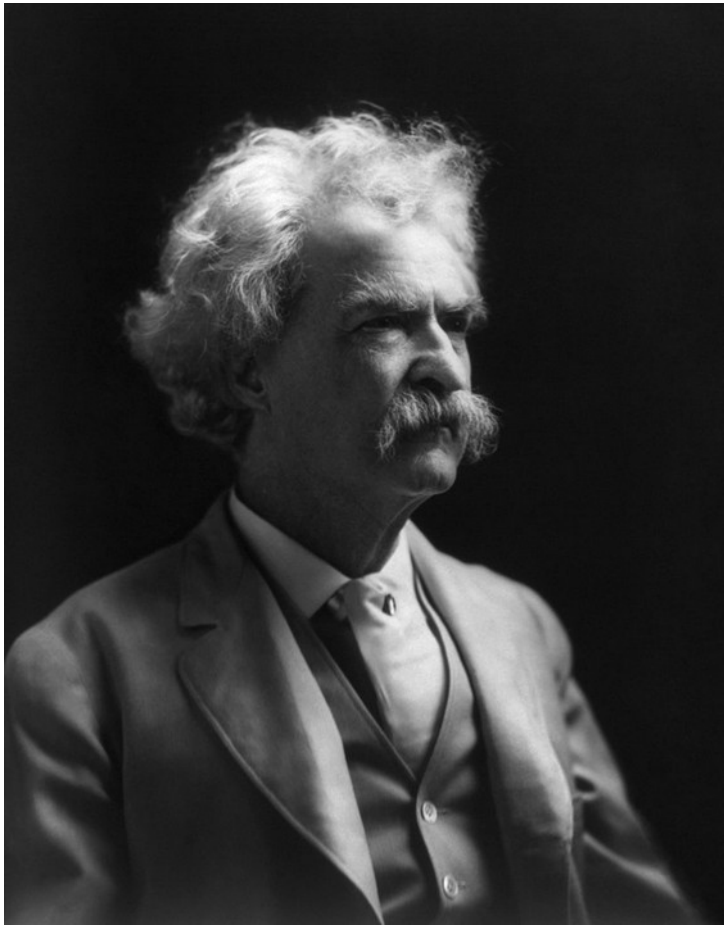
Machine Learning II

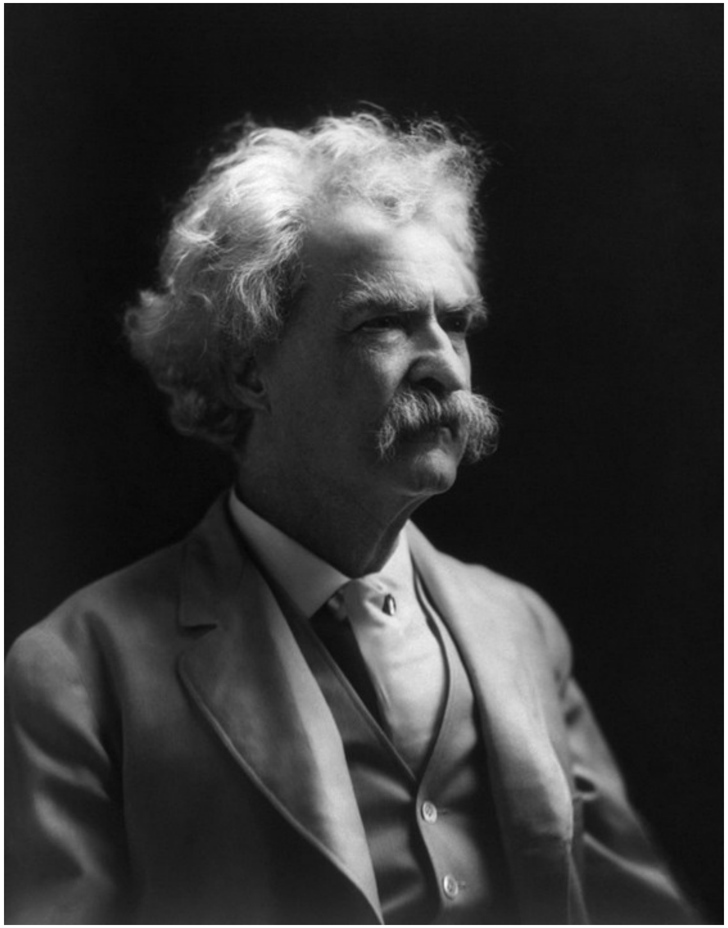
Techie Pizza #44267
Project Lesson 5

Michael Lyle



“Don’t use a five-dollar word when a fifty-cent word will do.”

