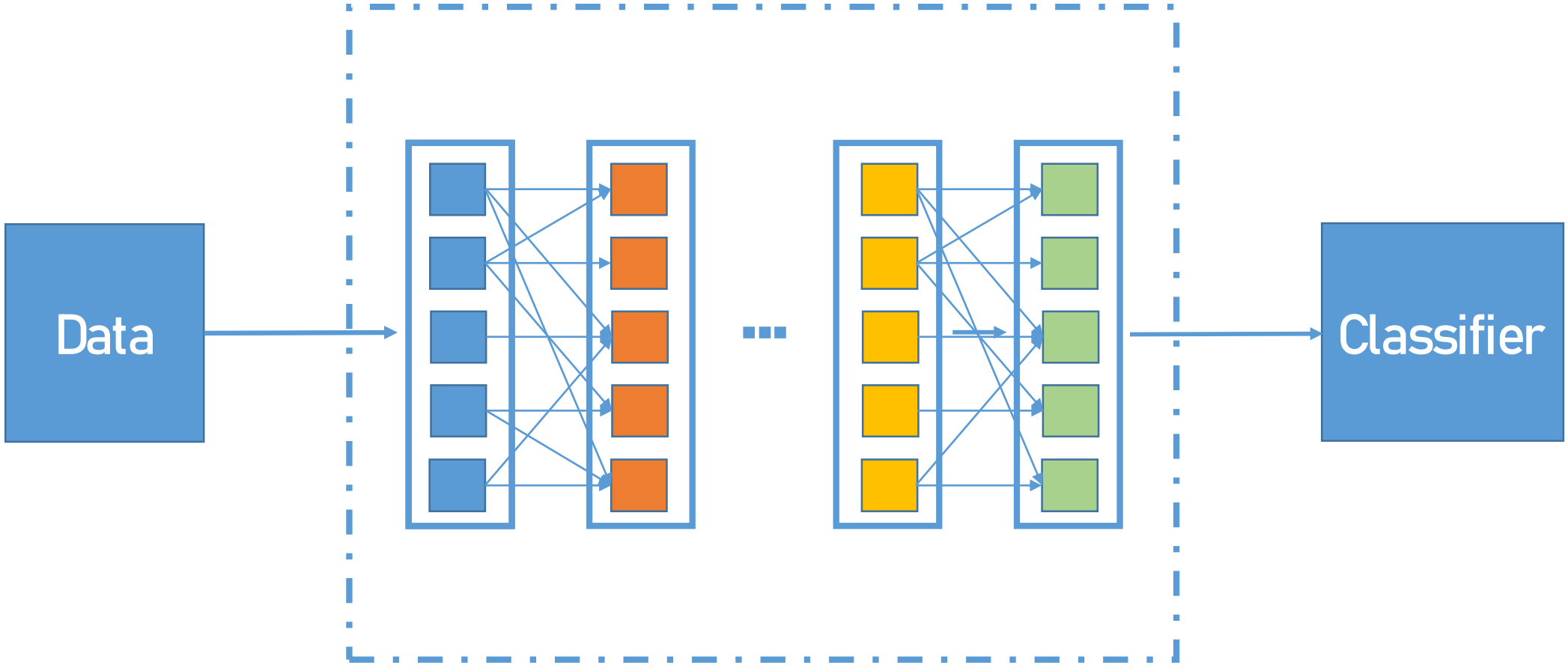
How is a neural network trained?

Problem: Classify sentiment as Positive/Negative

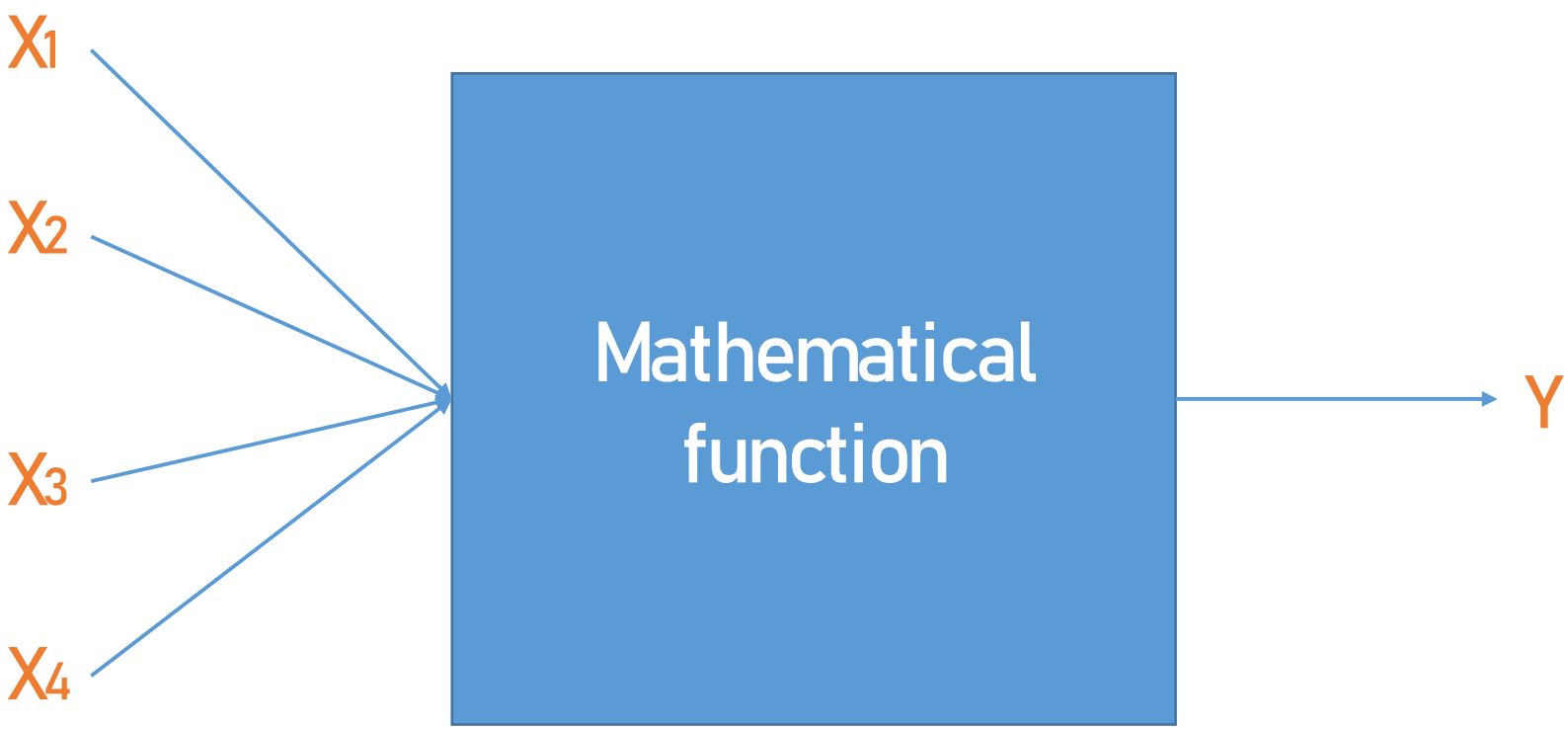
Problem: Classify sentiment as Positive/Negative

