Fatema Tuj Johora Faria

fatema.faria142@gmail.com | fatemafaria142.github.io | fatema-faria | fatemafaria142 | Fatema Faria
Fatema-Faria | 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 — *July 2019 – December 2023*
Ahsanullah University of Science and Technology (AUST) — *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) [90+ 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.ijdrr.2025.105800 **(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 **(Q1, IF: 2.2)**

- **Fatema Tuj Johora Faria**, Mukaffi Bin Moin, Zayeed 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 **(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 **(Q1, IF: 2.2)**
- **Fatema Tuj Johora Faria**, Laith H. Baniata, Mohammad H. Baniata, Mohannad A. Khair, Ahmed Ibrahim Bani Ata, Chayut Bunterngchit, 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. **(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. **(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]** **(Q1, IF: 65.8)**
- **Fatema Tuj Johora Faria**, Mukaffi Bin Moin, Ahmed Al Wase, Mehidi Ahmmed, Md Rabius 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]** **(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] **(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.
- **Fatema Tuj Johora Faria**, M. B. Moin, M. M. Rahman, M. M. A. Shanto, A. I. Fahim, & M. M. Hoque. **"Uddessho: 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.
- 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.
- **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 (TENSYMP)*, New Delhi, India, 2024, pp. 1-8, doi: 10.1109/TENSYMP61132.2024.10752197.
- **Fatema Tuj Johora Faria**, Mukaffi Bin Moin, Ahmed Al Wase, Md Rabius 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.
- **Fatema Tuj Johora Faria***, Mukaffi Bin Moin*, Mohammad Shafiul Alam*, Ahmed Al Wase, Md Rabius 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. — *July 2025 – Present*
Senior Application Developer — *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 and user-uploaded document support, incorporating custom guardrails to enable users to build personalized agents for retrieving information from e-learning training materials, recruitment processes, and internal financial document resource pages.
  - Orchestrated a Multimodal OCR Agent with a custom-built toolset to process diverse unstructured sources (scanned PDFs, images, DOCX, flowcharts, tables, diagrams), incorporating resilient fallback strategies and controlled retry mechanisms with rate limiting to ensure robust and fault-tolerant operation.

- Implemented session-based chat ensuring separate user-agent conversations with full history retention for context-aware interactions and a reset option to start anew.
- Integrated an update feature enabling users to modify existing agents by adding new files, editing leading questions, revising instructions and description, and deleting existing files from the agent's memory.
- 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.
- Developed an Action Tracking system to log API usage, capturing trigger events, execution time, input queries, and generated responses in the database for monitoring and analytics.
- Took leadership of the project while guiding and supporting junior application developers, promoting their technical growth and upholding high standards of code quality and overall team output.
- Achieved approximately 63% cost reduction by optimizing custom agent usage for 80 *Sales Managers* handling 50+ interactions per day, replacing the existing SharePoint Agent.
- Optimized RAG query search accuracy by 96%, significantly reducing manual effort.
  Tech Stack Used: Python, LlamaIndex, Azure OpenAI, Azure SQL, Azure Functions, AlloyDB for PostgreSQL, Azure App Service, OpenCV, React, FastAPI, Ragas

Dexian (Bangladesh) Limited. *May 2024 – July 2025*
Application Developer *Dhaka, Bangladesh*