- Mark Twain

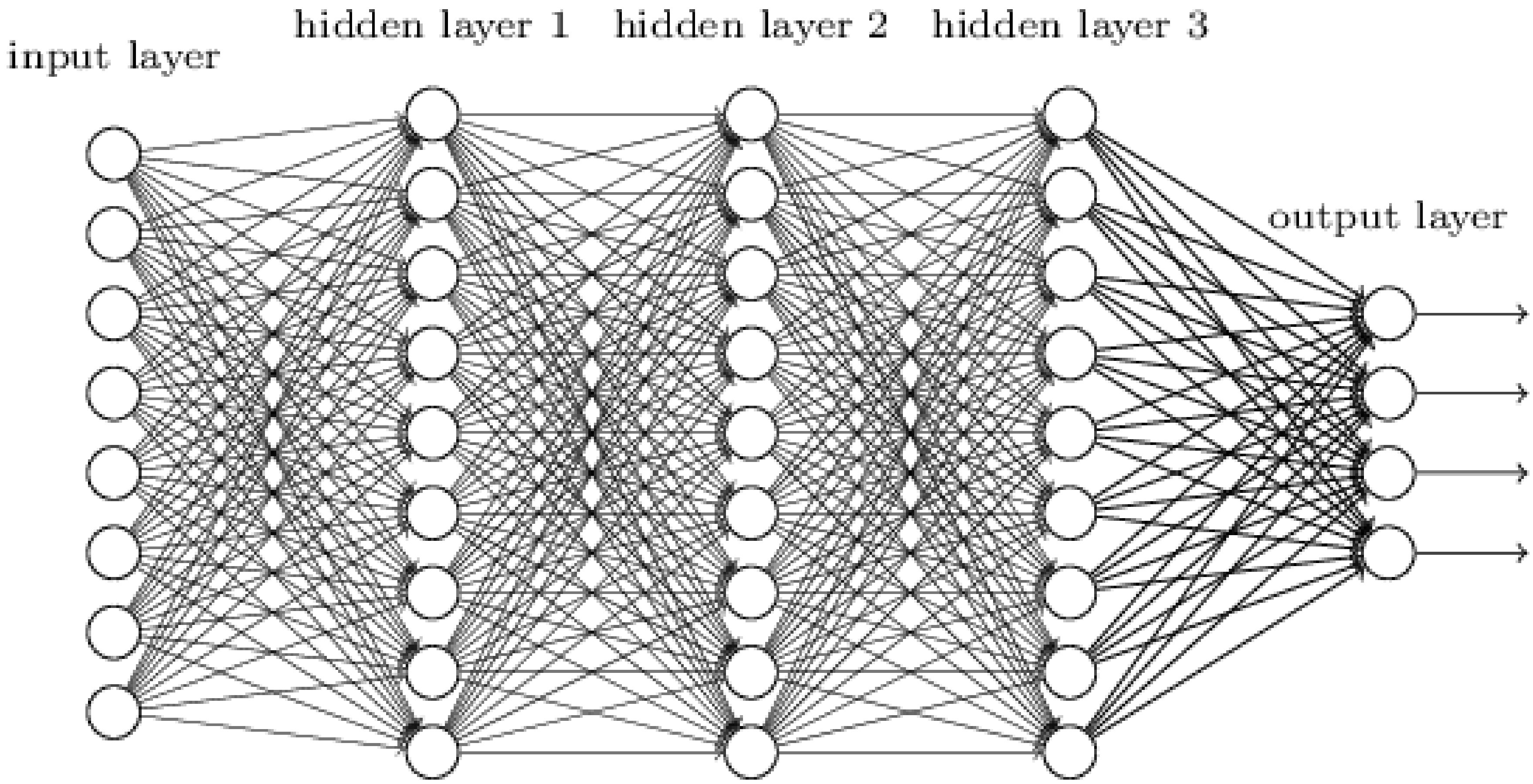


“Don’t use a five-dollar word when a fifty-cent word will do.”

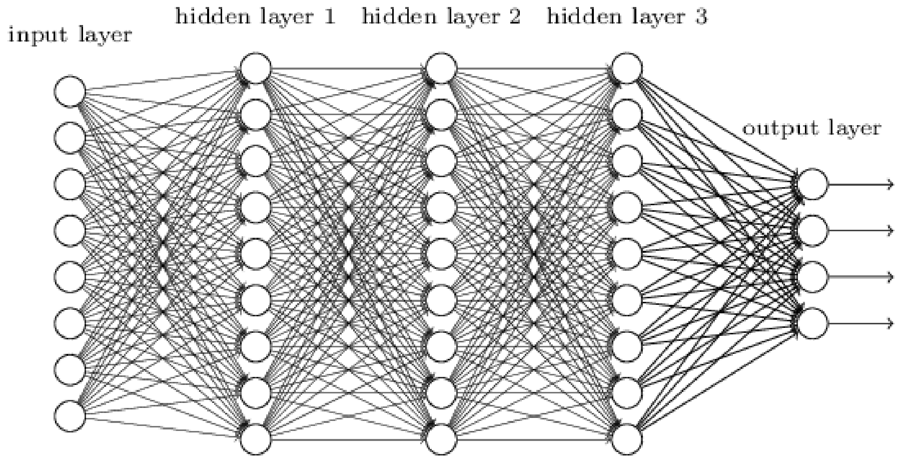
- Mark Twain

(But scientists like using five-dollar words; sorry about repeating them in this lesson!)

Dense Neural Network



Dense Neural Network



Every neuron is connected to every neuron in the previous layer.

This is a lot of connections. Each connection has its own different “weight” to learn. This makes training slow-- and risks overfitting.

