Introduction to Deep Learning

Goals for this week:

(1)Neural networks (a little linear algebra)
(2) Perceptron, activation function
(3)How is deep learning implemented? Gradient descent algorithm
(4) Implementation of deep learning in Python

Later this course:

Convolutional neural networks, recurrent neural networks, application of deep learning to image analysis, text analysis, and other areas.

- What is “learning”?
  - Improving performance through experience
  - Getting a computer to do well on a task without manually building in competence
- What is “deep”?
  - Learning using multi-layer neural networks

Visual data in business: some examples

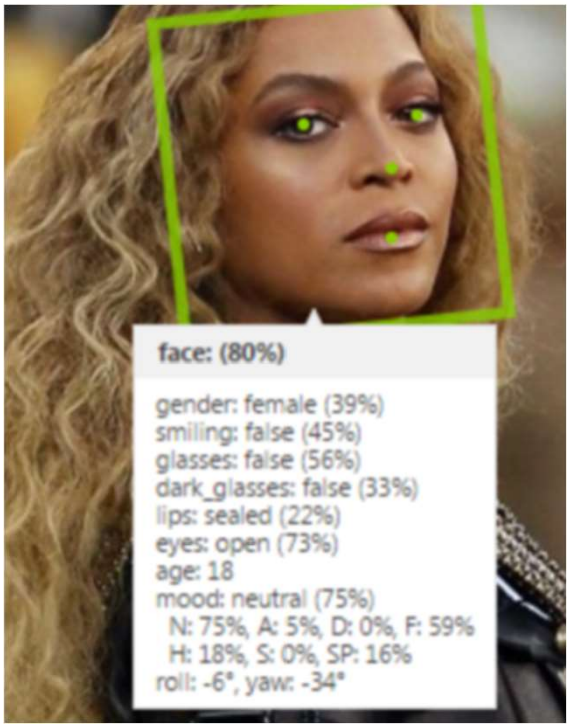
- Example 1: (**marketing more personal**)

McDonald's plans to open self-service kiosks equipped with cameras and image analysis software. They will use computer vision to identify customers' gender and age and then recommend items based on this visual data.

Example 2: Understanding consumer attention and sentiment

- Brand owners wish to use image analysis to see how advertising is perceived by viewers.

- Traditionally, this is done via surveys in which viewers express their sentiments about the content they see. It could be faulty however because humans are known for being inarticulate about their impressions, so a certain degree of misrepresentation is always here.



Small, inexpensive cameras mounted on or close to an outdoor signage, integrated with image analysis applications, could measure such metrics as:

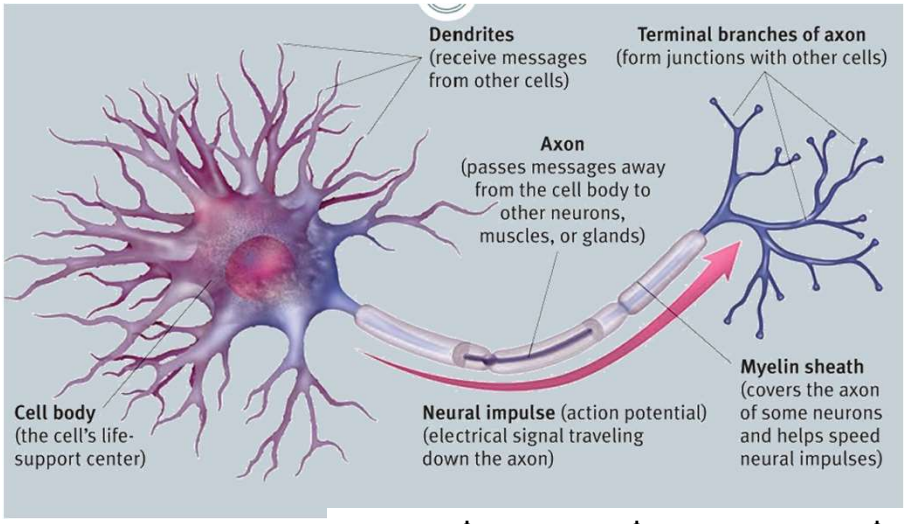
- How many people walked past the ad signage
- How many people stopped to look at it
- How long people looked at the advertisement
- The demographics of those who stopped to look at it

