Fuzzy Logic & Neural Networks (CS-514)

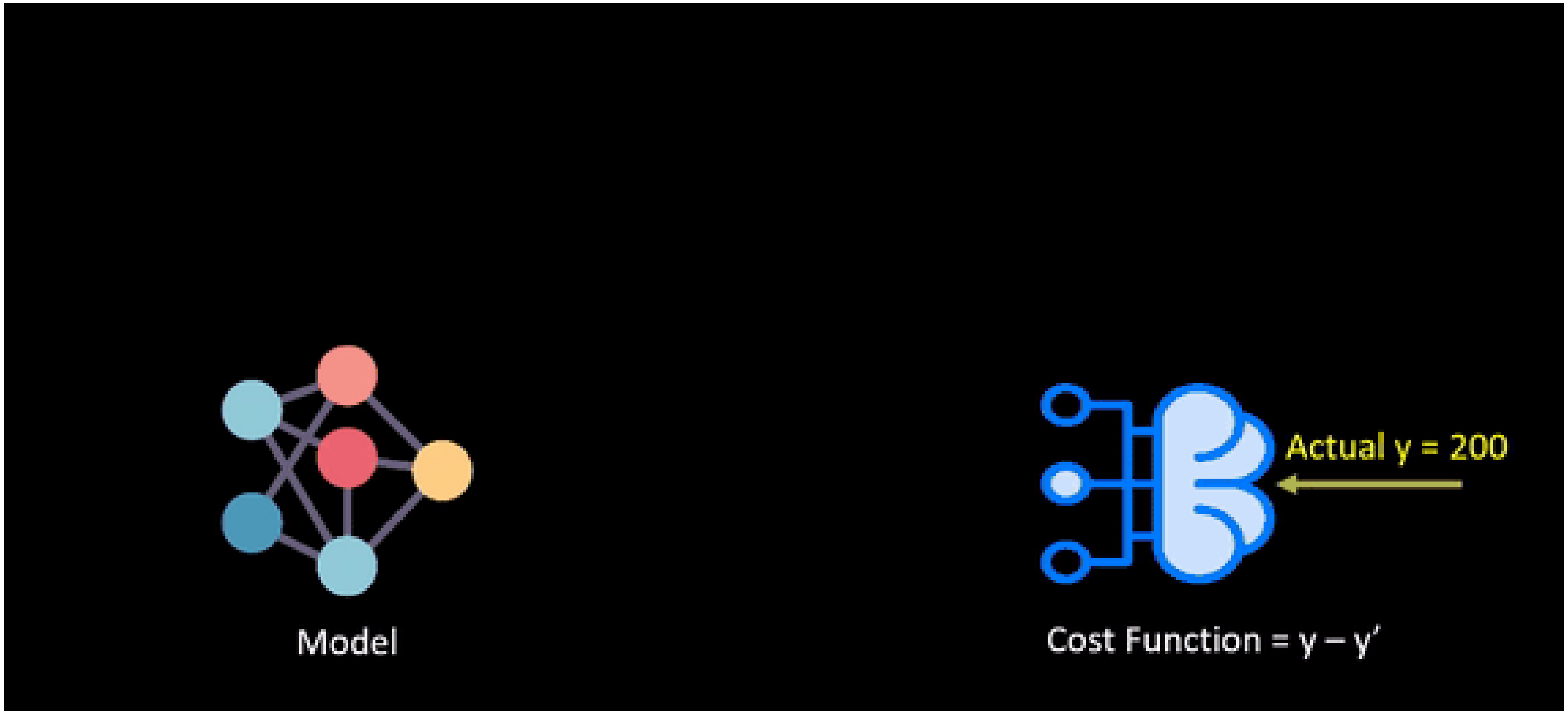
Dr. Sudeep Sharma

IIIT Surat

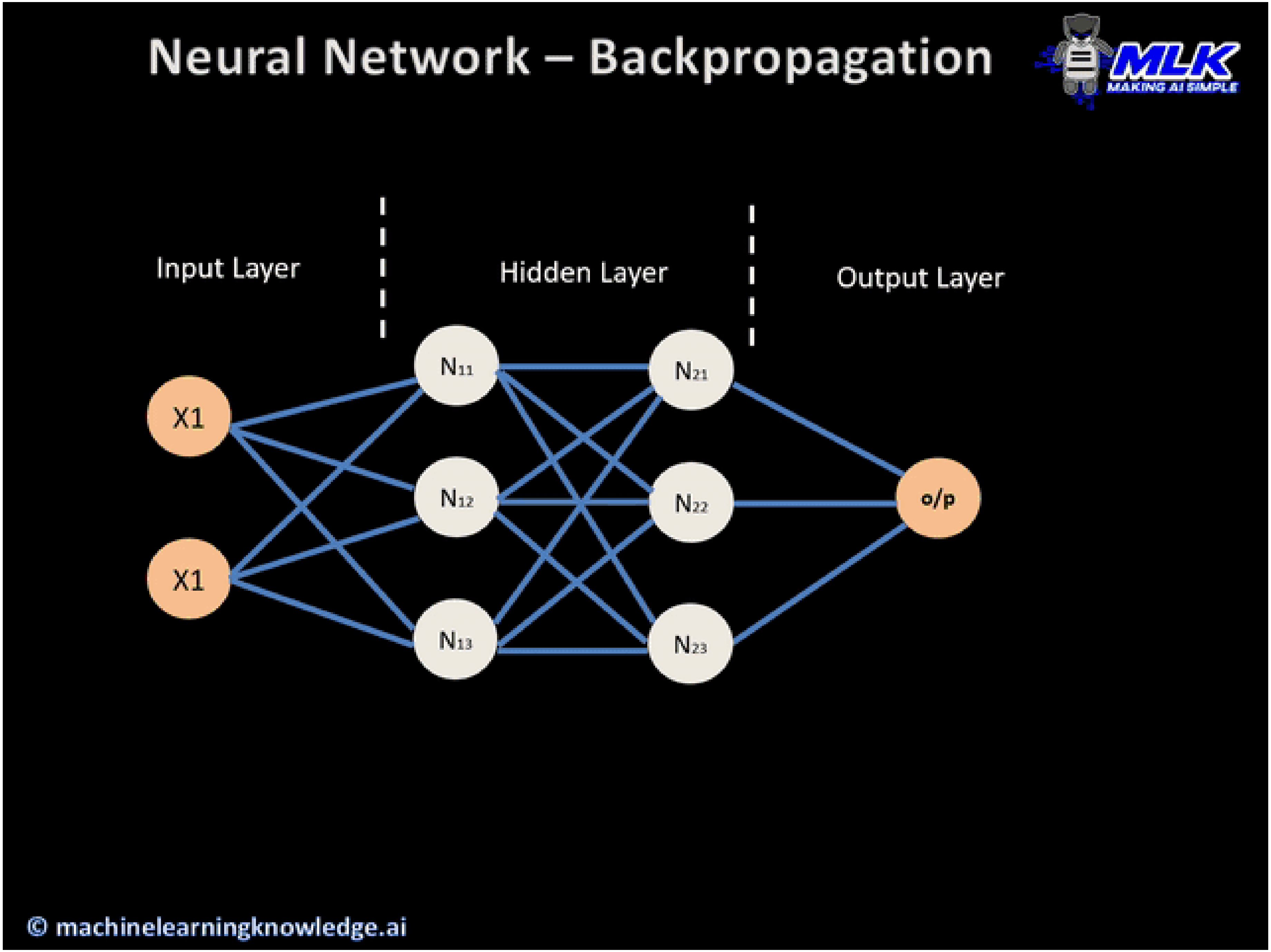
sudeep.sharma@iiitsurat.ac.in

Loss Function

- In neural network training, a loss function/cost function/objective function is a mathematical function.
- The goal of training a neural network is to minimize this loss function, to improving the accuracy of the model's predictions.