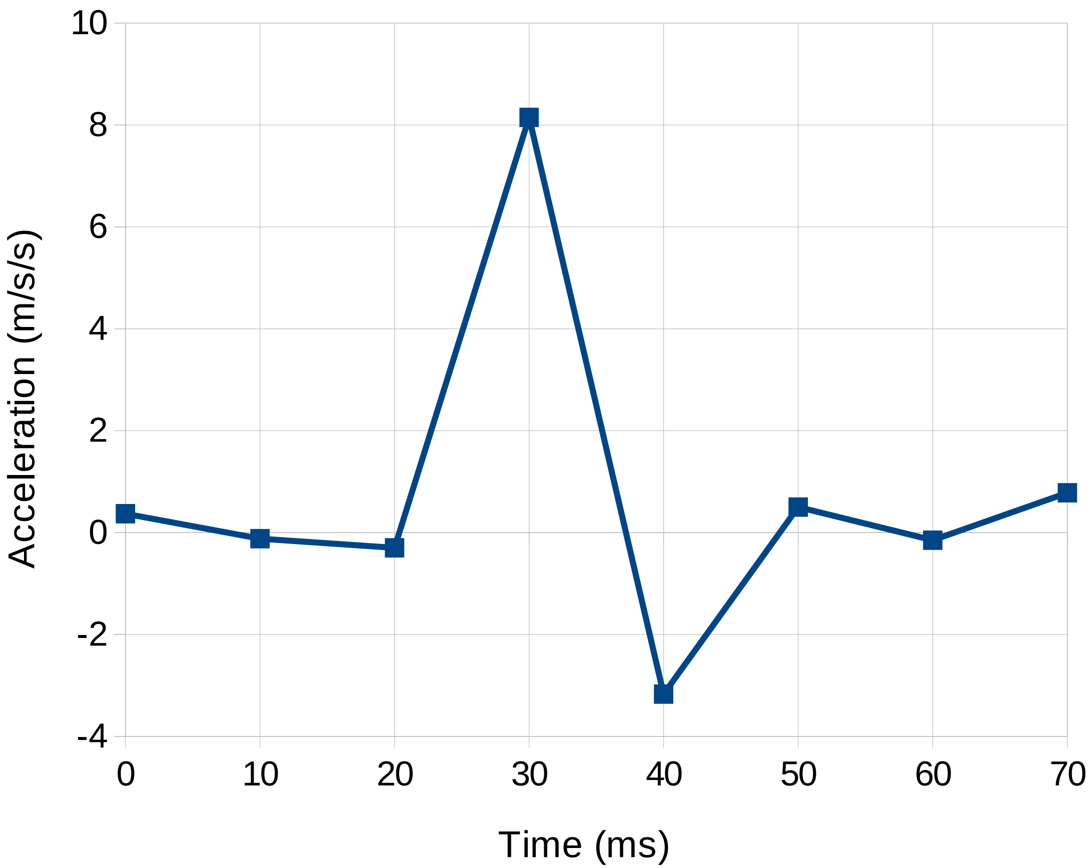
Time Series Data

- Measurements from an accelerometer arrive as time-series data

Time (ms)	Acceleration
0	0.37
10	-0.12
20	-0.30
30	8.15
40	-3.17
50	0.50
60	-0.15
70	0.78

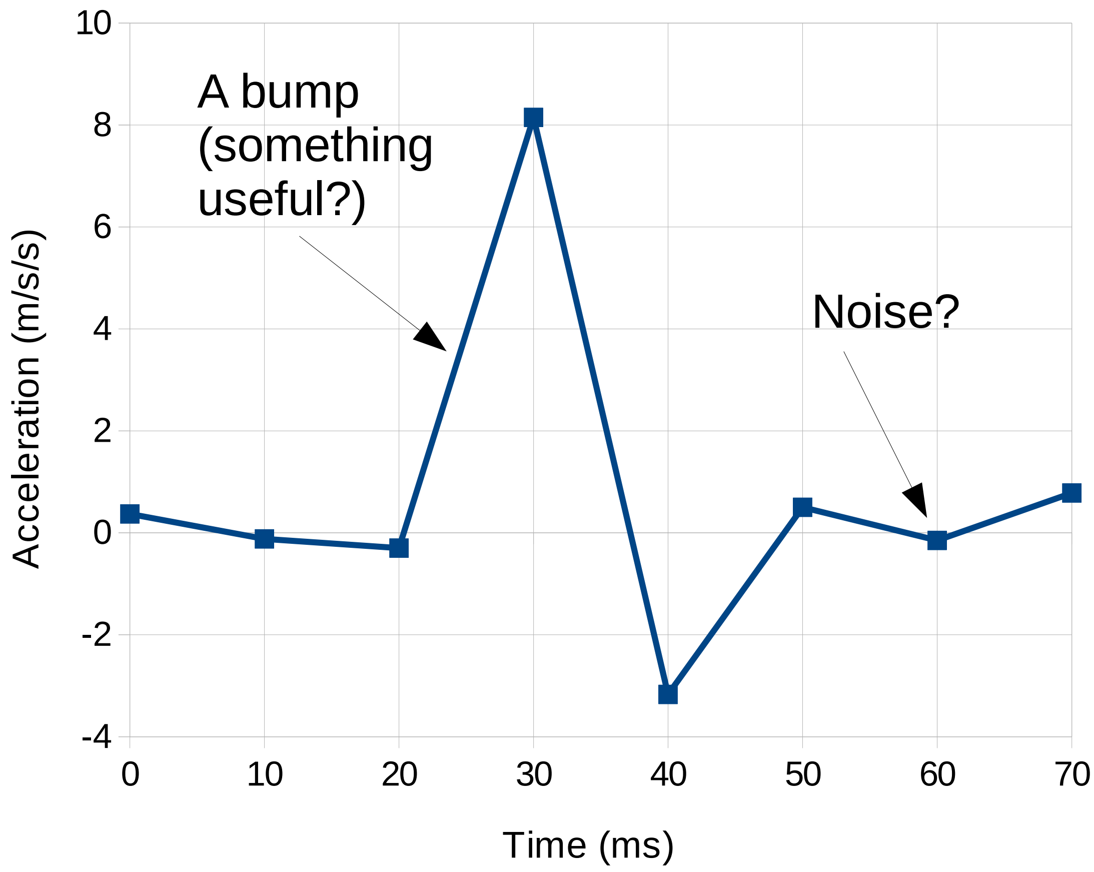
Graphing Time Series Data

Time (ms)	Acceleration
0	0.37
10	-0.12
20	-0.30
30	8.15
40	-3.17
50	0.50
60	-0.15
70	0.78



Graphing Time Series Data

Time (ms)	Acceleration
0	0.37
10	-0.12
20	-0.30
30	8.15
40	-3.17
50	0.50
60	-0.15
70	0.78



