Syeda Nahida Akter

1505027.sna@ugrad.cse.buet.ac.bd, nahidatisha2016@gmail.com
mobile: +8801821874052
Homepage, Github

RESEARCH INTERESTS

- Machine Learning
- Natural Language Processing
- Bioinformatics
- Distributed Deep Learning

EDUCATIONAL BACKGROUND

B.Sc. in Computer Science and Engineering Jan 2016 - Present
Bangladesh University of Engineering and Technology
Thesis: Geo-Distributed Machine Learning, Natural Language processing
Supervisor: Muhammad Abdullah Adnan
CGPA: 3.72 out of 4.00 (7/8 terms, Expected Completion: February, 2021)
Major CGPA: 3.87 out of 4.00 (7/8 terms)

PUBLICATIONS

WeightGrad: Geo-Distributed Data Analysis Using Quantization for Faster Convergence and Better Accuracy
Authors: Syeda Nahida Akter, Muhammad Abdullah Adnan
Proc. of the 26th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 2020), San Diego, CA, USA, August 23-27, 2020. [PDF]

RESEARCH EXPERIENCE

WeightGrad: Geo-Distributed Data Analysis Using Quantization for Faster Convergence and Better Accuracy [PDF][Promo Video]

Keywords: Distributed Deep Learning, Gradient Quantization

- We provide loss-aware weight-quantized network with quantized gradients to alleviate the limitations of state-of-the-art quantization.
- We propose a synchronous two-level structure to ensure global convergence within a fixed time interval.
- We attain $5.36\times$ speedup over baseline and $1.4 - 2.26\times$ speedup over the four state-of-the-art distributed ML systems without hurting the accuracy like traditional distributed systems. In fact, we obtain 1.06% gain in top-1 accuracy.

Status: Accepted for publication in ACM SIGKDD'20 as a full-track research paper.

Supervisor: Muhammad Abdullah Adnan (BUET)

Phylogenetic Tree Estimation Using Deep Neural Network

Keywords: Deep Learning in Bioinformatics, Phylogenetic Tree Estimation

- We design an end-to-end deep neural network to predict a species tree from a set of (embedded) gene trees. The model is designed to work well for a wide range of taxa.
- Our novelty is ensuring unique vector representation of each gene tree by incorporating an embedding layer. This way, different gene trees can be discriminated during the entire training process.

- We first learn distances between taxa using a simple neural network, then explore variational auto-encoders and different adversarial neural networks to summarize distances into a single species tree.

Status: In progress (expected completion: December 2020).

Supervisor: Md. Shamsuzzoha Bayzid (BUET)

Responding to the Stigma, Fear, Discrimination, and Misinformation Related to the COVID-19 Disease Outbreak

Keywords: NLP, Topic Modeling

- We use Natural Language Processing (NLP) and Social Network Analysis (SNA) to study collected anonymized Twitter data. I am working on the NLP section.
- I used state-of-the-art NLP techniques to analyze the following from social media posts: keyword gathering, frequency analysis, information extraction, automatic categorization and clustering, automatic summarization, and finding associations within the data.
- We are working on different topic modeling architectures (LDA, GSDMM, BTM, lda2vec, BERT) for short texts like tweets and experimenting with advanced classifier as well as DNN architectures.

Status: Preparing manuscript to be submitted to ACM CHI'21.

Supervisor: Ishtiaque Ahmed (University of Toronto)

BERT for Low Resource Languages: Bangla Q&A

Keywords: NLP, Topic Modeling, XLM-Roberta

- We develop a model for Bangla automated Question Answering (QA) using pretrained XLM-Roberta.
- We propose a new embedding layer with a topic modeling structure prior that to increase accuracy for context-based question answering system.

Status: In progress (expected completion: January 2021)

Supervisor: Muhammad Abdullah Adnan (BUET)

Malnutrition and Dehydration Detection From Body and Skin Images

Keywords: Computer Vision, CNN

- We develop a mobile-Based application that will capture image and video of specific body regions (face, lip, tongue, eye, fontanelle, urine and the video of skin turgor test) to detect a clinical sign of dehydration and malnutrition.
- We develop CNN model to extract features (e.g., body part detection, segmentation, color, border, shape and texture of skin) from images and videos.
- Build a classifier (SVM) to predict dehydration and malnutrition level (mild, moderate, acute).

- This work is a collaboration with a medical team from International Centre for Diarrhoeal Disease Research, Bangladesh [icddr,b]

Status: In progress (expected completion: February 2021)

Supervisor: Tanzima Hashem (BUET)

AWARDS AND HONORS

- **KDD Student Registration Award**, 2020
- **University Merit Scholarship**, 2019
- **Champion on BUET CSE Fest Inter-University Hackathon** (Two consecutive years 2018, 2019)

MAJOR PROJECTS

Machine Learning Based Credit Risk Prediction System for Commercial Banks 2019-20

A machine learning based banking software system (a collaboration with Bank Asia) which aims at predicting trustworthiness of a client who has applied for a loan by predicting credit risk. It also predicts the allowable loan amount that can be granted without causing any risk. I worked on the machine learning based backend system that included data cleaning, feature extraction, experimenting on different classifiers and regression models (random forest, decision tree, XGBoost). Our classifier gets 82% test accuracy (SOTA accuracy is 78% for the same dataset). [Repository]

ChatBot with Emotion January 2019

A chatbot which processes texts using Natural Language API of Google and interacts accordingly, it takes picture of user and analyzes user's facial emotion using FACE API of Google and suggests user specific genre of songs or videos depending on the mood, it also takes voice command using Speech-to-Text and replies back to user using Text-to-Speech APIs of Google, it also stores the location of the user and suggests local songs, videos, sites etc. [Repository]

Tori: Mental Health Care App using Machine Learning Jun 2018

Mental Health Care app using machine learning, it stores users activities on the mobile and interacts with users through a chatbot which collects textual responses from users and using NLP, tries to predict users' mood and conducts further interaction accordingly (e.g. sending songs, videos, sending suggestions to contact with close friends, informing users about time spent in mobiles each day etc.) [Repository]

REFERENCE

Supervisor: Muhammad Abdullah Adnan
Email: abdullah.adnan@gmail.com, adnan@cse.buet.ac.bd
Mobile No: +880 1552 336926