Feature 1	Feature 2	Sentiment
A1	B1	Positive
A2	B2	Negative
A3	B3	Positive
…	…	
An	Bn	Negative

Neural Networks Architecture



Understanding a Single Neuron



X_1, X_2, X_3, X_4 are features

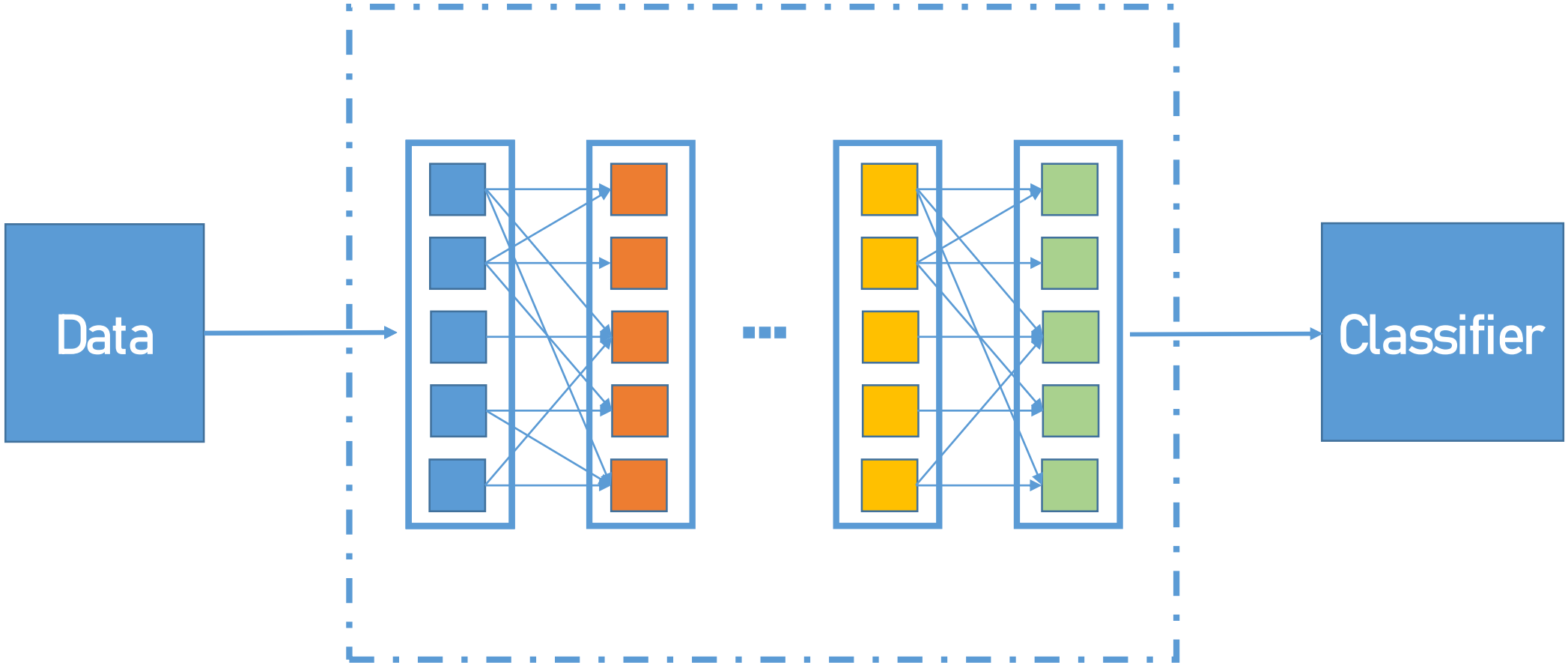
Problem: Classify sentiment as Positive/Negative

Feature 1	Feature 2	Sentiment
A1	B1	Positive
A2	B2	Negative
A3	B3	Positive
…	…	
An	Bn	Negative

