

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

## Education

### B. Sc. in Computer Science and Engineering

Ahsanullah University of Science and Technology

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

**CGPA:** 3.302 (Ranked 84<sup>th</sup> out of 145 students)

## Research Experience

### Remote Research Assistant

June 2024 – Present

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

- Investigated research on **“Dissecting the Reasoning Capabilities of Vision-Language Models in Medical Visual Question Answering: A Zero-shot Chain-of-Thought Approach”**. Supported by the National Institute of Health, South Korea (Project No. 2024ER080300), and the National Research Foundation of Korea (Grant No. NRF-2022R1A2C1005316), funded by the Ministry of Science and ICT.
  - 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 baseline 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 using intermediate fusion of SwiftFormer and mBERT, which led to a notable boost in performance, with the Weighted F1 Score improving from 64.3 to 73.28, significantly outperforming previous baseline models in multimodal Bangla sentiment analysis of memes.
- 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 (including 5-Shot, 10-Shot, and 15-Shot) using GPT-3.5 Turbo and Gemini 1.5 Pro for Bengali hate speech detection, which eliminated the need for extensive task-specific training data, enabled adaptability to low-resource scenarios, and achieved a 6.28% accuracy gain over baseline techniques across multiple datasets.
- Currently working on the research titled **“Towards Robust Chain-of-Thought Prompting with Self-Consistency for Remote Sensing VQA: A Cross-Model Empirical Study”**.

## Work Experience

### Dexian (Bangladesh) Limited.

Application Developer (AI/ML)

May 2024 – Present

Dhaka, Bangladesh

#### • Project 1: Org Info

- Implemented a multimodal agent for extracting organizational hierarchical information from organograms using in-context learning with tree-of-thought prompting, which incorporates multipath reasoning and Breadth-First Search (BFS) to resolve relational ambiguities and ensure accurate role placement, and stored the extracted hierarchy in a relational database

- o Designed an LLM-based agent that converted natural language queries into SQL using chain-of-thought (CoT) with self-consistency prompting, which enabled contextual reasoning to retrieve relevant organizational data and integrated the results into the “OrgChart front-end framework” for hierarchical visualization
- o Developed 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, roles, and goals, and generated context-aware natural language responses

**Technologies Used:** Python, LangChain, LangGraph, Azure OpenAI, OpenCV, Azure SQL, React JS, FastAPI

#### • **Project 2: RFPMatcher**

- o Developed a Retrieval-Augmented Generation (RAG) solution using CoT prompting to extract key information from Request for Proposal (RFP) documents
- o Built a Past Experience Matcher score that uses Automatic-CoT prompting and preset questions to extract requirements from RFPs, then matches them against a master database of prior proposal responses
- o 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
- o 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

**Technologies Used:** Python, LlamaIndex, Azure OpenAI, AlloyDB, CouchDB, React JS, FastAPI

#### • **Project 3: CaseAligner**

- o Built an LLM-based application using zero-shot prompting to generate PowerPoint presentations for case studies based on selected practice areas and industries
- o Implemented an interactive chat interface allowing users to query specific slide content and receive instant contextual responses
- o Developed comprehensive search functionality to locate information across all generated case studies
- o Created export capabilities for downloading slides in company’s official template

**Technologies Used:** Python, LlamaIndex, Azure OpenAI, React JS, FastAPI

#### • **Project 4: KnowledgeEngine**



- o Developed an LLM-based, multi-document RAG Q&A system for internal document information retrieval
- o Implemented a chat conversation interface with document page references for information sources
- o Maintained session-based dedicated knowledge bases to ensure user-specific context and data isolation
- o Built an admin panel enabling document upload and streamlined document management

