ABITATHA ROY

abitatharoy@cic.du.ac.in | github.com/AbitathaRoy | linkedin.com/in/abitatha | +91-9836962506

Headline

Undergraduate researcher in applied machine learning, with experience in cross-domain generalization, healthcare prediction, and dynamical systems. Interested in research-driven internships in ML and computational biology.

Education

B.Tech. Information Technology and Mathematical Innovations
Cluster Innovation Centre, University of Delhi (Expected 2027) (89.1%)
Relevant Coursework: Calculus, Linear Algebra, ODE, Linear Programming, Statistics, Data Science, Nanotechnology, *in silico* Biology, Genetics.
ISC (Class XII), Calcutta Public School, Kalikapur (2023) (97.0%)

Selected Experience

Student Intern 08/2025 – now
Cluster Innovation Centre, University of Delhi

- Analyzing urban meteorological and pollution datasets using ML techniques to model spatio-temporal trends.
- Developing predictive and exploratory models to derive actionable insights for environmental assessment.

Achievements

2nd Position – MathWorks CodeMatrix Sprint

- Secured 2nd Prize at the MathWorks Hackathon for a novel Gamified Fitness Tracker solution.
- Captures and processes real-time mobile sensor data and integrates machine learning to calculate and report on user kinematics (e.g. speed, distance) and calories burned and integrates gamification module.
- **Tech Stack:** MATLAB, Python

Selected Projects

Cross-Dataset Performance Evaluation of Predictive Algorithms for Phishing Website Classification
Tech Stack: Python, Scikit-learn, XGBoost, Pandas

- Implemented standardized URL feature extraction pipeline (17 features); identified optimal 9-feature subset via mutual information, achieving 84.55% cross-dataset accuracy with Random Forest.
- Conducted cross-dataset evaluation (Mendeley → PhiUSIIL) to study domain shift; implemented dataset harmonization to align feature distributions.

Predicting Tuberculosis Treatment Default Using Artificial Neural Networks
Tech Stack: Python, TensorFlow/Keras, Streamlit, Pandas

- Integrated three TB datasets; addressed domain variability via feature harmonization.
- Designed regularized ANN (dropout + L2) optimized for high recall in default prediction.
- Achieved ROC-AUC ≈ 0.75 and PR-AUC ≈ 0.89, prioritizing sensitivity to defaults.
- Deployed interactive Streamlit dashboard for real-time clinical risk assessment.

Implementation of Dynamical Circuits for Secure Communication via Chaotic Cryptography
Tech Stack: LTspice

- Designed and simulated Chua, Lorenz, and Rössler circuits for chaos-based secure signal transmission.
- Demonstrated synchronization-based encryption using low-cost analog components.
- Presented at Dynamics Day Delhi (DDD–XXI) Conference at BML Munjal University; developed structured lab manual for undergraduate and postgraduate dynamical systems education.

Technical report and reproducible code available upon request.

Selected Certifications

IBM Data Science Professional Certificate – Coursera (Ongoing)

Bioinformatics Specialization by UC San Diego – Coursera (Ongoing)

Skills

Core Areas: Probabilistic Modelling, Machine Learning, Artificial Neural Networks, Bioinformatics

Programming: Python, R, C, Java, SQL

Tools: TensorFlow/Keras, scikit-learn, MATLAB, Mathematica, LTspice, Pandas, NumPy, MS Excel