

Fatema Tuj Johora Faria

✉ fatema.faria142@gmail.com 🌐 fatemafaria142.github.io in fatema-faria 📧 fatemafaria142 🎓 Fatema Faria
📄 Fatema-Faria ID Fatema Tuj Johora Faria

Research Interests

Large Language Models, Large Multimodal Models, LLM Agents, Multimodal AI Agents, Human-LLM Interaction, NLP for Social Good, NLP for Low-Resource Languages, AI in Healthcare, Vision-Language Models, Trustworthy AI, and Computer Vision.

Education

B. Sc. in Computer Science and Engineering

Ahsanullah University of Science and Technology (AUST)

July 2019 – December 2023

Dhaka, Bangladesh

Undergraduate Thesis Title: Generative Adversarial Networks for Crop Disease: A Case Study with Potato Disease Classification and Instance Segmentation

Supervisor: Dr. Mohammad Shafiul Alam, Professor, Department of CSE, AUST

Co-supervisor: Khan Md Hasib, Assistant Professor, Department of CSE, Bangladesh University of Business and Technology

CGPA: 3.302 (Ranked 84th out of 145 students) 🔗 [Transcript](#)

Research Experience

Remote Research Assistant

June 2024 – Present



Supervisor: Dr. Laith H. Baniata, Research Professor, Gachon University, South Korea

- Led research on **“Towards Robust Chain-of-Thought Prompting with Self-Consistency for Remote Sensing VQA: An Empirical Study Across Large Multimodal Models”**. Supported by the National Research Foundation of Korea (Grant No. NRF-2022R1A2C1005316), funded by the Ministry of Science and ICT.
 - Proposed Zero-GeoVision, a zero-shot remote sensing framework employing LMMs (GPT-4o, Grok 3, Gemini 2.5 Pro, Claude 3.7 Sonnet) to interpret high spatial resolution (HSR) satellite images and answer six task-specific questions.
 - Designed CoT-GeoReason, a chain-of-thought extension of Zero-GeoVision, guiding LMMs to produce step-by-step reasoning before final answers using controlled decoding parameters to enable complex reasoning.
 - Developed Self-GeoSense, built on CoT-GeoReason by generating five independent CoT reasoning paths per input, aggregating answers via majority voting to improve consistency in remote sensing interpretation.
 - Implemented Geo-Judge, a two-stage evaluation framework for generative remote sensing VQA (Stage 1: automatic LMM-based labeling; Stage 2: human-in-the-loop review with feedback on model reasoning and predictions).
- Investigated research on **“Analyzing Diagnostic Reasoning of Vision-Language Models via Zero-Shot Chain-of-Thought Prompting in Medical Visual Question Answering”**. Funded by the National Institute of Health, South Korea (Project No. 2024ER080300), and the National Research Foundation of Korea.
 - Evaluated Large Vision-Language Models (Gemini 2.5 Pro, Claude 3.5 Sonnet, GPT-4o Mini) on the PMC-VQA benchmark, identifying limitations of zero-shot learning (ZSL), which often produced hallucinated answers to multiple-choice radiology questions due to insufficient clinical decision-making transparency.
 - Proposed a Zero-shot Chain-of-Thought (Zero-CoT) prompting framework to enable multi-step diagnostic reasoning by decomposing problems, analyzing visual and contextual cues, and generating stepwise explanations, achieving a 41.68% accuracy improvement over existing models.
- Carried out research on **“SentimentFormer: A Transformer-Based Multi-Modal Fusion Framework for Enhanced Sentiment Analysis of Memes in the Under-Resourced Bangla Language”**. Supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF), funded by the Ministry of Science and ICT under Grant No. NRF-2022R1A2C1005316.
 - Implemented the SentimentFormer framework by fusing SwiftFormer’s visual features with mBERT’s textual embeddings at an intermediate layer, which enhanced cross-modal interactions and improved the Weighted F1 Score from 64.3 to 73.28, surpassing all previously reported research baselines on the MemoSen dataset.
- Conducted research on **“Investigating the Predominance of Large Language Models in Low-Resource Bangla Language Over Transformer Models for Hate Speech Detection: A Comparative Analysis”**. Supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF), funded by the Ministry of Science and ICT under Grant No. NRF-2022R1A2C1005316.
 - Applied Zero-Shot and Few-Shot Learning approaches (5-shot, 10-shot, and 15-shot) using proprietary LLMs (GPT-3.5 Turbo and Gemini 1.5 Pro) for low-resource Bengali hate speech detection and improved accuracy by 6.28% over prior approaches across three datasets: BD-SHS, Bengali Hate Speech Dataset, and Bengali Hate Dataset.


