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, Large Vision Models, Computer Vision, Medical Imaging, Natural Language Processing, Generative Adversarial Networks, Explainable Artificial Intelligence, Machine Learning, Deep Learning and its applications, Internet of Things (IoT).

Education

July 2019 to Dec 2023

Ahsanullah University of Science and Technology, Dhaka-1208, Bangladesh

B. Sc. in Computer Science and Engineering

CGPA: 3.302 on a scale of 4.00 (83^{rd} in Merit Position Among 133 Students)

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

Work Experience

May 2024 to Present

Application Developer (AI/ML), **Dexian Bangladesh LTD.**, *Dhaka, Bangladesh.*

- Developed the **Agricultural Economics Chatbot**, an LLMs-based multi-document Q&A system for improving information retrieval in the agricultural sector. The chatbot aggregates data from multiple sources, provides timely answers, and enhances decision-making by delivering real-time insights specific to agricultural economics.
- Built **Dexian Innovation Insights**, an LLM-based multi-agent recommendation system that identifies emerging trends in a company's internal projects by analyzing project data and advancements in AI/ML technology. The tool features a recommendations module that guides decision-making based on the analysis.
- Designed **AgentDexi**, an LLMs-based multi-agent system and RAG solution that analyzes job demand across various companies to provide actionable insights. This innovative solution aims to assist technical recruiters in optimizing their hiring strategies by aligning recruitment efforts with current industry trends for better outcomes.
- Developed the **RFP Matcher**, an advanced RAG solution designed to extract key information and summaries from Request for Proposal (RFP) documents using specific prompts. The system evaluates responses to proposals and incorporates a Rubric Score for comparing bids. This functionality aids in determining potential win or loss outcomes.

Research Experience

June 2024 to Present

Research Assistant

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

- 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"**. This work was 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 NRF-2022R1A2C1005316.

Publications (* denotes equal contribution)

Conference Proceedings

- **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."** arXiv preprint arXiv:2409.09504 (2024). ****Accepted in ICITA 2024** [Preprint]**
- **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."** arXiv preprint arXiv:2407.19528 (2024). ****Accepted in IEEE Region 10 Symposium (TENSYMP)** [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). ****Accepted in ICCCNet 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). ****Accepted in ICCCNet 2024** [Preprint]**
- **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."** In 2023 IEEE 13th Annual Computing and Communication Workshop and Conference (CCWC), pp. 0820-0826. IEEE, 2023. **[Published Paper]**
- **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."** **[Submitted to an A* Rank Conference]**
- Saidur Rahman Sujon, Ahmadul Karim Chowdhury, **Fatema Tuj Johora Faria**, Mukaffi Bin Moin, and Faisal Muhammad Shah. **"Tackling Hallucination in Bengali NLP: Enhancing Paraphrase Generation, Reading Comprehension, and Formal Application Writing Using LLMs with Few-Shot Learning, Fine-Tuning, and RAG."** **[Submitted to an A* Rank Conference]**

Journals

- **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****
- **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."** ****Under Review in Online Social Networks and Media****
- **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."** ****Under Review in MDPI Mathematics****
- **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** [Preprint]**

- Mohammad Shafiul Alam*, **Fatema Tuj Johora Faria***, Mukaffi Bin Moin*, 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 Pattern Recognition and Image Analysis** [Preprint]**
- **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 Natural Language Processing Journal** [Preprint]**

Ongoing Research Projects

- Image-to-Text Generation for Agricultural Disease Diagnosis and Recommendations
- Multi-Modal Sentiment Analysis in Under-Resourced Bangla Language
- BanglaMedQA: A Comprehensive Benchmark Dataset for Medical Question Answering
- Mental Health Advice Generation in Low-Resource Bangla Language

Technical Skills

- **Programming Language:** Python, Java, C++
- **Web Development:** HTML5, CSS3, JavaScript, FastAPI, Flask, React, Streamlit
- **Database:** MySQL, MongoDB
- **Deep Learning Frameworks:** TensorFlow, Keras, PyTorch
- **Cloud Services:** Azure OpenAI, Azure Blob Storage, Azure Container Registry
- **Others:** LangChain, LlamaIndex, Vector Database, CrewAI, Prompt Engineering, OpenCV

Awards & Achievements

5th August, 2023 — **Poster Presentation**

- "**Classification of Potato Disease with Digital Image Processing Technique: A Hybrid Deep Learning Framework**", secured 1st position in **"RESEARCH SYMPOSIUM 2023: AN INTRA-AUST RESEARCH EXHIBITION"** organized by AUST Research and Publication Club. **[Poster Link]**

References

Dr. Mohammad Shafiul Alam
Professor, Department of CSE
Ahsanullah University of Science & Technology (AUST), Dhaka, Bangladesh.
Email: shafiul.cse@aust.edu

Faisal Muhammad Shah
Associate Professor, Department of CSE
Ahsanullah University of Science & Technology (AUST), Dhaka, Bangladesh.
Email: faisal.cse@aust.edu

Khan Md. Hasib
Assistant Professor, Department of CSE
Bangladesh University of Business and Technology (BUBT), Dhaka, Bangladesh.
Email: khanmdhasib@bubt.edu.bd