

Project Proposal

Problem Statement:

Global warming and climate changes have increased the probability of natural disasters like drought, hurricanes, floods etc. Natural disasters like droughts have intense adverse impacts on human and animal lives. Droughts are the second most costly weather events causing large amounts of economic and social losses. A drought can be defined as a period when certain geographical regions experience below-normal precipitation levels. This condition leads to reduced soil moisture or groundwater, diminished stream flow, crop damage, shortage of food and water. Drought conditions are known to last for weeks, months or years together. Thus, early prediction of drought would help in resource allocation and management, disaster preparedness and ultimately save human, plant and animal lives

Machine Learning Problem:

This problem can be viewed as a **classification** problem, more specifically *multi-class classification*. The target variable or classes has been divided into 6 categories as follows:

Class	None	D0	D1	D2	D3	D4
Description	No Drought	Abnormally Dry	Moderate Drought	Severe Drought	Extreme Drought	Exceptional Drought

Dataset:

We aim to use the dataset available on Kaggle: <https://www.kaggle.com/cdminix/us-drought-meteorological-data>
The dataset consists of weather data and soil data: with 21 feature variables. It consists of daily weather data from 2000 to 2020. The distribution of the dataset into train, validation, and test set is given below:

Split	Range (years)	Percentage
Train	2000 - 2009	47%
Validation	2010 - 2011	10%
Test	2012 - 2020	43%

Advantages of drought prediction:

Prediction of drought can help us to take measures at early stages to avoid it.

With timely measures we can minimize environmental impacts like degradation of land, deforestation, wildfire etc thus in turn helping us save our habitat and environment.

Drought can have adverse effects on the economy as well especially in the agricultural sector, hence having an early response time can ensure smooth economic activities and ensure to avoid adverse situations like water scarcity and famine risks.

Evaluation criteria:

We aim to reduce the false negative cases more in comparison to false positives, as facing a drought when we aren't prepared results in adverse effects than being prepared for a drought and not resulting in a drought.

Performance metric: Recall should be optimized to get less number of false negatives.

Submitted by,

Sahana Pandurangi Raghavendra (2176508)

Pooja Tank(2176991)

Pooja Nadagouda (2173740)