

SANJEEPAN SIVAPIRAN

✉ sanjeepan23@gmail.com  [sanjeepan23.github.io](https://github.com/sanjeepan23)  [Google Scholar](https://scholar.google.com/citations?user=sanjeepan23)  [sanjeepan23](https://orcid.org/sanjeepan23)  [in/sanjeepan](https://in.linkedin.com/in/sanjeepan)

RESEARCH INTERESTS

My research interest lies in the areas of Machine Learning, Responsible AI, Human-AI Symbiosis, Natural Language Processing, Health Informatics, and Large Language Models

EDUCATION

University of Moratuwa, Sri Lanka

Oct 2018 - Jun 2023

B.Sc. Engineering (Hons) in Computer Science & Engineering

First Class Honors - GPA 3.76/4

In Dean's Honors List

Kilinochi Central College, Sri Lanka



Jan 2015 - Aug 2017

G.C.E Advance Level Examination

Distinctions in Combined Mathematics, Physics and Chemistry

Ranked **295th** at National level out of 250, 000 students - Top 1%

PUBLICATIONS

- Party Extraction from Legal Contract Using Contextualized Span Representations of Parties. **Sanjeepan .S**, Charangan .V, and Uthayasanker .T : *RANLP23 - Recent Advances in Natural Language Processing*. 
- Sentiment Analysis in Dravidian Code-Mixed YouTube Comments and Posts. **Sanjeepan .S**, Charangan .V, and Uthayasanker .T : *Working Notes of FIRE 2021 - Forum for Information Retrieval Evaluation, CEUR, 2021*. 

RESEARCH EXPERIENCE

Information Extraction from Documents Using Transformer Models | PyTorch, Deep Learning

Undergraduate, University of Moratuwa

Aug 2022 - Jun 2023

Advisor: Dr. Uthayashanker Thayasivam

- Developed a Question-Answering(QA) system that can identify actual parties in a legal contract
- Modified the **RoBERTa** architecture to improve its learning capacity.
- Developed an open-source dataset of legal parties, encompassing 2,750 entities extracted from 1,000 documents.

Act-Sense: Human Activity Recognition using sensor data from wearables | LightGBM, Machine Learning

Undergraduate, University of Moratuwa

Jan 2023 - Jun 2023

Advisor: Dr. Kutula Gunasekara

- Trained a machine learning model (XGBoost) to identify Human Activity activities in real-time.
- Implemented a cross-platform mobile application to show the status of the activity

Text-based Economic Analyser for Sri Lankan Stock Market | Python, Stanford NLP, Machine Learning

Undergraduate, University of Moratuwa

Aug 2022 - Jun 2023

Advisors : Dr. Nisansa de Silva and Prof. Indika Perera

- Developed a text-based tool to analyze the Sri Lankan stock market
- The tool is built around ontology triples extracted from financial news and blog

Contextual Word embedding (word2vec, GloVe) for Lower Resourced Natural Languages

(Tamil) | Python, Word2vec, Machine Learning

Research Assistant, AAIUVU.org

Jun 2021 - Sep 2021

Advisor: Dr. Uthayashanker Thayasivam

- Created contextual word representations for the Tamil language words using word2vec and FastText algorithms.
- Implemented a web application that allows users to visualize words in a three-dimensional vector space.

Predicting Druggable Proteins from Primary Sequence (amino acid sequence) | XGBoost, Machine Learning

Undergraduate, University of Moratuwa

Feb 2023 - Jun 2023

Advisor: Dr. Charith Chitaranjan

- Developed a multi-class classification model using LightGBM to identify potentially druggable proteins
- Extracted and constructed relevant features from amino acid sequences to train the machine learning model.

Gastrointestinal(GI) tract anomaly detection | CNN, Deep Learning

Undergraduate, University of Moratuwa

Oct 2020 - Feb 2021

Advisor: Dr. Charith Chitaranjan

- Trained a classification model using KVASIR dataset to classify anomalies in the Gastrointestinal Tract using endoscopic imagery
- Implemented using Deep Convolutional Neural Network with transfer learning.

TEACHING EXPERIENCE

University of Moratuwa, Sri Lanka | Department of Computer Science and Engineering

Teaching Assistant

Aug 2022 - Jul 2023

- Modules: CS 3121 - Introduction to Data Science, CS 3630 - Deep Neural Networks
- Prepared and graded weekly quizzes, assignments, and mid-semester exams for the class of 85 students.
- Facilitated and supervised the laboratory sessions for the students.

PROFESSIONAL EXPERIENCE

H2O.ai

Machine Learning Engineer

Jul 2022 - Present

- Working as a core engineer in developing Enterprise Large Language Models (LLMs)
- Designed and implemented H2O LLM Data Studio to transform documents, audio, and text into LLM-compatible datasets
- Created a scalable, distributed inference pipeline for in-house MOJO model

H2O.ai

Machine Learning Engineer - Intern

Dec 2021 - Jul 2022

- Created a signature detection and verification system using PyTorch that is currently being used by the largest bank in Australia
- Implemented a machine learning model that can predict the credit score using non-conventional data, +15% better accuracy than existing model

ExentAI

Machine Learning Engineer - Intern

Sep 2021 - Nov 2021

- Created an end-to-end sentiment and emotion analysis system
- Deployed the models on a remote EC2 server for inference with low latency

RELEVANT COURSES

Machine Learning	A
Data Mining and Information Retrieval	A
Software Engineering	A
Bioinformatics	A
Intelligent Systems	A
Database Systems	A
Graph Theory for Computing	A+
Mobile Computing	A
Numerical Methods for Computer Science	A
Calculus for System modelling	A
Linear Algebra	A+
Differential Equations and Applications	A
Applied Statistics	A-
Research and Development Project	A+

AWARDS AND HONORS

Dean's Honors List - Outstanding academic performance in Semesters 04, 05, and 08.	2023
Google KickStart - 2020 - Got 3 rd in Sri Lanka in Round H	2020
Mahapola Higher Education Merit Scholarship - For outstanding performance in G.C.E A/L	2019
Nanajaya' Scholarship for A/L Achievers of 2020 - For outstanding performance in A/L.	2018
Best Patrol leader of the Year - Award for best patrol leader of 2nd Jaffna Scout Troop.	2014

SKILLS

Programming Languages	Python, Java, JavaScript, Bash, C++
Machine Learning	PyTorch, Sci-Kit Learn, Tensorflow, Keras
Web Development	HTML, CSS, NodeJS, ExpressJS, H2O-Wave
Mobile Development	React Native, Flutter, Dart
Databases	MySQL, MongoDB, Redis
Cloud	AWS, GCP, Heroku, NGINX, Jenkins
Tools	Git, Linux, Conda, Tmux, VSCode, Docker, FastAPI

OTHER

IELTS - 7.5

- Listening / Reading / Writing / Speaking : 7.5 / 8.5 / 6.5 / 6.5