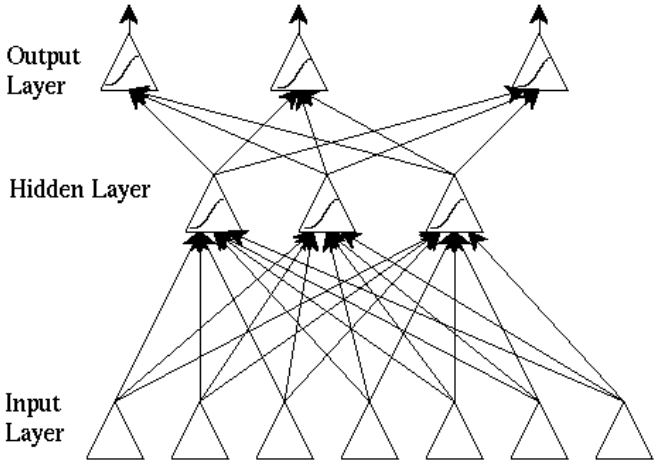
Why Do We Use A Neural Network?

- It has a remarkable ability to derive meaning from complicated or imprecise data.
- It can be used to extract patterns and detect trends that are too complex to be noticed by either humans or other computer techniques.
- Neural networks and conventional algorithmic computers are not in competition but complement each other.

Conceptually, what are artificial neural networks?

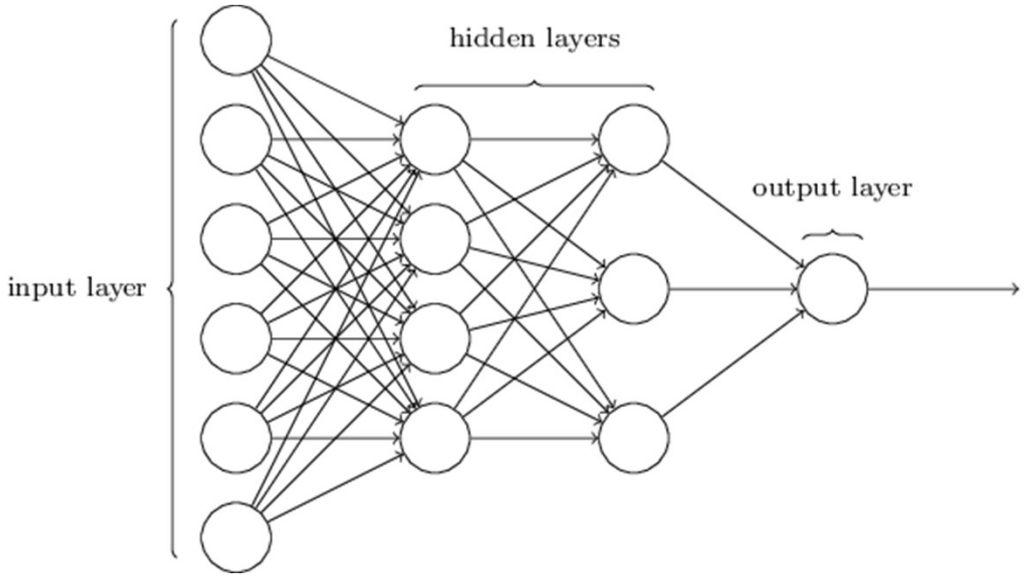
- A *neuron* receives a signal, processes it, and propagates the signal (or not)
- The brain is comprised of around 100 billion neurons, each connected to ~10k other neurons: 10^{15} synaptic connections
- *ANNs* are a *simplistic* imitation of a brain comprised of dense net of simple structures





The Architecture of Neural Networks

- These networks are called feedforward neural networks: no loops - information is always fed forward, never fed back



- There are other models of artificial neural networks in which feedback loops are possible - *recurrent neural networks*.

Elements of Neural Networks

- Neurons—each neuron or node is a function that takes the output from the layer ahead of it.
- The input layer and input neurons
- Hidden layers—these are full of many neurons and a neural network can have many hidden layers inside
- Output layer—this is where the result comes after the information is segmented through all the hidden layers
- Synapse—this is the connection between neurons and layers inside a neural network

How does a neuron do?

