# Fatema Tuj Johora Faria

fatema.faria142@gmail.com https://fatemafaria142.github.io/
**LinkedIn:** https://www.linkedin.com/in/fatema142/
**ResearchGate:** https://www.researchgate.net/profile/Fatema-Faria
**GitHub:** https://github.com/fatemafaria142

## *Research Interests*

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

## *Education*

**B. Sc. in Computer Science and Engineering** *July 2019 – December 2023*
*Ahsanullah University of Science and Technology* *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*
***CGPA:*** *3.302 (Ranked 84$^{th}$ out of 145 students)*

## *Research Experience*

**Remote Research Assistant** *June 2024 – Present*
***Supervisor:*** *Dr. Laith H. Baniata, Research Professor, Gachon University, South Korea*

- Currently working on the research titled **"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.*
  - Developed three frameworks: Zero-GeoVision, CoT-GeoReason, Self-GeoSense with five independent reasoning chains, and implemented Geo-Judge, a two-stage evaluation method assessing the quality of reasoning steps, and achieved higher performance across six types of questions compared to previous methods on the EarthVQA dataset.
- 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.*
  - Designed a zero-shot chain-of-thought prompting framework to guide Vision-Language Models (Gemini 2.5 Pro, Claude 3.5 Sonnet, and GPT-4o mini) in performing multi-step diagnostic reasoning on clinical images using the PMC-VQA benchmark, leading to a 41.68% improvement in accuracy 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 *SentimentFormer* by fusing SwiftFormer's visual features with mBERT's textual embeddings at an intermediate layer, which enhanced cross-modal interactions and led to a significant improvement in 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 GPT-3.5 Turbo and Gemini 1.5 Pro for low-resource Bengali hate speech detection, achieving a 6.28% accuracy improvement over prior approaches across three datasets: BD-SHS, Bengali Hate Speech Dataset, and Bengali Hate Dataset.

## *Work Experience*

**Dexian (Bangladesh) Limited.** *July 2025 – Present*
*Senior Application Developer* *Dhaka, Bangladesh*

- **Conversational Agent Platform for Legal Document Assistance: ShareFlow Agent**
  - Developed and deployed a Microsoft SharePoint-integrated custom ReAct Agentic RAG system that enables users to create personalized tools for retrieving information from long legal document repositories.
  - Implemented session-based chat functionality ensuring each user's conversations with individual agents are kept separate, with full history retention for context-aware interactions.
  - Designed a sharing functionality that allows users to share their agents either publicly with all app users or privately with a specific group of people for collaborative use.
  - Mentored junior application developers on coding standards, architectural best practices, and professional growth to ensure high-quality deliverables and overall team performance.
  - Achieved cost savings of approximately $14,400 per month by optimizing custom agent usage for 80 users with 50 interactions per day instead of using the existing SharePoint Agent.
    **Tech Stack Used:** Microsoft SharePoint, Python, LlamaIndex, Azure OpenAI, Azure SQL, AlloyDB, React, FastAPI