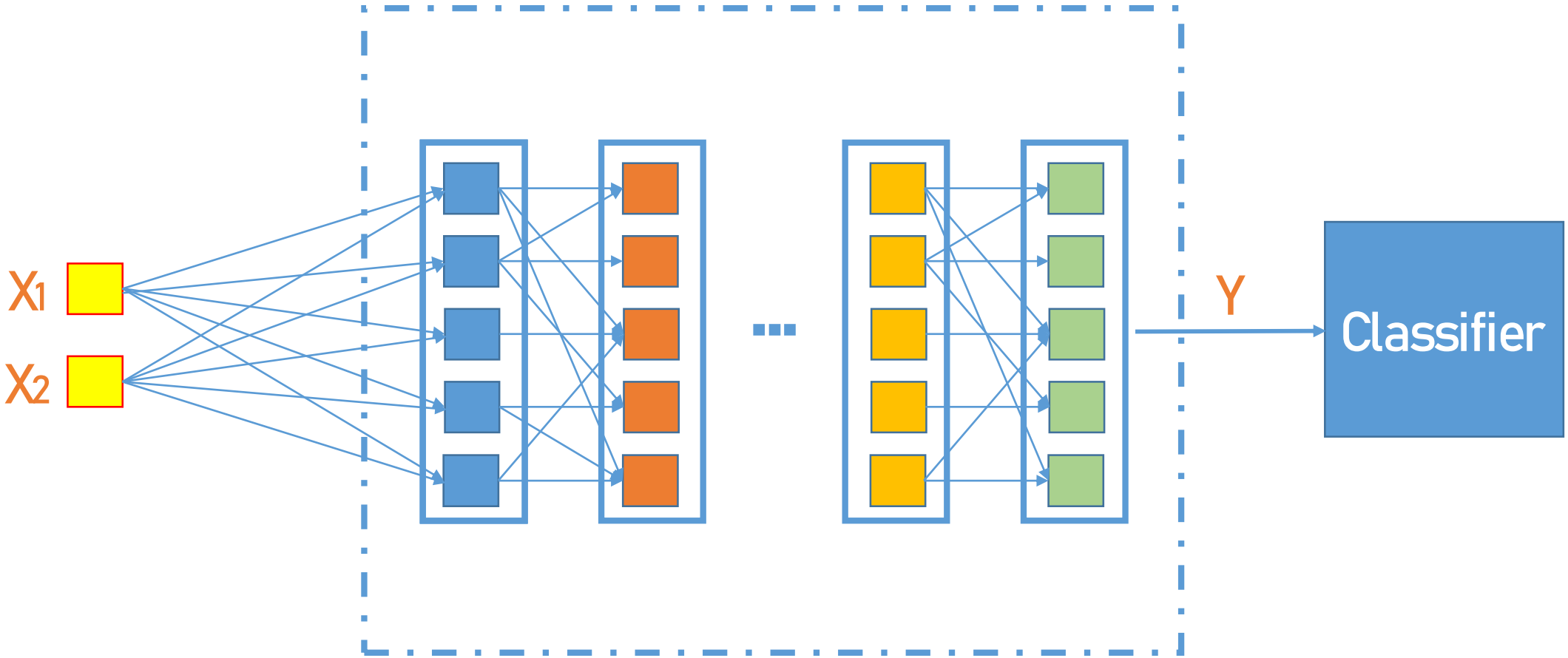
Problem: Classify sentiment as Positive/Negative

