

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

Boston, MA, USA

Indian Institute of Information Technology, Dharwad

2019 - 2023

Bachelor of Technology, Computer Science and Engineering

Karnataka, India

EXPERIENCE

Indian Institute of Technology, Bombay

May 2023 – Present

Research Associate

Mumbai, India

- Scaling the "Affect Aware Tutoring System Using Video Bots" to learn student engagement with adaptive feedback, improving accessibility and enhancing learning outcomes.
- Demonstrating leadership and project management skills by overseeing the development of a high-performance production application, designed to handle concurrent usage by over **10,000 users**.
- Engaged in extensive research on "Privacy Protection of Student Video Data through Stylegan-Based Deidentification in Diverse Learning Environments" exploring innovative approaches to safeguard student privacy and confidentiality within various educational settings.

Patenti Technology Solutions

Mar 2023 - May 2023

SDE Intern

Bangalore, India

- Created a comprehensive IP infringement search tool by utilizing OCR models and employing Django's Model-View-Template (MVT) architecture to develop the application.
- Deployed and maintained multiple production and testing environments on AWS EC2, S3, and Redis, utilizing Github Actions for CI/CD automation
- Developed numerous backend modules and executed an NLP pipeline to employ LLMs and construct a corpus for topic modeling and scientific term retrieval in patent documents.

Indian Institute of Information Technology, Dharwad

Oct 2022 - Jan 2023

Undergraduate Research Assistant (Under Prof. Kavi Mahesh and Dr. Abhilash C B)

Dharwad, India

- Led research efforts to explore semantic interestingness and conduct analysis of interesting patterns in various types of healthcare data, aiming to enhance rule representation.
- Additionally, facilitated the integration and utilization of transformer-based models for ontology-based associated rule mining, enabling advanced data analysis techniques in the healthcare domain.
- Engaged in building ontologies for diverse datasets, including US birth data and Karnataka Covid Patients data, to enhance data interoperability, semantic integration, and facilitate meaningful insights and analysis in these specific domains.

Camplus App, Varopro Private Limited

Feb 2021 - Oct 2022

Founder and CEO

Mumbai, India

- Led a team of more than **10 people** in the development of a campus management application catering to a user base of **800 individuals**, receiving grants and funding of more than **\$13000**.
- The startup, was incubated by New Gen IEDC IIIT Allahabad, Government of India, under the Startup India initiative, secured funding and garnered significant media attention.
- Engineered the application's multi-tenant backend leveraging a robust tech stack including AWS Lambda, DynamoDB, S3, SES, API Gateway, and the MEN (MongoDB, Express, Node) stack.

Google Summer of Code, PEcAn Project

Jun 2022 – Nov 2022 & Jun 2023 – Oct 2023

Contributor

Mumbai, India

- Created the REST APIs for the various PEcAn packages of data assimilation and meta-analysis in the R language.
- Enhanced the authentication of the existing REST APIs by incorporating robust API Key authentication and implementing efficient rate-limiting features.
- Leveraging 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.

Indian Institute of Technology, Bombay

Jul 2022 – Nov 2022

Research Intern

Mumbai, India

- Worked on a novel idea in the field of education technology, named "Affect Aware Tutoring System Using Video Bots"

- Built a learning management system that collects the click-stream log data of the student, simultaneously captures video and then predicts the user's affect state for realtime feedback.
- Developed an optimized transformer-based deep learning model **Vision Transformers** to predict the user's affect state. The model was trained on the huge DAiSEE dataset for approximately **300 hours**.

PROJECTS

Covid-19 Mortality Prediction | *Data Visualization, Machine Learning, Deep Learning* **May 2022**

- **Data analysis and processing:** Proficient in **cleaning, exploring, and transforming private hospital data** with **1 lakh data points** to better understand and extract information using **Pandas** and **Python**
- **Data Imbalance Mitigation:** Employed techniques such as SMOTE and undersampling to effectively tackle the challenge of imbalanced data.
- **Mortality Prediction:** Implemented **state-of-the-art machine learning models** to predict **COVID-19 patient mortality**, utilizing an ensemble approach and achieving an impressive **accuracy of 93%**

Protein Fold Recognition | *NLP, Transformers, Bio-Informatics* **Aug 2022**

- Implemented advanced **NLP techniques** to improve protein fold recognition for low similarity datasets such as **DD, EDD, TG, and SCOPe** baseline datasets 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**.
- The project has been documented and has been sent to **IEEE Transaction in Computational Biology and Bioinformatics** where it is currently under review.

SKILLS

Languages: Python, R, Java, C, C++, HTML/CSS, JavaScript, PHP

Tools: REST APIs, Docker, Docker, Jenkins, CI/CD, AWS, GCP, Kubernetes, Unix, Git

Database: SQLite, PostgreSQL, MySQL, NoSQL, MongoDB, DynamoDB, Redis

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

Frameworks: Node, React, Flask, Django, R Shiny, Flutter, Vue

PUBLICATIONS

An Effective Framework for the Prediction of Protein Folds using NLP and Evolutionary Features

- Developed a high-performance system for detecting malicious and benign QR codes.
- Paper currently under review in **IEEE/ACM Transactions on Computational Biology and Bioinformatics**.

Multi Modal Affect Aware Tutoring Systems | *Data Visualization, Machine Learning, Deep Learning*

- Developed a novel idea of "Multi-modal Affect Aware Tutor System using Video bots".
- Paper is documented and in procession at **IEEE Transactions in Learning Technologies Journal**.

Interestingness from COVID-19 Data: Ontology and Transformer-Based Methods

- Identify interesting patterns using ontology-based mining techniques and process them with transformer models for identifying the interesting rules from the mined corpora.
- Paper is accepted for publication at **ICON 2022, proceedings in ACL Anthology** (SCOPUS indexed)

Ontology-Based Semantic Data Interestingness Using BERT Models

- Implemented data curation technique with semantic integration and drove insights using ConstApriori algorithm for RDF data, optimizing decision-making.
- Implementation of Clinical BERT and Bio BERT model for identifying the most interesting rules using a cosine similarity measure - Semantic Interestingness.
- Paper published in **Taylor and Francis' Connection Science Journal** (SCI and SCOPUS Indexed).

ACHIEVEMENTS

- Director's Gold Medal for the Best Outgoing Student at IIIT Dharwad.
- 2 times Google Summer of Code (2022, 2023) Recipient
- Reviewer at ICCE and T4E conferences in Education Technology domain.
- Lead a team to the Grand Finals at Smart India Hackathon 2022
- Top 6 globally at International Accelerator Program, Austin, Texas
- Ex-Chairman of IEEE Computer Society of IIIT Dharwad.
- Ex-Secretary of IEEE Student Chapter of IIIT Dharwad.
- Open Source Contributor to Rucio (CERN), PEcAn Project, CircuitVerse.
- Founder and President of Velocity, Web and App Dev Club of IIIT Dharwad
- Senior Mentor of Data Science and Artificial Intelligence (DSAI) Society
- Web Lead at IEEE Bangalore Chapter and Postman Student Leader.