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

- **Organizational Intelligence Role Placement System: Org Info**
  - Implemented a multimodal agent for extracting organizational hierarchies from organograms using Tree of Thoughts (ToT) prompting to match and place consultant, employee, and sales representative roles against Bullhorn data.
  - Designed an LLM agent to translate natural language queries into optimized SQL for retrieving organizational data and integrated it into the **"OrgChart frontend framework"** for hierarchical visualization.
  - Developed an LLM-based agentic RAG-guided chat interface, **"OrgInfo Assistant"**, that allowed users to query and interact with specific organizational hierarchies and generated context-aware responses.
  - Reduced organizational hierarchy search time by 92% by eliminating the need to query the entire Bullhorn database, enabling faster access to relevant data.
    **Tech Stack Used:** Python, LangChain, LangGraph, Azure OpenAI, OpenCV, Azure SQL, React JS, FastAPI
- **Next-Gen Proposal Automation Engine: RFPMatcher**
  - Developed a RAG solution leveraging Chain of Thought prompting to accurately extract 12 critical data points such as project title, issuing organization, submission deadline, project scope, and submission method etc. from complex Request for Proposal (RFP) documents to optimize proposal analysis.
  - Built a Past Experience Matcher that uses Automatic-CoT prompting and preset questions to extract RFP requirements, match them with prior proposals, and predict win/loss outcomes with explanations.
  - Generated dynamic Tables of Contents (TOC) based on extracted key information and historical experience to streamline and structure the proposal writing process for new bids.
  - Reduced manual review time by 75%, increased proposal accuracy, and improved decision-making speed through automated extraction and predictive insights, empowering Proposal Managers to focus on strategic bid development.
    **Tech Stack Used:** Python, LlamaIndex, Azure OpenAI, AlloyDB, CouchDB, React JS, DeepEval, FastAPI
- **Smart Recruitment Analytics Tool: AgentDexi**
  - Designed an LLM-based multi-agent system to analyze job demand data and generate detailed Customer Intelligence Reports that highlight companies seeking candidates while providing actionable insights into market needs.
  - Developed a RAG solution that identified and tracked technological trends in job descriptions from external companies and helped uncover emerging skill requirements.
  - Created interactive graphical charts and dashboards that minimized analysis time by 80%, enabling technical recruiters to quickly interpret data insights and make timely, data-driven hiring decisions.
    **Tech Stack Used:** Python, LangChain, CrewAI, Azure OpenAI, React JS, FastAPI
- **Automated Presentation Insights Generator: CaseAligner**
  - Built and deployed an LLM-powered application that uses zero-shot prompting to generate PowerPoint presentations for case studies based on selected. practice areas and industries.
  - Implemented an interactive chat interface allowing users to query slide content and receive instant responses.
  - Designed export functionality to download slides formatted in the company's official presentation template.
  - 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 JS, FastAPI

## *Publications (* denotes equal contribution) [Google Scholar]*

### Journals

- **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**, 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."** **[Under Review in MDPI Remote Sensing (Q1)]**
- **Fatema Tuj Johora Faria**, Mukaffi Bin Moin, Busra Kamal Rafa, Swarnajit Saha, Md. Mahfuzur Rahman, Khan Md Hasib, and M. F. Mridha.**"BanglaCalamityMMD: A Comprehensive Benchmark Dataset for Multimodal Disaster Identification in the Low-Resource Bangla Language."** **[Under Review in International Journal of Disaster Risk Reduction (Q1)]**

- **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 IEEE Open Journal of the Communications Society (OJ-COMS) (Q1)]**
- **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)]** **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)]** **Preprint**

### Conference Proceedings

- **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, Md. Mahfuzur Rahman, Md. Morshed Alam Shanto, Asif Iftekher Fahim, and Md. Moinul 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.
- **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, Asif Iftekher Fahim, Pronay Debnath, and Faisal Muhammad Shah. **"Unraveling the Dominance of Large Language Models Over Transformer Models for Bangla Natural Language Inference: A Comprehensive Study."** arXiv preprint arXiv:2405.02937 (2024). **[Presented at ICCCNet 2024]** **Preprint**
- **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**
- Mukaffi Bin Moin, **Fatema Tuj Johora Faria**, Swarnajit Saha, Bushra Kamal Rafa, and Mohammad Shafiul Alam. **"Exploring Explainable AI Techniques for Improved Interpretability in Lung and Colon Cancer Classification."** arXiv preprint arXiv:2405.04610 (2024). **[Presented at ICCCNet 2024]** **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]**

## *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 JS, Streamlit |
| **Database** | MySQL, MongoDB |
| **Deep Learning Frameworks** | TensorFlow, Keras, PyTorch |
| **LLM Application Frameworks** | LangChain, LangGraph, LangSmith, LlamaIndex, DeepEval, CrewAI |
| **Cloud Services** | Azure OpenAI, Azure SQL Database, Azure App Service, Azure Blob Storage, Azure Boards |
| **Others** | Vector Database, Apache Airflow, Docker, OpenCV, GitHub, GitHub Copilot |

## *Awards & Achievements*

**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**