Time Series Data

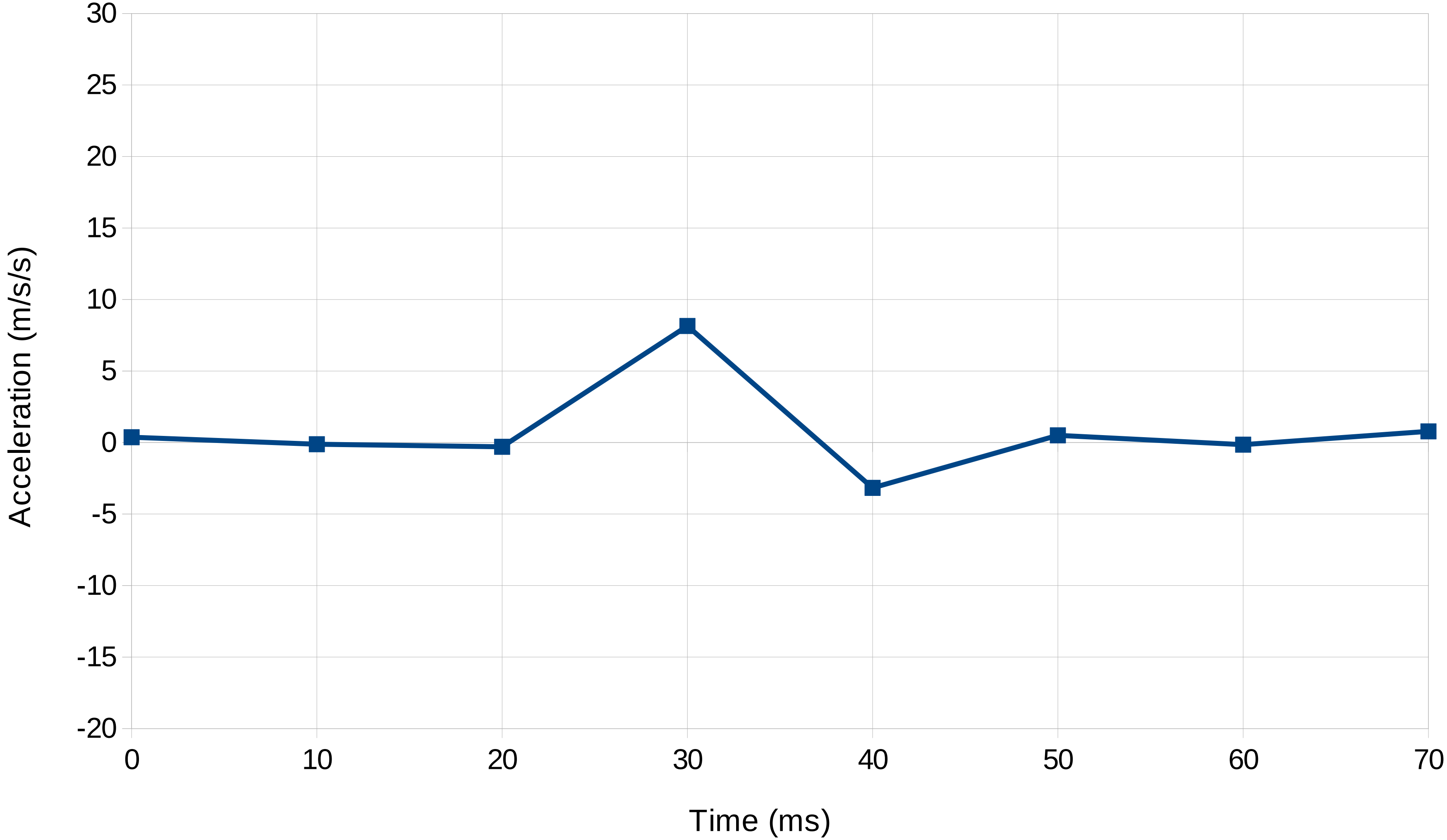
- If we record 10 seconds of data, with 100 measurements per second, that’s 1,000 measurements; each is an input
- If we have a big dense layer using this data, that is 1,000,000 weights (1,000 neurons each connected to 1,000 inputs)
- Small computers like in current scooters can handle neural networks with 25,000 weights

5th Grade Math
6th Grade Math
Pre-Algebra
Algebra
Geometry
Algebra II
Trigonometry
Pre-Calculus
Calculus
Linear Algebra
Differential Equations
Multivariate/Vector Calculus
Real & Complex Analysis
Group Theory
...

Convolutions

- Convolutions are usually studied during a Differential Equations class, but we can get the “gist” now!
- Convolutions are a way of filtering data-- to smooth it out or exaggerate features
- We make a recipe for the transformation we want-- called a convolution kernel
- Then we follow the recipe for each entry in our data table
- Kernels can be any size, but for these examples size=3

Our Data



Acceleration (m/s/s)
30
25
20
15
10
5
0
-5
-10
-15
-20
0
10
20
30
40
50
60
70
Time (ms)

Convolutions - Smooth

Take the average of each measurement, the measurement before, and the measurement after

Time (ms)	Acceleration
0	0.37
10	-0.12
20	-0.30
30	8.15
40	-3.17
50	0.50
60	-0.15
70	0.78

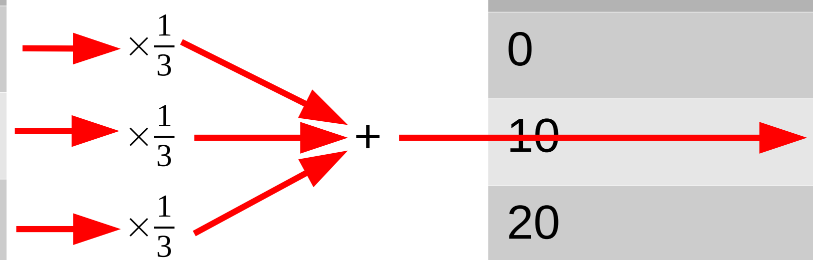
$$\begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}$$

