Debjoy Saha

Patel Hall, IIT Kharagpur, West Bengal, India
(+91)8448508987 sahadebjoy10@gmail.com Google Scholar github.com/Debjoy10

Education

Indian Institute of Technology, Kharagpur **Jun. 2018 – Mar. 2023**
B.Tech + M.Tech in Electronics and Electrical Communication with a minor in Computer Science *West Bengal, India*

- Cumulative Grade Point- 9.35/10.0, Additional Cumulative Grade Point- 9.37/10.0, after 8/10 semesters.
- Among the top 6% students at IIT Kharagpur and ranked 7th out of 60 in my class.

Relevant Coursework

- Algorithms and Data Structures
- Digital Image Processing
- Cyber-physical Systems
- Machine Learning
- Digital Signal Processing
- Embedded Machine Learning
- Natural Language processing
- Digital Speech Processing
- Digital Electronic Circuits
- Graph Machine Learning
- Computer Architecture
- Operating Systems

Experience

KLA Tencor Software **May. 2022 – July. 2022**
Machine Learning Intern *Chennai, India*

- Implemented various Neural Architecture Search (NAS) algorithms for improving defect classification from chip images.
- Improved search efficiency by incorporating weight-sharing techniques (DARTS and RSWS) and evaluated transferability of obtained architectures across similar datasets.

Language Technology Group, Universität Hamburg **Jun. 2021 – Dec. 2021**
Research Intern *Hamburg, Germany*

- Working with Dr Timo Baumann at the Language Technology Group at Universität Hamburg.
- Built the first German audio-visual-text dataset using clips extracted from morning talk shows.
- Incorporating video knowledge into neural speech synthesis models to improve lip synchrony for dubbing.
- Performed adversarial attacks and interpretability analysis on neural audio-visual synchrony evaluation models.

Aerial Robotics Laboratory, IIT Kharagpur **Mar. 2019 – Present**
Software and AI team member *IIT Kharagpur, India*

- Responsible for seamless functioning of the group under the guidanace of Prof. Somesh Kumar, IIT Kharagpur.
- Developed various image processing tools for object recognition and image understanding using OpenCV.
- Worked on SLAM, Localisation, Path-Planning and State estimation for aerial robots in ROS framework.

Projects

Dialogue State Tracking and Representation Learning | *Conversational AI* **Aug. 2021 – Nov. 2021**

- Worked with Dr. Pawan Goyal on my bachelor thesis on conversation-AI at CNERG Lab, IIT Kharagpur.
- Designed prompt-based few-shot learning techniques for state-tracking of task-oriented dialogue systems using language models obtaining similar slot accuracy as existing few-shot approaches with fewer parameters. (Thesis) (Presentation)
- Experimented with a novel distillation-enhanced pretraining framework using Discourse Mutual Information (DMI)-based pretraining for improving representation learning tasks for dialogue systems. (Thesis) (Presentation)

Offensive Language Identification in Dravidian Languages | *Text Classification, Transformers* **Dec. 2020**

- Trained multilingual text classification transformer models (XLM-Roberta, mBERT) for offensive language detection.
- Developed a novel BERT-CNN fusion architectures and a Genetic-Algorithm based ensembling strategy.
- Ranked overall first in the EACL shared task on Offensive Language detection in Dravidian Languages. (Code)

Drone Racing Optimisation | *Trajectory Planning, Reinforcement Learning* **Jun. 2020 – Nov. 2020**

- Deployed minimum-snap trajectory generation with Policy Gradients-based waypoints optimisation for a warehouse traversal quadrotor on an environment built on Microsoft AirSim, obtaining 40% improvement in path smoothness.
- Reached National Finals in Flipkart GRID robotics competition and presented our solution to an expert panel. (Code)

Detection of Hateful Memes in Social Media | *Multimodal machine learning* **Feb. 2020 – Dec. 2020**

- Working with Dr. Animesh Mukherjee on a multimodal machine learning project at CNERG Lab, IIT Kharagpur
- Performed pretraining and finetuning of multimodal attention-based models like Visual-BERT on Image-Text data.
- Experimented with self-supervised and weak-supervised training techniques like negative supervision and weak labels.
- Performed score calibration using Monte-Carlo Dropout and obtained improvements in AUROC metric.