Publications (* denotes equal contribution) [95+ citations]

Journals

- Fatema Tuj Johora Faria**, Mukaffi Bin Moin, Busra Kamal Rafa, Swarnajit Saha, Md. Mahfuzur Rahman, Khan Md Hasib, and M. F. Mridha. (2025). **“BanglaCalamityMMD: A Comprehensive Benchmark Dataset for Multimodal Disaster Identification in the Low-Resource Bangla Language,”** *International Journal of Disaster Risk Reduction*, 130, 105800. <https://doi.org/10.1016/j.ijdr.2025.105800> (Elsevier, Q1, IF: 4.5)
- Fatema Tuj Johora Faria**, Laith H. Baniata, Ahyoung Choi, and Sangwoo Kang. **“Towards Robust Chain-of-Thought Prompting with Self-Consistency for Remote Sensing VQA: An Empirical Study Across Large Multimodal Models,”** *Mathematics* 2025; 13(18):3046. <https://www.mdpi.com/2227-7390/13/18/3046> (MDPI, Q1, IF: 2.2)

- **Fatema Tuj Johora Faria**, Mukaffi Bin Moin, Zayed Hasan, Md. Arafat Alam Khandaker, Niful Islam, Khan Md Hasib, and M.F. Mridha. 2025. **“MultiBanFakeDetect: Integrating Advanced Fusion Techniques for Multimodal Detection of Bangla Fake News in Under-Resourced Contexts,”** *International Journal of Information Management Data Insights*, 5 (2): 100347. <https://doi.org/10.1016/j.jjimei.2025.100347> (Elsevier, Q1, IF: 15.53)
- **Fatema Tuj Johora Faria**, Laith H. Baniata, Ahyoung Choi, and Sangwoo Kang. **“Analyzing Diagnostic Reasoning of Vision–Language Models via Zero-Shot Chain-of-Thought Prompting in Medical Visual Question Answering,”** *Mathematics*. 2025; 13(14):2322. <https://doi.org/10.3390/math13142322> (MDPI, Q1, IF: 2.2)
- **Fatema Tuj Johora Faria**, Laith H. Baniata, Mohammad H. Baniata, Mohammad A. Khair, Ahmed Ibrahim Bani Ata, Chayut Bunternghchit, and Sangwoo Kang. 2025. **“SentimentFormer: A Transformer-Based Multimodal Fusion Framework for Enhanced Sentiment Analysis of Memes in Under-Resourced Bangla Language,”** *Electronics* 14, no. 4: 799. <https://doi.org/10.3390/electronics14040799>. (MDPI, Q2, IF: 2.6)
- **Fatema Tuj Johora Faria**, Laith H. Baniata, and Sangwoo Kang. **“Investigating the Predominance of Large Language Models in Low-Resource Bangla Language over Transformer Models for Hate Speech Detection: A Comparative Analysis,”** *Mathematics* 2024, 12, 3687. <https://doi.org/10.3390/math12233687>. (MDPI, Q1, IF: 2.2)
- **Fatema Tuj Johora Faria**, Mukaffi Bin Moin, Md. Mahfuzur Rahman, Khan Md Hasib, Md. Jakir Hossen, and M. F. Mridha. **“MindSpeak-Bangla: A Human–LLM Collaborative Dataset for Chain-of-Thought Adaptation in Bangla Mental Health Advice Generation.”** [Under Review in *World Psychiatry*] (Wiley, Q1, IF: 65.8)
- **Fatema Tuj Johora Faria**, Mukaffi Bin Moin, Ahmed Al Wase, Mehidi Ahmmed, Md Rabiul Sani, and Tashreef Muhammad. **“Vashantor: a large-scale multilingual benchmark dataset for automated translation of Bangla regional dialects to Bangla language,”** arXiv preprint arXiv:2311.11142 (2023). [Under Review in *Array*] (Elsevier, Q1, IF: 4.5)  Preprint
- **Fatema Tuj Johora Faria**, Mukaffi Bin Moin, Pronay Debnath, Asif Iftekher Fahim, and Faisal Muhammad Shah. **“Explainable Convolutional Neural Networks for Retinal Fundus Classification and Cutting-Edge Segmentation Models for Retinal Blood Vessels from Fundus Images,”** arXiv preprint arXiv:2405.07338 (2024). [Under Review in *Journal of Visual Communication and Image Representation*] (Elsevier, Q1, IF: 3.1)  Preprint