Time (ms)	Smoothed
0	
10	-0.02
20	2.57
30	1.56
40	1.83
50	-0.94
60	0.38
70	

Convolutions - Smooth

Take the average of each measurement, the measurement before, and the measurement after

Time (ms)	Acceleration
0	0.37
10	-0.12
20	-0.30
30	8.15
40	-3.17
50	0.50
60	-0.15
70	0.78



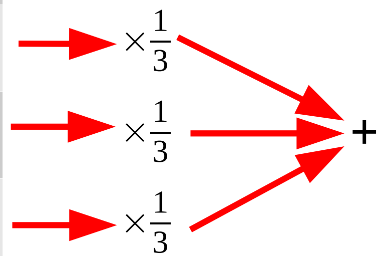
Time (ms)	Smoothed
0	
10	-0.02
20	2.57
30	1.56
40	1.83
50	-0.94
60	0.38
70	

$$\begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}$$

Convolutions - Smooth

Take the average of each measurement, the measurement before, and the measurement after

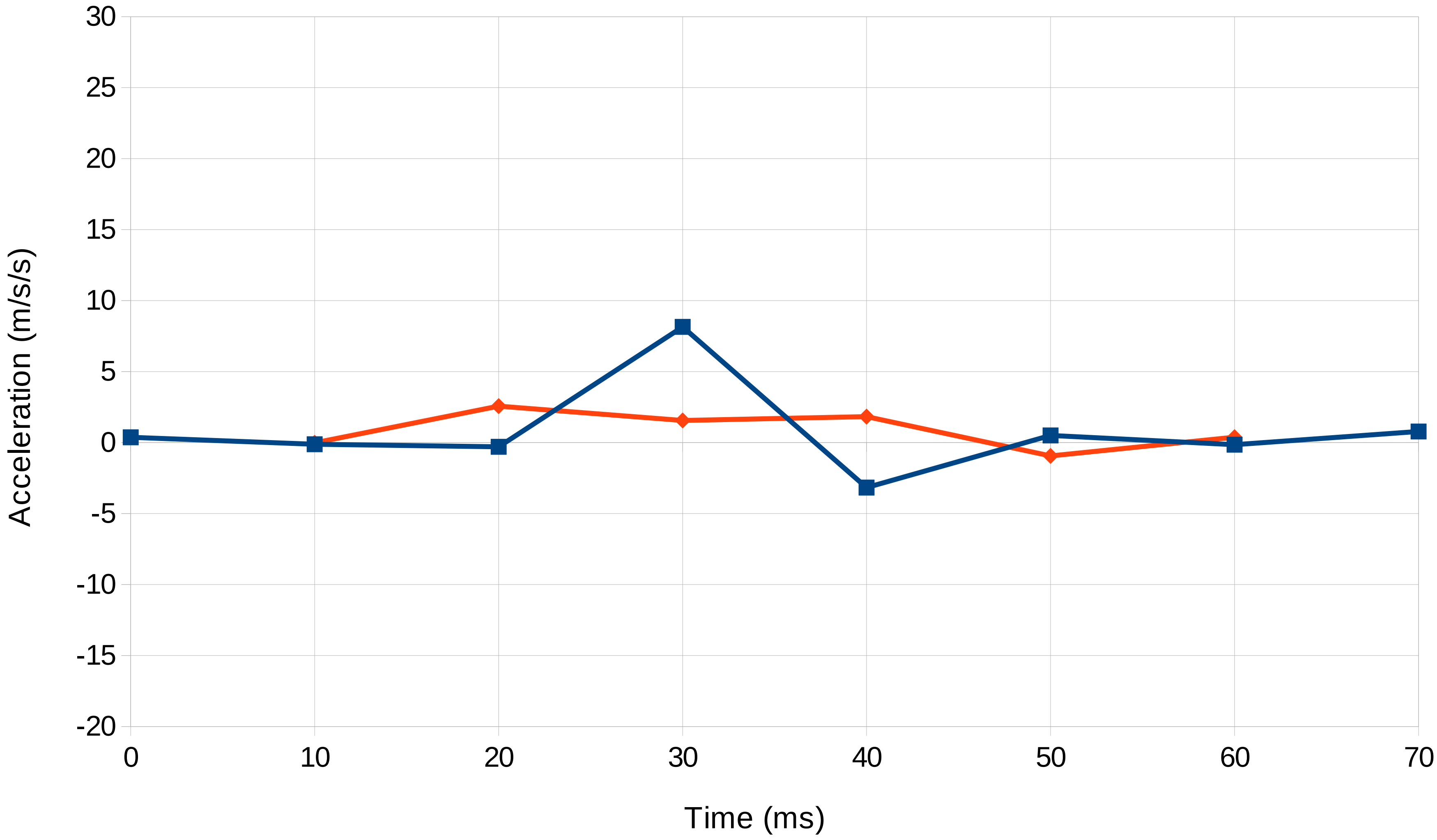
Time (ms)	Acceleration
0	0.37
10	-0.12
20	-0.30
30	8.15
40	-3.17
50	0.50
60	-0.15
70	0.78



Time (ms)	Smoothed
0	
10	-0.02
20	2.57
30	1.56
40	1.83
50	-0.94
60	0.38
70	

$$\begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}$$

Our Data, Smoothed



Convolutions - Exaggerate

Take each measurement times 3, minus the measurement before and minus the one after

Time (ms)	Acceleration
0	0.37
10	-0.12
20	-0.30
30	8.15
40	-3.17
50	0.50
60	-0.15
70	0.78

