

Arup Sankar Roy

+91 9679713408 | arupsankarroy11@gmail.com | [LinkedIn](#) | [GitHub](#) | [Portfolio](#)

Education

University Institute Of Technology (UIT),Burdwan

Burdwan,West Bengal, India

Bachelor of Engineering in Computer Science and Engineering . current CGPA : 7.38

Aug. 2022 – Jul, 2025

Experience

Data Science Intern

August 2023 – August 2023

CodersCave Virtual Internship Program – CVIP,India

- **Conducted EDA** on **Global Terrorism Database** using Numpy, Pandas, Matplotlib, Seaborn, and Plotly. Utilized Choropleth maps, Bar Plots, Area Plots, Pie Charts, and Heat Maps for **insightful visualizations**, unveiling **security implications**.
- Performed **Email Spam Detection** utilizing **NLTK library**. Employed **text preprocessing**, feature extraction, and **machine learning techniques**. Achieved accurate classification, highlighting potential **spam emails**.

Machine Learning Intern

August 2023 – August 2023

TechnoHacks EduTech, India

- **Diabetes Prediction:** Achieved **88% accuracy** using the **Random Forest Classifier** with **data visualization**.
- **Fraud Transaction Detection:** Attained **97% accuracy** using **CatBoost**, handling an imbalanced dataset with the **Imblearn library**.
- **Wine Quality Prediction:** Reached **91% accuracy** and **0.93 precision** with **AdaBoost** and **K Nearest Neighbor** algorithms.

Featured Projects

Energy Consumption Prediction Flask Application

August 2023 – September 2023

Used Python Libraries : Numpy,Pandas,Scikit-Learn,Flask.

- Fetched data from kaggle, processed and performed EDA .Created a model with **accuracy 99%** and saved in pickle format.
- Prepared **flask** application to run the model.and Generated an user friendly UI. [Github Link](#)

CAPM : Capital Asset Pricing Model Web-App

March 2023 – March 2023

Used Python Libraries : Numpy , Pandas , Streamlit , Plotly.

- Used **YFinance** Library to access the **financial data** available on **Yahoo Finance** processed and performed EDA.
- Prepared **streamlit** application to run the model. [Github Link](#)

Skills

Data Analytics and Visualization Libraries : Numpy,Pandas,Matplotlib, Plotly,Streamlit,Seaborn,Scipy.

Machine Learning Libraries : Scikit-Learn,PyTorch , OpenCV,NLTK.

Programming Languages : Python,Java, C, HTML, CSS,YAML,Flask.

System Experience : Linux, Windows.

Softwares : Microsoft Power BI,Microsoft Excel,Matlab.

Databases : MySQL.

Achievements

Secured the 65th rank out of 600+ participants in Geeks for Geeks problem-solving at UIT, Burdwan.

Lead the team (Tensors) secured rank 5 out of 20+ teams in 'CodeBird HackWitz 2k22'.

Profile Links



[Geeks For Geeks](#)



[Github](#)



[Linked In](#)