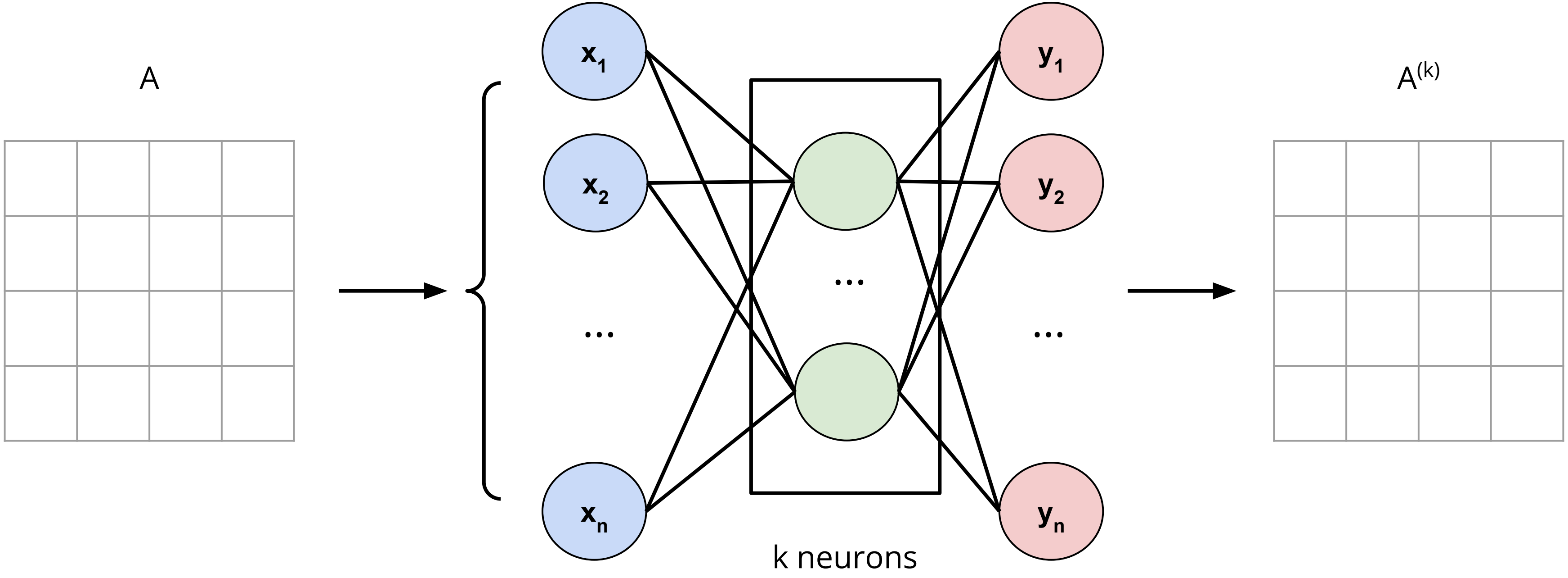
Advanced Neural Networks

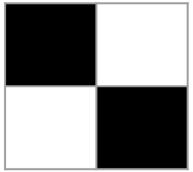
Boston University CS 506 - Lance Galletti

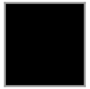
Neural Networks - Auto Encoders

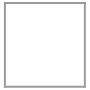


Logistic Regression Revisited

Given a 2 x 2 grid where each cell a_{ij} can take on one of two colors c_1 and c_2, find a function that can identify the following diagonal pattern:



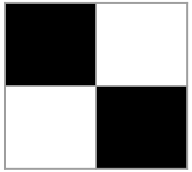
 $= c_1 = -1$

 $= c_2 = 1$

Logistic Regression Revisited

That is, find **f** such that

$$\mathbf{f}\left(\begin{array}{|c|c|}\hline a_{00} & a_{01} \\ \hline a_{10} & a_{11} \\ \hline\end{array}\right) = \begin{cases} ✔ & \text{if } \begin{array}{|c|c|}\hline a_{00} & a_{01} \\ \hline a_{10} & a_{11} \\ \hline\end{array} = \begin{array}{|c|c|}\hline c_1 & c_2 \\ \hline c_2 & c_1 \\ \hline\end{array} = \text{[pattern]} \\ ✘ & \text{otherwise} \end{cases}$$



We can define: ✔ = 1 and ✘ = 0

Logistic Regression Revisited

We can assign weights to each cell

w_1	w_2
w_3	w_4

Logistic Regression Revisited

$$\begin{bmatrix} w_1 & w_2 \\ w_3 & w_4 \end{bmatrix} \otimes \begin{bmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{bmatrix} = w_1 a_{00}$$

Logistic Regression Revisited

$$\begin{bmatrix} w_1 & w_2 \\ w_3 & w_4 \end{bmatrix} \otimes \begin{bmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{bmatrix} = w_1 a_{00} + w_2 a_{01}$$

Logistic Regression Revisited

$$\begin{bmatrix} w_1 & w_2 \\ w_3 & w_4 \end{bmatrix} \otimes \begin{bmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{bmatrix} = w_1a_{00} + w_2a_{01} + w_3a_{10}$$

Logistic Regression Revisited

$$\begin{bmatrix} w_1 & w_2 \\ w_3 & w_4 \end{bmatrix} \otimes \begin{bmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{bmatrix} = w_1 a_{00} + w_2 a_{01} + w_3 a_{10} + w_4 a_{11}$$

Logistic Regression Revisited

We can assign weights to each cell

w_1	w_2
w_3	w_4

such that:

$$\underbrace{w_1 a_{00} + w_2 a_{01} + w_3 a_{10} + w_4 a_{11}}_{\begin{bmatrix} w_1 & w_2 \\ w_3 & w_4 \end{bmatrix} \otimes \begin{bmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{bmatrix}} = b \quad \text{if diagonal pattern found}$$

Logistic Regression Revisited

For example:

$$\begin{array}{|c|c|} \hline w_1 & w_2 \\ \hline w_3 & w_4 \\ \hline \end{array} = \begin{array}{|c|c|} \hline -2 & 2 \\ \hline 2 & -2 \\ \hline \end{array}$$

What value b do we get when applied to the diagonal pattern?

$$\begin{array}{|c|c|} \hline -2 & 2 \\ \hline 2 & -2 \\ \hline \end{array} \otimes \begin{array}{|c|c|} \hline -1 & 1 \\ \hline 1 & -1 \\ \hline \end{array} = ?$$