Conference Proceedings

- **Fatema Tuj Johora Faria**, Moin, M. B., Fahim, A. I., Debnath, P., & Shah, F. M. (2025)., **“Unraveling the Dominance of Large Language Models Over Transformer Models for Bangla Natural Language Inference: A Comprehensive Study,”** *Proceedings of Fourth International Conference on Computing and Communication Networks. ICCCN 2024. Lecture Notes in Networks and Systems* (Vol. 1396). Springer, Singapore. doi: [10.1007/978-981-96-6124-4_2](https://doi.org/10.1007/978-981-96-6124-4_2).
- **Fatema Tuj Johora Faria**, M. B. Moin, M. M. Rahman, M. M. A. Shanto, A. I. Fahim, & M. M. Hoque. **“Uddeshho: An Extensive Benchmark Dataset for Multimodal Author Intent Classification in Low-Resource Bangla Language,”** *Proceedings of International Conference on Information Technology and Applications (ICITA 2024)*, *Lecture Notes in Networks and Systems*, vol. 1248, Springer, Singapore, 2025. doi: [10.1007/978-981-96-1758-6_32](https://doi.org/10.1007/978-981-96-1758-6_32).
- Moin, M. B., **Fatema Tuj Johora Faria**, Saha, S., Rafa, B. K., Alam, M. S. (2025). **“Exploring Explainable AI Techniques for Improved Interpretability in Lung and Colon Cancer Classification,”** *Proceedings of Fourth International Conference on Computing and Communication Networks. ICCCN 2024. Lecture Notes in Networks and Systems*, vol 1396. Springer, Singapore. doi: [10.1007/978-981-96-6124-4_1](https://doi.org/10.1007/978-981-96-6124-4_1).
- **Fatema Tuj Johora Faria***, Mukaffi Bin Moin*, Rabeya Islam Mumu, Md Mahabubul Alam Abir, Abrar Nawar Alfy, and Mohammad Shafiul Alam., **“Motamot: A Dataset for Revealing the Supremacy of Large Language Models Over Transformer Models in Bengali Political Sentiment Analysis,”** *2024 IEEE Region 10 Symposium (TENSYP)*, New Delhi, India, 2024, pp. 1-8, doi: [10.1109/TENSYP61132.2024.10752197](https://doi.org/10.1109/TENSYP61132.2024.10752197).
- **Fatema Tuj Johora Faria**, Mukaffi Bin Moin, Ahmed Al Wase, Md Rabiul Sani, Khan Md Hasib, and Mohammad Shafiul Alam. **“Classification of potato disease with digital image processing technique: a hybrid deep learning framework,”** *2023 IEEE 13th Annual Computing and Communication Workshop and Conference (CCWC)*, Las Vegas, NV, USA, 2023, pp. 0820-0826, doi: [10.1109/CCWC57344.2023.10099162](https://doi.org/10.1109/CCWC57344.2023.10099162).
- **Fatema Tuj Johora Faria***, Mukaffi Bin Moin*, Mohammad Shafiul Alam*, Ahmed Al Wase, Md Rabiul Sani, and Khan Md Hasib. **“PotatoGANs: Utilizing Generative Adversarial Networks, Instance Segmentation, and Explainable AI for Enhanced Potato Disease Identification and Classification,”** arXiv preprint arXiv:2405.07332 (2024). [Under Review in *IEEE i-COSTE 2025*]  Preprint
- **Fatema Tuj Johora Faria**, Mukaffi Bin Moin, Asif Iftekher Fahim, Pronay Debnath, and Faisal Muhammad Shah. **“Bangla MemeEvidence: A Multimodal Benchmark Dataset for Explanatory Evidence Detection in Bengali Memes.”** [Under Review in *2025 9th International Conference on Vision, Image and Signal Processing*]