X_1	X_2	Y
A1	B1	Positive
A2	B2	Negative
A3	B3	Positive
…	…	
An	Bn	Negative

Neural Networks Architecture

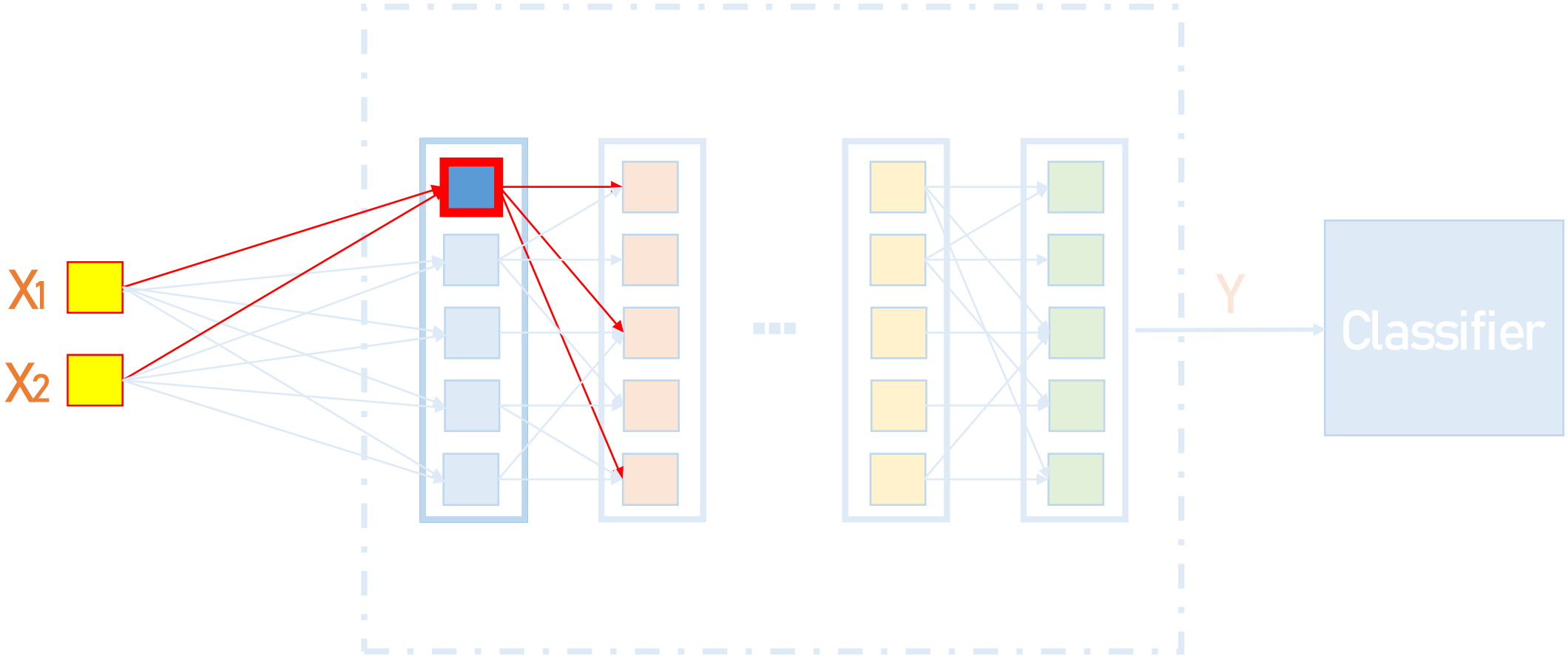


Neural Networks Architecture



All features are connected to all neurons in the first layer

