

FATIMA JAHARA

✉ fatimajahara23@gmail.com, fatima@workera.ai ☎ (+880)1631886022

fatimajahara.com

🌐 linkedin.com/in/fatimajahara 🐦 @FatimaJahara1 🏠 github.com/FatimaJahara

AREA OF INTEREST

My interest lies in the broad intersection of Machine Learning and Natural Language Processing, focusing on but not limited to multimodality (mostly vision), multilinguality and cross-linguality, embodied AI, compositionality, semantics, reasoning, and commonsense intelligence. I am equally committed to enhancing robustness, explainability, and fairness in language models and conversational AI across different domains. My interest extends to the intersection of Computer Vision, Human-Computer Interaction (HCI), and NLP in enhancing user-centric experiences.

EDUCATION

Chittagong University of Engineering and Technology, Bangladesh

February 2016 - June 2021

B.Sc. in Computer Science and Engineering

Advisor: Dr. Mohammed Moshui Hoque

Thesis: "Classification of Bangla News Articles by using Multilayer Perception (MLP)" - Development of a hierarchical classification dataset and system for Bangla news articles and headlines using Multi-Layer Perceptron.

GPA: 3.79 (with Honors) / 4.00

Merit Position: 5th among 120

Awards: Dean's Award (for graduating with Honors)

PUBLICATIONS

Published (All published papers can be found [here](#))

- **F. Jahara** O. Sharif, and M. M. Hoque, "Automatic Categorization of News Articles and Headlines using Multi-layer Perceptron," in International Conference on Intelligent Computing & Optimization. Springer, 2021.
- **F. Jahara**, A. Barua, M. A. Iqbal, A. Das, O. Sharif, M. M. Hoque, and I. H. Sarker, "Towards pos tagging methods for Bengali language: A comparative analysis," in International Conference on Intelligent Computing & Optimization. Springer, 2020, pp. 1111–1123.
- D. Talukder and **F. Jahara**, "Real-Time Bangla Sign Language Detection with Sentence and Speech Generation," 2020 23rd International Conference on Computer and Information Technology (ICCIT), 2020.
- D. Talukder, **F. Jahara**, S. Barua and M. M. Haque, "OkkhorNama: BdSL Image Dataset For Real-Time Object Detection Algorithms," 2021 IEEE Region 10 Symposium (TENSYPMP), 2021.

RESEARCH EXPERIENCE

Fatima Al-Fihri Predoctoral Fellowship

April 2023 - Present

Predoctoral Fellow

- Working closely with the NLP Lab at the University of California, Santa Barbara, on "Multilingual Conceptual Coverage in Text-to-Image Models" to understand the semantic and reasoning capabilities of text-to-image models across different languages.
- Focusing on benchmarking consistency and faithfulness in text-to-image models.

CUET NLP Lab, CUET, Bangladesh

June 2019 - June 2021

Undergraduate Student Researcher

- **Dataset Development:** Developed datasets tailored for low-resource languages: (1) BEmoD (Bangla Emotion Dataset) for emotion detection and (2) a comprehensive dataset for fine-grained news categorization.
- **Comparative Analysis of POS Tagging:** Performed a thorough comparative analysis of Parts-of-Speech tagging methodologies based on execution time and accuracy.
- **Fine-Grained Classification of Bangla News:** Proposed a benchmark study on fine-grained categorization of news data for Bangla for categorizing news into four primary categories and 26 fine-grained categories. Experimented with 170 distinct models exploring classical deep learning, pre-trained BERT, and ensemble (average and weighted) models.

PROFESSIONAL EXPERIENCE

Workera

February 2021 - Present

Palo Alto, California, United States

Senior Assessment Developer

January 2023 - Present

- Serve as the Primary Subject Matter Expert developing assessments for Generative AI (tech and non-tech), Diffusion Models, Large Language Models, and others.
- Engineer prompts to integrate LLMs (ChatGPT, Bing, Claude, Bing Chat) into the assessment development pipeline.
- Implemented an adaptive time allocation system for the Workera GOAT™ (Generative Ontology & Authoring Tool).

Assessment Developer

February 2021 - January 2023

- Developed assessments and learning plans as the Primary Subject Matter Expert for Natural Language Processing, Computer Vision, Automatic Speech Recognition, TensorFlow, and PyTorch. Collaborated on MLOps, Responsible AI, AI Communication, Digital Literacy, Data Analytics, and Web Applications.
- Led a Role Delineation Study as a Committee Co-Chair and an SME to align domains with evolving industry roles.
- Partnered on cross-departmental projects focusing on score calibration, skills alignment, learning pathway standardization, and beta testing for Computer Adaptive Tests (CATs).