Logistic Regression Revisited

Any other pattern will have a value lower:

$$\begin{bmatrix} -2 & 2 \\ 2 & -2 \end{bmatrix} \otimes \begin{bmatrix} -1 & -1 \\ 1 & -1 \end{bmatrix} = ?$$

Logistic Regression Revisited

Equivalently we can decide to move the value b to the left of the equation in order for the weighted sum to reveal a diagonal pattern at 0:

$$w_1 a_{00} + w_2 a_{01} + w_3 a_{10} + w_4 a_{11} + b = 0 \quad \text{if diagonal pattern found}$$

Logistic Regression Revisited

We could then find a function $\boldsymbol{\sigma}$ to apply to the result of this sum in order to get probabilities of being diagonal:

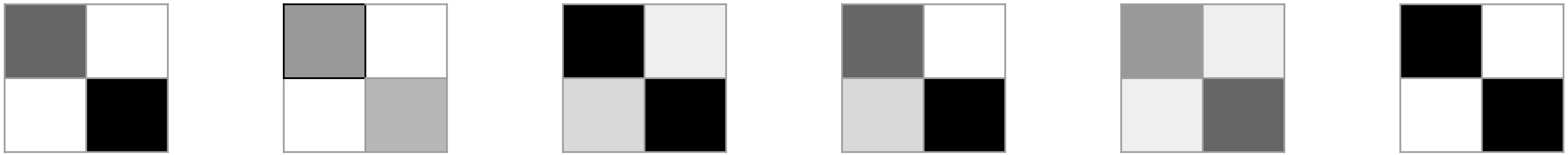
$$\boldsymbol{\sigma}(w_1 a_{00} + w_2 a_{01} + w_3 a_{10} + w_4 a_{11} + b) > \tfrac{1}{2} \text{ if } w_1 a_{00} + w_2 a_{01} + w_3 a_{10} + w_4 a_{11} + b > 0$$

Logistic Regression Revisited

$$\sigma(x) = \frac{1}{1 + e^{-x}}$$

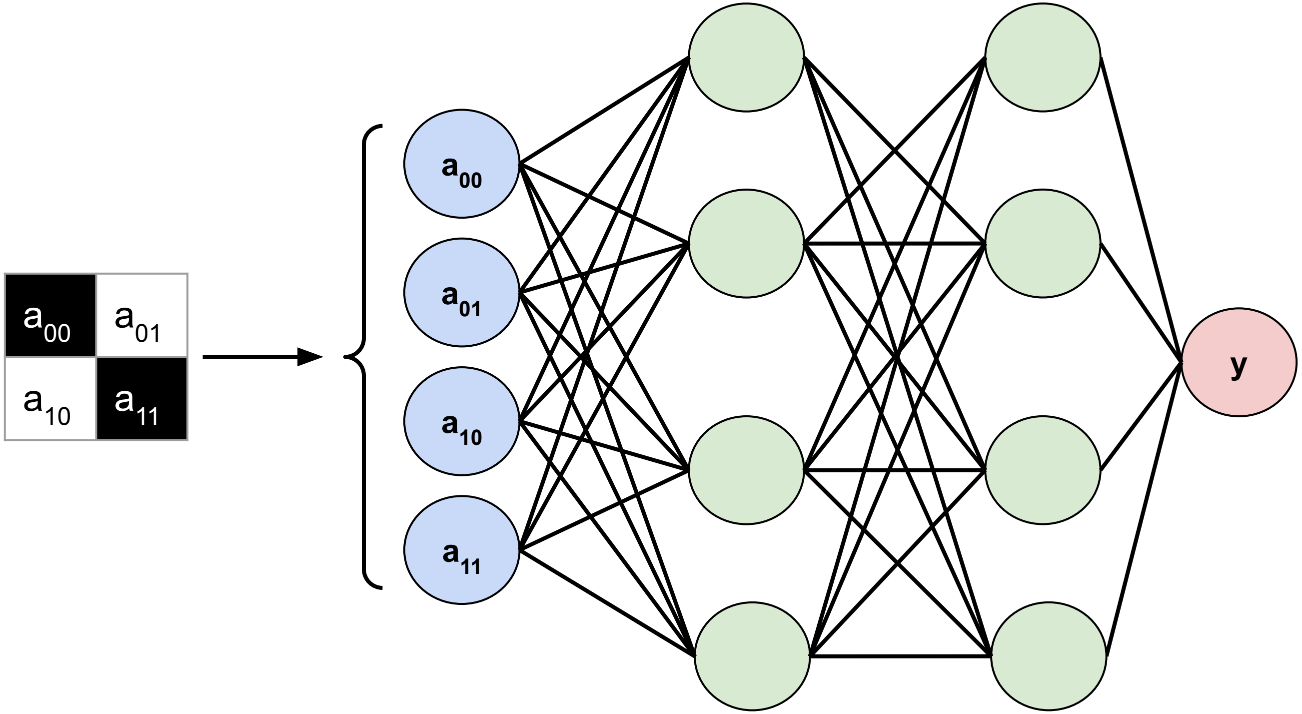
When **σ** is the logit^{-1} (also called sigmoid) function, this is Logistic Regression.

So for each cell we're looking to learn a weight w_i that makes **σ** larger for diagonal patterns. The bias term b lets us account for systemic dimming or brightening of cells (i.e. when the data is not normalized).