Neural Networks Architecture

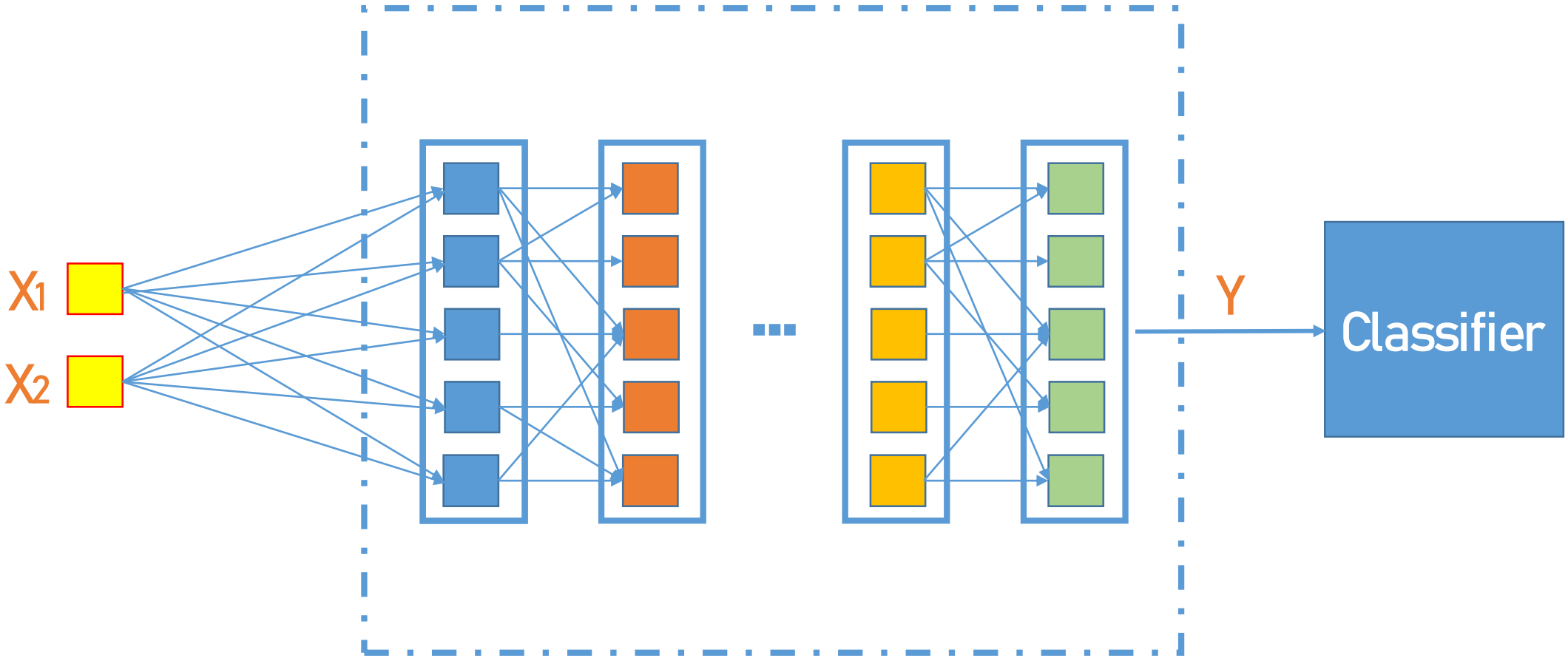


A neuron in the entire network

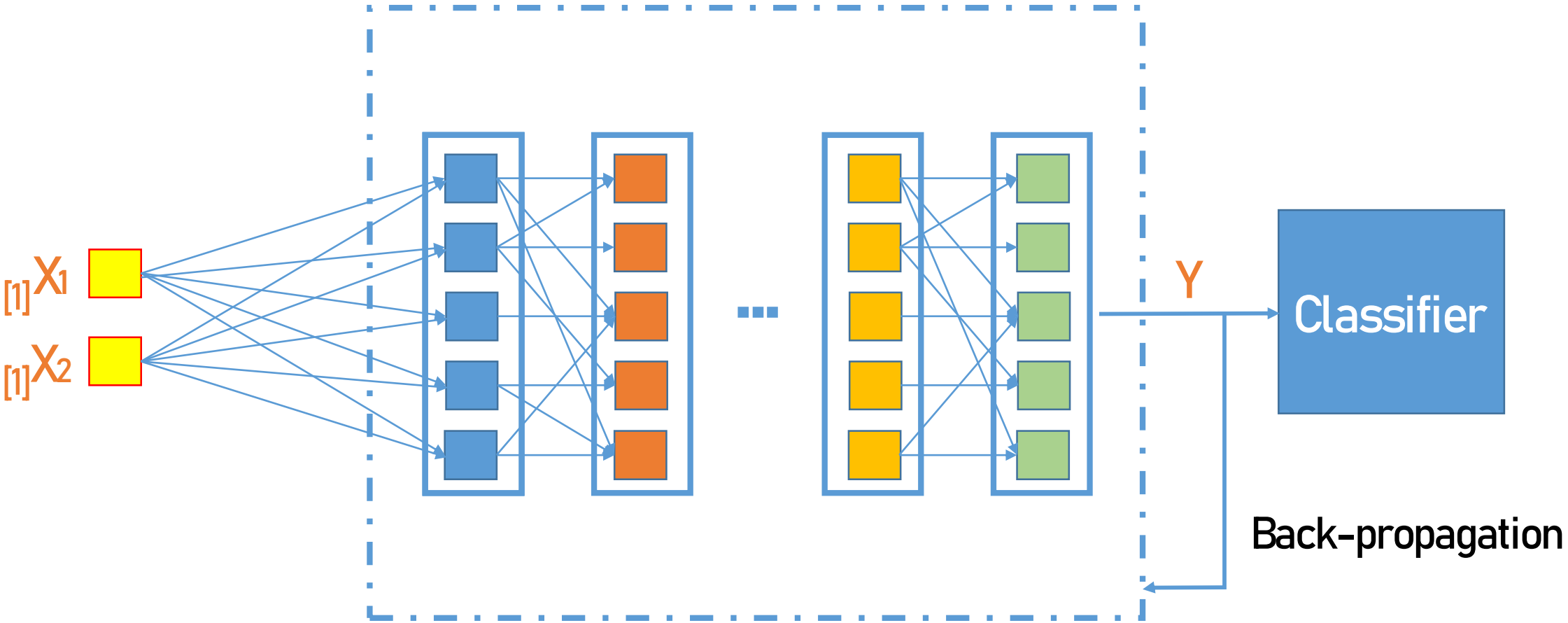
Problem: Classify sentiment as Positive/Negative