DeepLearning.ai

December 2020 - June 2021

Mentor, Alpha Testing Consultant

- Assisted in designing functional requirements for the **Practical Data Science Specialization** courses.
- Created and executed test cases for courses, identifying potential coding workbook errors pre-release.
- Identified bugs and provided feedback on course effectiveness and engagement to refine content and structure.

Brain Station 23

August 2019 - August 2019

Trainee Software Developer

- Collaborated on developing the "University Enrollment Log System" to facilitate student-course assignments utilizing Django, Ajax, D3.js, HTML, CSS, and JavaScript.
- **Key Contributions:** Assisted in the backend with Django, contributed to the frontend for an intuitive UI/UX, and participated in iterative testing to refine functionalities.

SELECTIVE PROJECTS [[GITHUB](#)]

- **Bahon:** Engaged in developing "Bahon," a specialized vehicle dataset tailored for Bangladesh to facilitate real-time vehicle detection with a decreased response time using state-of-the-art object detection architectures while laying the groundwork for automating traffic surveillance processes in Bangladesh. (ongoing)
- **Multimodal Emotion Detection:** Working on developing a multimodal emotion detection framework that integrates visual, auditory, and textual data to identify and analyze emotional expressions. (ongoing)
- **Emotion Detection using LLMs:** Working on developing a text-based emotion detection framework to identify emotions using pre-trained and fine-tuned large language models. (ongoing)
- **FoodFond:** Implemented a recommendation engine for restaurants leveraging latent factor collaborative filtering grounded on ratings, reviews, and personalized preferences.

SELECTIVE HONORS AND ACHIEVEMENTS

Predoctoral Fellowship , Fatima Al-Fihri Predoctoral Fellowship	2023
Dean's Award , Chittagong University of Engineering and Technology (CUET)	2021
Undergraduate Merit Scholarship , Chittagong University of Engineering and Technology (CUET)	2019
National Girls' Programming Contest , Bangladesh ICT Division - 2 nd Runners up	2017
Higher Secondary School Certificate (HSC) General Scholarship Award , Bangladesh Education Board	2015
Junior School Certificate (JSC) General Scholarship Award , Bangladesh Education Board	2010

LEADERSHIP AND PROFESSIONAL ACTIVITIES

Workera - Community Leader	February 2021 - May 2021
Coursera - Beta Tester	April 2021 - Present
IEEE, Bangladesh - Member	June 2019 - Present
International Conference on Machine Learning (ICML) - Volunteer	2022
IEEE CUET Women In Engineering, Bangladesh - Vice Chair	November 2019 - November 2020
National High School Programming Contest - Volunteer	2017

SKILLS

Languages:	[Spoken] Bangla (Native), English (Fluent), Hindi & Urdu (Intermediate) [Others] Python, MySQL, SQLite, C, C++, Java, HTML5, CSS3, JavaScript, Ajax
Frameworks & Libraries:	TensorFlow, Keras, PyTorch, Scikit-learn, NumPy, Pandas, Matplotlib, Seaborn, Django Rest Framework, Amazon Web Services, Rasa
Fields:	Machine Learning, Deep Learning, Natural Language Processing, Computer Vision, Data Analysis, Data Science, Web Development, Data Visualization, Quantum Computing
Interpersonal Skills:	Teamwork and Collaboration, Leadership, Adaptability, Work Ethic, Problem Solving, Empathy and Understanding, Communication
Technologies:	GIT, Latex, Web Scraping, ChatGPT, Bing-Chat, Bard, Claude-2

SELECTIVE TRAINING AND WORKSHOPS

Qubit by Qubit 2020-2021 Introduction to Quantum Computing- *IBM Quantum and The Coding School*
Training on Machine learning and AI with Python- *ICT Ministry and Hi-Tech Park Authority of Bangladesh*
Computational Linguistics & Bangla Language Processing- *IEEE Bangladesh Section*
Ethical Hacking and Blockchain Technology- *IEEE Bangladesh Section*

SELECTIVE ONLINE COURSES

Deep Learning Specialization - *Coursera*
DeepLearning.AI TensorFlow Developer Specialization - *Coursera*
Natural Language Processing with Classification and Vector Spaces - *Coursera*
Analyze Datasets and Train ML Models using AutoML - *Coursera*
DeepLearning.AI TensorFlow Developer Specialization - *Deeplearning.ai*
Data Analyst in Python - *Dataquest.io*

REFERENCES

Dr. Mohammed Moshiul Hoque, Professor, Dept. of CSE, CUET (moshiul_240@cuet.ac.bd)
Dr. Scott Frohn, Senior Manager of Assessment and Psychometrics, Workera (scott@workera.ai)
Michael Saxon, Ph.D. student, Computer Science (NLP Lab), UC Santa Barbara (saxon@ucsb.edu)