Crop Yield Prediction using LSTM

Here we have used the **LSTM (Long Short-Term Memory)** model, which is based on the **Recurrent Neural Network (RNN)** architecture. This method allows us to analyze historical data to predict crop yields for future seasons. By leveraging temporal dependencies, this approach enhances the ability of our machine learning model to provide high-level yield predictions.

Mathematical Structure of the LSTM Unit

Let the input features at time t be defined as:

1. Y_t: Crop yield at time t
2. R_t: Rainfall
3. T_t: Temperature
4. S_t: Soil Nutrient
5. F_t: Fertilizer used
6. H_t: Humidity
7. pH_t: Soil pH level

Our goal is to predict the crop yield for the future (Y_{t+1}) using the past n time steps.

LSTM Mathematical Intuition

Input at each time step:

- **Input vector (x_t):** A vector containing the 7 features listed above.
- **Hidden state (h_{t-1}):** Information from the previous time step.
- **Cell state (C_{t-1}):** The long-term memory from the previous time step.

1. Forget Gate

Decides how much information from the previous cell state C_{t-1} should be discarded.

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f)$$

2. Input Gate

Decides which new information will be stored in the cell state.

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C)$$

3. Update Cell State

The updated memory combining past records and new information.

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t$$

4. Output Gate

Determines the current hidden state which is used for the prediction.

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o)$$

$$h_t = o_t * \tanh(C_t)$$

Crop Yield Prediction Layer

After passing through the LSTM layers, the final prediction is calculated:

$$\hat{Y}_{t+1} = W_y h_t + b_y$$

Where:

- h_t: The hidden state of the LSTM.
- $\hat{Y}_{t+1}$: The predicted crop yield for the next season.
- W_y, b_y: Learned weights and bias for the output layer.

Loss Function and Model Training

The model minimizes the difference between the actual yield (Y) and the predicted yield ($\hat{Y}$) using the **Mean Squared Error (MSE)**:

$$L = \frac{1}{N} \sum_{i=1}^{N} (Y_i - \hat{Y}_i)^2$$

Where:

- N: Number of samples.
- Y_i: Actual yield.
- $\hat{Y}_i$: Predicted yield.