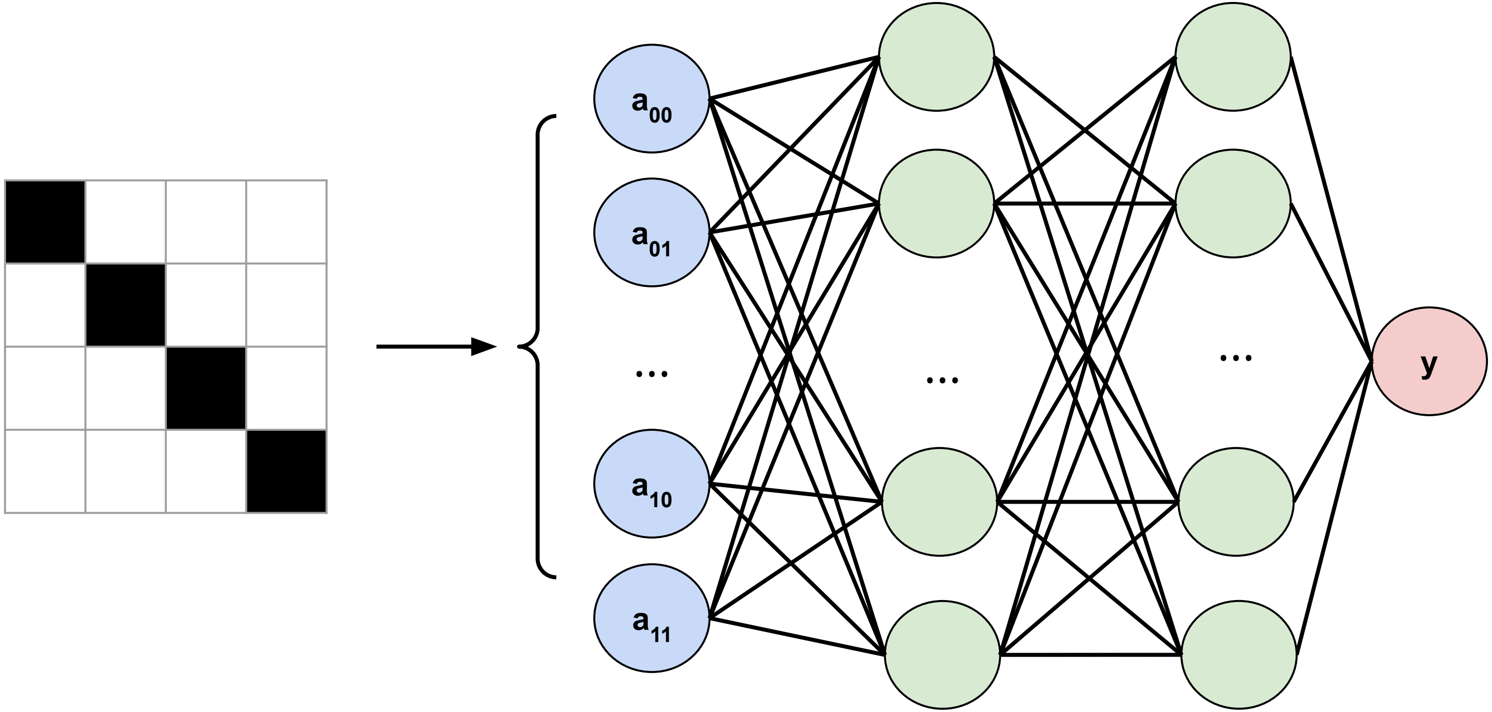


.92 .81 .95 .73 .68 .99

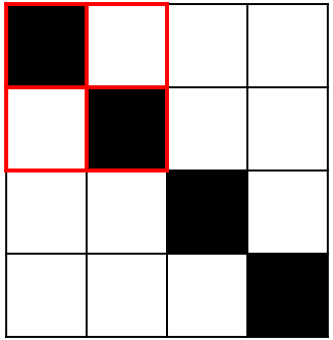
Neural Networks - Convolutional Neural Networks



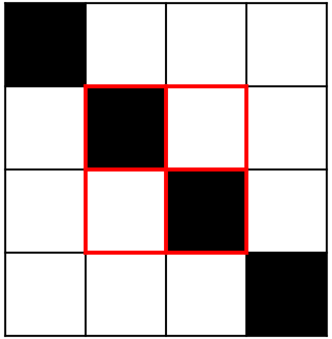
Neural Networks - Convolutional Neural Networks



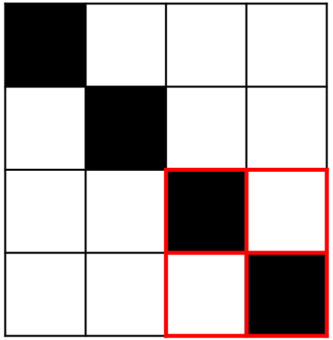
Neural Networks - Convolutional Neural Networks



Neural Networks - Convolutional Neural Networks



Neural Networks - Convolutional Neural Networks



Neural Networks - Convolutional Neural Networks

Recall: Our network learns weights for each cell

w_1	w_2
w_3	w_4

Neural Networks - Convolutional Neural Networks

$$\begin{bmatrix} w_1 & w_2 \\ w_3 & w_4 \end{bmatrix} \otimes \begin{bmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{bmatrix} = w_1 a_{00}$$

Neural Networks - Convolutional Neural Networks

$$\begin{array}{|c|c|} \hline w_1 & w_2 \\ \hline w_3 & w_4 \\ \hline \end{array} \otimes \begin{array}{|c|c|} \hline a_{00} & a_{01} \\ \hline a_{10} & a_{11} \\ \hline \end{array} = w_1 a_{00} + w_2 a_{01}$$

Neural Networks - Convolutional Neural Networks

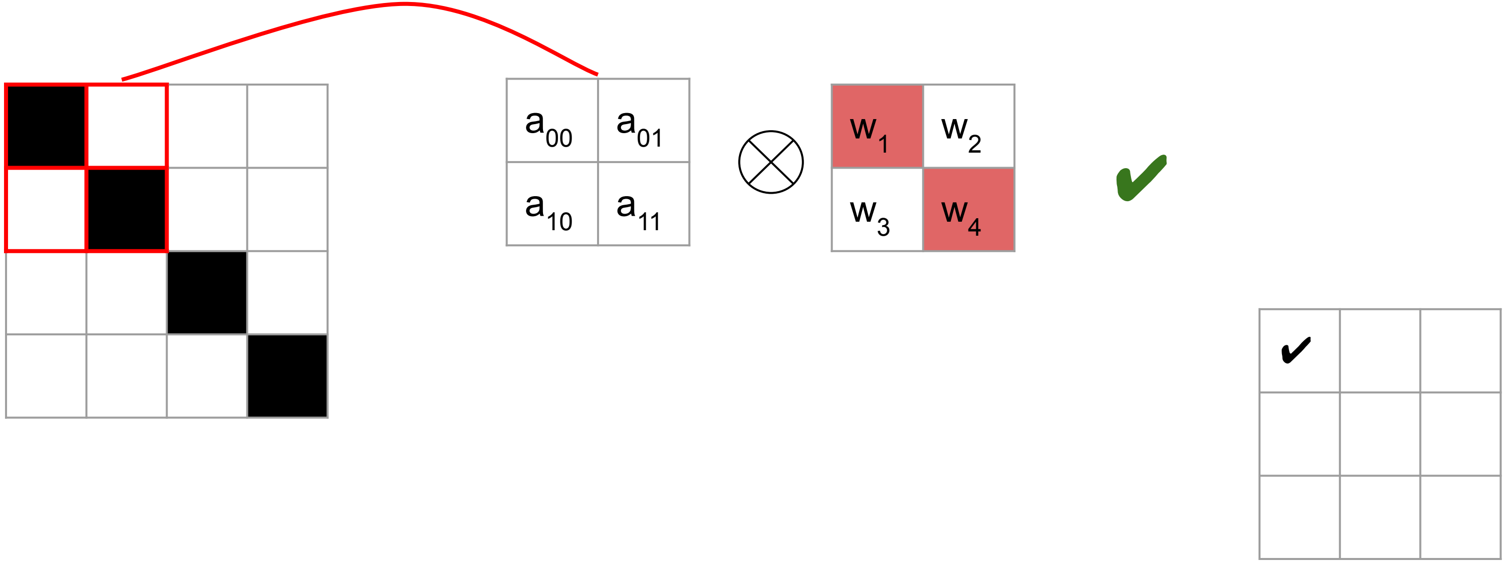
$$\begin{bmatrix} w_1 & w_2 \\ w_3 & w_4 \end{bmatrix} \otimes \begin{bmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{bmatrix} = w_1 a_{00} + w_2 a_{01} + w_3 a_{10}$$