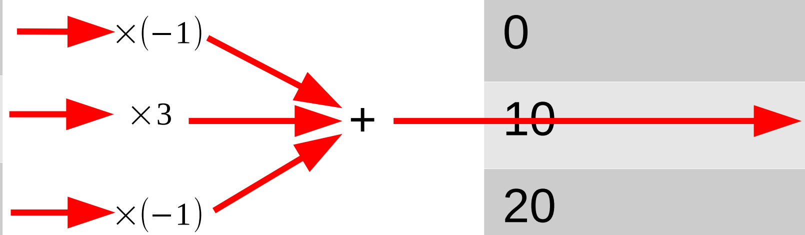
$$\begin{bmatrix} -1 & 3 & -1 \end{bmatrix}$$

Time (ms)	Exaggerated
0	
10	-0.43
20	-8.93
30	27.92
40	-18.16
50	4.82
60	-1.73
70	

Convolutions - Exaggerate

Take each measurement times 3, minus the measurement before and minus the one after

Time (ms)	Acceleration
0	0.37
10	-0.12
20	-0.30
30	8.15
40	-3.17
50	0.50
60	-0.15
70	0.78



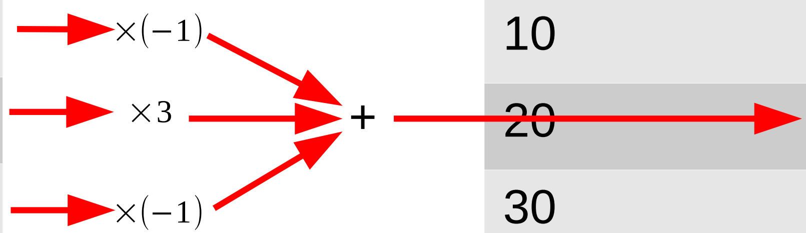
Time (ms)	Exaggerated
0	
10	-0.43
20	-8.93
30	27.92
40	-18.16
50	4.82
60	-1.73
70	

$$\begin{bmatrix} -1 & 3 & -1 \end{bmatrix}$$

Convolutions - Exaggerate

Take each measurement times 3, minus the measurement before and minus the one after

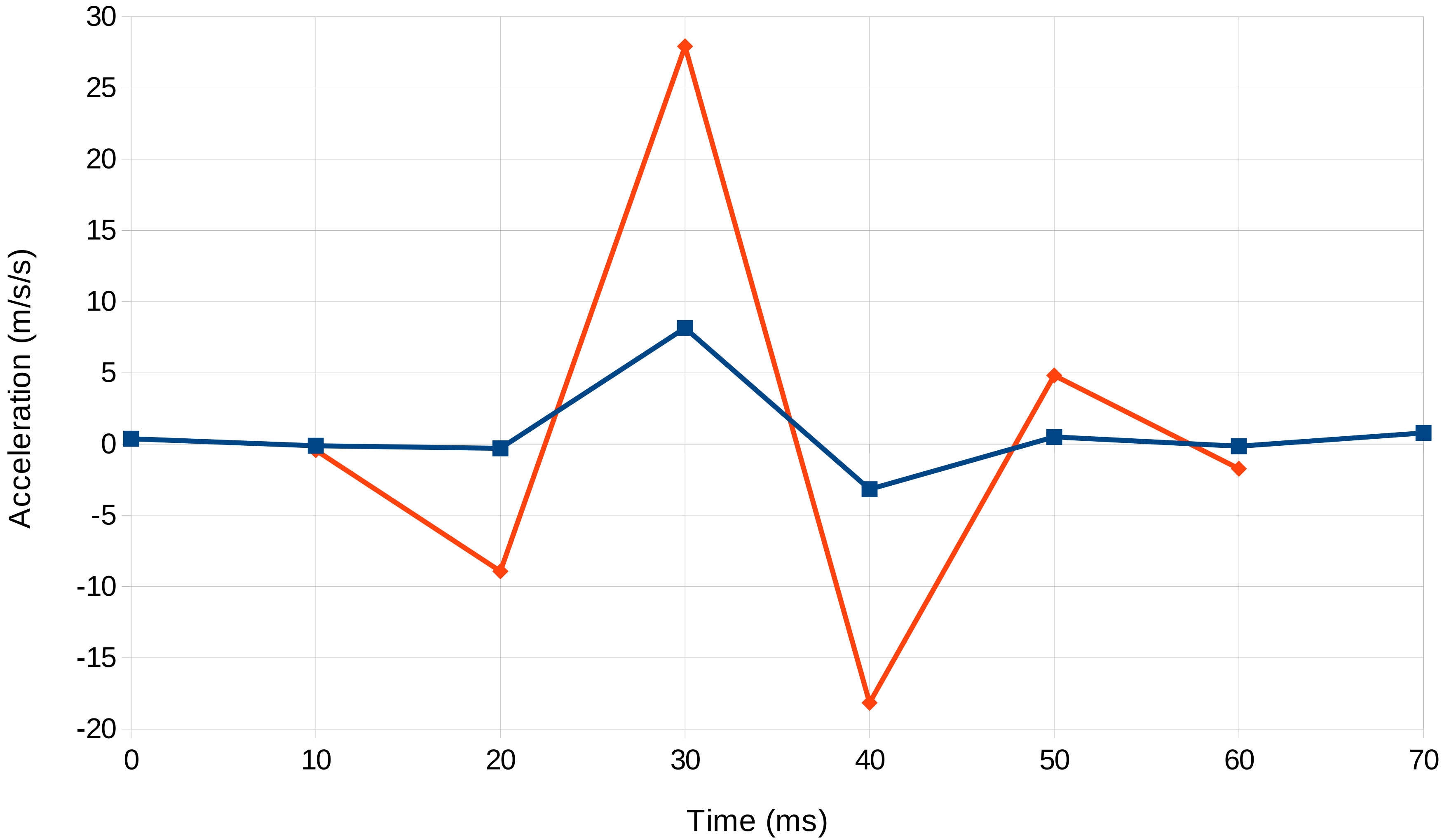
Time (ms)	Acceleration
0	0.37
10	-0.12
20	-0.30
30	8.15
40	-3.17
50	0.50
60	-0.15
70	0.78



$$\begin{bmatrix} -1 & 3 & -1 \end{bmatrix}$$

Time (ms)	Exaggerated
0	
10	-0.43
20	-8.93
30	27.92
40	-18.16
50	4.82
60	-1.73
70	