Work Experience

Dexian (Bangladesh) Limited.

Senior Application Developer

July 2025 – Present

Dhaka, Bangladesh

• Conversational AI Agent Platform for Large-Scale Document Interaction: ShareFlow Agent

- Constructed and deployed a ReAct-based Agentic RAG system integrated with Microsoft SharePoint to retrieve information from selected files across resource pages, along with functionality for processing user-uploaded documents, incorporating custom guardrails that enable users to build personalized agents.
- Orchestrated a multimodal OCR agent with a custom toolset for autonomous content extraction from diverse unstructured sources (scanned PDFs, images, DOCX files, flowcharts, tables, and diagrams), incorporating resilient fallback strategies and controlled retry mechanisms with rate limiting to ensure fault-tolerant performance.

- o Implemented session-based chat with full history retention for context-aware interactions and a reset option, and integrated it with Microsoft Bot Framework to enable interactive conversations within Microsoft Teams.
- o Enabled users to input up to six leading questions to facilitate interactions, and automatically generated at least three questions based on the agent's instructions and description if the user provided none or fewer.
- o Integrated an update feature that enables users to modify existing agents by adding new files, editing leading questions, revising instructions and descriptions, and deleting existing files, along with a real-time file tracking system on the SharePoint site that updates from the agent's memory.
- o Designed an agent-sharing functionality enabling users to share agents publicly with all app users, with individual users, or privately with groups, with automated email notifications to keep collaborators informed.
- o Took leadership of the project while guiding and supporting junior application developers in system design, promoting their technical growth, and upholding high standards of code quality and overall team output.
- o Achieved ~63% operational cost reduction by optimizing custom agent usage for 80 *Sales Managers* handling 50+ interactions per day, replacing the existing SharePoint Agent (Microsoft 365 Copilot Agent).
- o Optimized RAG query search accuracy by 96%, reduced manual effort, lowered token usage costs by 42%, and enabled users to upload or select up to 30–50 files with the ShareFlow Agent, surpassing the Microsoft Copilot (general/free version) limitation of 20 files per conversation.

Tech Stack Used: Python, LlamaIndex, Azure OpenAI (GPT-4.1, GPT-4o, text-embedding-3-large), Microsoft Bot Services, Azure SQL, Azure Functions, AlloyDB for PostgreSQL, Azure App Service, OpenCV, React, FastAPI, Ragas

Dexian (Bangladesh) Limited.