Neural Networks - Convolutional Neural Networks

$$\begin{bmatrix} w_1 & w_2 \\ w_3 & w_4 \end{bmatrix} \otimes \begin{bmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{bmatrix} = w_1a_{00} + w_2a_{01} + w_3a_{10} + w_4a_{11}$$

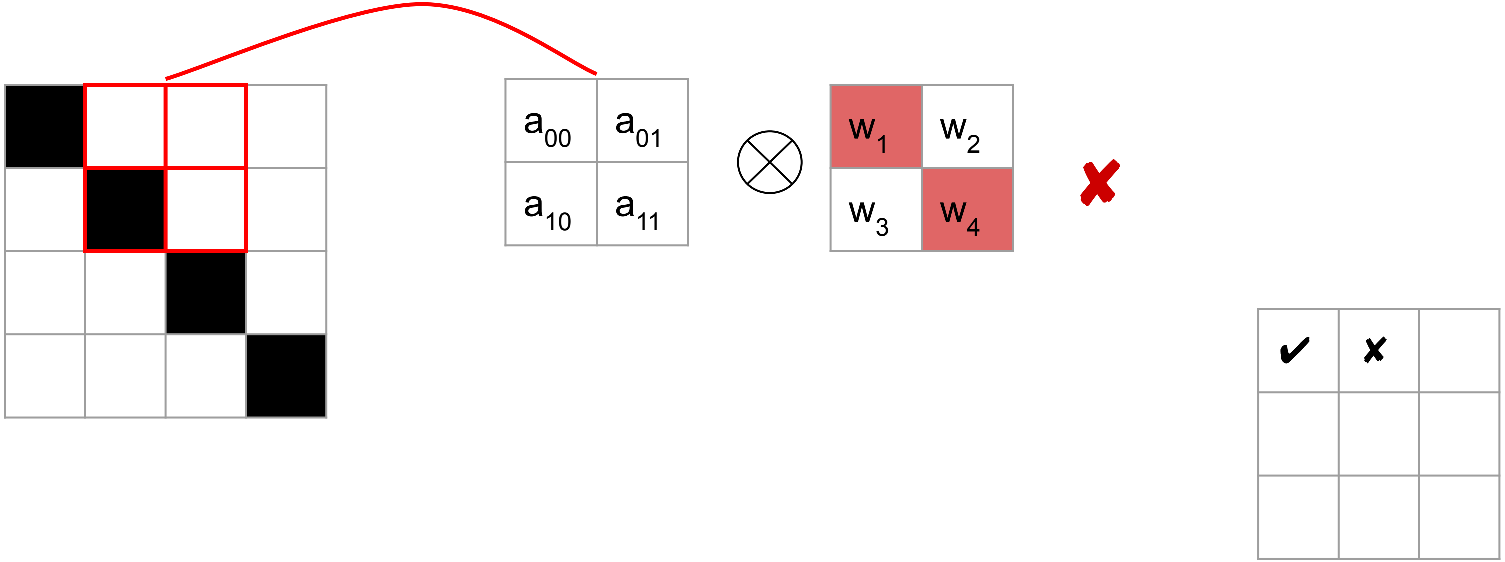
Neural Networks - Convolutional Neural Networks

Knowing this, what happens if we slide this filter across the larger diagonal?



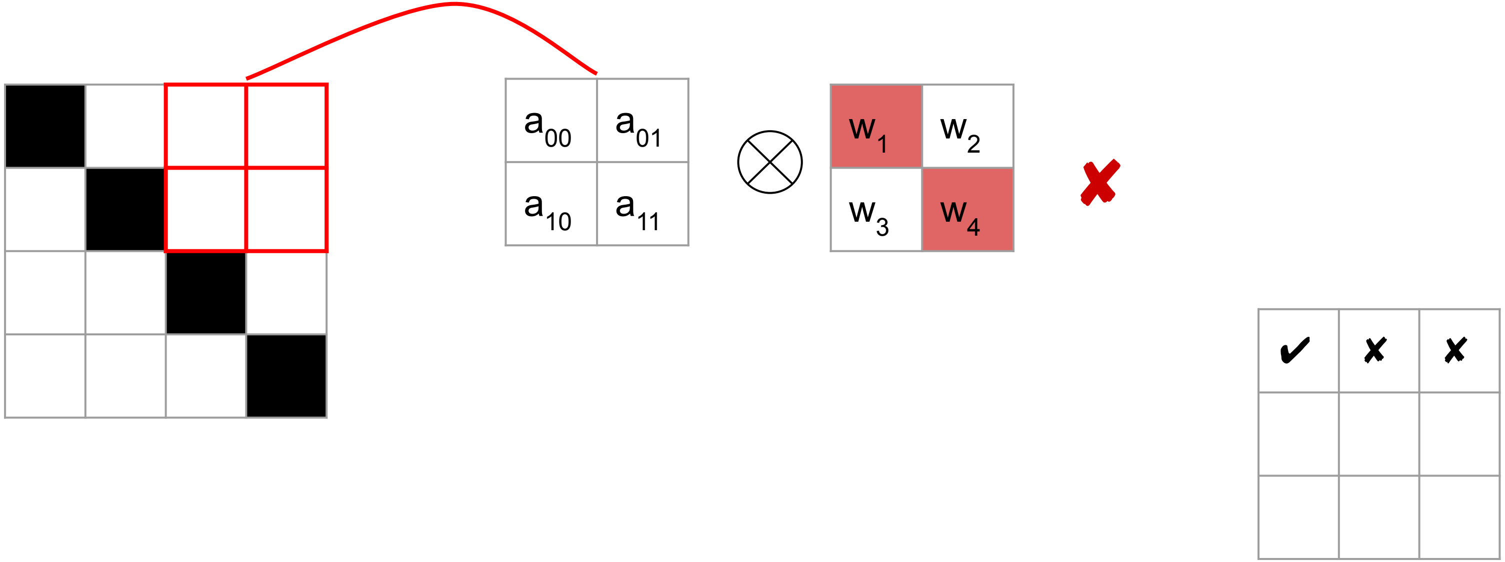
Neural Networks - Convolutional Neural Networks

Knowing this, what happens if we slide this filter across the larger diagonal?