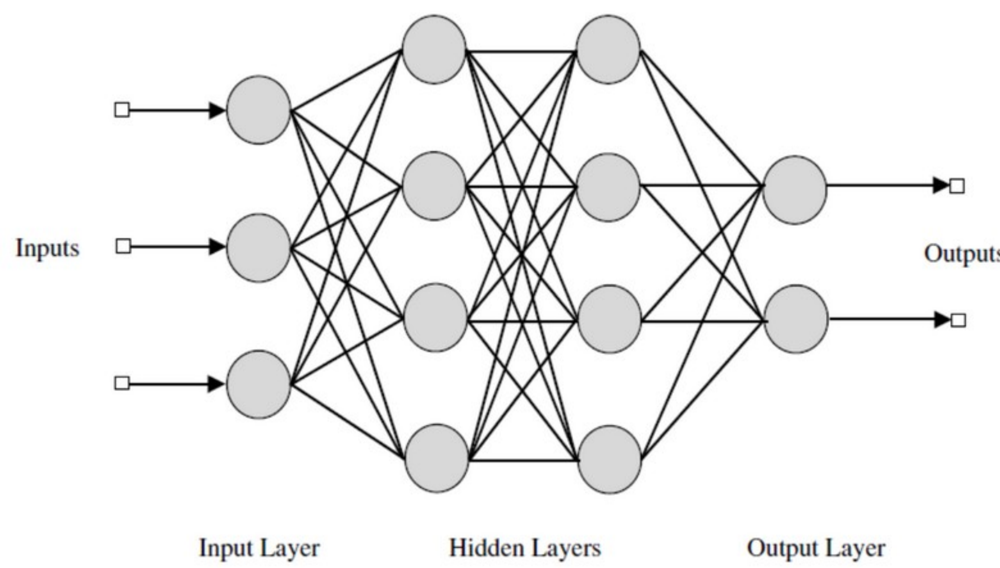
Our Data, Exaggerated



Training an artificial neural network

Remember this slide?

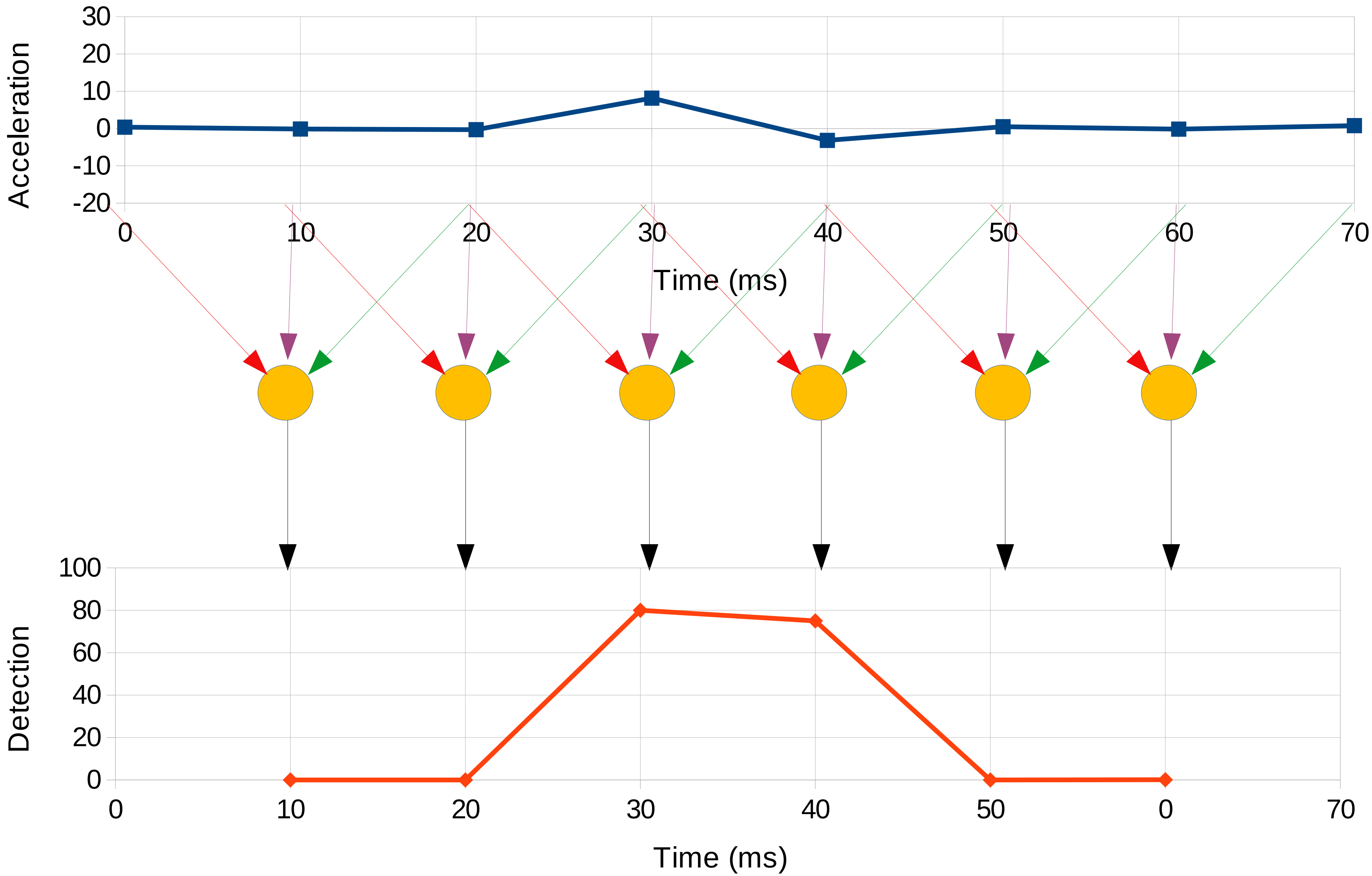
1) Start with example data and a set of "correct answers."
2) Adjust how strong the connections are to make the neural network produce closer to the output we want. ("Training")
3) Repeat. A lot.
4) For some problems, we may get a result that's as good as a human, or even better!



