ANN Mathematical Process from Scratch for Paddy Yield Prediction

Dataset

- **Samples:** 2,789
- **Input Features (X):** 45 (Climatic, Soil, Farm Management, Land area, County)
- **Target (y):** Paddy yield

Data Preprocessing:

1. Missing & duplicate values removed
2. Wind direction: Sin-Cos encoding
3. Management inputs: Input i = ci * Hectares; excluded if redundant
4. Feature scaling: Z-score normalization for continuous variables (Xij-np.mean(Xj)/ np.std(Xj))
5. Categorical: One-hot encoding

Research Hypothesis: ANN can predict paddy yield accurately by capturing nonlinear relationships among climate, soil, and agronomic factors.

Artificial Neural Network Network Architecture

Forward Propagation

$X*W1 = Z1 \Longrightarrow A1 = \tanh(Z1) \Longrightarrow$

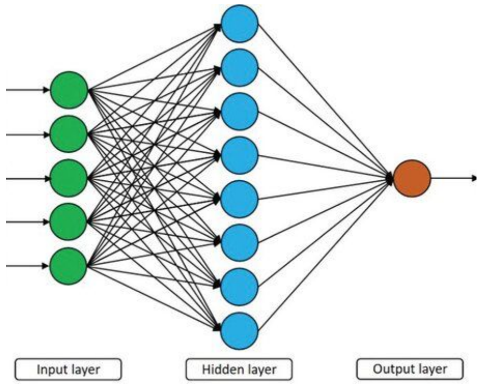
$\Longrightarrow A1 * W2 = Z2 \Longrightarrow y\_pred = \text{linear}(Z2)$

Parameter Update

$W1 := W1 - \eta\, \partial L / \partial W1$

$W2 := W2 - \eta\, \partial L / \partial W2$

Inputs (X)



Loss Function

$MSE = 1 / n \sum (yi - y\_pred)^2$ for (i = 1 to n)

Backpropagation & Parameter update

$\partial L / \partial W2 = \partial L / \partial y\_pred \cdot \partial y\_pred / \partial Z2 \cdot \partial Z2 / \partial W2$

$\partial L / \partial W1 = \partial L / \partial y\_pred \cdot \partial y\_pred / \partial Z2 \cdot \partial Z2 / \partial A1 \cdot \partial A1 / \partial Z1 \cdot \partial Z1 / \partial W1$

ANN - Experimental Setup & Training Conditions

$X\_train \in [-2.09, 2.00]$
$y\_train \in [-1.85, 1.77]$
$\eta = 0.01 / 0.001$

Standard Normal: $W \sim N(0,1)$
Xavier Initialisation: $W = randn(\cdot)\sqrt{(1 / n\_inputs)}$ and $\sqrt{(1 / n\_hidden)}$

$Z_1 \in [-1.44, 1.42]$
$Y\_pred \in [-0.43, +0.43]$
Batch Gradient Descent
Epoch: 1000

Exp No	Architecture	A1	Output	Train MSE Scaled (Forward Pass)	Train MSE Scaled (After one step Gradient descent)	Test MSE (Scaled)	RMSE (Kg)	Relative RMSE (%)
B. 1	43 / 1/ 1	tanh	linear	1.0581	1.054	1.0144	9311.2	41.00%
B. 2	43 / 1/1	Sigmoid	linear	1.0343	1.0336	0.9954	9223.4	40.61%
B. 3	43 / 1/1	Relu	linear	1.0541	1.0506	1.0048	9266.89	40.80%
B. 4	43 / 1/1	Linear	linear	1.098	1.0868	1.0525	9484.35	41.76%
B. 5	43 / 1/1	Linear	Tanh	1.092	1.082	1.0473	9460.96	41.65%
Exp No	**Architecture**	**A1**	**Output**	**Train MSE Scaled (Epoch 0)**	**Train MSE Scaled (Final step)**	**Test MSE (Scaled)**	**RMSE (Kg)**	**Relative RMSE (%)**
6	43 / 1/ 1	Relu	linear	1.0581	0.1014	0.1023	2957.2	13.02%
7	43 / 1/1	Linear	linear	1.098	0.0138	0.014	1093.81	4.82%
8	43 / 5 / 1	tanh	linear	1.0668	0.0149	0.0136	2771.95	4.74%
9	43 / 5 / 1	linear	tanh	1.095	0.099	0.0899	2771.95	12.20%
10	43 / 30/ 1	tanh	linear	1.1159	0.011	0.0113	980.88	4.32%
11	43 / 30/ 1	Sigmoid	linear	1.8264	0.014	0.0139	1091.42	4.81%
12	43 /20/ 10 / 1	Tanh / tanh	linear	1.0521	0.0118	0.01225	1023.48	4.51%
13	43 /30 / 1	linear	linear	1.2378	0.0127	0.0125	1031.85	4.54%
14	43 / 30/ 1	tanh	linear	TensorFlow benchmark		0.0107	954.63	4.20%
15	43 /20/1	sigmoid	linear	Models (for comparison)		0.0097	910.3	4.01%

Gradient Descent Convergence for Top Models & Actual vs Predicted Plot for Best Model