Pragyan Jyoti Dutta

pragyan-portfolio.netlify.app | github.com/cyber-prags | linkedin.com/in/pragyan-jyoti-dutta
| duttapragyanjyoti@gmail.com | +44-7407747069

About Me

As a Master's student in Data Science and AI at the University of Liverpool, complemented by a Data Science Diploma from IIT Madras and a Bachelor's in Physics from Tezpur University, my academic journey is characterized by a potent combination of intensive training and real-world application. My exposure encompasses spearheading groundbreaking projects in healthcare and astronomy, evidenced by my authorship in an IEEE Xplore-published research paper. Professionally, my tenure as a Data Engineer at DaaS and a Machine Learning researcher at Spartificial has fortified my skills in both teamwork and autonomous problem-solving. I thrive in leadership positions, steering teams through intricate projects with a commitment to lucid communication and adaptability in dynamic environments. My method integrates analytical thoroughness with a profound insight into team interplay, ensuring that both technical hurdles and collective goals are adeptly addressed.Eager to embrace new challenges in data-centric roles such as Data Analyst, Data Scientist, or Data Engineer, I am on the lookout for opportunities that not only promise professional growth but also allow me to contribute significantly to the organization.

Education

University of Liverpool, United Kingdom — Sept 2023 - Aug 2025
MSc. in Data Science and Artificial Intelligence — GPA: 2:1 degree(predicted)

Indian Institute of Technology, Madras, India — Jan 2022 - Sept 2023
Online Diploma in Data Science — GPA: 7.75/10.0

Tezpur University, India — Oct 2020 - Jun 2023
BSc. in Physics — GPA: 8.06/10.0

Delhi Public School, Digboi, India — May 2018 - Jun 2020
Higher Secondary — Percentage: 95.67%

Work Experience

Data Engineering Intern — Aug 2022 - Feb 2023
Developer As a Service(DaaS) - Team

- Conducted data mining using OCRs from textual image data.
- Engineered the data using Regex and other cleaning methods on unstructured pixelated image legal documents
- Collaborated with my team to create a full stack ecosystem for advanced legal research services
- Employed strong communication and leadership skills to foster collaboration with clients and team members, resulting in the successful delivery of customized solutions. This initiative led to an impressive 80% increase in efficiency for legal document processing within the firm.

Machine Learning Research Intern — Jun 2022 - Sept 2022
Spartificial — Online, India

- Training: Learned about RNN, Neural Networks, TensorFlow, OpenCV, Image Segmentation etc
- Research: Optimized pulsar detection in imbalanced datasets using ML techniques.

Research Internship — Nov 2021 - Dec 2021
Society for Space Education and Research — Online, India

- Worked on a project titled, "Indirect methods to determine the fundamental properties of a Stellar-mass black hole and the discrepancies found in the measurement."

Data Science Intern — Sept 2021 - Oct 2021
The Sparks Foundation — Online, India

- Performed Exploratory Data Analysis on a bunch of datasets
- Learned about Supervised and Unsupervised Learning

Summer Research Internship — Aug 2021 - Oct 2021
Naxxatra Sciences and Collaborative Research — Online, India

- Worked on a project titled, "Analyzing the forest cover of India and land usage pattern of the last decade using Python" that helped the stakeholder to better understand the forest cover of India and change their policies accordingly

Projects

Blog generator using Generative AI Jan 2024 - Present
Personal Project
- Engaging in a project harnessing Generative AI capabilities through LangChain and Meta's LLaMA 2 model, aiming to develop an advanced blog generator that crafts compelling content tailored to user inputs.

Tomato Disease Classification App using Deep Learning Dec 2023 - Present
Personal Project
- Developing a comprehensive full-stack application leveraging Deep Learning with TensorFlow, orchestrating diagnosis of tomato plant diseases from images. Implementing ReactJS and NodeJS for seamless deployment across web and mobile platforms.

Stocktible - A web app to predict stock prices Dec 2023 - Jan 2024
Personal Project
- Utilized advanced libraries including TensorFlow, fbprophet, and yfinance to meticulously extract data from financial websites and accurately forecast stock prices through sophisticated predictive modeling.

Sentiment Prediction on Movie Reviews Jun 2023 - Aug 2023
MLP Project , IIT Madras Click to view
- Worked on a Natural Language Processing project that aimed at predicting the sentiment of a movie based on its review by various reviewers using Machine Learning models

Online Grocery Store WebApp- GROCIFY Jun 2023 - Aug 2023
MAD-1 Project , IIT Madras Click to view
- Created "Grocify," a community-centered e-commerce grocery web app utilizing Flask, HTML, CSS, and Bootstrap, innovatively merging traditional grocery shopping with online convenience. Drove personalized shopping experiences and precise promotions through advanced analytics, elevating user satisfaction and product excellence.