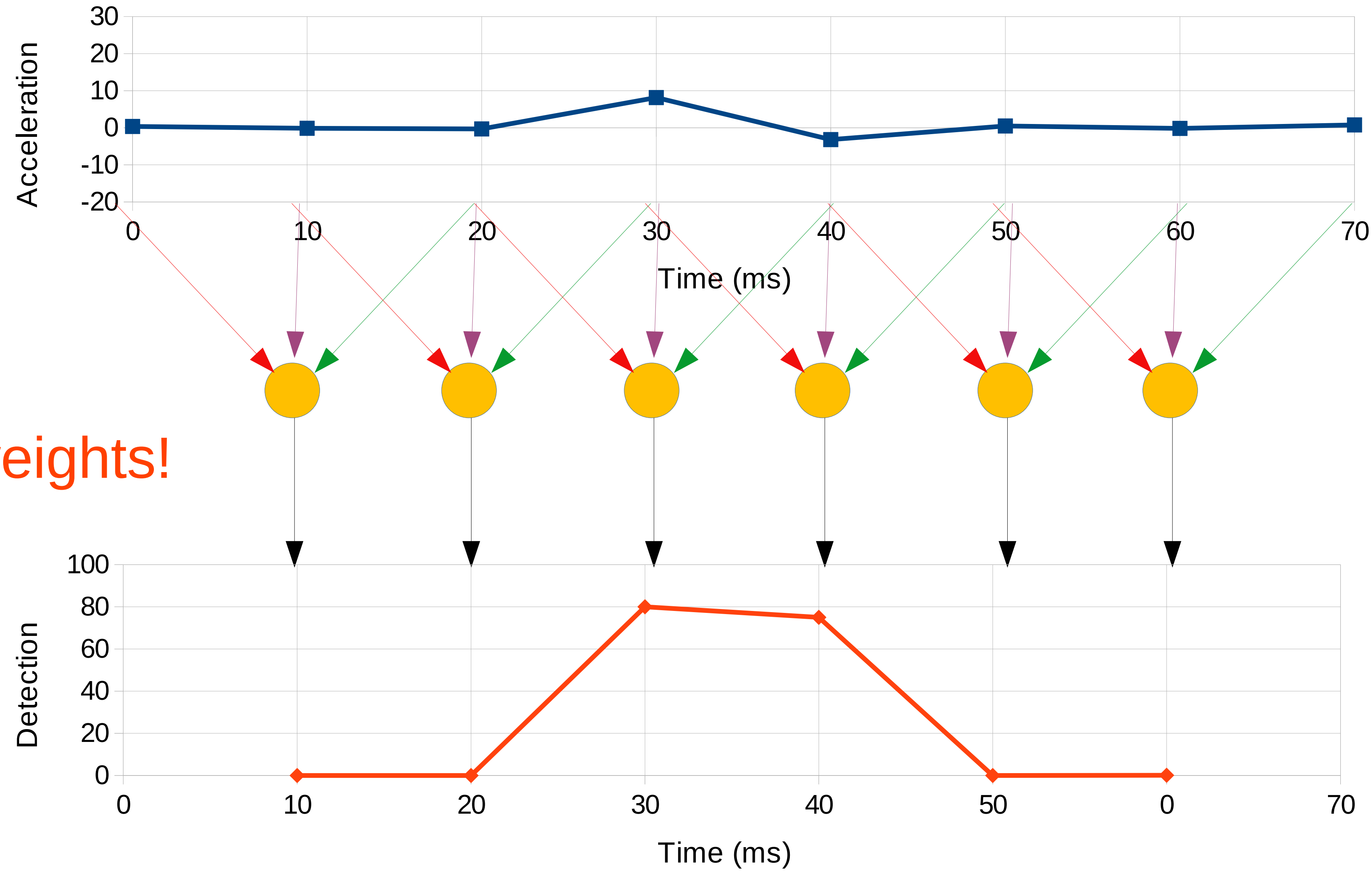
Convolutional Neural Network

- A convolutional layer is a neural network layer that performs convolutions
- We don’t need to know the exact convolution we want: training will find it for us
  - This means we don’t need to take Differential Equations first!
  - Also the computer can find better convolutions than people usually can.
- Hopefully it simplifies the data in ways that make life easier for the later layers

Convolutional Neural Network



Convolutional Neural Network



3 weights!

Summary

- Time series data measures how values from a sensor change over time.
- Convolutional neural networks are good at matching patterns in time-series data.
- Convolutional layers are much more efficient (fast to train, fast to “run”) than dense layers, but are limited to spotting local patterns.