MD. SHARIAR KABIR

✉ shariar1405076@gmail.com 🌐 shariar076.github.io in Shariar076 ⌂ Shariar076 ℊ Google Scholar

RESEARCH INTERESTS

My interests broadly span the areas of NLP, Multimodal ML, Data Science, and Conversational AI. Currently, I am working on:

- Understanding how LLMs reflect the bias and social epistemology embedded in training data.
- Hallucination reduction through the formation of collective truth among LLMs in structured agent interaction.
- Investigating how visual and auditory cues modulate collective reasoning in ideologically diverse LMMs.

PUBLICATIONS

Do Words Reflect Beliefs? Evaluating Belief Depth in Large Language Models. **Under Review**
Shariar Kabir, Kevin Esterling, and Yue Dong. *arXiv*

  In this paper, we show that when it comes to political matters, LLM behavior is guided by topic-specific belief depth rather than ideology and propose a framework to distinguish genuine from shallow beliefs and assess their consistency.

Automatic Speech Recognition for Biomedical Data in Bengali Language.
Shariar Kabir, Nazmun Nahar, Shyamasree Saha, and Mamunur Rashid. *arXiv*

  This paper presents the development of a prototype Automatic Speech Recognition (ASR) system specifically designed for Bengali biomedical data.

SynthNID: Synthetic Data to Improve End-to-end Bangla Document Key Information Extraction. **EMNLP Workshop**
Syed Mostofa Monsur, Shariar Kabir, and Sakib Chowdhury.

  In this paper, we have introduced SynthNID, a system to generate domain-specific document image data for training OCR-less end-to-end Key Information Extraction systems.

AgnoSVD: Dynamic Resource Allocation for Serverless Workloads Using Collaborative Filtering. **Under Review**
Shariar Kabir and Muhammad Abdullah Adnan. *paper*

  In this paper, we present AgnoSVD, an approach for predicting the optimum resource configuration for an incoming workload using collaborative filtering and Singular Value Decomposition (SVD).

EDUCATION

Bangladesh University of Engineering and Technology (BUET) **2019 - Present**
MSc in Computer Science & Engineering - Part time *CGPA (Predicted): 3.75/4.00*

  Thesis: Dynamic Resource Allocation for Workloads in Serverless Architecture using Collaborative Filtering. Under the supervision of Dr Muhammad Abdullah Adnan.

  Coursework: Bioinformatics Algorithms, Distributed Computing Systems, Data Mining, Data Management in the Cloud, Advanced Database Systems, Advanced Artificial Intelligence.

Bangladesh University of Engineering and Technology (BUET) **2015 - 2019**
BSc in Computer Science & Engineering *CGPA: 3.53/4.00*

  Thesis: Active Learning on Big Data; A research on how active learning can be applied to big data in a distributed cloud computing system. Under the supervision of Dr Muhammad Abdullah Adnan. [report] [ppt]

  Selected Coursework: Machine Learning, Pattern Recognition, Computer Graphics, Artificial Intelligence, Digital Image Processing, Data Structures, Database, Operating Systems.

WORK EXPERIENCE

Celloscope Limited **September 2020 – Present**
Senior AI Research Engineer (NLP, Conversational AI and Vision R&D) *Dhaka, Bangladesh*

MedAI Pvt. Limited **August 2021 – November 2024**
Data Scientist - Part time (Data Science, Conversational AI, and Speech) *Cambridge, UK*

GRP, ICT Division **May 2019 – August 2020**
DevOps (Containerization, CICD and Cloud Monitoring) *Dhaka, Bangladesh*

SKILLS

Soft-Skills:	Communication, Collaboration, Presentation, Technical Writing, Quick Prototyping
Languages:	Python, Shell, C, C++, LaTeX, Java, Scala, SQL, TypeQL, JavaScript, Assembly
Libraries:	Pandas, Datasets, PyTorch, ScikitLearn, OpenCV, Transformers, SpaCy, SQLAlchemy, OpenGL
Model Experience:	LLaMA, GPTs, Adapters, LoRA, QLoRA, DIET, BERT, DeepSpeech, Whisper
Frameworks:	LangChain, RASA, Flask, FastAPI, Spark, PySpark, OpenWhisk, PyQT
Tools & Services:	Colab, Huggingface, Anaconda, VSCode, AWS EC2, AWS lambda, AWS Cognito, NginX, Keycloak, Docker, Ansible, Grafana, PostgreSQL, TypeDB, Socket.IO, ffmpeg