X_1	X_2	Y
A1	B1	Positive
A2	B2	Negative
A3	B3	Positive
…	…	
An	Bn	Negative

Neural Networks Architecture



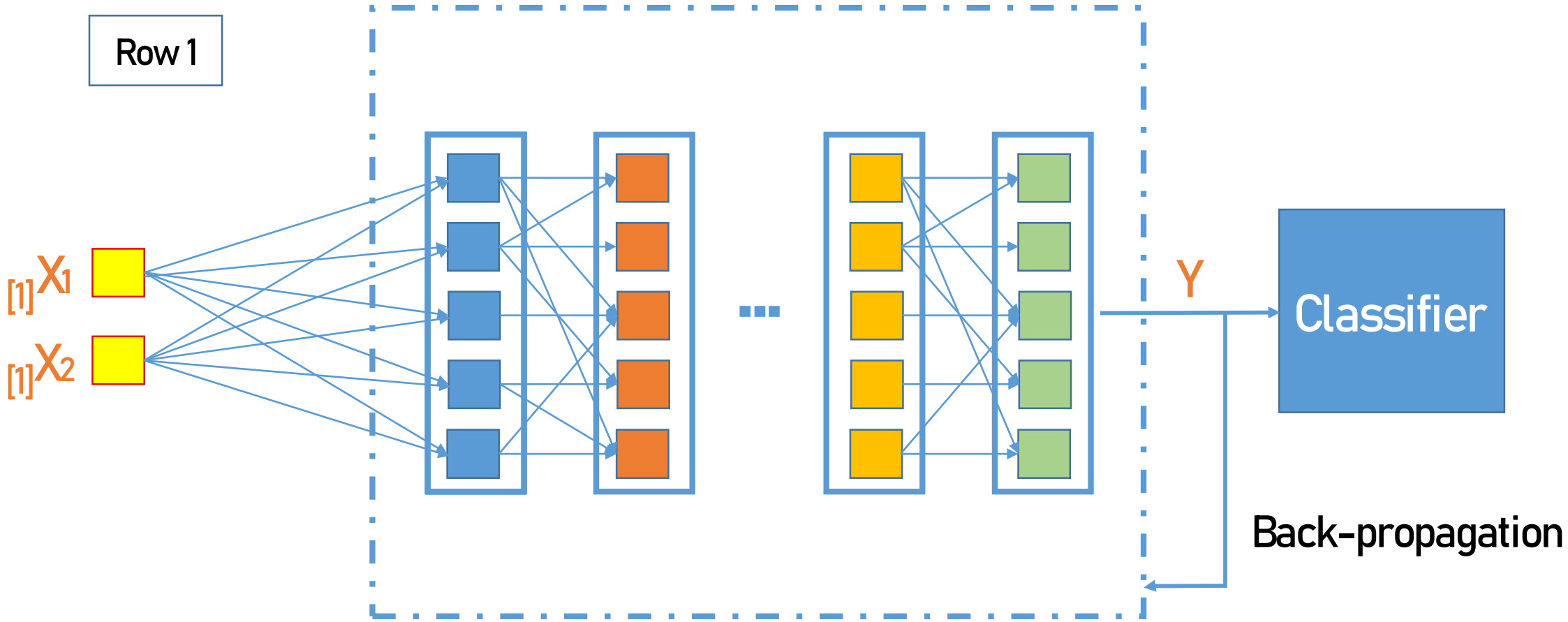
Neural Networks Architecture



Training Process

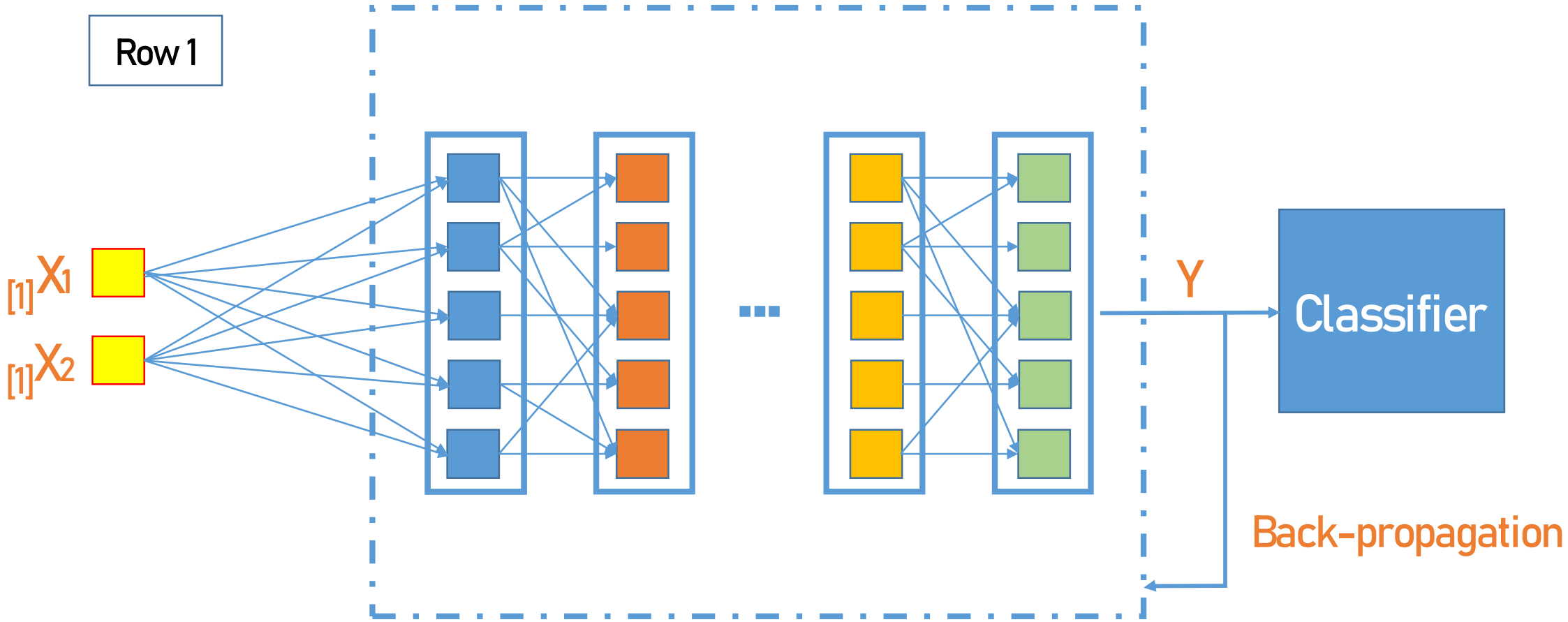
- Initialize all the weights in the network before training
- Iterate over the dataset and pass one row after another through the network
- After each row is passed, do the following:
  - Calculate the loss
  - Change the value of the weights to minimize the loss
  - The process of calculating gradients which are used to change the weights is called "back-propagation"
- Iterate through the dataset "n" number of times
- "n" is called the number of epochs

