**Readme file:**

**Weather Forecasting Using Hybrid Machine Learning Algorithms**

**Project Overview**

This project focuses on improving weather forecast accuracy using hybrid machine learning algorithms. It integrates models such as CatBoost-Logistic Regression, XGBoost-KNN, and Gaussian Process-Decision Tree to predict key weather parameters, including rainfall, temperature, and relative humidity. The framework leverages multi-label classification techniques on a decade-long dataset from Gannavaram to achieve high predictive performance.

**Features**

* **Hybrid Machine Learning Models:** Combines the strengths of different algorithms for better accuracy.
* **Multi-label Classification:** Predicts multiple weather parameters simultaneously.
* **Data-Driven Insights:** Utilizes real-world data to improve forecasting reliability.
* **High Accuracy:** Achieved 100% accuracy in rainfall prediction with the Gaussian Process model.

**Technologies Used**

* **Machine Learning Algorithms:** CatBoost, XGBoost, Gaussian Process, Logistic Regression, Decision Tree, Random Forest, K-Nearest Neighbors (KNN)
* **Data Processing:** Python, Pandas, NumPy
* **Visualization:** Matplotlib, Seaborn
* **Evaluation Metrics:** Accuracy Score, F1 Score, Confusion Matrix, Classification Report

**Dataset**

* **Source:** Gannavaram Weather Department
* **Duration:** 2013 - 2023
* **Features:** Rainfall (mm), Maximum and Minimum Temperatures, Relative Humidity (RH%) at different times
* **Data Split:** 80% Training (2013-2021), 20% Testing (2022-2023)

**Project Workflow**

1. **Data Collection:** Gathered a decade-long weather dataset.
2. **Data Preprocessing:** Cleaned, normalized, and handled missing values.
3. **Model Development:** Implemented standalone and hybrid machine learning models.
4. **Evaluation:** Assessed model performance using various metrics.
5. **Optimization:** Fine-tuned models to enhance prediction accuracy.

**Key Learnings**

* Applied hybrid machine learning algorithms for complex forecasting tasks.
* Gained expertise in data preprocessing, feature engineering, and model optimization.
* Developed skills in multi-label classification and model evaluation techniques.

**Results**

* **Rainfall Prediction:** 100% accuracy with Gaussian Process.
* **Temperature & RH% Prediction:** High accuracy with CatBoost, XGBoost, and hybrid models.

**How to Run the Project**

1. **Clone the Repository:**

git clone [repository\_link]

1. **Navigate to the Project Directory:**

cd weather-forecast-ml

1. **Install Dependencies:**

pip install -r requirements.txt

1. **Run the Notebook:** Open Weather.ipynb using Jupyter Notebook or any compatible IDE.

**Future Work**

* Incorporating real-time data for dynamic weather prediction.
* Enhancing model performance with advanced deep learning techniques.
* Expanding the dataset to include more geographical regions for broader applicability.

**Contributors**

* Naga Charitavya Madala
* Sai Durga Saradhi Pranu Deepak Tallapudi
* Venkata Srikari Malladi
* Durga Mahesh Muthinti
* Srilatha Tokala
* Murali Krishn Enduri

**License**

This project is licensed under the MIT License - see the LICENSE file for details.