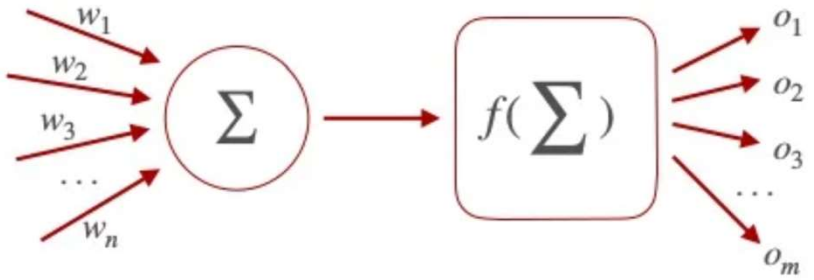
- A neural network is a bundle of neurons connected by synapses. Talking about the artificial one, the role of neurons are played by the units that perform calculations. Each of these “neurons”:
  - receives data from the input layer;
  - processes it performing simple calculations with it;
  - and then transmits it to another “neuron”.

Conceptual mathematical model

- Receives input from n sources
- Computes weighted sum

$$h_1 = x_1 w_1 + x_2 w_2 + \ldots x_n w_n$$

- Passes through an activation function
- Sends the signal to m succeeding neurons



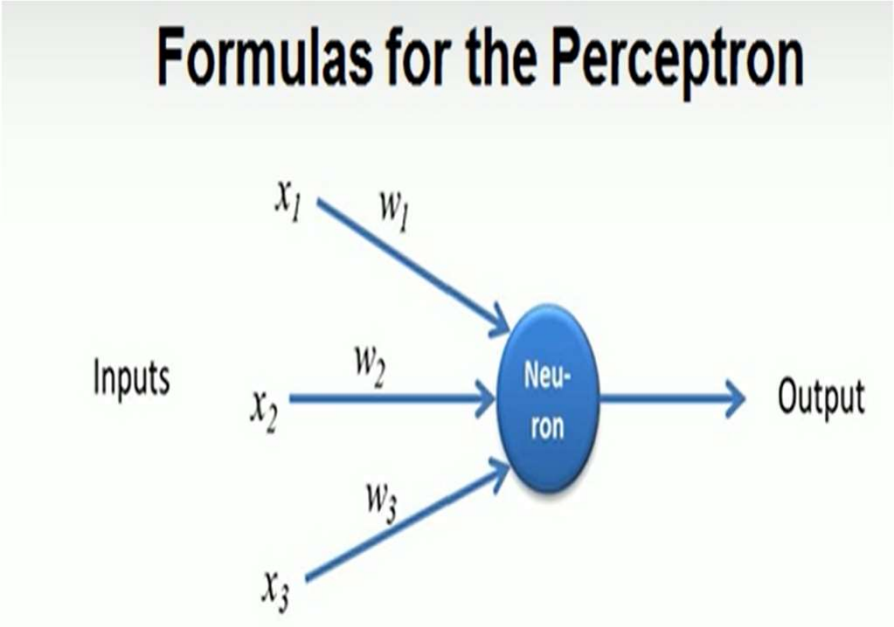
A typical neural network works like this:

- it receives data through the input layer of neurons;
- the data is processed by the neurons and passed to the next layer with the help of synapses each of which has its own coefficient;
- the next layer of neurons receive the information that is the sum of all of all data for neural networks, which are multiplied by the weight coefficients (each by its own);
- the resulting value is substituted into the activation function, resulting in the formation of output information;
- information is passed on until it reaches the final exit.

What is a Perceptron?

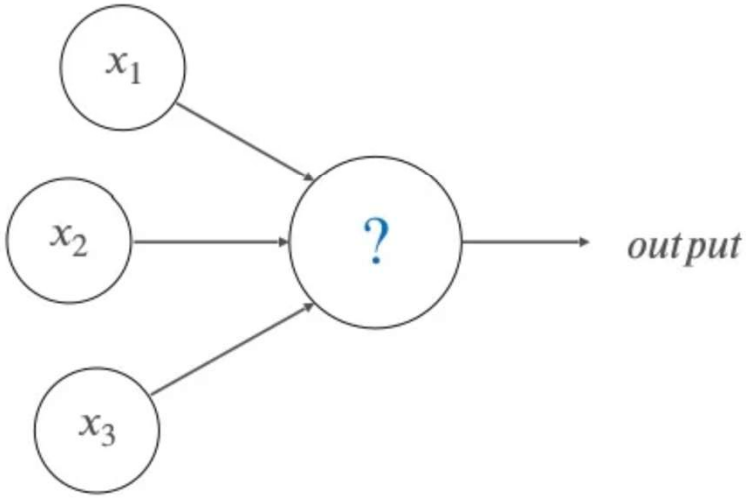
Perceptron is the basic unit of a neural network.

