

NIHAR SANDA

+1 (617) 959-5688

sanda.n@northeastern.edu

niharsanda

koolgax99

koolgax99.github.io

EDUCATION

Northeastern University - Khoury College of Computer Sciences

2023 - 2025 (Expected)

MS in Computer Science — GPA: 4/4

Boston, MA, USA

Indian Institute of Information Technology, Dharwad

2019 - 2023

Bachelor of Technology, Computer Science and Engineering — GPA: 8.5/10

Karnataka, India

SKILLS

Languages: Python, R, Java, CUDA, C++, C, JavaScript, Typescript, OWL, RDF, HTML, CSS

Libraries: Numpy, PyTorch Matplotlib, Seaborn, Scikit, Pandas, Keras, TensorFlow, Dlib, SciPy, Jax

Technologies: Parallel Programming, RESTful services, Flask, Django, Docker, Jenkins, CI/CD, R Shiny

Database: SQLite, PostgreSQL, MySQL, MS SQL, MongoDB, NoSQL, DynamoDB, Redis.

EXPERIENCE

Institute for Software Integrated Systems, Vanderbilt University

May 2024 - Present

Data Science Co-op

Nashville, Tennessee

- Developed a **Human Activity Tracking** system for collaborative learning environments, integrating state-of-the-art **Kalman filters** and **triplet loss-based ReID models**, achieving 92% mAP across 10+ hours of video data from 10 classrooms.
- Engineered a **multimodal fusion technique for emotion recognition**, attaining 87% accuracy in affect mapping for the NSF AI Engage Institute's project, impacting 1000+ students across 5 universities.
- Architected **scalable real-time inference endpoints** processing **100 requests/second** for Emotion Recognition and YOLO models on a 20-node Kubernetes cluster, reducing latency by 40% and maintaining 99.9% uptime.

Indian Institute of Technology, Bombay

May 2023 – Sep 2023

Applied Research Associate

Mumbai, India

- Engineered an **"Affect Aware Tutoring System"** processing 10,000+ daily facial images and 1M click-stream events, achieving **95% accuracy** in real-time affect prediction using a **custom vision transformers model**, boosting student engagement by 30%.
- Implemented a **privacy-preserving federated learning framework** for educational data, reducing exposure risk by 99.9% while maintaining **97% model performance** across 100 institutions and 500,000+ student records.
- Optimized **multi-modal deep learning pipeline** for affect recognition, reducing inference time by 40% and improving **F1-score** from **0.82 to 0.91**, serving 50,000+ daily active users.

Google Summer of Code, PEcAn Project

2022 and 2023

Student Intern

Bengaluru, India

- Developing the **various PEcAn packages** of **data assimilation and meta-analysis** for Carbon and Land data.
- Leveraged **R Shiny** to create a **robust and user-friendly dashboard**, empowering users to generate dynamic **SDA (State Data Assimilation)** and forecasting graphs for various researchers around the world.
- Enhanced the **authentication of the existing REST APIs** by incorporating robust **API Key authentication** and implementing efficient **rate-limiting features**.

PROJECTS

Realtime Person Tracking and Reidentification in Embodied Learning Environment | Computer Vision

Jan 2024

- Engineered a **high-performance person re-identification (Re-ID)** system using a custom dataset and deep learning pipeline in **PyTorch**, designed for **real-time tracking** in closed-room classroom settings.
- Achieved **92% mAP accuracy** by fine-tuning pre-trained models (**TriNet**, **OSNet**) and applying **ensemble learning**.
- Integrated **YOLOv8** with **DeepSORT** for real-time **multi-object tracking**, utilizing **Kalman** and **Particle Filters** for precise motion prediction, trajectory estimation, and minimized **identity switches**.

Protein Fold Recognition | NLP, Transformers, Bio-Informatics

August 2022

- Implemented advanced **NLP techniques** to improve protein fold recognition for low similarity baseline datasets such as **DD**, **EDD**, **TG**, and **SCOPE**, encompassing diverse amino acid-based **protein sequences** and their **corresponding folds**.
- Extracted **features** by utilizing evolutionary **PSSM** and **HMM profiles** of protein sequences, and concatenating them with global **Convolutional** and **Skip Bi-gram features**.
- Implemented **BERT** and **ESM** by **Meta transformer-based models** for classification and achieving an impressive accuracy **exceeding 93%** across all datasets, **surpassing the previous 85% accuracy**.

ACHIEVEMENTS AND LEADERSHIP

- Director's Gold Medal for the Best Outgoing Student at IIIT Dharwad for 2023 Batch
- Led a team to the Grand Finals at Smart India Hackathon 2022 and also won many National Level hackathons
- 2 times Google Summer of Code (2022, 2023) Recipient at PEcAn Project
- Open Source Contributor for Rucio(CERN), CircuitVerse, PEcAn Project with many accepted PRs.

RESEARCH PAPERS AND PUBLICATIONS

- An Effective Framework for the Prediction of Protein Folds using Natural Language Processing and Evolutionary Features** - IEEE/ACM Transactions on Computational Biology and Bioinformatics
- Interestingness from COVID-19 Data: Ontology and Transformer-Based Methods** - Proceedings in ACL Anthology
- Ontology-Based Semantic Data Interestingness Using BERT Models** - Taylor and Francis' Connection Science Journal