Warehouse Inventory Management using UAVs | *Computer vision, OCR* **May. 2019 – Oct. 2019**

- Developed systems for autonomous warehouse inventory management using UAVs for the International Micro Aerial Vehicles indoor competition (IMAV-2019). (Code)
- Developed the OCR tools and for accurate detection of alphanumeric codes using google's Tesseract library and QR-codes using Zbar library for inventory management of boxes in a warehouse, obtaining 96% detection.
- This project won us the first place in IMAV which also made us the first team from India to ever win at IMAV.

Technical Skills

Languages: Python, C++, C, SQL, MATLAB
Technologies/Frameworks: Linux, GitHub, PyTorch, Tensorflow, Scikit-Learn

Publications

A Deep Dive Into Neural Synchrony Evaluation for Audio-visual Translation

ACM International Conference on Multimodal Interaction (ICMI), 2022

- We present a comprehensive analysis of SyncNet as an evaluation tool for neural audio-visual synchrony.

Merkel Podcast Corpus: A Multimodal Dataset Compiled from Angela Merkel's Weekly Podcasts

Language Resources and Evaluation Conference (LREC), 2022

- We introduce the first single speaker corpus in the German language consisting of audio, visual and text modalities.
- We present a novel method for semi-automatically extracting target speaker clips using forced alignment, active speaker recognition and face similarity obtaining 94% extraction accuracy and perfect precision. (Paper) (Poster)

Ensembling strategies for Transformer-based Offensive language Detection

Workshop on Speech and Language Technologies for Dravidian Languages, EACL 2021

- We present a Genetic Algorithm-based ensembling approach on multilingual Transformer-based models for Offensive language detection in Dravidian languages in a low-resource setting. (Paper) (Poster)

Warehouse Management Using Real-Time QR-Code and Text Detection

International Micro Aerial Vehicles competition and conference (IMAV), 2019

- We develop computer vision tools for efficient inventory management of a warehouse using quadrotors (Paper)

Talks

Self-Supervised and Weak-Supervised Techniques to improve Hateful Memes Detection

Hateful Memes Challenge Session, Conference on Neural Information Processing Systems (NeurIPS), 2020

- Discussed techniques to improve classification of hateful memes using multimodal Transformer models. (Presentation)

Achievements

- Was selected for **WISE** (Working Internships in Science and Engg.) funded by DAAD, Germany.
- **Ranked 1st**, in the shared task on Offensive Language Identification in Dravidian Languages, **EACL**, 2021.
- **Ranked 11th**, out of 3173 participants, in the Hateful Memes Detection challenge, **NeurIPS**, 2020.
- **National Finalists**, top 9 out of 250 teams, at the **Flipkart GRID Robotics Competition**, 2019.
- **Won** the **IMAV-2019 Indoor Competition** held in Madrid, Spain, among 14 teams from 11 countries.
- **Gold medalist**, among 20 teams, **Inter-Hall Data Analytics Competition** at IIT Kharagpur, 2019
- Secured the **452nd national rank** out of 50K students, **KVPY-SX Fellowship Examinations**, 2018.
- Secured the **375th national rank** out of 50K students, **KVPY-SA Fellowship Examinations**, 2017.
- **Qualified** in the **RMO** (Regional Mathematics Olympiad) from New Delhi, India, 2017.
- **Qualified** in the **NTSE** (National Talent Search Examination), 2016.

Extracurricular

- **Teaching Assistant** at Digital Signal Processing Lab, Dept. of Electronics, IIT Kharagpur. (2022-present)
- **Mentor** with the SWG (Student Welfare Group), guiding a group of junior students from the dept. of Electronics and electrical communication engineering in their academic and non-academic ventures. (2021-present)
- **Mentor** at the IEEE Certified Winter Workshop on Image Processing at IIT Kharagpur, tutored 90 fresher and sophomore students on beginner-level topics in digital image processing in the spatial domain and graph theory. (2019)
- **Keyboardist** for the hall band at Patel Hall, IIT Kharagpur. (2018-2019)