The perceptron is a network(neural network) that takes a number of inputs, carries out some processing on those inputs and produces an output



An Example of a Perceptron Model

Simplified (binary) artificial neuron



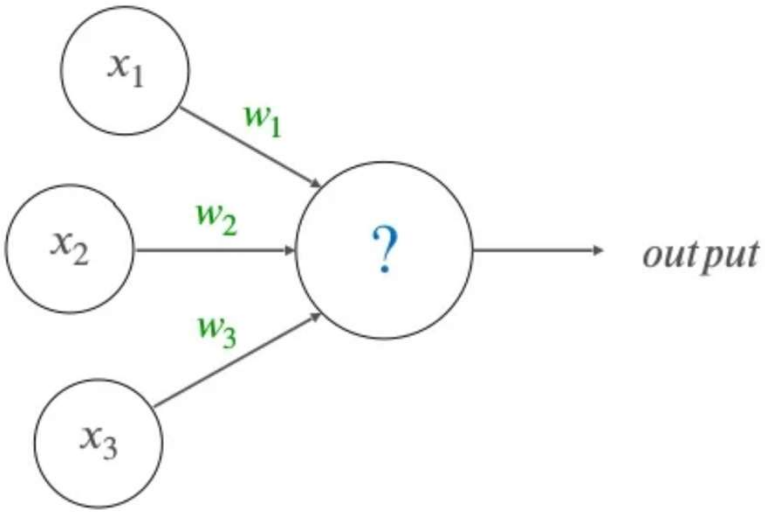
Do I snowboard this weekend?

$x_1 \rightarrow$ *Is the weather good?*

$x_2 \rightarrow$ *Is the powder good?*

$x_3 \rightarrow$ *Am I in the mood to drive?*

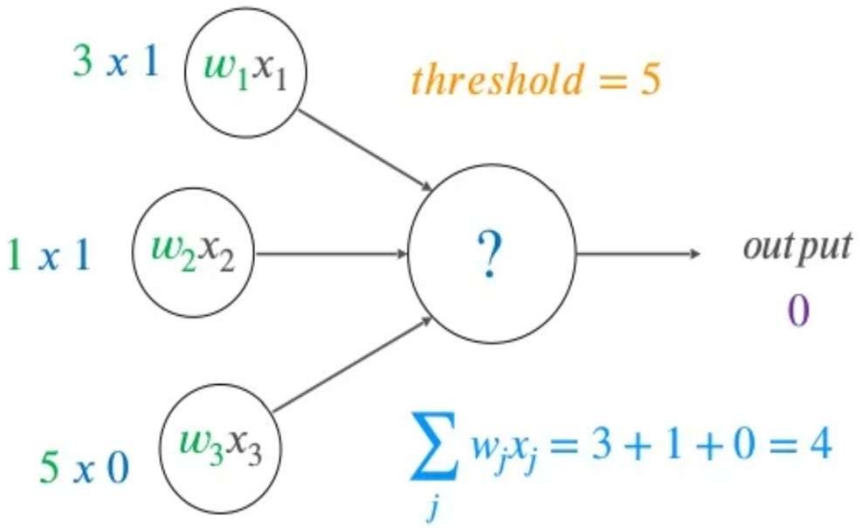
Simplified (binary) artificial neuron with weights



$$output = \begin{cases} 0, & \sum_{j=0}^{n} w_j x_j \leq threshold \\ 1, & \sum_{j=0}^{n} w_j x_j > threshold \end{cases}$$

Simplified (binary) artificial neuron; *add weights*

Persona: Après-ski'er



Do I snowboard this weekend?

