




NIHAR SANDA

✉ sanda.n@northeastern.edu  [nihar-sanda-07ab6118b](https://www.linkedin.com/in/nihar-sanda-07ab6118b)  github.com/koolgax99  koolgax99.github.io

Education

Northeastern University - Khoury College of Computer Sciences <i>MS in Computer Science</i>	2023 - 2025 (Expected) <i>Boston, MA, USA</i>
Indian Institute of Information Technology, Dharwad <i>Bachelor of Technology, Computer Science and Engineering</i>	2019 - 2023 <i>Karnataka, India</i>

Experience

Indian Institute of Technology, Bombay <i>Research Associate</i>	May 2023 – Present <i>Mumbai, India</i>
--	---

- 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 <i>SDE Intern</i>	March 2023 - May 2023 <i>Banglore, India</i>
--	--

- **LookupTM Tool:** Creating a comprehensive IP infringement search tool by **utilizing OCR models** and employing Django's Model-View-Template (MVT) architecture to develop the application.
- **Devops Operations:** Deployment and Maintenance of multiple **production and testing environments** on AWS EC2, S3, and Redis, utilizing Github Actions for CI/CD automation
- **NLP Pipeline:** Developed 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 <i>Undergraduate Research Assistant (Under Prof. Kavi Mahesh and Dr. Abhilash C B)</i>	October 2022 - January 2023 <i>Dharwad, India</i>
--	---

- 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 <i>Founder and CEO</i>	February 2021 - October 2022 <i>Mumbai, India</i>
---	---

- **Leadership: Led a team** of more than **10 people** in the development of a **campus management application** catering to a user base exceeding **800 individuals**, receiving **grants and funding** of more than **\$13000**.
- **Incubation:** 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**.
- **Backend Development:** 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 <i>Contributor</i>	June 2022 – November 2022 <i>Mumbai, India</i>
---	--

- Contributed to the creation of new R language's API for PEcAn packages of data assimilation and meta-analysis.
- Enhanced the authentication of the existing REST APIs by incorporating robust **API Key authentication** and implementing efficient **rate-limiting features**.
- Collaborated with several scholars from across the world to comprehend the functionality and use of the PEcAn API

Indian Institute of Technology, Bombay <i>Research Intern</i>	July 2022 – November 2022 <i>Mumbai, India</i>
---	--

- Worked on a novel idea in the field of education technology, named "Affect Aware tutoring system "
- 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 feedback
- Developed an optimized transformer-based deep learning model to predict the user's affect state. The model was trained on the huge DAiSEE dataset for approximately 300 hours.

Projects

Camplus App | *Flutter, AWS Lambda, S3, DynamoDB, API Gateway, Python* **February 2021**

- Tracks the various systems on the campus such as the Mess System, Delivery System, Canteen, Gate Tracker etc.
- Developed various modules in the backend using the MVC architecture in MEN stack, AWS Lambda and API Gateway.
- Connecting the REST APIs with the Flutter UI and integrating secure authentication of the Application.

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

- Pre-processing raw data from a private hospital in Karnataka with care while keeping essential features.
- Using SMOTE and undersampling techniques to address the issue of data imbalance.
- implemented state-of-the-art machine learning models to predict COVID-19 patient mortality, utilizing an ensemble approach. Conducted a comprehensive analysis and comparison of results, achieving an impressive accuracy of 93%.
- Using the **ML models** to build an ensemble model and a thorough comparison and analysis of all the results inferred.

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

- Enhanced Protein Fold Recognition for Low Similarity Datasets of protein sequences and its classes
- Feature Extraction of the data using PSSM and HMM profiles of the protein sequence and combining them with Sequence Graph Transform (SGT) and Bi-gram features.
- Implemented transformer-based models like BERT to make the model more robust.

Technical Skills

Languages: Python, R, C, C++, HTML/CSS, JavaScript, Bash

Technologies: APIs, SQL Databases, NoSQL Databases, RDF, OWL, Docker, AWS, Nginx, Unix, Git

Tools: Machine Learning, DL, CV, NLP, EDA

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

Frameworks: MERN, Flask, Django, AWS, Flutter

Publications

Sqr: High Performance Secure Qr System to Filter Against Malicious Qr Codes Using Deep Learning Techniques

- Developed a high-performance system for detecting malicious and benign QR codes.
- Paper currently under review in Expert Systems with Applications Journal (SCI and SCOPUS Indexed).

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 is set to be published in 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).

Leadership / Extracurricular

- Director's Gold Medal for the Best Outgoing Student at IIIT Dharwad.
- 2 times Google Summer of Code (2022, 2023) Recipient
- 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.