Debapriya Tula

debapriya.t17@iiits.in | LinkedIn | GitHub | Google Scholar | Portfolio

EDUCATION

Indian Institute of Information Technology, Sri City, Andhra Pradesh (Aug 2017 – May 2021)

Bachelor of Technology in Computer Science and Engineering Cumulative GPA: **9.35/10**

EXPERIENCE

Machine Learning Engineer at Tata Consultancy Services (TCS) R&I (Aug 2021 – Present)

- Develop ML/DL models to understand and predict user behavior for services related to **General Electric**'s HealthCare segment.
- Build and deploy ML models using AutoML toolkits for big datasets (**10-20 GB**), with modules for statistical data analysis and output explainability.

NLP Software Development Intern at LimeChat (Jan 2021 – June 2021)

- Redesigned the **order tracking** system to make it more seamless and fault-tolerant.
- Redesigned Limechat's **FAQ management** and **Utterance management** systems and deployed them as core features in **5 weeks**.
- Set up the chatbot for **Nissan**, LimeChat's biggest client undertaking to date.

Computer Vision Research Intern at IIT Delhi (May 2020 – July 2020)

- Designed an efficient pipeline for the problem of motion segmentation of fish in **underwater scenarios** solved as an **unsupervised** learning task.
- Replicated a few research papers in the course of exploring the transferability of other models to the problem.

Research Intern at Tezpur University (May 2019 – June 2019)

- Maximize stacking regions to find the most stable secondary structure(s) of RNA using concepts from graph theory.
- Awarded the **best paper** at **ICCCIoT, 2020.**

PUBLICATIONS

Published

- **Tula, Debapriya**, Prathyush Potluri, Shreyas Ms, Sumanth Doddapaneni, Pranjal Sahu, Rohan Sukumaran and Parth Patwa. "Bitions@DravidianLangTech-EACL2021: Ensemble of Multilingual Language Models with Pseudo Labeling for offence Detection in Dravidian Languages." *DRAVIDIANLANGTECH* (2021). https://www.aclweb.org/anthology/2021.dravidianlangtech-1.42
- Sen Piyali, **Tula Debapriya**, Ray S.K., Satapathy S.S. Estimating RNA Secondary Structure by Maximizing Stacking Regions. In: Mandal J., Mukhopadhyay S., Roy A. (eds) Applications of Internet of Things. Lecture Notes in Networks and Systems, vol 137. Springer, Singapore. https://doi.org/10.1007/978-981-15-6198-6_15
- **Tula, Debapriya**, Shreyas Ms, Viswanatha Reddy, Pranjal Sahu, Sumanth Doddapaneni, Prathyush Potluri, Rohan Sukumaran and Parth Patwa. "Offense Detection in Dravidian Languages using Code-Mixing Index based Focal Loss and Cosine Normalization". *Accepted* at SN Computer Science Journal. https://link.springer.com/article/10.1007/s42979-022-01190-1

Under Review

- Basha SH, **Tula Debapriya**, Vinakota Sravan Kumar, Dubey Shiv Ram. "Target Aware Network Architecture Search and Compression for Efficient Knowledge Transfer." *Under review* at Applied Intelligence Journal. https://arxiv.org/abs/2205.05967
- Suvendra K Ray, Ruksana Aziz, Piyali Sen, Pratyush Kumar Beura, Saurav Das, **Debapriya Tula**, Madhusmita Dash, Nima Dondu Namsa, Ramesh Chandra Deka, Edward J Feil, Siddhartha Sankar Satapathy. "Incorporation of transition to transversion ratio and nonsense mutations, improves the estimation of the number of synonymous and non-synonymous sites in codons". *Under review* at DNA Research (Journal).

SKILLS AND TOOLS

Machine Learning, **Deep Learning**, **Computer Vision**, **Natural Language Processing**, Data Science, **Python**, C, C++, MATLAB, Javascript, **Git**, Docker, SQL, NoSQL, **Bash**, Rasa. **Pytorch**, **Keras**, FastAI, **Sklearn**, **Numpy**, **Pandas**, **Seaborn**, LaTeX

PROJECTS

Content-Based Image Retrieval (Oct 2020 – June 2021)

- Developed a **multi-loss** model for retrieving relevant images from large datasets.
- Optimized the model using **self-attention** and an angular-based loss for a **curriculum-based sampling**.

Speech Emotion Recognition (Sept 2020 – Dec 2020)

- Applied augmentation to speech signals, extracted MFCC features and trained a Random Forest Classifier for identifying emotion from speech directly.
- Accuracy obtained on datasets -> On **RAVDESS** : **73.5 %,** On **TESS**: **98.6 %**

Solingo (Feb 2020 – April 2020)

- An app that recognises handwritten math expressions from captured images.
- Implemented an **attention-based** model for expression recognition, including LaTeX transformation of the input, using **Pytorch** and **Selenium**.
- Accuracy on CROHME dataset - **73 %**

Gringotts (Sept 2019 – Dec 2019)

- Provide a vault to securely store secrets like passwords, keys (GPG/SSH), securely transfer data among people etc.
- Made a face-recognition system using **Keras** and **OpenCV** to validate the person using Gringotts.

StackOverFlow API-recommender (Sept 2019 – Dec 2019)

- Provide an API recommender for Java APIs for questions asked on StackOverflow.
- Crawled all questions and their respective answers for the given question ids, using Selenium and BeautifulSoup.

Speech Dereverberation (Sept 2018 – Dec 2018)

- Led a team of 4 to build a system to remove reverb(echo) from sound signals by predicting the reverb's contribution in the present signal, using NumPy and audioread.
- Similarity with original signal: **65-75** %

VOLUNTEERING

AI Student Ambassador at Intel (Oct 2019 – Present)

- Organize hands-on sessions and paper reading sessions on topics related to AI/ML. Encourage students to work on AI/ML projects and assist them.
- Designed an interpolation method that reduces the frame rate in videos followed by frame reconstruction for efficient (internet) data usage.
- Average interpolation error on **Visual Tracker Benchmark (VTB) dataset - 12.6 %**

OTHER DETAILS

- **Language Proficiency:** English, Hindi, Odia, Assamese
- **Hobbies:** Reading books, playing the guitar, singing, playing table tennis and badminton.