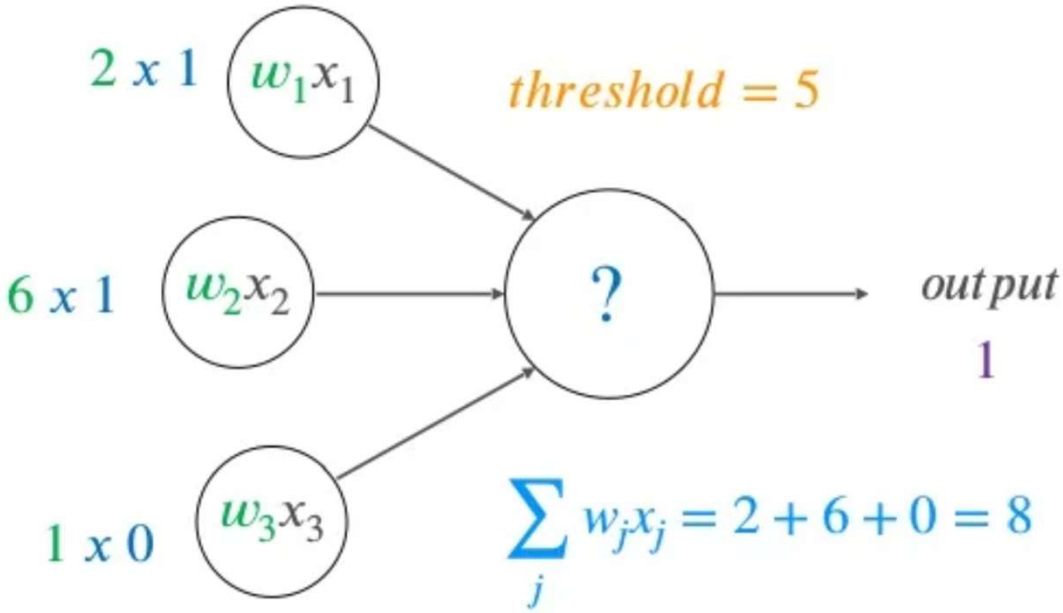
$x_1 = 1$ (*good weather*)
$w_1 = 3$

$x_2 = 1$ (*a lot of powder*)
$w_2 = 1$

$x_3 = 0$ (*driving sucks*)
$w_3 = 5$

Simplified (binary) artificial neuron; *add weights*

Persona: Shredder



Do I snowboard this weekend?

$x_1 = 1$ (*good weather*)

$w_1 = 2$

$x_2 = 1$ (*a lot of powder*)

$w_2 = 6$

$x_3 = 0$ (*driving sucks*)

$w_3 = 1$

Perceptron needs to take into account the bias

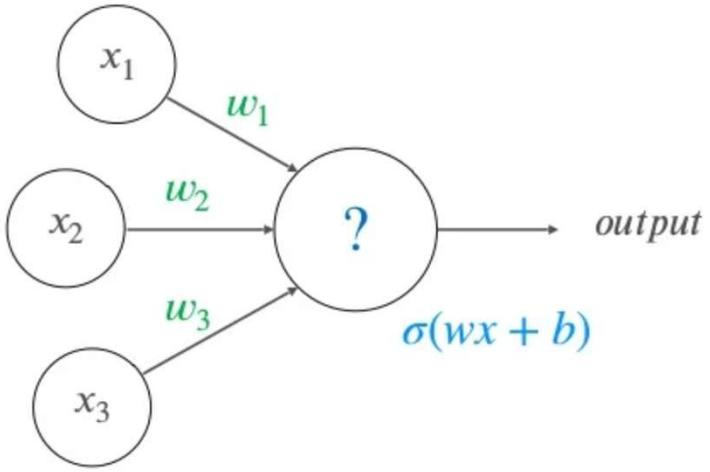
$$output = \begin{cases} 0, & wx + b \leq 0 \\ 1, & wx + b > 0 \end{cases}$$

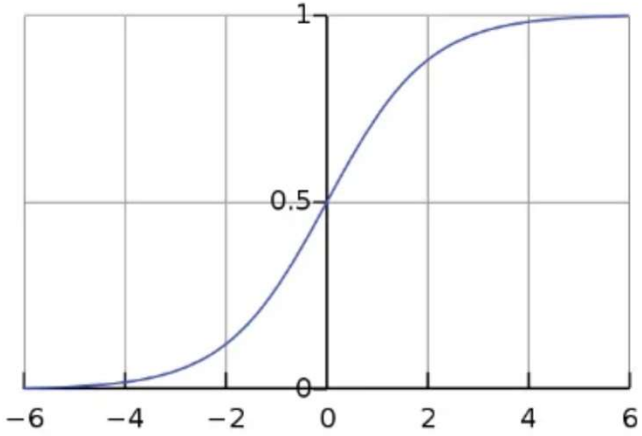
where b is how easy it is to get the perceptron to fire

e.g. Shredder has a strong positive bias to go to Whistler while Après-Ski'er bias is not as strong

