

## 2) Thunderstorm Forecasting with MLFlow Tracking (Climatology)

### # What are Thunderstorms?

- A rain-bearing cloud that also produces lightning.  
All thunderstorms are dangerous. Every thunderstorm produces lightning.

#### Impact

Impact on Humans  
Impact on Animals  
Impact on Properties  
Impact on Nature

#### Benefits

Production of Nitrogen  
Production of Ozone  
Replenish Underground  
Water Table Reservoir  
To maintain Earth's Electric  
balance.

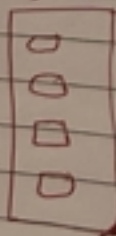
### # 2 Deployment

- 1 local
- 2 cloud deployment -- Free Cost

### # Assignment

- 1) From the doc provided, read and understand the information given about data.
- 2) Read all accuracy metrics from document

### # Python File



#### modules

logic1  
• .py

logic4  
• .py

main  
• .py

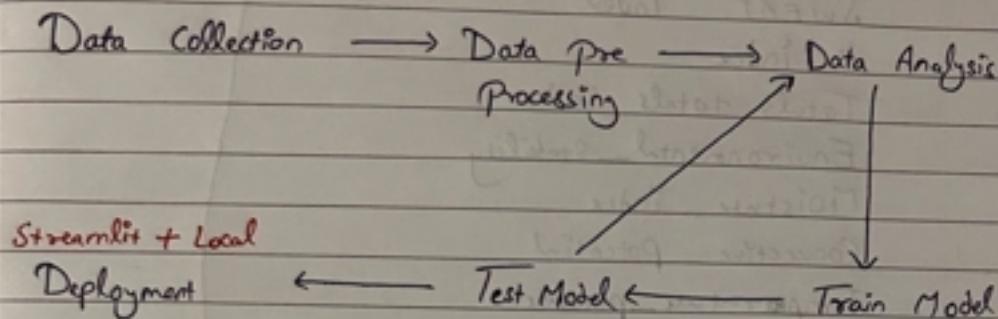
#### Python Files

logic2  
• .py

logic3  
• .py

#

## Machine Learning Life Cycle



Streamlit + Local  
Deployment

## OUR LOGIC

#

### Development Process

- Raw Data
  - Processing & Merging as one Table
- Combining
- Model Training
- Model Testing

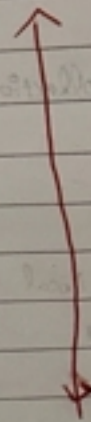
#

### Data Transformation + Feature Engineering

- 1) 'PRECIPITABLE WATER' → 'Moisture Indices'
- 2) 'CAPE' → 'Convective potential'
- 3) '1000-500 THICKNESS' → 'Temperature Pressure'
- 4) 'PLCL' → 'Moisture Temperature profiles'
- 5) 'Showalter Index' → 'Environmental stability'
- 6) 'LIFTED index' → 'SWAT index'
- 7) Same as before → 'K index'
- 8) Totals totals index

Only After taking Suggestion of SME

Variable	Dependent
SWEAT	Index
K	Index
Totals	totals index
Environmental	stability
Moisture	index
Convective	potential
Temperature	pressure
Moisture	Temperature profiles
Tt	Dependent

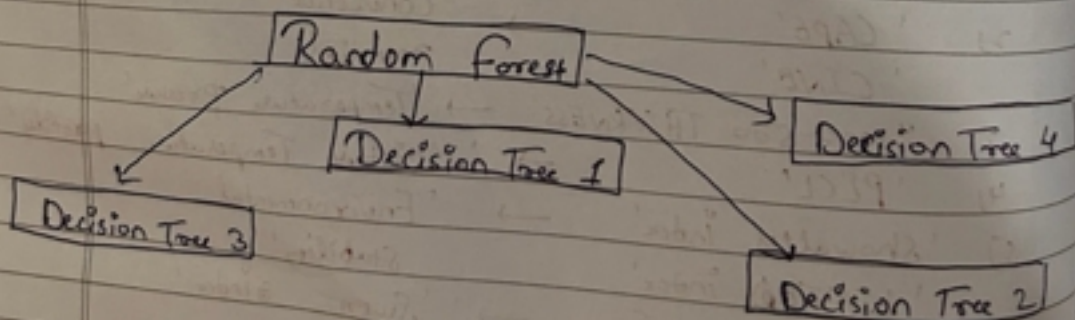


## # Model Training + Vibe coding

Model training includes <sup>the</sup> following steps:

- Data Preparation
- Model Selection
- Data Split
- Model Training

## # Random Forest Classifier





- Random forest is an ensemble learning method that combines Multiple decision Trees to make Predictions.

- It creates a collection of Decision Trees, where each Tree is Trained on a random subset of the Data & features.

- The final prediction determined by aggregating the predictions of all individual Trees.

## ML Flow Tracking (To Land on the best Model)

