Drought Prediction: Leveraging Meteorological and Soil Data

A Comparative Analysis of Techniques to Forecast Drought Events

Taha Abdullah EE-014
Muhammad Waqas EE-024

Project Overview

What this project is about:

- Predicting drought occurrence using meteorological and soil data.
- Analyzing feature importance in drought prediction.

Why we are doing it:

- To understand the environmental factors contributing to droughts.
- To develop better drought preparedness strategies using ML.

Objectives:

- Perform comparative analysis of ML techniques (e.g., logistic regression, decision tree, ANN).
- Correlation between features

Insight of Meteorological and Soil Data

There are 18 features in meteorological and 32 features in soil data. Some are shown below:

Meteorological Data:

T2M: Temperature at 2 Meters (C)

PRECTOT: Precipitation (mm day-1)

PS: Surface Pressure (kPa)

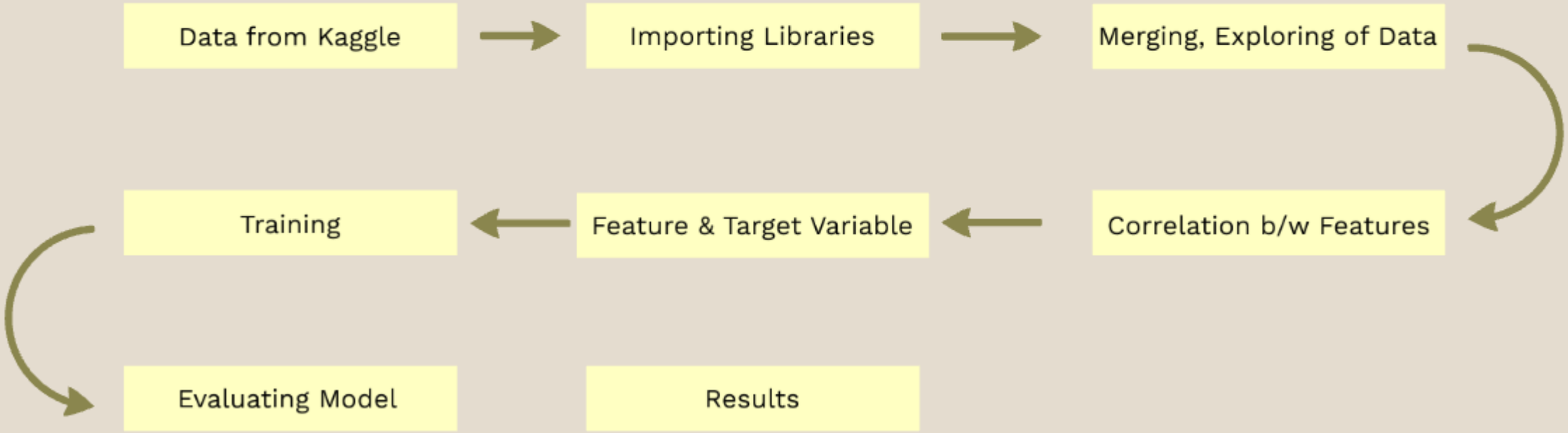
Soil Data:

SQ1: Nutrient availability

WAT_LAND: Mapped Water Bodies

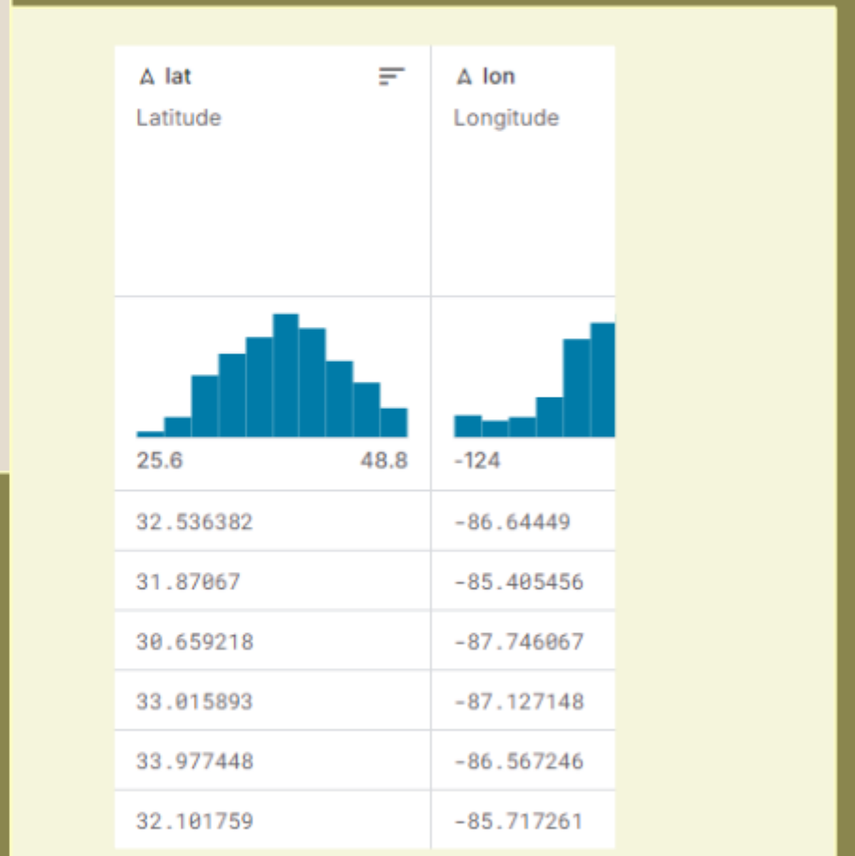
Elevation: Median elevation (meters)

Data Pipeline

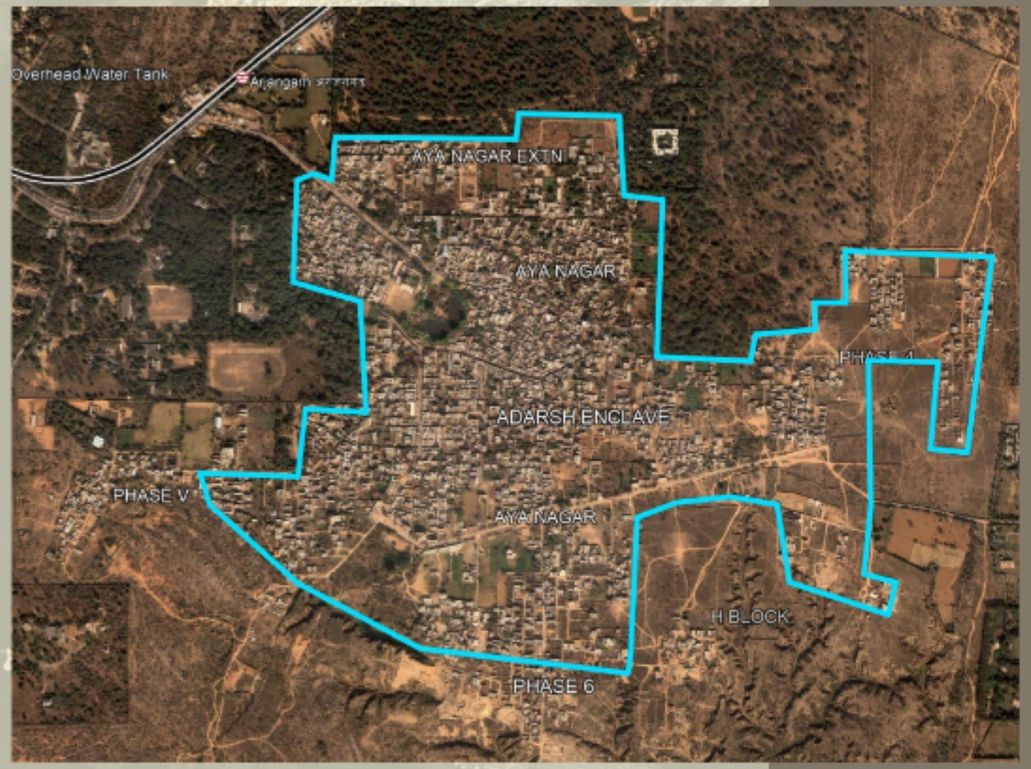


Data Acquisition from Kaggle

Data is sourced from Kaggle's diverse datasets related to meteorological and soil conditions.



A lat Latitude	A lon Longitude
25.6 48.8	-124
32.536382	-86.64449
31.87067	-85.405456
30.659218	-87.746067
33.015893	-87.127148
33.977448	-86.567246
32.101759	-85.717261



Importing Libraries

Essential libraries such as pandas, NumPy, and Scikit-learn are imported to facilitate data manipulation, visualization, and machine learning tasks.

