SAMPOORNA P.

+91 87770 99837
sampoornap0172@gmail.com ◇ LinkedIn ◇ Github

EDUCATION

BTech in Computer Science and Engineering, 2021 - 2025
Maulana Abul Kalam Azad University of Technology
GPA: 9.7/10 (Department Rank: 1/102)
Bachelor's Project : Mathematical Word Problem Solving using Transformers

High School Diploma, 2011 - 2021
DPS Newtown
96.25%

PUBLICATIONS

1. A Hybrid Deep Learning Framework to Diagnose Sleep Apnea using Electrocardiogram Signals for Smart Healthcare
Sampoorna Poria, Ahona Ghosh, Biswarup Ganguly, Sriparna Saha*
Intelligent Data Analytics for Bioinformatics and Biomedical Systems (Accepted) [pre-print]

2. Internet Security Issues in Robotics for BCI: Opportunities & Challenges
Sampoorna Poria, Sriparna Saha*
3rd International Conference on Network Security and Blockchain Technology (Submitted)

EXPERIENCE

Research Intern Indian Institute of Technology, BHU Aug '23 - Present

- Currently Working under Dr. Anil Kumar Singh in the NLP lab on designing a framework capable of translating code written in Java to its equivalent python representation.
- Previously, worked on finetuning a deep learning framework for Neural Machine Translation of low-resource languages especially Hindi-Bhojpuri.
- Learning more about morphological analysis, Parts-of-Speech tagging and other MT tools like MGIZA, Trankit and OpenNMT.

Data Science Volunteer Omdena x BCS Sri Lanka Aug '23 - Present

- Worked with the Data Collection and Exploratory Data Analysis Teams to help train a language model capable of identifying inaccuracies in GESI (Gender Equality and Social Inclusion) conversations in Sri Lankan media.
- The project aims to develop a tool that can be used by CSOs, journalists and other stakeholders to support their efforts in promoting GESI in Sri Lanka.

Feature Importance Evaluation for Detection of Freezing of Gait (Parkinson's Disease) Jul'23-Dec'23
(Guide: Prof. Sriparna Saha), MAKAUT

- Implemented SHAP (SHapley Additive exPlanations) for feature importance evaluation in Python. Understood Explainable AI (XAI) and it's importance in interpretability of black-box models.
- Processed and formatted CSV data for analysis, contributing to accurate gait freezing detection..
- Skills: Explainable AI, Bash Scripting, Data Formatting and Analysis, Interpretable ML, LIME, SHAP

Deep learning framework to identify/classify sleep apnea from ECG signals Apr '23 - June '23
(Guide: Prof. Sriparna Saha), MAKAUT and (Prof. Biswarup Ganguly), NIT Silchar

- Contributed on a research project designing a hybrid deep learning framework to identify and classify sleep apnea from single channel ECG signals into four classes.

- Worked on pre-processing the ECG signal data via Discrete Wavelet Transform to remove noise and unwanted signals while simultaneously learning about the implementation of Autoencoders and Bi-LSTMs.
- First-authored a research paper, plotted necessary graphs and analysed accuracy(=95%) of the model.
- Skills: Deep Learning, Signal Processing, CNN-RNN architecture, bi-LSTMs, matplotlib, Autoencoders, Data Pre-processing, Academic Writing

ContextCue Aug 2023 - Present
RoundtableML (Open Source)

- Contributing to the development of an **open-source language learning chrome extension** that uses **NLP techniques** to selectively substitute words in a web page while maintaining **user-configurable settings** for optimal engagement.
- **Skills: Python, LLMs for token classification, gRPC**

PROJECTS

ShakespeareGPT Generatively Pretrained Transformer for generating Shakespearean-style quotes. GitHub

- Explored advance NLP concepts and developed a PyTorch-based GPT model from scratch, including its components such as Multihead Attention.

NaiveContextCue GitHub

- Implemented a language learning platform using streamlit as the lower-bound baseline for the ContextCue project. The program translates all noun phrases in the text provided as input. It uses spacy to get the noun phrases and the translate library to translate.

IrishMusicGenerator Recurrent Neural Network for Music Generation GitHub

- Trained a model on a dataset of Irish folk songs, to learn the patterns in raw sheet music (in ABC notation) and then use this model to generate new music of the same kind. The model is based off the LSTM architecture, where a state vector is used to maintain information about the temporal relationships between consecutive characters.

TaylorLyrics Music Generation using Markov Chains GitHub

- Developed a Python-based lyrics generator using markov chains, which utilized the Spotify API to extract the top 10 songs of a particular artist and the lyricsgenius API to obtain the lyrics.
- Implemented a markov chain algorithm that incorporated bigrams of words, tokenization using nltk, and transition probabilities of words. The model was able to successfully generate lyrics similar to the songwriter's style. Deployed the model as a Flask application and built a simple frontend.

SKILLS

Languages	Java, Python, MATLAB, C, SML, Racket, Verilog, HTML, MySQL
Frameworks	PyTorch, TensorFlow, Flask, Scikit-Learn, NLTK, spaCy, streamlit
Tools	Emacs, VSCode, DrRacket, Git/Github, Linux, Excel

TEST SCORES

GRE	330 (Quantitative : 170/170 — Verbal : 160/170) (February, 2024)
TOEFL	119/120 (March, 2024)

EXTRA-CURRICULAR ACTIVITIES

- Department Rank 1, MAKAUT, WB.
- Class Representative for the following courses : Digital Electronics, DSA, Theory of Computation, Software Engg.).

- Inter-college Hackathon lead, came among top 3.
- GDSC AI-ML Core Member(Sep '23 - Present), Codechef MAKAUT Chapter Core Member(Dec '21 - Mar '23), WomenTech Network Member(Jul '23 - Present)
- Mentored over 15 undergraduate women pursuing engineering.
- ISI B.Stat Qualified Part 1 (among top 300/ 50,000)
- Scholar Badge Holder (for securing 90%+ in school every year)