Source: Internet



Source: Internet

Loss Function

Mean Squared Error (MSE) Loss Function

- MSE loss function is used for Regression Problems.

$$MSE = \frac{1}{N}\sum_{i=1}^{N}\left(y_i - \hat{y}_i\right)^2$$

- MSE measures the average squared difference between the actual target values (y_i) and the predicted values $(\hat{y}_i)$.
- Lower MSE indicates better performance.

Loss Function

Mean Absolute Error (MAE) Loss Function

- MAE loss function is used for Regression Problems.

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|y_i - \hat{y}_i\right|$$

- MAE measures the average absolute difference between actual (y_i) and predicted values $(\hat{y}_i)$.

- It is less sensitive to outliers compared to MSE.

Loss Function

Example

MAE is less sensitive to outliers compared to MSE.

Data Point	Actual y	Predicted $\hat{y}$
1	10	12
2	15	14
3	14	13
4	18	20
5 (Outlier)	100	80

$$\text{Errors} = [10 - 12, 15 - 14, 14 - 13, 18 - 20, 100 - 80] = [-2, 1, 1, -2, 20]$$

Loss Function

Example

MAE is less sensitive to outliers compared to MSE.

Data Point	Actual y	Predicted $\hat{y}$
1	10	12
2	15	14
3	14	13
4	18	20
5 (Outlier)	100	80

$$\text{MAE} = \frac{1}{5}(|-2| + |1| + |1| + |-2| + |20|) = \frac{1}{5}(2 + 1 + 1 + 2 + 20) = \frac{26}{5} = \mathbf{5.2}$$

Loss Function

Example

MAE is less sensitive to outliers compared to MSE.

Data Point	Actual y	Predicted $\hat{y}$
1	10	12
2	15	14
3	14	13
4	18	20
5 (Outlier)	100	80

$$\mathrm{MSE} = \frac{1}{5}((-2)^2 + 1^2 + 1^2 + (-2)^2 + 20^2)$$
$$= \frac{1}{5}(4 + 1 + 1 + 4 + 400) = \frac{410}{5} = \mathbf{82}$$

Loss Function Derivative

(MSE) Loss Function Derivative

$$\frac{\partial}{\partial \hat{y}_i} MSE = \frac{\partial}{\partial \hat{y}_i}\left(\frac{1}{N}\sum_{i=1}^{N}\left(y_i - \hat{y}_i\right)^2\right)$$

$$\frac{\partial}{\partial \hat{y}_i} MSE = \frac{1}{N}\frac{\partial}{\partial \hat{y}_i}\left(y_i - \hat{y}_i\right)^2$$

$$\frac{\partial}{\partial \hat{y}_i} MSE = \frac{2}{N}\left(y_i - \hat{y}_i\right)\frac{\partial}{\partial \hat{y}_i}\left(y_i - \hat{y}_i\right)$$

$$\frac{\partial}{\partial \hat{y}_i} MSE = -\frac{2}{N}\left(y_i - \hat{y}_i\right)$$

Loss Function Derivative

(MAE) Loss Function Derivative

$$\frac{\partial}{\partial \hat{y}_i} MAE = \frac{\partial}{\partial \hat{y}_i}\left(\frac{1}{N}\sum_{i=1}^{N}\left|y_i - \hat{y}_i\right|\right)$$

$$\frac{\partial}{\partial \hat{y}_i} MAE = \frac{1}{N}\frac{\partial}{\partial \hat{y}_i}\left|y_i - \hat{y}_i\right|$$

$$\frac{\partial}{\partial \hat{y}_i} MAE = \frac{1}{N}\begin{cases} -1 & \left(y_i - \hat{y}_i\right) > 0 \\ 1 & \left(y_i - \hat{y}_i\right) < 0 \end{cases}$$