- **Organizational Intelligence Role Placement System: Org Info**
  - 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.
  - 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.
  - 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).
  - 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.
  - 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, OpenCV, Azure SQL, React, FastAPI
- **Next-Gen Proposal Insights Automation Engine: RFPMatcher**
  - Architected a RAG solution utilizing Chain-of-Thought prompting to accurately extract 12 critical data points, such as project title, issuing organization, submission deadline, project scope, and submission method, from complex Request for Proposal (RFP) documents to optimize proposal analysis.
  - Performed semantic chunking of documents, storing 5–6 QA pairs with a summary per chunk, which allowed the system to perform targeted semantic searches within relevant chunks rather than the entire document, reducing search space by 65%, enabling faster retrieval, and generating accurate answers with 90% improved precision.
  - Orchestrated a Past Experience Matcher system integrating Automatic Chain-of-Thought prompting, in-context learning, and custom stakeholder-defined questions to extract RFP requirements and match them against a master database of prior successful Responses for Proposal.
  - Enabled the system to generate Yes/No decisions with detailed explanations of how similar requirements were addressed in the past, aiding in the prediction of potential win/loss outcomes for new proposals.
  - Generated dynamic Tables of Contents (TOC) by extracting key requirements from RFPs and relevant insights from historical proposal responses, thereby streamlining the proposal writing process for new bids.
  - Implemented session-based chat functionality linked to each user's generated TOC, retaining conversation history and providing an editing system for responses before integrating them into proposal writing.
  - 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, AlloyDB for PostgreSQL, CouchDB, Azure App Service, FastAPI
- **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.

- Introduced 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, React, FastAPI

- **Smart Recruitment Analytics Tool: AgentDexi**
  - Constructed a RAG system to identify tech trends through scraping and analyzing job postings from external companies.
  - Designed an LLM-based multi-agent system to generate detailed Customer Intelligence Reports, highlighting companies seeking candidates and providing actionable insights into market needs.
  - Created interactive charts and dashboards to interpret data insights, compare company demands, track skill gaps, benchmark roles, and optimize hiring strategies through data-driven decisions.
  - Empowered users to customize agents through a user interface by defining Role, Goals, and Backstory, while the system autonomously selected optimal tools to generate interview questions and engagement strategies for relevant positions.
  - Minimized analysis time by 80%, allowing *Technical Recruiters* to accelerate decision-making, optimize candidate sourcing, improve interview planning, and enhance overall hiring outcomes.

    Tech Stack Used: Python, LangChain, CrewAI, Azure OpenAI, AlloyDB for PostgreSQL, Azure App Service, JobSpy, React, FastAPI
- **Personalized Assistant for Internal Knowledge Discovery: KnowledgeEngine**
  - Orchestrated an LLM-based RAG system employing chain-of-thought reasoning to identify relevant internal legal documents across multiple sources and generate accurate, context-aware answers by synthesizing retrieved information.
  - Engineered semantic chunking of documents by storing 3–4 QA pairs with summaries per chunk, enabling targeted searches within relevant sections, reducing search space by 60%, speeding up retrieval from long documents, and improving answer precision by 95%.
  - Managed a session-based chat interface that delivers responses with precise document page references and maintains conversation history, including reset options.
  - Generated suggestive questions from user-uploaded documents to guide users and help them explore related information effectively.
  - Developed a user interface enabling dynamic management of the RAG knowledge base, allowing users to delete or update documents in real time.
  - Took leadership of the project and evaluated the performance of the RAG system against a custom gold-standard dataset to measure accuracy, relevance, and retrieval consistency.

    Tech Stack Used: Python, LangChain, Azure OpenAI, AlloyDB for PostgreSQL, Azure App Service, React, FastAPI, Ragas

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, Ragas
Vector Database	ChromaDB, FAISS
Cloud Services	Azure OpenAI, Azure SQL Database, Azure App Service, Azure Blob Storage, Azure Boards, Azure Functions, AlloyDB for PostgreSQL
Others	Prompt Engineering, Context Engineering, Docker, CrewAI, OpenCV, GitHub, Github Copilot, Apache Airflow, Hugging Face Transformers

Awards & Achievements

1. Achievement Award — *14th August 2025*
DEXIAN (BANGLADESH) LIMITED – H1 2025 REFLECTION — *Dhaka, Bangladesh*
Recognized for **Excellence in Strategic Leadership, Agile Project Delivery, and Results-Driven Execution**

2. Poster Presentation — *5th August 2023*
RESEARCH SYMPOSIUM 2023: AN INTRA-AUST RESEARCH EXHIBITION — *Dhaka, Bangladesh*
"Classification of Potato Disease with Digital Image Processing Technique: A Hybrid Deep Learning Framework," secured 1st 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**