SELECTED ACADEMIC PROJECTS

wQFMSpark – Performance Analysis of Species Tree Estimation Using wQFM in a Distributed System: Species tree estimation from Gene Trees is crucial in **Phylogeny**. Quartet-based estimation techniques like ASTRAL, QMC, and QFM are widely used, but some struggle with scalability. This project aims to redesign wQFM for scalability by distributing taxa and quartets across computational nodes exploring various clustering strategies. [report] [code]

3PC: Implementing the 3 Phase Commit: The goal of this project is to implement a consistent **distributed** music "playlist" using the three-phase commit protocol. A distributed playlist is an un-ordered list of song_name and URL pairs. The system can consist of two or many devices and ensures *consistency*, by satisfying the *ACID* property. [report] [code]

Modelling The Rehearsal Scheduling Problem: This problem originated at Lancaster University in the 1970s. It is reported to have been devised by a member of staff in the Management Science department, who was a member of an amateur orchestra and formalized the problem whilst waiting to play during a rehearsal. The goal of this project was to explore different ways to model this problem using CSP. [report]

Implementation of ML algorithms from scratch: Implemented different machine-learning and deep-learning algorithms from scratch using basic Python libraries like Numpy and Pandas. Examples: Naive Bayes, Alternating Least Squares (ALS), Decision Tree, PCA & EM, Multilayer Perceptrons, Neural Networks, Template Matching etc. [repo1] [repo2]

RECENT INDUSTRY PROJECTS

Voice Banking Chatbot: Bangladesh's pioneering Voice-based AI Chatbot for seamless banking activities, serving hundreds of thousands of real users. Agrani Bank is one of the largest state-owned banks in Bangladesh, with a huge number of customers who have limited access to information. Agrani Voice Banking makes banking services accessible to everyone. It is powered by Bengali ASR and a finetuned NLU engine for performing banking tasks in a natural language-driven way. It can behave *dynamically* based on the input messages by the user.

FAQ Voicebot: A pioneering tool in the Bengali language designed to provide FAQ services over the web or phone using Asterisk IVR. This project uses a **vector index search** over a domain-specific FAQ dataset to fetch the most similar answers related to user queries. The system utilizes our Bengali speech-to-text and text-to-speech tools to transcribe users' queries and provide audio answers, ensuring seamless and accurate interactions.

SynthCases Creator and Disease Recommendation System: A recommendation system using an ensemble of classifiers to predict diseases from patient symptoms. Trained on synthetic data mirroring real-world demographics, it considers risk factors, family, and medical history. The multi-layer pipeline predicts disease probabilities from symptoms, filters using a prevalence table by ethnicity, and then incorporates risk factors for final predictions.

ASR System for Patient Symptoms: ASR system for understanding medical symptoms spoken by patients in the Bengali language. We trained the **DeepSpeech** model from scratch using audio data collected from consented users using our audio data collection portal. We finetuned the model for a noisy environment, using the 13 domain augmentations provided by DeepSpeech. This model performed poorly when the user said any out-of-vocabulary words. Therefore, we finetuned a **Whisper** (tiny) model which was pre-trained using the Bangla Mozilla Common Voice Dataset.

End-to-end Bengali Document Key Information Extraction Using Synthetic Data: In this work, we propose a simple synthetic document image generation framework for Bengali documents. We show that the generated data improves the performance of the extraction model on real datasets, and the system is easily extendable to generate other types of scanned documents for a wide range of document understanding tasks. Short paper accepted in BLP Workshop at EMNLP'23.

Document AI: In this work, we fine-tuned a large multimodal model for extracting key information from documents using both synthetic and real Bengali data. This model excels at key information extraction (KIE) on a variety of scanned and unstructured documents, surpassing several benchmarks, including commercially available OCR products.

Awards & Achievements

Global Health Equity Challenge Award   **2024**
MIT Solve

AmarDoctor by MedAI has been selected as one of the top six solvers out of 2200+ participants worldwide for its innovative approach to accessible healthcare. [link]