Neural Networks Architecture



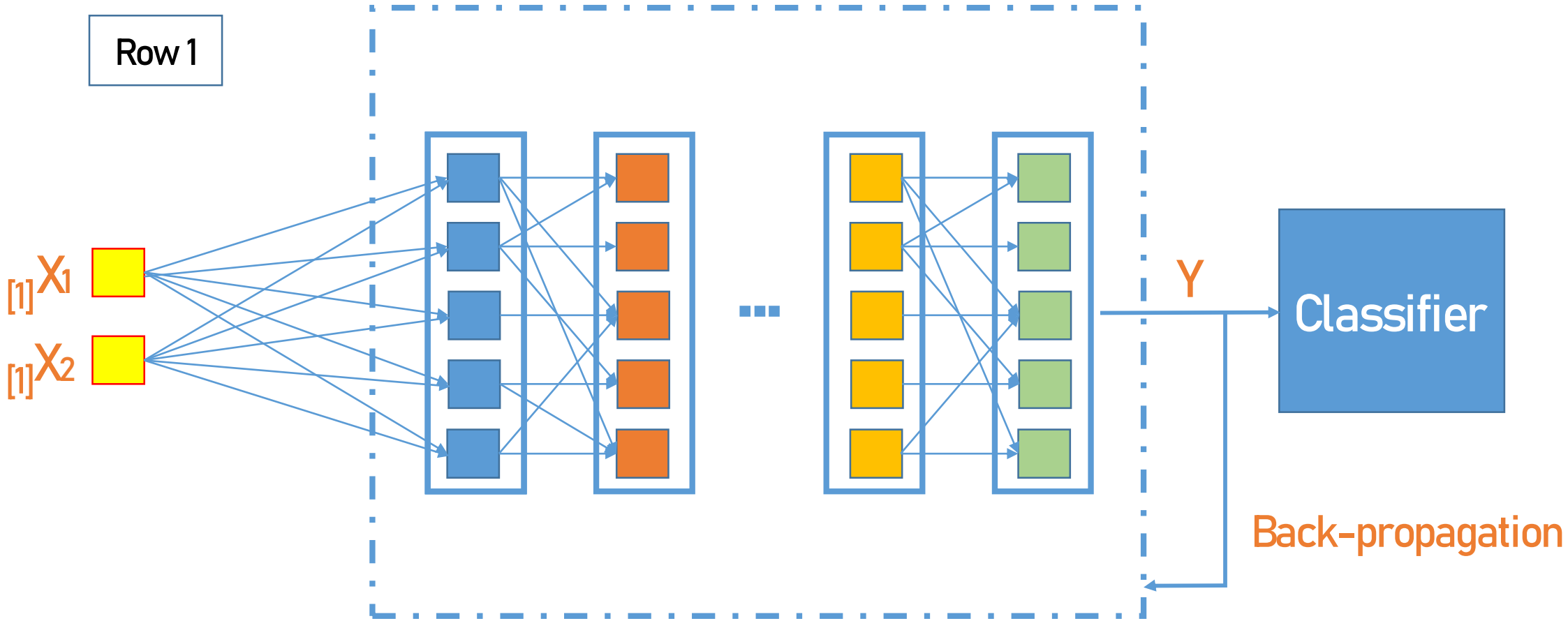
W: Weight tensor for the entire network

Neural Networks Architecture



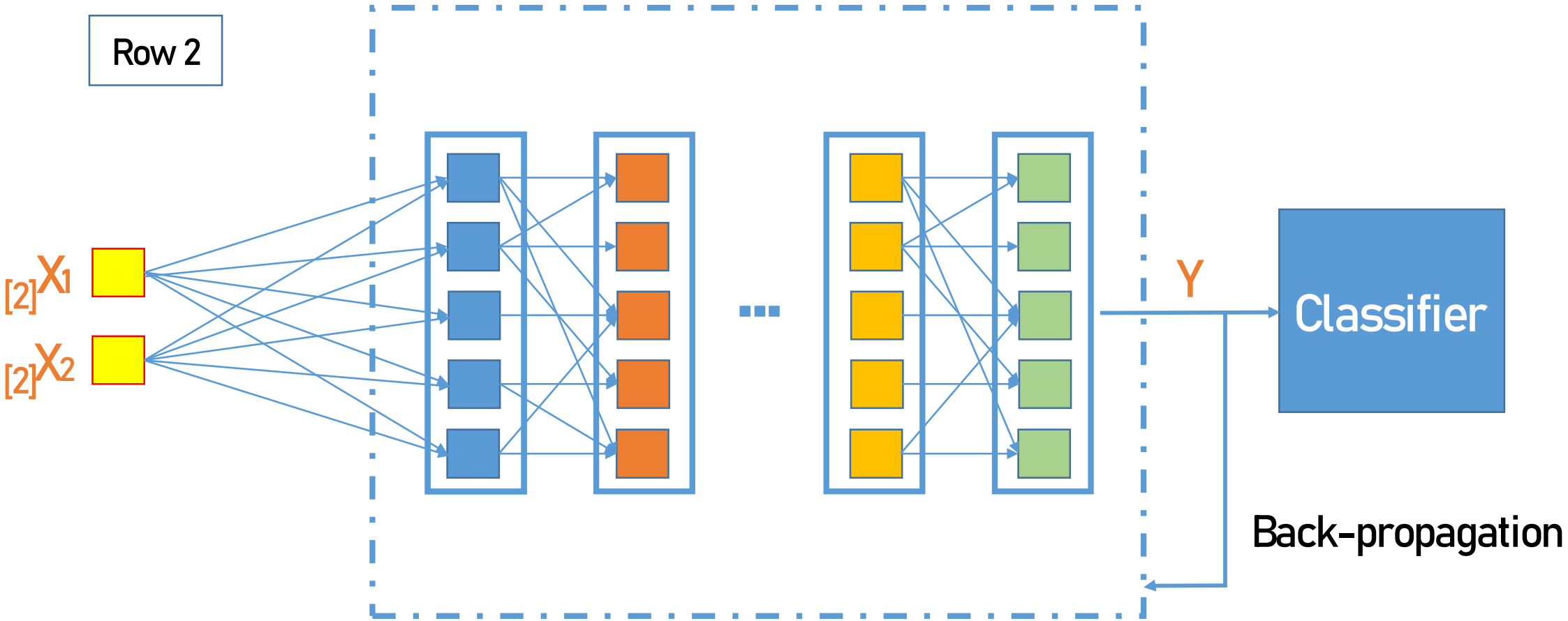
W: W [Back-propagation calculates gradients & then those are subtracted]

