

SHARIAR KABIR

✉ shariar1405076@gmail.com 🌐 shariar076.github.io in Shariar076 🔄 Shariar076 ID 0009-0006-6749-8570

Research Interests

NLP, Conversational AI, Data Science, Health Informatics, Distributed Computing, HCI and ML.

Publications

1. Monsur, Syed Mostofa, **Shariar Kabir**, and Sakib Chowdhury. “SynthNID: Synthetic Data to Improve End-to-end Bangla Document Key Information Extraction.” Proceedings of the First Workshop on Bangla Language Processing (**BLP-2023**). 2023. [[paper](#)]
2. **Shariar Kabir**, Nazmun Nahar, Shyamasree Saha, Mamunur Rashid. “Automatic Speech Recognition for Biomedical Data in Bengali Language.” **arXiv preprint arXiv:2406.12931** (2024). [[paper](#)].
3. **Shariar Kabir** and Muhammad Abdullah Adnan. “AgnoSVD: Dynamic Resource Allocation for Serverless Workloads Using Collaborative Filtering.” **Manuscript under review**. [[paper](#)]

Education

Bangladesh University of Engineering and Technology (BUET) **2019 - 2024**

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 Health Limited **August 2021 – Present**

NLP and Data Scientist - Part time (ML, Speech and Data Science)

Remote

GRP, ICT Division **May 2019 – August 2020**

DevOps

Dhaka, Bangladesh

Skills

Research: Design Study, Technical Writing, System Design, Quick-Prototyping
Languages: Python, Java, Shell, JavaScript, Scala, C, C++, L^AT_EX, SQL, TypeQL, Assembly
Libraries: Pandas, Datasets, ScikitLearn, OpenCV, Transformers, SpaCy, TensorFlow, PyTorch, SQLAlchemy, OpenGL, CTranslate2
Model Experience: LLaMA, Falcon, Mixtral, Adapters, LoRA, QLoRA, DIET, BERT, DeepSpeech, Whisper
Frameworks: LangChain, RASA, Flask, FastAPI, Spark, PySpark, OpenWhisk, PyQT
Tools & Services: Colab, Huggingface, ffmpeg, NginX, Keycloak, Docker, Ansible, AWS EC2, AWS lambda, AWS Cognito, Grafana, PostgreSQL, TypeDB, Socket.IO

Recent Industry Projects

FAQ Voicebot: A pioneering tool in the Bengali language designed to provide FAQ services over the web or phone using Asterisk IVR. This project implemented a vector index search over a domain-specific FAQ dataset to fetch the most similar answers related to user queries. The system utilized 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 says any out-of-vocabulary words. Therefore we finetuned a **Whisper** (tiny) model which was trained using Bangla Mozilla Common Voice Dataset.

Voice Banking Chatbot: Bangladesh's pioneering Voice-based AI Chatbot using **RASA** 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 very little 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.

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 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.

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](#)]

Solving The Hidato Puzzle Using CSP: Hidoku (or Hidato) puzzles, were invented by the mathematician Gyora Benedek. Surprisingly, these puzzles have turned out to be quite difficult for modern SAT solvers. Recently the inventor of this game proposed a more complex *Beehive Hidato*. The goal of this project was to explore ways to solve this puzzle using CSP. [[report](#)]

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](#)]

References

Dr. Muhammad Abdullah Adnan

✉ abdullah.adnan@gmail.com

Professor in CSE, Bangladesh University of Engineering and Technology (BUET)

Dr. Mamunur Rashid

✉ m.rashid.1@bham.ac.uk

Assistant Professor in Bioinformatics, University of Birmingham

Dr. Shyamasree Shaha

✉ shyama.saha@medaihealth.com

Chief Technical Officer, MedAI Health Limited