Sigmoid Neuron

The more common artificial neuron



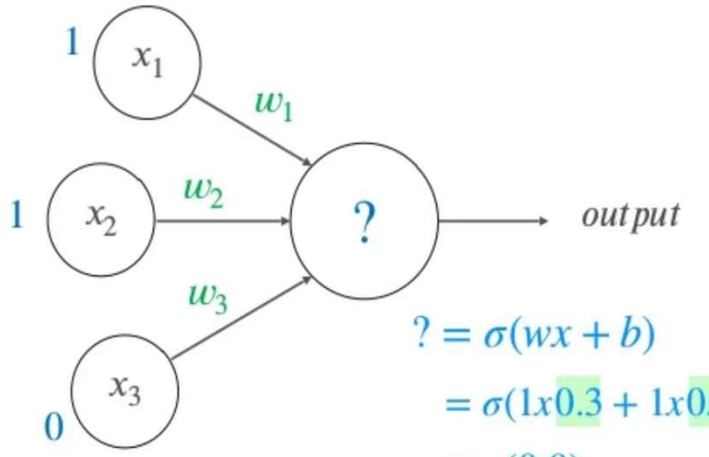


Instead of [0, 1], now (0…1)

Where output is defined by $\sigma(wx + b)$

Sigmoid Neuron

Persona: Shredder



$$? = \sigma(wx + b)$$
$$= \sigma(1x0.3 + 1x0.6 + 0x0.1)$$
$$= \sigma(0.9)$$
$$= 0.7109$$

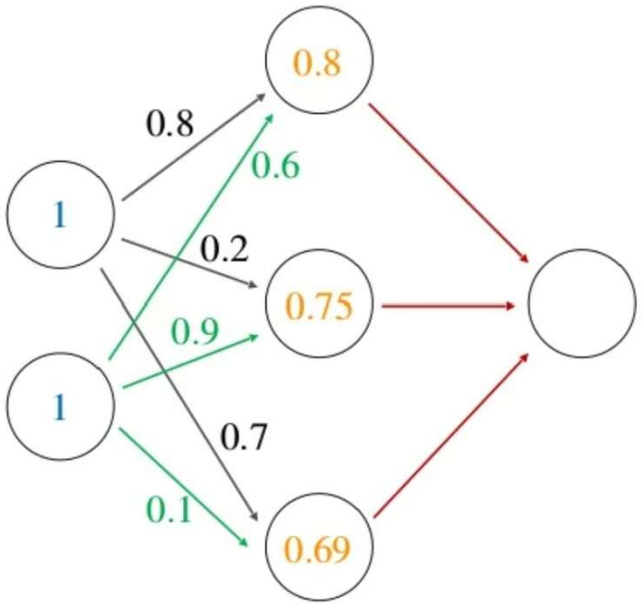
Do I snowboard this weekend?

$x_1 = 1$ (*good weather*)
$w_1 = 0.3$

$x_2 = 1$ (*a lot of powder*)
$w_2 = 0.6$

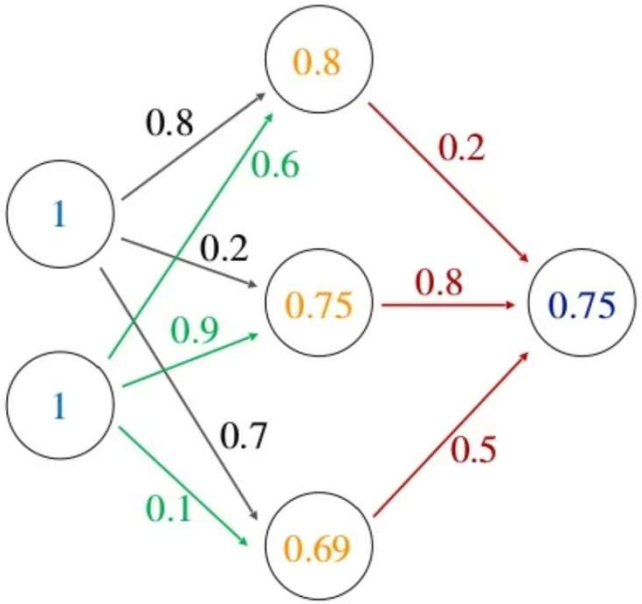
$x_3 = 0$ (*driving sucks*)
$w_3 = 0.1$

Simplified two-layer ANN



$$h_1 = \sigma(1x0.8 + 1x0.6) = 0.80$$
$$h_2 = \sigma(1x0.2 + 1x0.9) = 0.75$$
$$h_3 = \sigma(1x0.7 + 1x0.1) = 0.69$$

Simplified two-layer ANN



$$out = \sigma(0.2x0.8 + 0.8x0.75 + 0.5x0.69)$$

$$= \sigma(1.105)$$

$$= 0.75$$