

Pratibimb Swain

[LinkedIn](#) | [GitHub](#) | [LeetCode](#)

+91 7303560017

pratibimb.swain@gmail.com

Aspiring Data Scientist with strong foundations in computer science and hands-on experience in NLP, deep learning, and data-driven solutions. Seeking opportunities to apply technical expertise to real-world challenges and drive innovation.

EDUCATION

B.TECH(Computer Engineering)	2024-2028	Delhi Technological University, New Delhi	8.80 CGPA
BS (Data Science and Application)	2024-2028	Indian Institute of Technology, Madras	9.00 CGPA
AISSCE/CBSE (Class XII)	2024	Delhi Public School, Vasant Kunj	95.8%
AISCE/CBSE (Class X)	2022	Delhi Public School, Vasant Kunj	97.2 %

WORK EXPERIENCE

Open Source Contributor — GirlScript Summer of Code (GSSoC) 2025

July 2025- September 2025

Project: BMC Product Reviews Sentiment Analysis

Python, NLTK, Regular Expressions

- Developed a robust NLP preprocessing module incorporating text cleaning, stopword removal, and POS tagging with NLTK.
- Enhanced the text normalization process for more accurate downstream analysis

ACADEMIC PROJECT

NLP-based Mental Health Chatbot:

Tensorflow, Model Serialization, Tensorflow.js, Matplotlib

- Built a conversational AI chatbot leveraging the **Universal Sentence Encoder with two dense layers** for intent classification and sentiment analysis.
- Trained on **16,000 samples (2,000 validations)**, achieving **87.4% training accuracy** and **84.6% validation accuracy**.
- Implemented deployment pipeline: exported to **TensorFlow.js** for web-based accessibility, with scope for **cloud integration (AWS/Azure)**.

Inventory Optimization and Data Analytics project:

Excel, Python

- Analyzed inventory data across **multiple SKUs** for a medical devices supplier to assess stock-out.
- Identified **43%+ products** operating below reorder thresholds and flagged high-risk stock-out items.
- Proposed optimized reorder points with the potential to reduce stock-out incidents.

Income Classification Model:

Scikit-Learn, Pandas, Numpy, Matplotlib

- Built a binary classifier to predict income ($>50K$ or $\leq 50K$) using the UCI Adult dataset.
- Implemented and compared models like **Logistic Regression, Random Forest** and **XGBoost**, achieving **87.5% accuracy** and **0.72 F1 score**.

TECHNICAL SKILLS

- Programming Languages**-Python, SQL
- Core CS Concepts**-Object-Oriented Programming, Data Structures and Algorithms, Operating Systems
- Technologies**-ML, DL, NLP, LLM, RAG
- Frameworks**- Scikit-learn, Matplotlib, TensorFlow, Keras, OpenCV, NLTK, LangChain, Hugging Face
- Databases and Tools**-MySQL, SQLite, ChromaDB, GitHub, Git, Excel
- Certifications**-Machine Learning and Deep Learning Specialization (Stanford University), Generative AI Specialization(IBM)

ACADEMIC ACHIEVEMENTS AND AWARDS

- Rating of **1624** on Leetcode
- Secured **First Position** in a Coding competition consisting of DSA questions
- Secured **Second Position** in Financial Cricket Competition
- Completed **360+** problems on LeetCode