Classifying Pulsar Stars using Machine Learning and Neural Networks June 2022 - Sept 2022
Dr. Susheela Dahiya , UPES Dehradun Click to view Report
- Worked under the guidance of Dr. Susheela Dahiya to classify pulsar stars using Machine Learning

Indirect methods to determine the fundamental properties of a Stellar-mass black hole Nov 2021 - Dec 2020
SSERD, Online Click to view Report
- Analyzed the various measurable parameters of a Stellar Mass Black Hole
- Gave a review of the discrepancies found in the various measurement techniques being used currently

Chronic-Kidney-Disease-Prediction-using-Machine-Learning April 2021
Independent Project Click to view Notebook
- Predicted the chances of a person having Chronic Kidney Disease considering various parameters
- Used Logistic Regression, KNN, Decision Trees, and RandomForest algorithms

Other Projects
- **FB-add-campaign-analysis :** Analyzed FB ad campaigns data and suggested ways to optimize them
- **Analyzing India's Forest Cover of the last decade :**

Technical Skills

Programming languages: C++, Python, Java, R
ML/AI: Pytorch, Numpy, Pandas, Matplotlib, Scikit-learn, Keras
Web Technologies: HTML, Django, React
Miscellaneous: MySQL, Git, Shell, Latex, PowerBI, Tableau, MS-Office, Excel, Office

Soft Skills

Communication: demonstrated excellent communication skills to facilitate collaboration and successfully deliver customized solutions in team projects.

Adaptability: Displayed adaptability by taking on various roles, from data mining to research, in different projects for DaaS, and research projects for Spartificial, Naxxatra.

Problem-Solving: Proficiently utilized problem-solving skills in tasks like data analysis, machine learning, and project development.

Teamwork: Successfully collaborated with team members on various projects, ensuring the attainment of shared objectives and the delivery of high-quality work.

Leadership: Orchestrated and led cross-functional teams to deliver successful solutions, resulting in a 20% improvement in project efficiency. Served as the team leader for my DaaS (Data as a Service) team, overseeing a 4-member group, and provided strategic direction that led to a 80% increase in efficiency for the firm. Additionally, I assumed leadership of the research group at Spartificial, where I guided a team of 3 researchers.

Time Management: Effectively managed time to meet project deadlines and ensure project efficiency. Successfully balanced concurrent commitments, including pursuing a B.Sc. in Physics at Tezpur University and completing a Diploma in Data Science from IIT Madras, demonstrating exceptional time management abilities.

Resilience: demonstrated resilience in handling challenges and contributing to the successful completion of projects.

Cultural Sensitivity: Recognized the importance of cultural sensitivity in diverse work environments in DaaS and Spartificial, contributing to effective teamwork and project outcomes

Publications

IEEE Xplore: Health Risk Detection through Web App using Machine Learning April 2022
Publication link

- The main motive of this project was to help the general people who are unaware of the technologies and can easily use it through the web app present online which has all the features to detect the disease at an early stage using ML techniques of Naive Bayes and RandomForests.

Relevant Certifications

Matillion Data Productivity Cloud, Foundations Certificate Link
Matillion Academy

- Obtained foundational knowledge, practical skills, and a functional understanding of Data Engineering and various ETL and ELT tools available in the market.
- Gained a foundational knowledge of building data pipelines in Matillion Data Productivity cloud.

Building a Data Warehouse using Matillion Certificate Link
Matillion Academy

- Gained insights about data warehousing and knowhow of building Data Warehousing pipelines in Matillion Data Productiviry Cloud and integrate it with Snowflake.

Generative AI with Large Language Models Certificate Link
DeepLearning.AI

- Obtained foundational knowledge, practical skills, and a functional understanding of generative AI, gaining insights into the latest research and how companies leverage cutting-edge technology for value creation.
- Benefited from expert instruction by AWS AI practitioners actively involved in building and deploying AI for real-world business applications.

Deep Learning with PyTorch: Object Localization Certificate Link
Coursera

- Developed essential skills through this course, including the ability to create customized datasets for localization tasks, augment data effectively for improved model performance, and utilize pre-trained models efficiently, also mastered the creation of training functions and evaluators, streamlining the training process.