Neural Networks Architecture



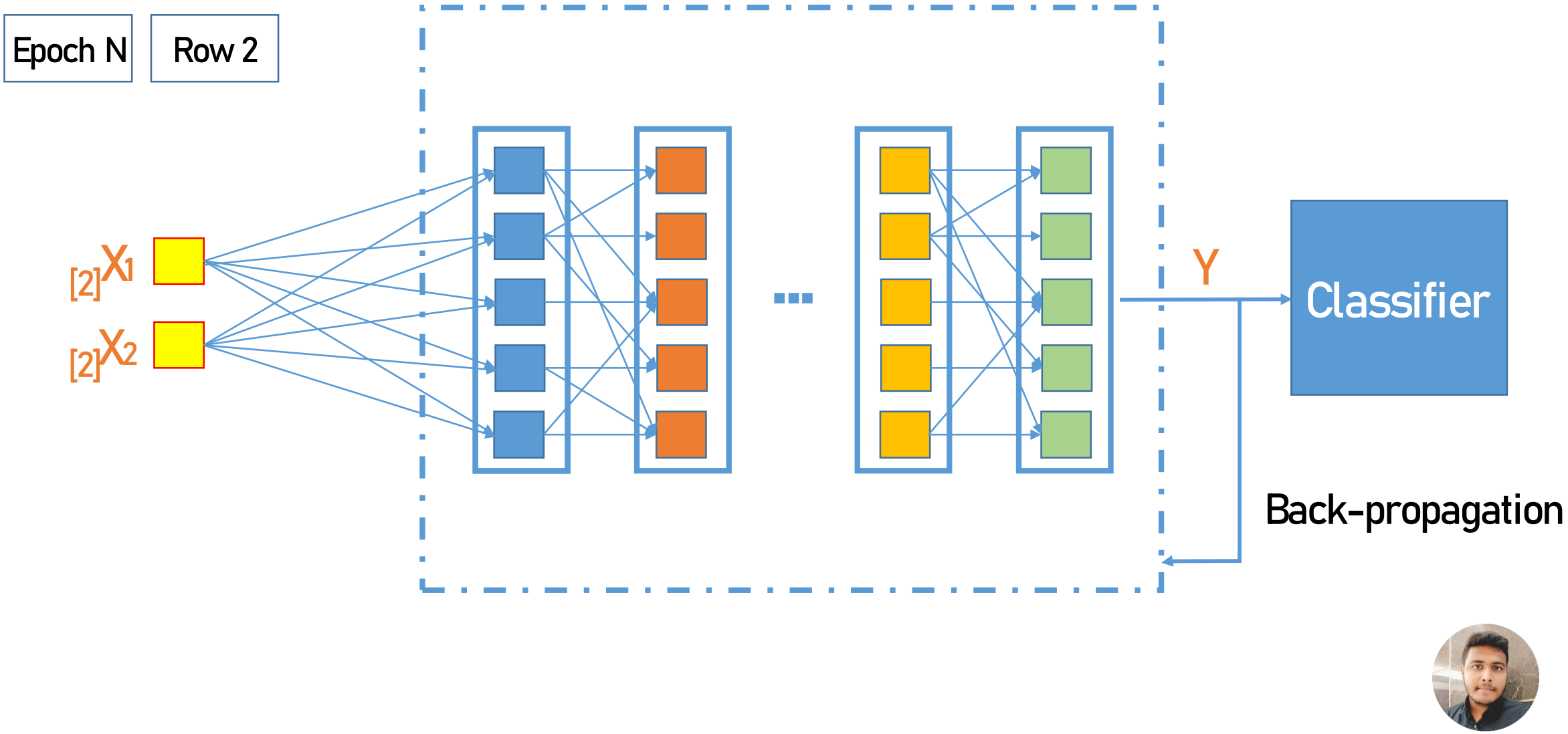
$$W = W - dW$$

Neural Networks Architecture

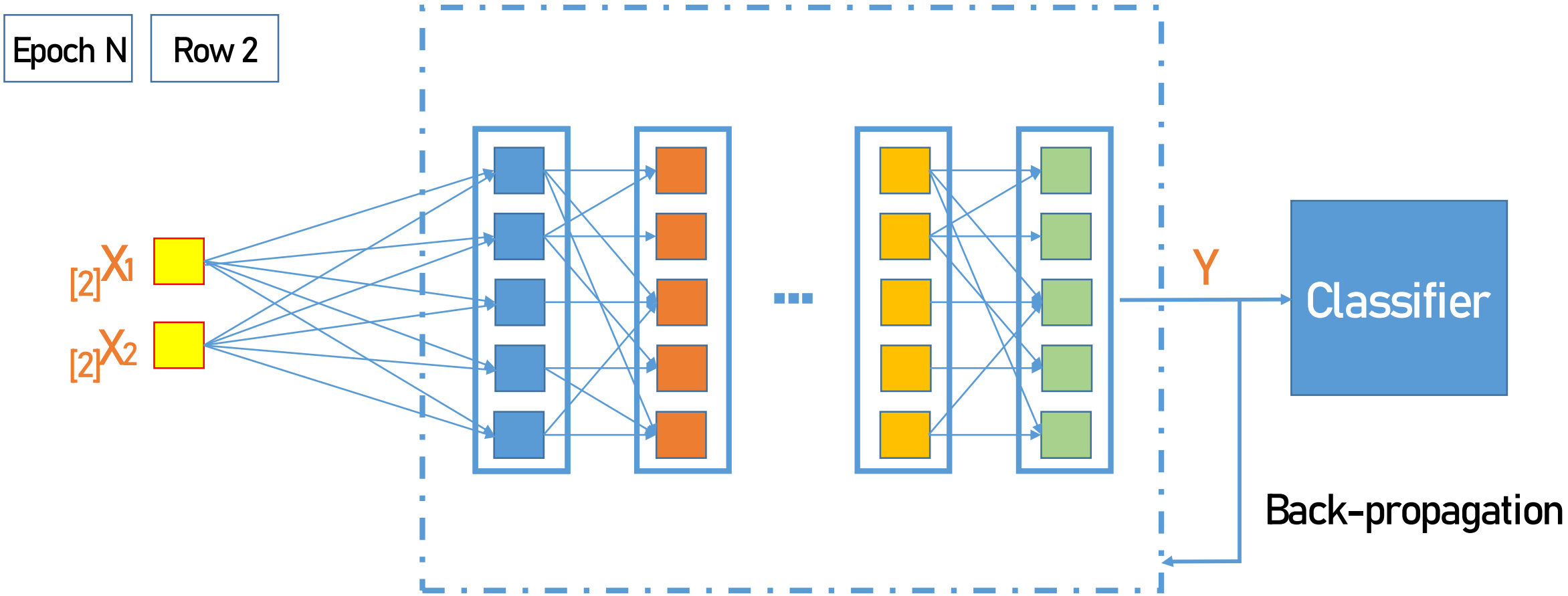


W: Weight tensor for the entire network

Neural Networks Architecture

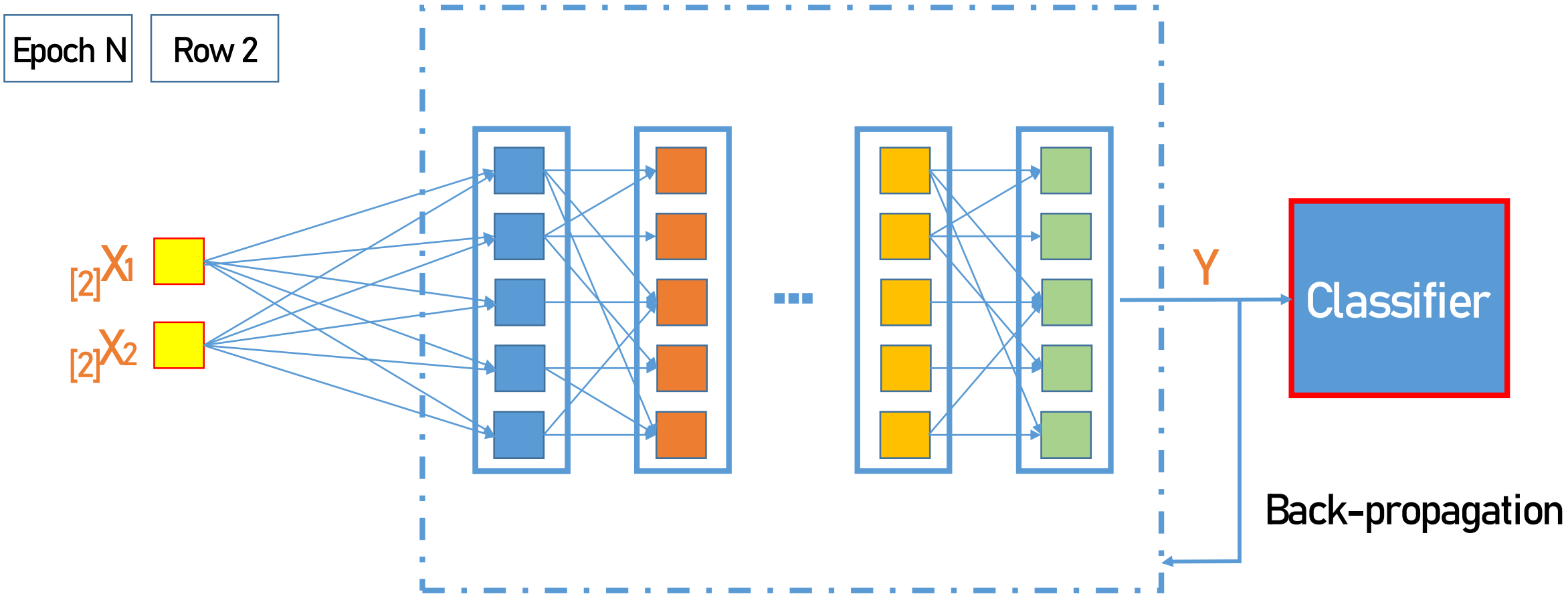


Neural Networks Architecture



Epoch is a term used to indicate the number of passes through a dataset

Neural Networks Architecture



Epoch is a term used to indicate the number of passes through a dataset