**Technologies Used:** Python, LlamaIndex, Azure OpenAI, AlloyDB, React JS, FastAPI

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

#### **Journals .....**

- o **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: 2.3)**
- o **Fatema Tuj Johora Faria**, Laith H. Baniata, Mohammad H. Baniata, Mohannad A. Khair, Ahmed Ibrahim Bani Ata, Chayut Bunterngrchit, 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)**
- o **Fatema Tuj Johora Faria**, Mukaffi Bin Moin, Zayeed Hasan, Md Arafat Alam Khandaker, Niful Islam, Khan Md Hasib, and M. F. Mridha. **“MultiBanFakeDetect: Integrating Advanced Fusion Techniques for Multimodal Detection of Bangla Fake News in Under-Resourced Contexts.”** **[Accepted in International Journal of Information Management Data Insights] (Q1, IF: 19.2)**
- o **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)]**
- o **Fatema Tuj Johora Faria**, Laith H. Baniata, Ahyoung Choi, and Sangwoo Kang. **“Dissecting the Reasoning Capabilities of Vision-Language Models in Medical Visual Question Answering: A Zero-shot Chain-of-Thought Approach.”** **[Under Review MDPI Mathematics (Q1)]**

- **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**
- **Fatema Tuj Johora Faria**, Mukaffi Bin Moin, Ahmed Al Wase, Mehidi Ahmmmed, 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 Neural Computing and Applications \(Q1\)\]](#)  **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 (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, Md Mahfuzur Rahman, Md Morshed Alam Shanto, Asif Iftekher Fahim, and Md Moinul Hoque. **“Uddeshho: An Extensive Benchmark Dataset for Multimodal Author Intent Classification in Low-Resource Bangla Language.”** arXiv preprint arXiv:2409.09504 (2024). [\[Presented at 18<sup>th</sup> International Conference on Information Technology and Application \(ICITA 2024\)\]](#)  **Preprint**
- **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). [\[Presented at ICCNet 2024\]](#)  **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). [\[Presented at ICCNet 2024\]](#)  **Preprint**
- **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, Asif Iftekher Fahim, Pronay Debnath, and Faisal Muhammad Shah. **“BanglaMemeEvidence: A Multimodal Benchmark Dataset for Explanatory Evidence Detection in Bengali Memes.”** [\[Under Review in 2025 9th International Conference on Vision, Image and Signal Processing\]](#)
- Saidur Rahman Sujon, Ahmadul Karim Chowdhury, **Fatema Tuj Johora Faria**, Mukaffi Bin Moin, and Faisal Muhammad Shah. **“Enhancing Bangla NLP Tasks with LLMs: A Study on Few-Shot Learning, RAG, and Fine-Tuning Techniques”** [\[Under Review in 2025 IEEE 9th ICSECS\]](#)

**Ongoing Research Projects**

- MindSpeak-Bangla: A Human-LLM Collaborative Dataset for Mental Health Support in Low-Resource Bengali
- BanglaMedQA: A Dataset for Adapting Zero-Shot CoT Reasoning in Bengali Medical Question Answering

**Technical skills**

<b>Programming Languages</b>	Python (NumPy, SciPy, Matplotlib, Pandas, Seaborn), Java, C++
<b>Web Development</b>	HTML5, CSS3, JavaScript, FastAPI, Flask, React, Streamlit
<b>Database</b>	MySQL, MongoDB
<b>Deep Learning Frameworks</b>	TensorFlow, Keras, PyTorch
<b>LLM Application Frameworks</b>	LangChain, LangGraph, LlamaIndex
<b>Cloud Services</b>	Azure OpenAI, Azure SQL Database, Azure App Service, Azure Blob Storage
<b>Others</b>	Vector Database, Apache Airflow, Docker, CrewAI, OpenCV, GitHub

**Awards & Achievements**

<b>Poster Presentation</b>	5 <sup>th</sup> August 2023
RESEARCH SYMPOSIUM 2023: AN INTRA-AUST RESEARCH EXHIBITION	Dhaka, Bangladesh
<b>Classification of Potato Disease with Digital Image Processing Technique: A Hybrid Deep Learning Framework</b> , secured 1 <sup>st</sup> position in <b>“RESEARCH SYMPOSIUM 2023”</b> organized by AUST Research and Publication Club.  <b>Poster</b>	