**🌧 Ahmedabad Real-time Forecast Model — RUVISION**

**👉 Overview**

This repository runs an operational rainfall forecasting model for **Ahmedabad** using GFS reanalysis data and observed IMD rainfall. The model has two major components:

1. **Data Download & Processing**
2. **Forecast Modelling (Lead Day 1–3) & Plotting**

**📌 Variables of Interest**

|  |  |
| --- | --- |
| **Code** | **Description** |
| U1000 | U-component of wind at 1000 mb |
| V1000 | V-component of wind at 1000 mb |
| PREC | Precipitation rate from GFS |

* **Observed Rainfall**: IMD gridded daily rainfall at 0.25° resolution  
  📖 *Reference*: Pai et al., MAUSAM (2014)
* **Training Period**: 2015–2023
* **Testing / Real-time Forecast**: 2024–present

**⏱️ Initialization Details**

* **Initialization Hour**: 06 UTC (11:30 AM IST)
* **Forecast Hours**:
  + Lead Day 1: 15–36 hrs.
  + Lead Day 2: 39–60 hrs.
  + Lead Day 3: 63–84 hrs.
* **Grid Size**: 5x5 region over Ahmedabad
* **Forecast Method**: Two-stage **Censored Quantile Regression** (CQR)

**🎯 Forecast Objective**

To forecast **daily rainfall in Ahmedabad** for:

* Tomorrow
* Day after tomorrow
* Third day from initialization

**🧠 Model Setup Workflow**

**1. Data Collection**

* Literature study for variable selection
* Download historical GFS data (0.25°) from NCEP/NOMADS
* Restrict lat/lon to Ahmedabad city region

**2. Data Processing**

* Save historical data (for each lead day) into Excel
* Pre-processing functions:
  + preprocess\_wind() for wind (U1000, V1000)
  + preprocess\_precipitation() for precipitation (rate → accumulation)
* Convert GFS hourly to daily:
  + **Precipitation**: sum hourly values
  + **Wind**: average hourly values

**3. Modelling**

* For each lead day:
  + Combine U1000, V1000, PREC
  + Normalize & apply PCA
  + Save transformed features in .pkl
  + Run **GLM-based CQR** for quantile forecasting

**⚙️ Real-time Operations**

* Create base path & folders for each variable
* Construct download URLs from NOMADS
* Download GFS .pygrb2 data (06z run)
* Pre-process and split into 3 lead days
* Append new data to historical Excel
* Run quantile regression forecast
* Output Excel with daily quantile predictions
* Plot 3 next days forecasts with current date initialisation

**🚀 Automation**

* Code refactored & modularized
* Deploy on **HPC cluster**
* Scheduling daily at **5 PM IST**
* Output forecasts hosting on website

**🛠️ Additional Tasks**

* ✅ GitHub Code Review
* ✅ Experimented with 100+ variable combinations with past data
* ✅ Scoring Metrics: Accuracy, F1, Recall, HSS, CQVSS
* ✅ Selected best-performing variable set
* ✅ Validation of Real-time Forecasts with Zomato Weather Union Data