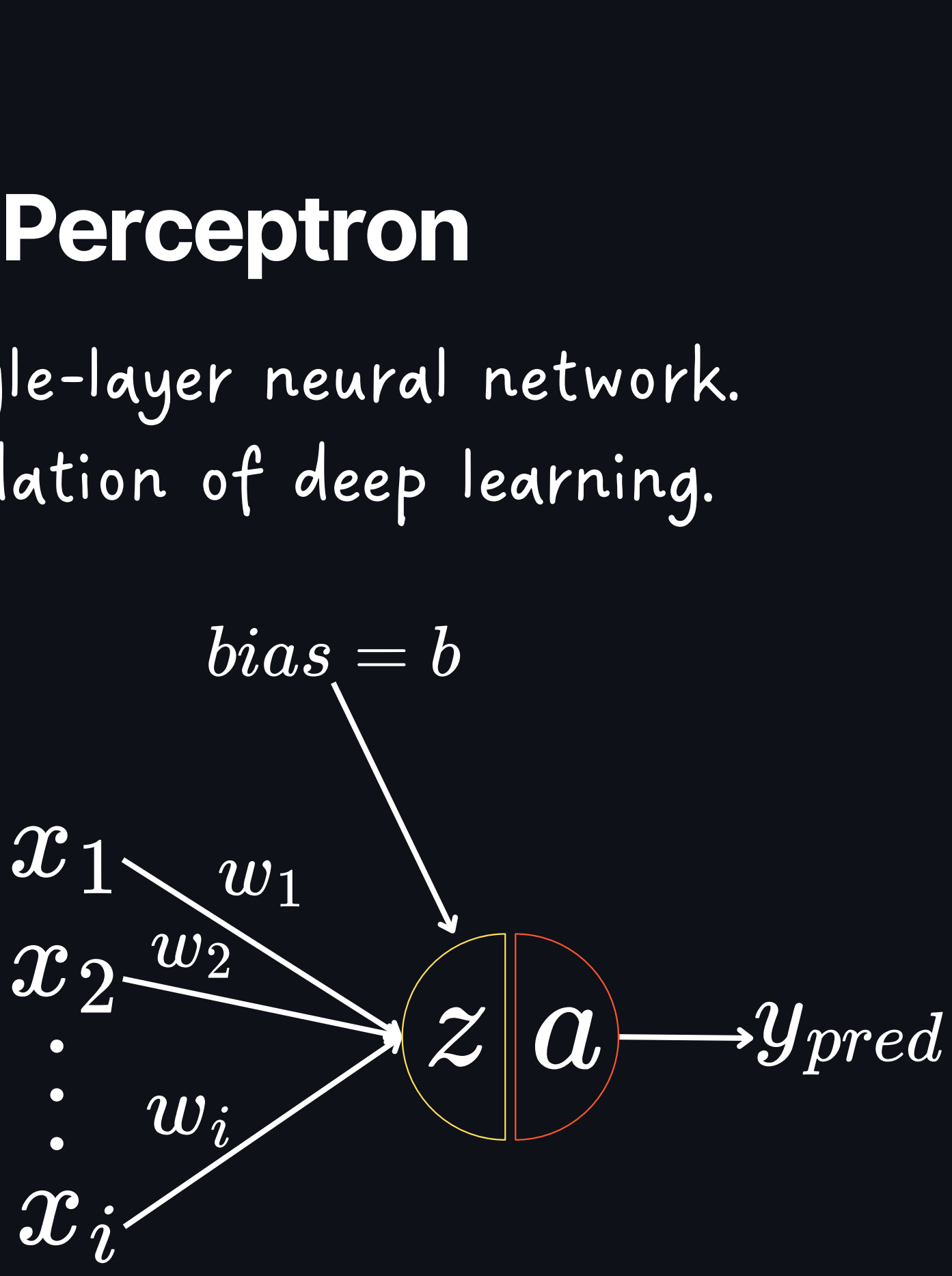
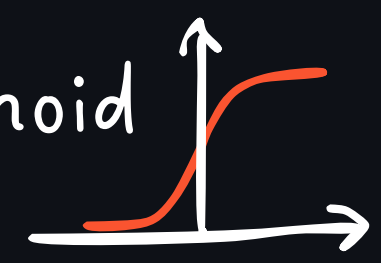
The Perceptron

A single-layer neural network.
Foundation of deep learning.



$$z = w_1x_1 + w_2x_2 + \ldots + w_ix_i + b$$

$$a(z) = \frac{1}{1 + e^{-z}}$$



To sum up: $y_{pred} = a(\vec{w} \cdot \vec{x} + b) = p_i$

Log Loss Error Function

Begin from maximizing likelihood L to minimize Log Loss.

Likelihood: $$L = \prod_{i=1}^{n} p_i^{y_i} (1 - p_i)^{(1-y_i)}$$

⋮ Some math...

$$LogLoss = -\frac{1}{N} \sum_{i=1}^{N} (y_i \log(p_i) + (1 - y_i) \log(1 - p_i))$$

Gradient Descent

Adjust the weights $\vec{w}$ and bias b to minimize the log loss through gradient descent.

Update Rule:

$$w_i = w_i - \alpha \cdot \frac{\partial L}{\partial w_i}$$

Iteration: This process is repeated for many iterations or until the cost function (Log Loss) converges to a minimum value.

Thanks for Reading

I hope this helps!