SYED MOSTOFA MONSUR

mostofamonsur9396@gmail.com dv66.github.io (+880)1775016325

Research Interests

NLP, Conversational AI, Interactive Systems, Human-Centered Design

Publications

EMNLP'23 BLP Workshop — **Syed Mostofa Monsur**, Md Shariar Kabir and Sakib Chowdhury. SynthNID: Synthetic Data to Improve End-to-end Bangla Document Key Information Extraction. [paper]

UIST'22 — Md Ehtesham-Ul-Haque, **Syed Mostofa Monsur** and Syed Masum Billah. *Grid-Coding: An Accessible, Efficient, and Structured Coding Paradigm for Blind and Low-Vision Programmers.* Proceedings of the 35th ACM UIST **[Best Paper]** [paper][video]

LREC'22 — **Syed Mostofa Monsur**, Sakib Chowdhury, Md Shahrar Fatemi and Shafayat Ahmed. *SHONGLAP: A Large Bengali Open-Domain Dialogue Corpus.* [paper]

NSysS'20 — **Syed Mostofa Monsur** and Muhammad Abdullah Adnan. *Distributing Active Learning Algorithms.* In Proceeding of the *7th International Conference on Networking, Systems and Security (NSysS), December 22–24, 2020, Dhaka* [paper]

Education

Bangladesh University of Engineering and Technology (BUET) — Dhaka, Bangladesh
Bachelor of Science in Computer Science and Engineering, CGPA 3.54/4.00 (Final 2 years 3.64/4.00) — *Feb. 2015 – Apr. 2019*

Thesis: Analysis of Distributed Active Learning Algorithms. [pdf]

Selected Coursework: Machine Learning (A+), AI (A+), Simulation and Modeling (A+), OS(A), Databases(A), Algorithms (A+), Data Structures(A+) etc.

Professional Employments

Celloscope — Dhaka, Bangladesh
Lead AI Research Engineer, AI Team (NLP, Speech and Vision R&D) — *Jun. 2023 – Present*
AI Research Engineer, NLP and Speech (R&D) — *Sept. 2020 – Jun. 2023*

GRP, ICT Division — Dhaka, Bangladesh
Software Engineer, Backend — *Jul. 2019 – Sept. 2020*

Industry Projects

Agrani Voice Banking — Led the development of *Agrani Voice Banking*, Bangladesh's pioneering Voice-based AI Banking Chatbot for seamless banking activities, serving thousands of real users. *Agrani Bank* is Bangladesh's one of the largest state-owned banks with a huge number of users 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.

Document AI — Finetuned large multimodal document question answering model with Bengali synthetic and real data. The model can perform question answering on various types of scanned/unstructured documents beating several benchmarks including commercially available OCR products.

NLU ChatBots — Developed task-oriented bots for businesses; intent detection, slot-filling using Dual Intent and Entity Transformer; training data augmentation using unsupervised paraphrasing model (*t5*-based, fine-tuned on *TaPaCo*).

Bengali Speech Tools — Collected and pre-processed 400+ hrs of Bengali audio and transcription. Trained end-to-end high-quality ASR models. Trained industry-grade TTS for Bengali language with 40+ hours of curated data and improved generated audio quality with Vocoders (naturalizing audio) Integrated with Natural Language driven User Interfaces including speech-driven chatbots. Developed industry-grade speaker verification system using ensemble of pre-trained unispeech-sat, wavlm and ecapa-tdnn.

Research Projects

End-to-end Bengali document key information extraction using synthetic data. We propose a simple synthetic document image generation framework for Bengali documents. We fine-tune end-to-end models and report the key information extraction performances on real datasets. Short paper accepted in BLP Workshop at EMNLP'23.

Novel UX paradigm for code editors to increase accessibility. We propose a new coding paradigm, Grid-Coding, specially designed to increase accessibility for blind and low-vision programmers. It introduces non-visual programming on grid structure instead of text editors. Through participatory sessions and evaluation with 12 BLV programmers, it proved to enhance code navigation, context understanding, syntax error detection, unique usage patterns and overall programming productivity compared to regular text editors.

Bengali Open-Domain Dialogue Data Collection and Speaker Bias Detection. Collected raw dialogue audio from political talk-shows, debates. Performed speech-to-text, diarization, speaker-role labeling. Fine-tuned BanglaBERT with our dataset for speaker bias detection task. Paper accepted in LREC'22.

Active Learning algorithms for distributed systems. In this work, we propose distributed implementations of state-of-the-art Active Learning algorithms and perform various analyses on them. The algorithms are tested with real datasets on a spark cluster with data distributed on a distributed file system HDFS.

Awards and Achievements

Kaggle OOD ASR 31st position globally (1st in Bangladesh) - organized by Google, received award $3000.
Best Paper Award in ACM UIST, 2022
Champion in Inter-school Science Olympiad
Honorable Mention in Regional Math Olympiad, High School Level
Gold Medal for excellent academic records in High School
Champion in Environment Day Art Competition

Skills

Research	Study Design, Technical Writing, System Design, Quick-Prototyping
Tools & OS	Huggingface, PyTorch, Spacy, Docker, AWS, GCP, Linux
Languages & Frameworks	Python, Java, C++, SQL, Shell, LaTeX, Spring, FastAPI, React
Databases	PostgreSQL, Redis, BerkeleyDB

Selected Academic Projects

Phylogenetic Tree generation using DL. Inception V3 to construct a Phylogenetic Tree Hierarchy from the similarity information provided by the DNN model. [report][code]

ML algorithms from Scratch. Implemented multi-layer perception, EM+PCA, Adaboost, Alternating Least Squares (Recommender) from scratch in Python [link1][link2]

Microcontroller Footstep Counter. A footstep counter setup where signals generated from piezoelectric sensors in shoes are processed using an embedded ATMega32. The signal is then sent via a bluetooth module to an android phone – the app shows number of footsteps in real-time. [video]

Auto-Traffic Controller with Pi. This is an automated traffic control system that uses an image processing backend built on OpenCV which runs on a Raspberry Pi module. Roads are opened depending on traffic density. [video]

Red Drop. An app to find blood donors nearby using GoogleMaps API. [code]

XtremeEngine. This tool indexes and makes large unstructured datasets searchable using a key-value store library BerkeleyDB. [code]

OS Tools Implementation. File, memory management, threading implementation using NachOS. [code]

Reference

Dr. Muhammad Abdullah Adnan	Professor, Department of CSE, BUET
Dr. Faizul Bari	University of Waterloo, Chief Technology Officer at SSCL