Name: Harsh Sagar| Roll No: BS19B013

Indian Institute of Technology Madras [**Github**](https://github.com/harshs21/Machine-Learning-Projects.git) **Linkedin**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **EDUCATION** | | | | | | |
| Program | | Institution | | CGPA/% | | Year of Completion |
| Dual Degree in Biological Sciences | | Indian Institute of Technology, Madras | | 7.94 | | 2019-2024 |
| XII (GSEB) | | Best High School, Ahmedabad | | 80% | | 2019 |
| X (GSEB) | | Best High School, Ahmedabad | | 87.17% | | 2017 |
| **SCHOLASTIC ACHIEVEMENTS** | | | | | | |
| * **Only** student from **2019 batch** to achieve **Panasonic Scholarship Program** * **Gold Medal** for Class Rank 1, cleared 1st Level and secured **Int. Rank 3359** in 2nd Level of National Science Olympiad in Class X * Secured International **Rank 859** in International Science Olympiad in **Class IX** conducted by **Science Olympiad Foundation (SOF)** | | | | | | |
| **RELEVANT COURSES AND SKILL** | | | | | | |
| * Principle of Economics | | | * Computational Economics | | * Climate Economics | |
| * Fundamentals of Operation Research | | | * Series and Matrices | | * Biostatistics | |
| **WORK EXPERIENCE** | | | | | | |
| **Data Science Intern at FN MathLogic Consulting Services, Gurgaon** *May’23 – July ‘23* | | | | | | |
| *Developing a Conversational Question Answering System using Large Language Models (LLMs)* | | | | | | |
| * Explored **LLM memory retention** via various **model finetuning methods** on **ICICI Lombard Policy data** for answering user’s query * Implemented **Transfer Learning** on **Microsoft DialoGPT** LLM & **Reinforcement Learning** from **Human Feedback** finetuning method | | | | | | |
| *Document Parsing for Question Answering using LLMs with Langchain framework* | | | | | | |
| * Retrieved chunks from **ICICI policy docs** that are alike to user query using Langchain. **Flan-T5 LLM** embeddings were used for answering | | | | | | |
| **Machine Learning Intern at Street Style Store Company, Delhi** *May’22 – Aug’22* | | | | | | |
| * Applied different machine learning models for improving user’s experience by analysing chats of user & agent | | | | | | |
| **Data Science Intern at Supratech Lab, Gujarat** Nov*’21 – Dec’21* | | | | | | |
| * Identified crucial **biomarker** **genes** for enhancing **IVF** **treatment** using machine learning algorithm that got 82% test accuracy * Reduced **57k** gene features to **6k**, using **PCA**, **t-SNE** for dimensionality reductions, and **T-test** for feature significance. | | | | | | |
| **RESEARCH EXPERIENCE** | | | | | | |
| **ML Research Intern**  **IIT Varanasi**  [**Research Paper**](https://link.springer.com/article/10.1007/s12161-022-02379-z)  *June’21 – July’21* | **Machine Learning for Milk Foam Analysis, Guide: Prof. Abhishek Dhoble** | | | | | |
| * Worked on Classification of **Surfactants** to study milk foam quality as demand for **Cappuccino foams** has increased * **Randomforest** achieved **0.955 roc-auc score** & **88.1% test accuracy**, also applied algorithms eg **SVM, XGBoost etc** * Explored **Casein** reaction with **surfactants**, this study also got published as the **research** **paper** in **Springer Nature** | | | | | |
| **DDP PROJECT**  **IIT Madras**  *July’23- Ongoing* | **Evaluating ML models for Chest X-ray of Diseases, Guide: Prof. Ganapathy Krishnamurthi** | | | | | |
| * Applied **Transfer Learning** on **Ensemble** of **CNN** models the chest x-ray images of Pneumonia and Normal person * Implemented Image Data Generator for increasing training data to avoid **overfitting** & obtained **93.7%** accuracy | | | | | |
| **COURSEWORK** | | | | | | |
| **Finance Plus, Finance & Economics Club IIT Guwahati** | | | | | | |
| * Completed assignments based on concept of **asset** evaluation like future cash flow from lease, rental payments & capital appreciation * Applied statistical concepts to real-world situations, such as estimating **population variance** for quantifying investment risk | | | | | | |
| **Real Estate Auction System, [HS4007: Computational Economics]** | | | | | | |
| * Created **payoff matrix** of **hawk-dove** game to simulate contest for maximizing player payoff in **real estate** trades using **game theory** * Assigned bids using **Gaussian distribution** or **compounding** prior values & evaluated based on final bid price to highest valuation of bidder | | | | | | |
| **Forecasting CO2 Emission and Evaluating its Impacts on Climate Change, [HS4870: Climate Economics]** | | | | | | |
| * **Assessed** the overall financial burden posed by each metric ton of **CO2** released by assigning monetary values to repercussions * **Polynomial Regression** was used to **forecast** per capita annual carbon emissions for the years **2021-2030** using past records since 1960 | | | | | | |
| **PROJECTS** | | | | | | |
| ***Trading Strategy on Time Series data*** | | | | | | |
| * Meticulously designed stock trading strategy and backtest on **TATASTEEL** data from 2021. Calculated the **P/L** of entire backtest at end * **Bought** the stock if current closing price (**CP**) exceed both Support1 (**S1**) & Simple Moving Average(**SMA**) that ensured potential uptrend * **Sold** the stock if current CP was below S1 & 20-period SMA that enhanced risk management. Fixed **1% Profit Target** & **0.5% Stop Loss** | | | | | | |
| ***Trading Strategy using Momentum*** | | | | | | |
| * Developed trading strategy based on momentum indicators using the historical **Closing prices (CP)** of stocks for a subset of the **S&P500** * Implemented functions to resample prices, compute **log returns** and generated **trading** **signal** using **Momentum** for **1 month** period * Evaluated the performance of the resulting portfolio through hypothesis testing using **T-test** and calculated **p vlaues** | | | | | | |
| ***Stocks & Crypto Currencies Price Prediction Using LSTM*** | | | | | | |
| * Developed Long Short-Term Memory (**LSTM**) models with Keras to predict **Closing price** values based on over a decade of trade data * While training, past **19** days values used as input to predict next day’s value. Achieved mean squared error (**MSE**) of **0.465** on testdata | | | | | | |
| **POSITION OF RESPONSIBILITY** | | | | | | |
| **Computer Vision & Intelligence Club Project Member, Shaastra 2022** *June’21 – April’22* | | | | | | |
| * Worked on **YOLO v5 model** for detecting circuit components and the **mAP** score of the model evaluated with 25 images was **92.7%** | | | | | | |
| **Biogen Super Coordinator, Shaastra 2022** *July’21 – Jan’22* | | | | | | |
| * **Supervised** all the coordinators in **theme ideation** & related events to be organized related to biotechnology under the Biogen Team * **Collaborated** with other teams of Shaastra which help organizing event and also assist them with the **publicity** of the event | | | | | | |
| **Public Relations Volunteer, Shaastra 2020** *Sept’19 – Feb’20* | | | | | | |
| * Reached out to various **social media** accounts as a part of PR activity for our social campaign named **“BLINK “.** * Planned and executed **“Walk in the Dark”** and various activities for our stall at KV ground and spread **campaign motto** | | | | | | |
| **EXTRA-CURRICULAR ACTIVITIES** | | | | | | |
| **Sports** | * Selected for NSO Fitness program and also participated in Samanvay Marathon 2019 | | | | | |