Application Developer

May 2024 – July 2025

Dhaka, Bangladesh

• Organizational Intelligence Role Placement System: Org Info

- o Implemented a multimodal agent to extract organizational hierarchy from organograms using tree-of-thought (ToT) prompting, incorporating multipath reasoning and Breadth-First Search (BFS) to ensure accurate role placement, and storing the hierarchy in a database after cross-checking with existing data and mapping it to Bullhorn records.
- o Designed an LLM-based agent that converts natural language queries into optimized SQL (text-to-SQL) using chain-of-thought (CoT) with self-consistency prompting, retrieves relevant organizational data, and integrates the results into the “OrgChart front-end framework” for hierarchical visualization.
- o Engineered a dynamic LLM-based agentic RAG-guided chat interface called “OrgInfo Assistant” that allowed users to interact with specific organizational hierarchies using predefined query types, specialized roles, and goals, where user queries were first converted to SQL (text-to-SQL), executed on a temporary organization-specific database, and the retrieved results were converted to natural language (SQL-to-text).
- o Developed a 7-day summarization of organizational activities by extracting relevant notes on placements, submissions, and communication logs from the Bullhorn database, and set up scheduled jobs to update a SQL database every 7 days with new and updated organizational data.
- o Deployed and optimized organizational hierarchy search for *Account Managers* by eliminating full Bullhorn database queries, reducing search time by 92%, and enabling faster access to relevant data.

Tech Stack Used: Python, LangChain, LangGraph, Azure OpenAI (GPT-4.1, GPT-4o-Mini, text-embedding-3-small), Prompt Engineering, Azure SQL, Azure App Service, OpenCV, React, FastAPI, WebSocket, Docker

• Next-Gen Proposal Insights Automation Engine: RFPMatcher

- o Architected a RAG solution utilizing a temporary vector database and Chain-of-Thought prompting with 12 specialized analytical prompts to systematically extract key information, such as client details, scope of work, deliverables, and submission timelines, from Request for Proposal (RFP) documents.
- o Orchestrated a comprehensive master database by processing historical Request for Proposal (RFP) responses using 38+ predefined questions (23 deliverable-focused and 15 experience-based), incorporating metadata such as detailed generated answers, chunk summaries, manual classification of RFP responses (winning or losing), and question categories (Deliverables or Experience), while storing vector embeddings to enable semantic search capabilities.
- o Designed a Past Experience Matcher system that analyzes new RFP documents by dynamically extracting three key requirements (core client needs or expectations), generates both binary questions (Yes/No — whether the company has handled similar situations before) and descriptive questions (open-ended questions to explore how the company addressed similar challenges or implemented solutions) for each chunk, then queries a master database, performs semantic similarity matching to identify relevant past experiences.
- o Developed a Table of Contents (TOC) generation pipeline that analyzes extracted RFP key information against predefined section libraries (10 standard and 12 non-standard sections), employing GPT-4 to intelligently select and prioritize relevant sections based on project requirements.
- o Implemented a section-based conversational AI platform that generates initial content for each TOC section using extracted key information and past experience data, and enables iterative refinement through targeted chat interactions where users can request modifications, while maintaining conversation history.
- o Applied a set of evaluation metrics (Contextual Precision, Contextual Recall, Response Relevancy, Faithfulness, and Factual Correctness) to measure the overall performance of the RAG application on custom-created gold standard datasets for key information extraction and past experience matching.
- o Reduced manual review time by 75% and accelerated decision-making through automated extraction and predictive insights, enabling *Proposal Managers* to focus on strategic bid development.

Tech Stack Used: Python, LlamaIndex, Azure OpenAI (GPT-3.5 Turbo, GPT-4, text-embedding-3-large), Prompt Engineering, AlloyDB for PostgreSQL, CouchDB, Azure App Service, React, FastAPI, DeepEval, Docker

• **Automated Presentation Insights Generator: CaseAligner**

- Designed and deployed an LLM-powered application that repurposes existing client-facing PowerPoint presentations for case studies, using zero-shot prompting to transform them into new practice areas and industries, enabling Sales Managers to rapidly generate domain-specific demo presentations.
- Implemented an interactive chat interface linked to separated slides, allowing users to query and modify specific slide content, with session-wise conversation history stored for reference.
- Created a summarization feature for all existing case studies, highlighting technology stacks, work types, key benefits, and outcomes to assist users in quickly identifying relevant content.
- Developed comprehensive search functionality to locate information across all generated case studies, including references to specific slide numbers for precise navigation.
- Added export functionality to download slides in the company's official presentation template.
- Built a user interface for users to download and edit the knowledge base of case studies used for generating new ones.
- Significantly accelerated demo preparation by reducing slide crafting time by approximately 90%, enabling Salespersons to focus more on client engagement and closing deals.