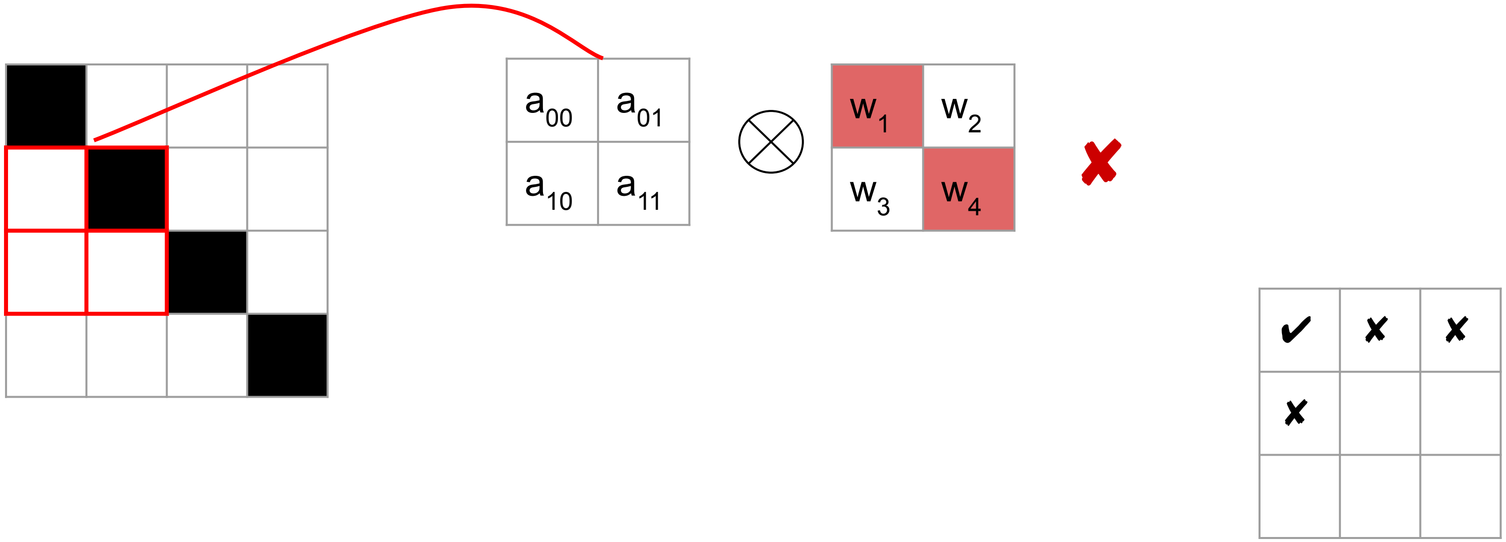
Neural Networks - Convolutional Neural Networks

Knowing this, what happens if we slide this filter across the larger diagonal?



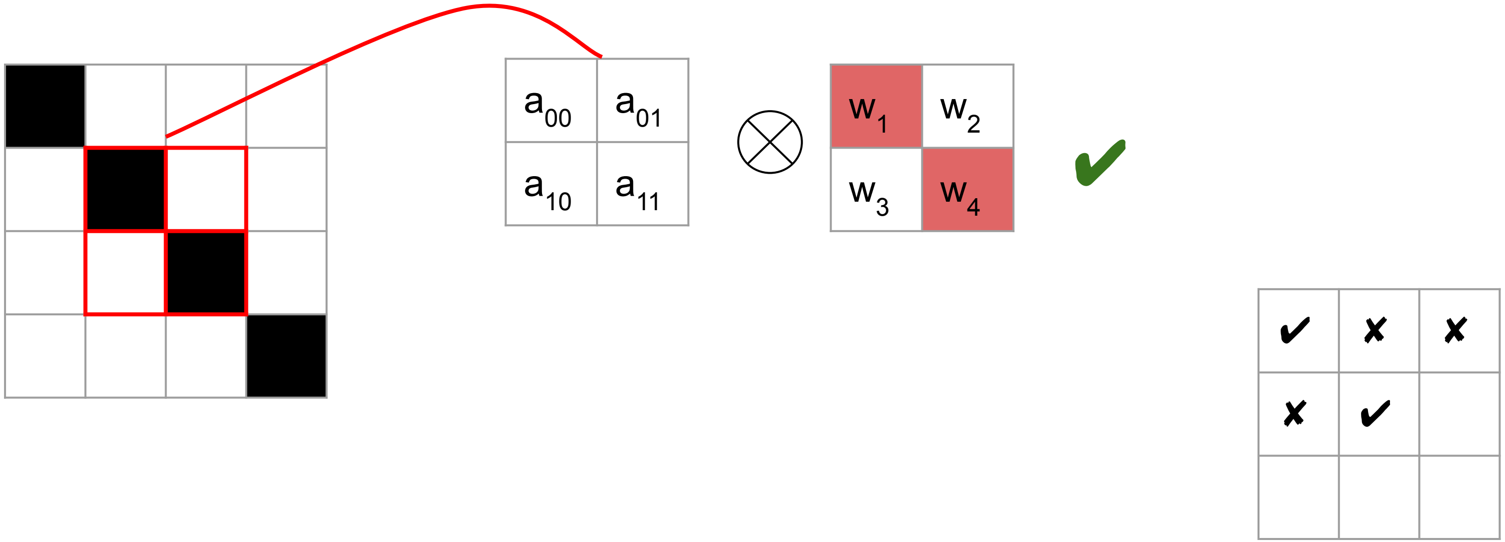
Neural Networks - Convolutional Neural Networks

Knowing this, what happens if we slide this filter across the larger diagonal?



