SHEIKH FARHAN

B.Tech.

CSE - DS

Haldia Institute Of Technology

Portfolio

+917584930131

farhan.sheikh.new@gmail.com

FarhanSheikh-01

Sheikh Farhan

farhansheikh-01.github.io

EDUCATION

- **Haldia Institute Of Technology** *November 2022 - June 2026*

  Bachelor of Technology - Computer Science and Engineering(DS); GPA: 7.65 Purba Medinipur, India

SKILLS

- **Programming Languages::** C , Java , Python

- **Libraries & Frameworks::** Mathplotlib, Seaborn, Pandas, Numpy, Scikit-Learn, Tensorflow, Keras

- **Development::** Django, Flask, Docker, MySQL

- **Data Science & Machine Learning::** Supervised and Unsupervised Learning, Feature Engineering, Natural Language Processing (NLP), Deep Learning(CNN, ANN, LSTM)

- **Version Control::** Git, GitHub

- **Deployment::** Render, Railway , AWS(beginner)

EXPERIENCE

- **Training and Internship Program**

  Euphoria Jan 2025 – Feb 2025

  - **Technologies**: Machine Learning, AWS Cloud (EC2, S3, API Gateway), Django
  - **Roles and Responsibilities**:
    * Implemented **5+ Machine Learning algorithms**, optimizing classification accuracy by **12%**.
    * Deployed ML models on **AWS EC2**, reducing sentiment analysis processing time by **30%**.
    * Managed **1.6M+ tweets dataset**, applying NLP techniques such as tokenization, lemmatization, and stopword removal.
    * Developed a full-stack Django-based web application, improving **real-time sentiment classification** efficiency.
    * Built a project titled **"Sentiment Analysis on Tweets"**, scraping **10K+ tweets** using `ntscraper`, classifying sentiments with an **80.06% accuracy model**.

PROJECTS

- **Sentiment Analysis of Tweets** *Jan 2025 - Feb 2025*

    * Collected a dataset of over 1.6 lakh tweets.
    * Preprocessed text using **Natural Language Processing (NLP)** techniques such as tokenization, stopword removal, and lemmatization, reducing vocabulary size by 40% for better generalization.
    * Implemented multiple ML models for sentiment classification, achieving **80.06% accuracy** using algorithms like **Naïve Bayes, Support Vector Machines (SVM), and Logistic Regression**.
    * Built and deployed a **Django-based web application**, hosting it live on **Render**. Visualized insights with Matplotlib & Seaborn, identifying sentiment trends across **10+ industry sectors**.

Tea Sickness Analysis and Management System Aug 2024 – Dec 2024

    * Aimed to identify the seven most common diseases in tea agriculture in India.
    * The system uses a **Convolutional Neural Network (CNN)** to classify these diseases with an accuracy of **84.27%**.
    * This system has the potential to increase farmers' yields by 25% and reduce an annual crop loss of 147 million. Additionally, offers a **management system** to enhance modern agricultural techniques.
    * Developed a Django application and hosted it live on Railway.

CERTIFICATIONS

- **Data Analysis with Python**, IBM · Course | *Nov 2024*
- **Supervised Machine Learning: Regression and Classification**, Coursera | *Aug 2024*

ACHIEVEMENTS & PARTICIPATIONS

- ,Vice Captain of the Departmental Football and Volleyball Teams

- , Active Kaggle Contributor