Tech Stack Used: Python, LlamaIndex, Azure OpenAI (GPT-3.5 Turbo), React, FastAPI, Azure App Service, Docker

• **Smart Recruitment Analytics Tool: AgentDexi**

- Formulated a RAG system to identify technological trends and track the top 20 most in-demand skills by extracting information from job descriptions scraped from external company postings.
- Developed an automated data analysis pipeline with interactive graphs and charts to compare company-wise data, enabling *Technical Recruiters* to make data-driven talent acquisition decisions.
- Orchestrated a web Research Agent that autonomously performs web searches using the Google Search API and Tavily Search API to extract company information from diverse sources, retrieving the top 10 ranked URLs with contextual preview summaries and filtering results based on relevance to support downstream data extraction and analysis tasks.
- Implemented a Website Scraping Agent utilizing ScraperAPI to extract targeted data from discovered URLs, processing raw HTML/text content into structured data, and built a custom RAG pipeline that filters contextually relevant information and generates structured outputs, including tables, organizational reports, and business intelligence summaries.

Tech Stack Used: Python, LangChain, Langfuse, CrewAI, Azure OpenAI (GPT-4o, text-embedding-3-small), ChromaDB, JobSpy, React, FastAPI

Ongoing Research Projects

- BanglaMedQA: A Dataset for Adapting Zero-Shot Chain-of-Thought Reasoning in Bengali Medical Question Answering
- Cross-Cultural Moral Bias Detection in Story Understanding: Analyzing Intentionality and Fairness in LLM Judgments
- Breaking Silence: A Jailbreaking Prompt Framework for Generating Sensitive and Controversial Narratives in Bangla

Technical skills

Programming Languages	Python (NumPy, SciPy, Matplotlib, Pandas, Seaborn), Java, C++
Web Development	HTML5, CSS3, JavaScript, FastAPI, Flask, React, Streamlit
Database	MySQL, MongoDB
Deep Learning Frameworks	TensorFlow, Keras, PyTorch
LLM Application Frameworks	LangChain, LangGraph, LlamaIndex
LLM Evaluation Frameworks	LangSmith, Langfuse, Ragas, DeepEval
Vector Database	ChromaDB, FAISS
Cloud Services	Azure OpenAI, Azure SQL Database, Azure App Service, Azure Blob Storage, Azure Boards, Azure Functions, AlloyDB for PostgreSQL, Microsoft Bot Services
Others	Prompt Engineering, Context Engineering, Docker, CrewAI, GitHub, Github Copilot, OpenCV, Websocket, Apache Airflow, Hugging Face Transformers

Awards & Achievements

1. Achievement Award DEXIAN (BANGLADESH) LIMITED – H1 2025 REFLECTION	14 th August 2025 Dhaka, Bangladesh
🏆 Recognized for Excellence in Strategic Leadership, Agile Project Delivery, and Results-Driven Execution	
2. DL Sprint 2.0 - BUET CSE Fest 2023 Bengali Document Layout Analysis Competition	5 th August 2023 Dhaka, Bangladesh
🏆 Led “Team AustPhoenix” and secured 10 th position out of 94. 🔗 Overview	
3. Poster Presentation RESEARCH SYMPOSIUM 2023: AN INTRA-AUST RESEARCH EXHIBITION	14 th June 2023 Dhaka, Bangladesh
🏆 “Classification of Potato Disease with Digital Image Processing Technique: A Hybrid Deep Learning Framework,” secured 1 st position in “RESEARCH SYMPOSIUM 2023” organized by AUST Research and Publication Club. 🔗 Poster	

Reviewer Experience

- Scientific Reports [📄 Certificate](#)
 - Language Resources and Evaluation [📄 Certificate](#)
- Discover Mental Health [📄 Certificate](#)
 - Cluster Computing [📄 Certificate](#)