edunet foundation

Topic: Storm and Heavy Rainfall Prediction

College: Lokmanya Tilak College of Engineering

Students: Kushal Nishad
Saurav Patil
Ruturaj Desai

Learning Objectives

- Understand the significance of weather prediction in disaster prevention
- Learn how to preprocess and analyze weather datasets
- Apply machine learning algorithms to forecast heavy rainfall and storms
- Evaluate model performance using real-world metrics
- Explore deployment of AI-based weather prediction systems

Tools and Technology used

- Languages: Python
- Libraries: Pandas, NumPy, scikit-learn, XGBoost, Matplotlib, Seaborn
- Tools: Google Colab
- Datasets: IMD, NOAA, Kaggle
- Icons/Graphics: www.freepik.com

Dataset Overview

- Meteorological Data: Weather conditions (real-time or historical)
- Storm Event Data: Official storm reports
- Rainfall Measurements: Precipitation in mm or inches
- Satellite Data (optional): For cloud coverage, wind patterns, etc.

Methodology

1. Data Collection (weather records from IMD/NOAA)
2. Data Cleaning & Preprocessing
3. Exploratory Data Analysis (EDA)
4. Feature Selection and Engineering
5. Model Training: Logistic Regression, Random Forest, XGBoost
6. Evaluation using accuracy, F1 score, confusion matrix
7. (Optional) Deployment using a web interface

Alogrithms used:

1. Logistic Regression
- Baseline model for binary classification (rain/storm vs. no rain/storm)
- Fast, interpretable, and good for linearly separable data

2.Decision Tree
- Splits data based on feature thresholds
- Easy to understand, but may overfit

3.Random Forest
- Ensemble of decision trees
- Reduces overfitting, works well with complex data

4.XGBoost
- Boosting-based model
- High accuracy, handles missing values, widely used in competitions

5. Support Vector Machine (SVM)
- Finds the optimal hyperplane for class separation
- Effective in high-dimensional space

6. K-Nearest Neighbors (KNN)
- Classifies based on closest data points
- Simple, but slower with large datasets

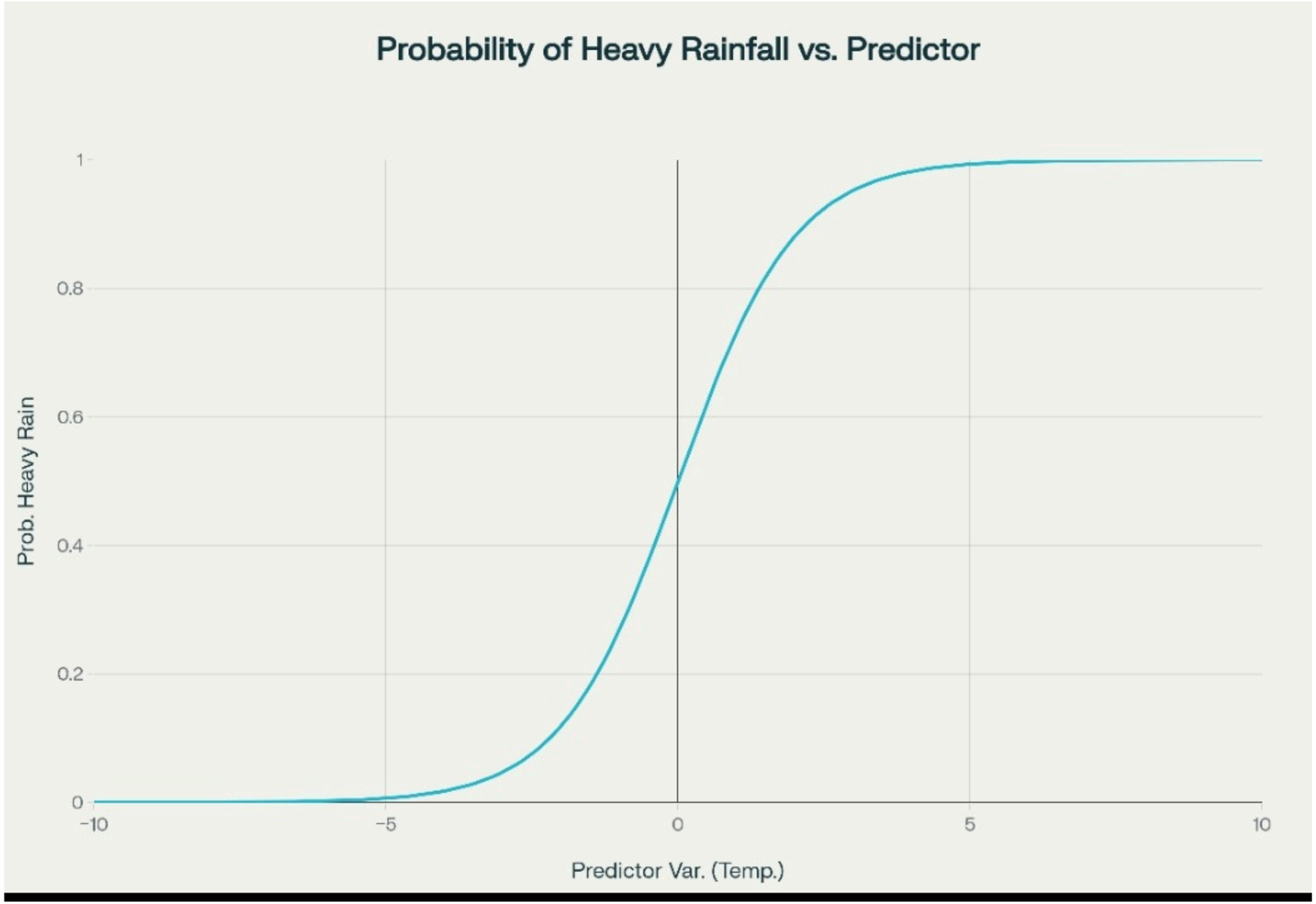
Problem Statement:

Unpredictable rainfall and storms cause massive damage to life, agriculture, and infrastructure. Traditional systems often lack real-time accuracy. The challenge is to build an AI model that can forecast these events early and reliably.

Solution:

- A machine learning-based solution that:
- Predicts storm or heavy rain likelihood from historical weather inputs
- Uses classification (e.g., logistic regression, random forest)
- Generates alerts for potential disasters with minimal false positives

Screenshot:


Probability of Heavy Rainfall vs. Predictor
Prob. Heavy Rain
1
0.8
0.6
0.4
0.2
0
−10
−5
0
5
10
Predictor Var. (Temp.)

Conclusion:

The AI/ML model provides a data-driven approach to predicting heavy rainfall and storms. It supports early warning systems, urban planning, and agricultural decision-making. With further training and integration into live weather systems, it can significantly reduce risk and improve preparedness.

GitHub repository link of the project

git@github.com:Soro2006/edunet.git