Data Merging and Exploration

Merging meteorological data with soil data is crucial for comprehensive analysis. This step involves exploring the dataset's integrity, identifying duplicates, and pre-processing the information to ensure consistency and relevance to the prediction models.

Drought Categories Explained

Drought has 5 categories include meteorological, agricultural, and hydrological droughts. Each category is determined by specific indicators such as precipitation deficits, crop yield impacts, and streamflow levels, providing a holistic view of drought conditions.

Category	Description	Possible Impacts
D0	Abnormally Dry	Going into drought: ▪ short-term dryness slowing planting, growth of crops or pastures Coming out of drought: ▪ some lingering water deficits ▪ pastures or crops not fully recovered
D1	Moderate Drought	▪ Some damage to crops, pastures ▪ Streams, reservoirs, or wells low, some water shortages developing or imminent ▪ Voluntary water-use restrictions requested
D2	Severe Drought	▪ Crop or pasture losses likely ▪ Water shortages common ▪ Water restrictions imposed
D3	Extreme Drought	▪ Major crop/pasture losses ▪ Widespread water shortages or restrictions
D4	Exceptional Drought	▪ Exceptional and widespread crop/pasture losses ▪ Shortages of water in reservoirs, streams, and wells creating water emergencies

Model Training Techniques

Model training involves applying different algorithms such as logistic regression, decision tree, and artificial neural network on the dataset to establish predictive relationships. This includes data preprocessing steps like checking for duplicate values and scaling data to ensure clean input for models.

SEE CODE

Key Features Influencing Drought

Identifying meteorological variables such as **temperature**, **precipitation**, and **soil moisture** content is vital for understanding drought impacts. These features significantly influence drought events, enabling more accurate predictions by analyzing their relationships with drought severity.

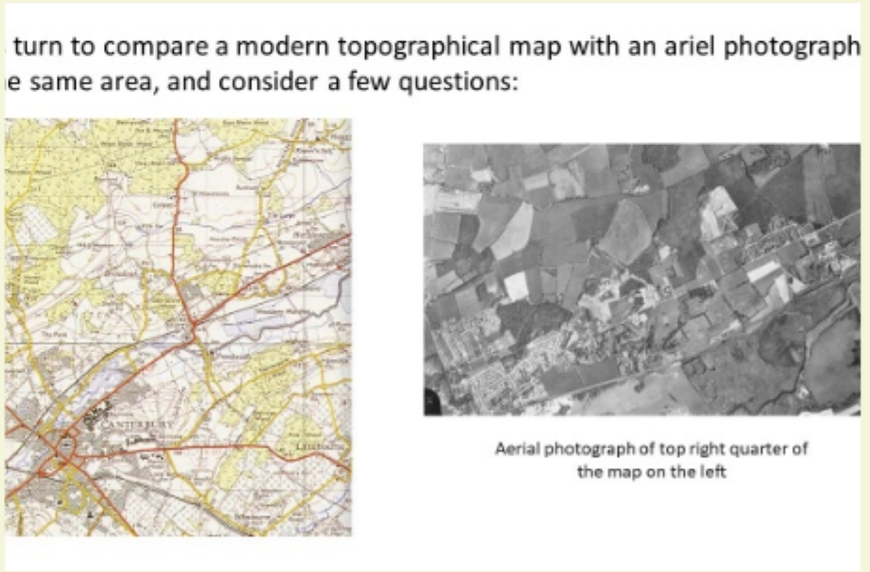
Name	Validation Accuracy	Test Accuracy
Decision Tree	75 %	69 %
LOG REG	28 %	0.2 %
ANN	56 %	76 %

Key Findings

Machine learning models demonstrated varying accuracy rates, emphasizing the need for model selection based on specific forecasting scenarios. Our analysis revealed that certain meteorological factors, such as temperature and rainfall patterns, have a more significant impact on drought occurrences than soil moisture levels. Furthermore,

Conclusion and Future Work

Future work will focus on enhancing model algorithms such as Long Short Time Memory.
Moreover, we can achieve higher accuracy by doing feature engineering and use of hyperparameters.



Q&A Session


THERE'S THIS THING CALLED GOOGLE
YOU CAN ASK IT YHE SAME QUESTION YOU JUST ASKED