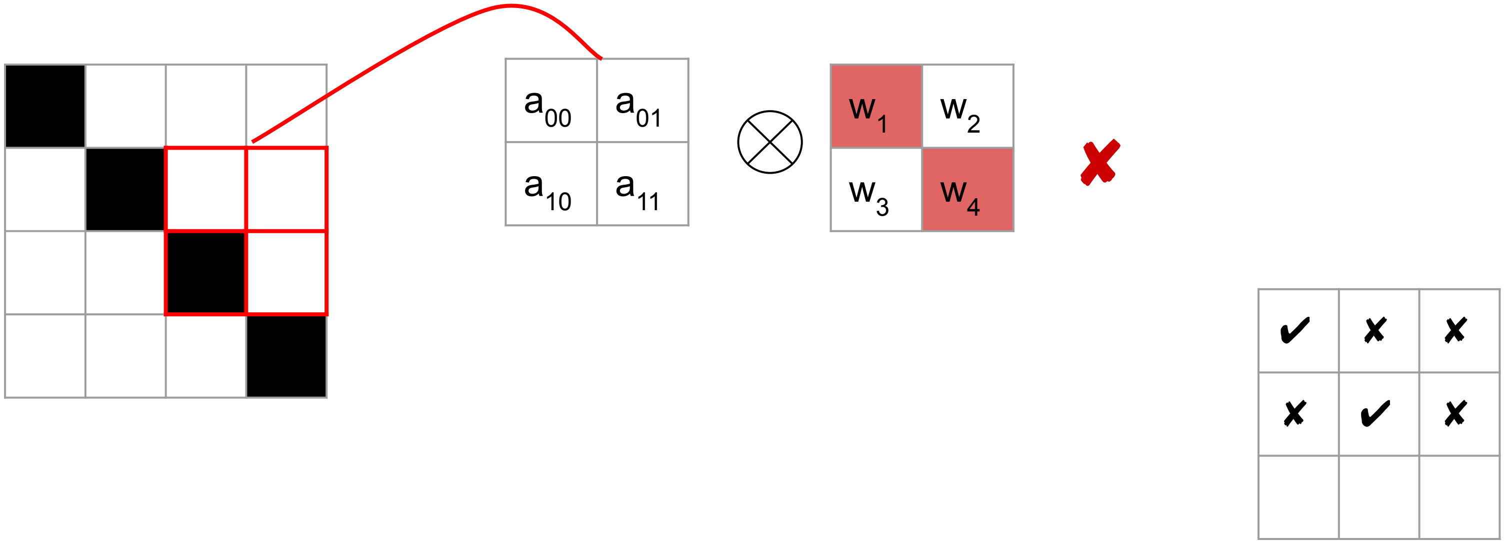
Neural Networks - Convolutional Neural Networks

Knowing this, what happens if we slide this filter across the larger diagonal?



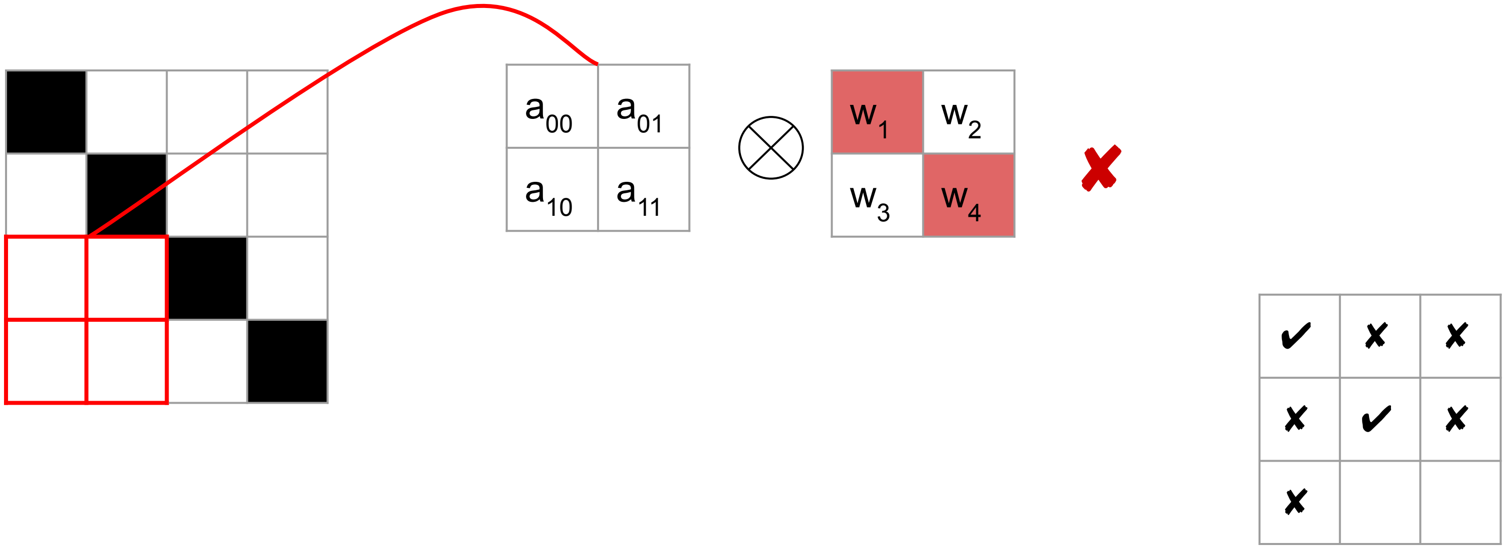
Neural Networks - Convolutional Neural Networks

Knowing this, what happens if we slide this filter across the larger diagonal?



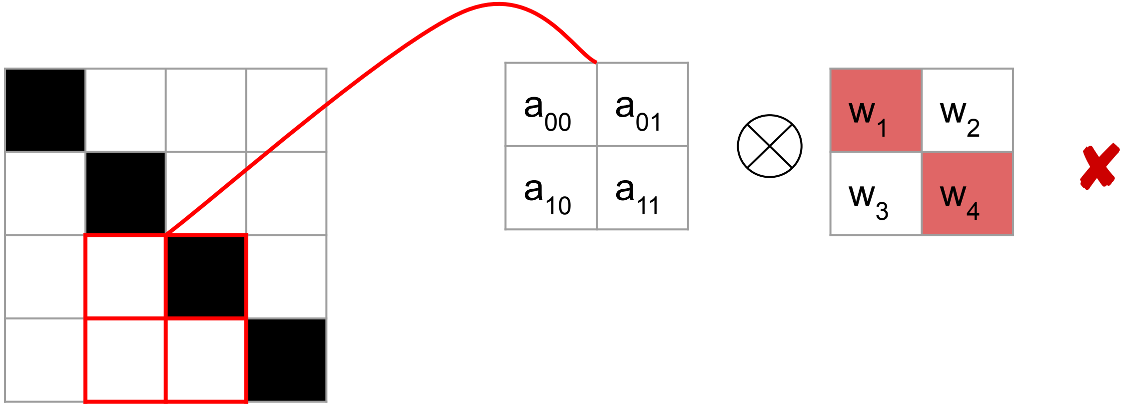
Neural Networks - Convolutional Neural Networks

Knowing this, what happens if we slide this filter across the larger diagonal?



