a real-time AI-based cross-platform mobile application that simplifies care through caretakers. Led the app's backend and frontend development as a full-stack engineer.
Tech Stack: React Native, Node.js, Flask, MongoDB, Firebase, LLaMA-Index, PyTorch.
Acknowledgements: Champion of the UIU CSE Project Show, Software Engineering Laboratory; Champion of the Hult Prize OnCampus round in UIU; Selected for the Hult Prize Summit in Boston.

IPBlocks (github.com/S-M-J-I/ipblocks-blockchain)
Developed a web3 dApp that helps users protect, license, and commercialize Intellectual Property (IP) through blockchain technology. Led the development of the app's backend, frontend, and blockchain development as a full-stack engineer.
Tech Stack: HTML, CSS, JavaScript, PHP, PostgreSQL, Solidity, Truffle, Ganache.
Acknowledgements: Gold Award (Champion) at the International Blockchain Olympiad 2023.

BhaShammo (https://github.com/S-M-J-I/bhashammo-dgt-regional-bengail-2-ipa)
Solving the unique problem of IPA Transcription of Bengali Regional Dialects using Dialect Guided Tokens.
Tech Stack: Python, Hugging Face, PyTorch.
Acknowledgements: Champion of the Bhashamul national NLP datathon.

SKILLS

Programming: C, C++, Java, Python, JavaScript, PHP, SQL, Solidity.
Machine Learning: PyTorch, Tensorflow, Scikit-Learn, Matplotlib, Numpy, Pandas, NLTK, HuggingFace.
Frameworks/Libraries: React.js, Node.js, Spring Boot, FastAPI, Firebase.
DevOps Tools: Git, GitHub, Docker.
Soft Skills: Communication, Leadership, Design Thinking, Public Speaking, Team Player.

ACHIEVEMENTS

Winner, Harvard Health Systems Innovation Lab Hackathon 2025 (Dhaka Hub) 2025
Academic Excellence Scholarship 2020 — 2024
Gold Award (Champion), International Blockchain Olympiad (IBCOL) 2023 2023
Champion, Bhashamul: Bengali Regional Text to IPA National NLP Datathon 2024
Champion, Hult Prize OnCampus Round 2024
Summit Participant, Hult Prize Summit in Boston, USA 2024
Silver Award(1^{st} Runner Up), Bangladesh Blockchain Olympiad (BCOLBD) 2023 2023
1^{st} Runner Up, Intra-University Deep Learning Sprint 2023
Finalist, ICT Innovation Grant (for NurtureAid) 2023
Champion, Software Engineering Lab 2023
Champion, Database Management Systems Lab 2022
Champion, Advanced Object Oriented Programming Lab 2022
The Daily Star Award 2017

EXTRA CURRICULARS

UIU Computer Club - served as the General Secretary 2022 — 2024
UIU Entrepreneur Development Forum - served as the Organizing Secretary 2020 — 2022
Triumph of IBCOL 2023 and the Future of Blockchain in Bangladesh by Jamuna TV - interview link.
Trainer, Hands-on Workshop on Automatic Speech Recognition by Bengali.AI and IUT Computer Society - workshop link.
Trainer, Hands-on Workshop on building a dApp from scratch by UIU Computer Club - workshop link.

REFERENCES

Dr. Swakkhar Shatabda
Professor of Computer Science and Engineering
BRAC University, Dhaka, Bangladesh
E-mail: swakkhar.shatabda@bracu.ac.bd

Dr. Salekul Islam
Professor of Electrical and Computer Engineering
North South University, Dhaka, Bangladesh
E-mail: salekul.islam@northsouth.edu