Neural Networks - Convolutional Neural Networks

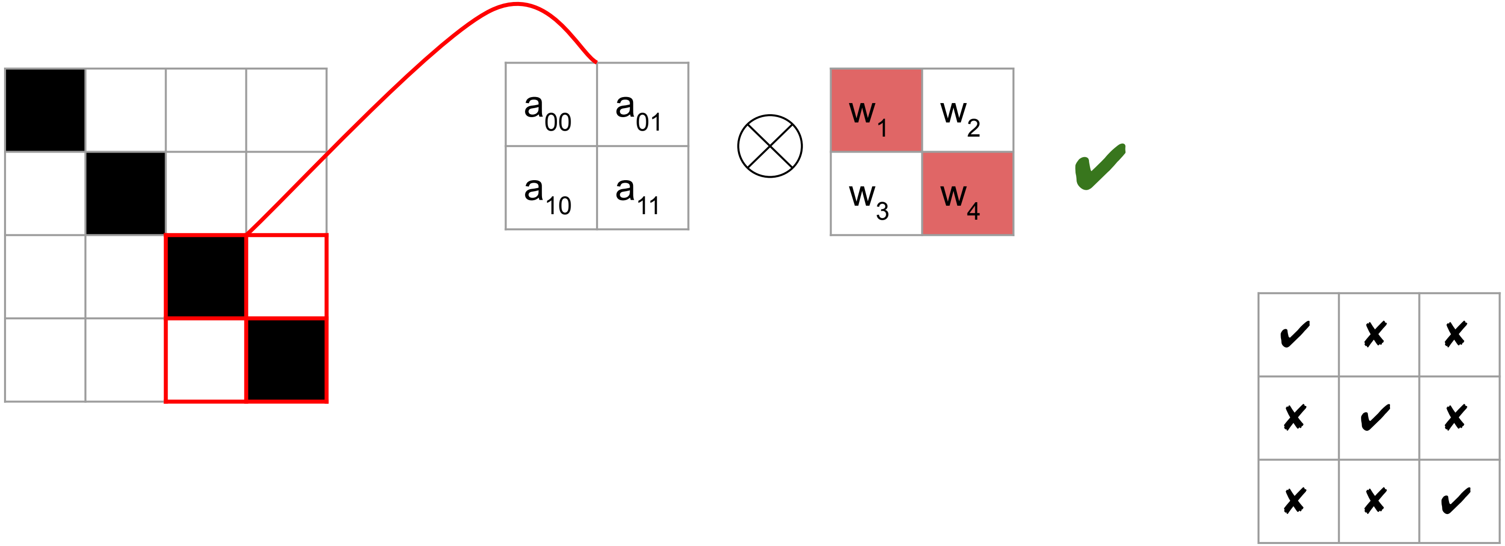
Knowing this, what happens if we slide this filter across the larger diagonal?



✔	✘	✘
✘	✔	✘
✘	✘	

Neural Networks - Convolutional Neural Networks

Knowing this, what happens if we slide this filter across the larger diagonal?



Neural Networks - Convolutional Neural Networks

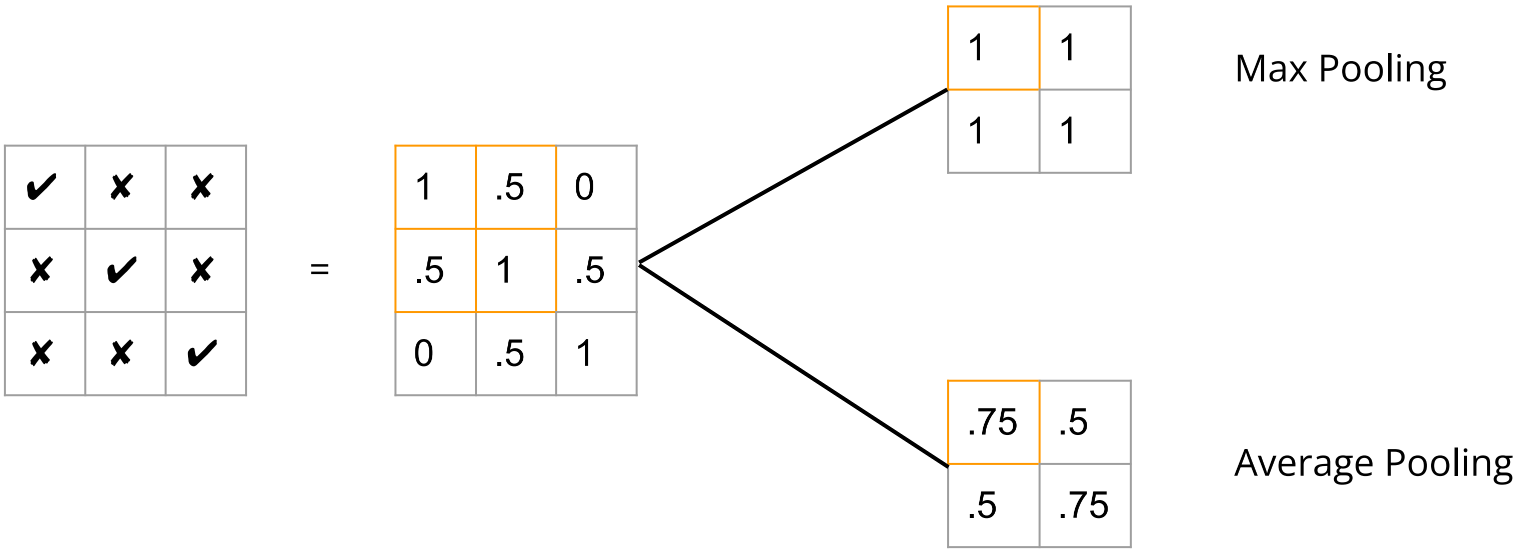
Creating such a filter allows us to:

1. Reduce the number of weights
2. Capture features all over the image

The process of applying a filter (or kernel) is called a convolution

Neural Networks - Convolutional Neural Networks

To reduce the weights even further, another phase is done after convolution called Pooling:



Neural Networks - Convolutional Neural Networks

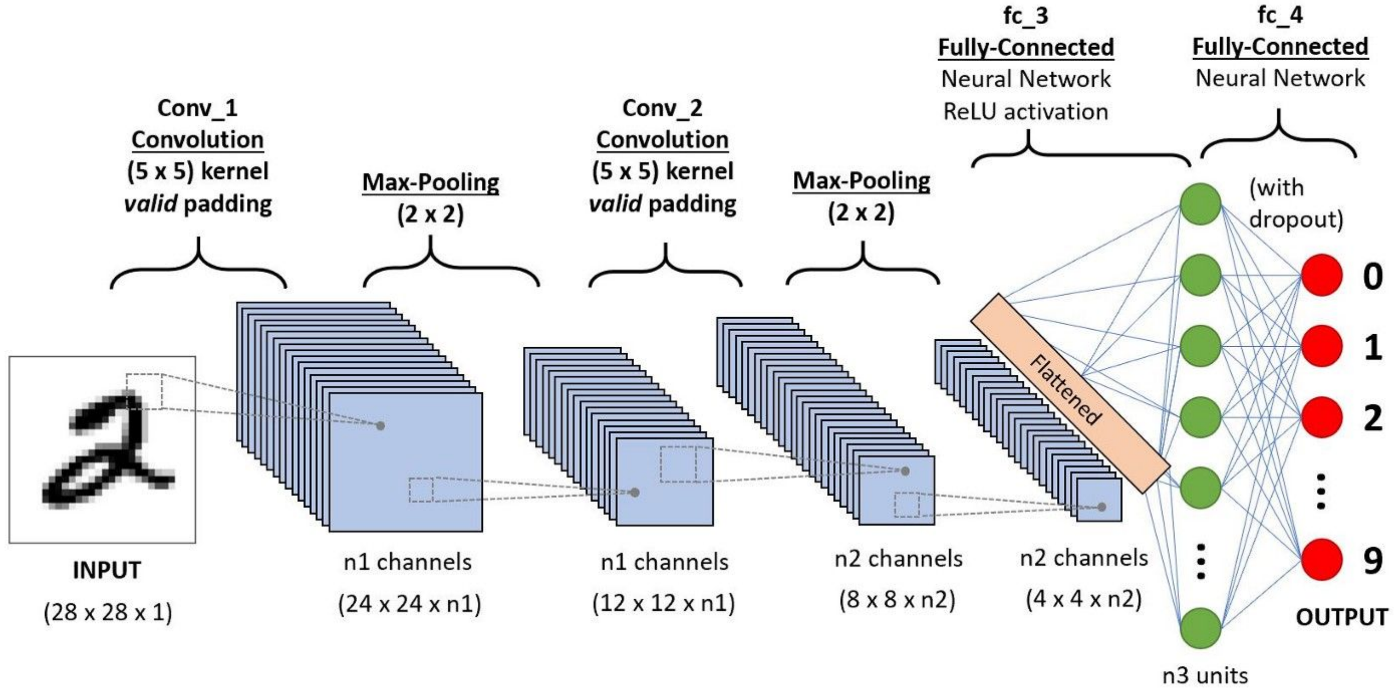


Image from https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53

Neural Networks - Convolutional Neural Networks

Main application: Computer vision

Recurrent Neural Networks

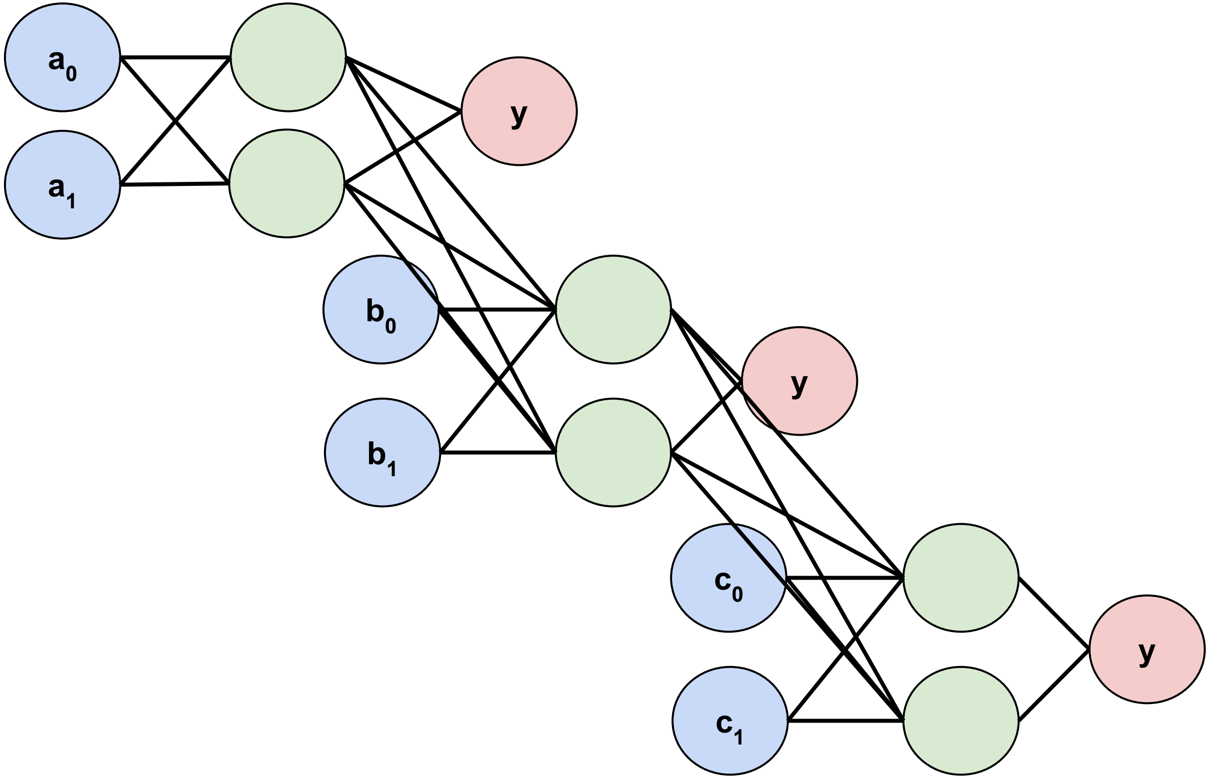
Handling sequences of input.

Intuition: What a word is / might be in a sentence is easier to figure out if you know the words around it.

Applications:

1. Predicting the next word
2. Translation
3. Speech Recognition
4. Video Tagging

Recurrent Neural Networks



Intro to Neural Networks

https://medium.com/@gallettilance/list/introducing-neural-networks-d74f0dc25400