



The Underpinnings of ML

- How algorithms work

- How algorithms work
- Explore important algorithms

- How algorithms work
- Explore important algorithms
 - Internal working

- How algorithms work
- Explore important algorithms
 - Internal working
 - Using real-life examples

- How algorithms work
- Explore important algorithms
 - Internal working
 - Using real-life examples
- Concepts

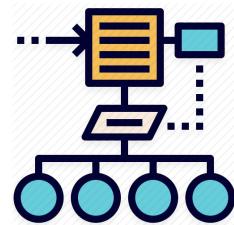
- How algorithms work
- Explore important algorithms
 - Internal working
 - Using real-life examples
- Concepts
 - Linear Regression

- How algorithms work
- Explore important algorithms
 - Internal working
 - Using real-life examples
- Concepts
 - Linear Regression
 - Decision Trees

- How algorithms work
- Explore important algorithms
 - Internal working
 - Using real-life examples
- Concepts
 - Linear Regression
 - Decision Trees
 - Neural Networks



Data



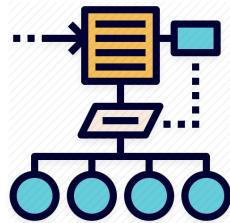
Algorithm



Model



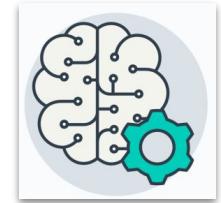
Data



Algorithm



Live Data



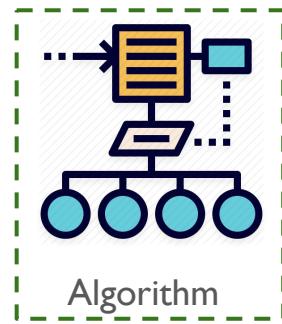
Model



Prediction



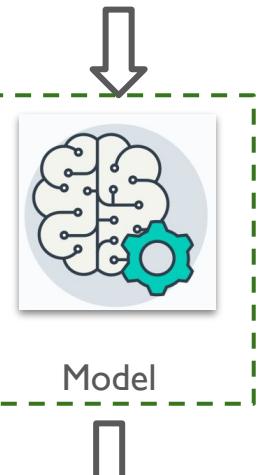
Data



Algorithm



Live Data



Model



Prediction

Predicting salary based on the number of years
of experience

2000

2000, 4000

2000, 4000, 6000

2000, 4000, 6000, _____

2000, 4000, 6000, 8000

Let's formulate the machine learning problem..

Historical Data

Years of Experience	Salary
1	2000
2	4000
3	6000

Let's formulate the machine learning problem..



Let's formulate the machine learning problem..

Historical Data

Years of Experience	Salary
1	2000
2	4000
3	6000

Training

Unknown Salary	
Years of Experience	Salary
4	???

Model

Let's formulate the machine learning problem..

Historical Data

Years of Experience	Salary
1	2000
2	4000
3	6000

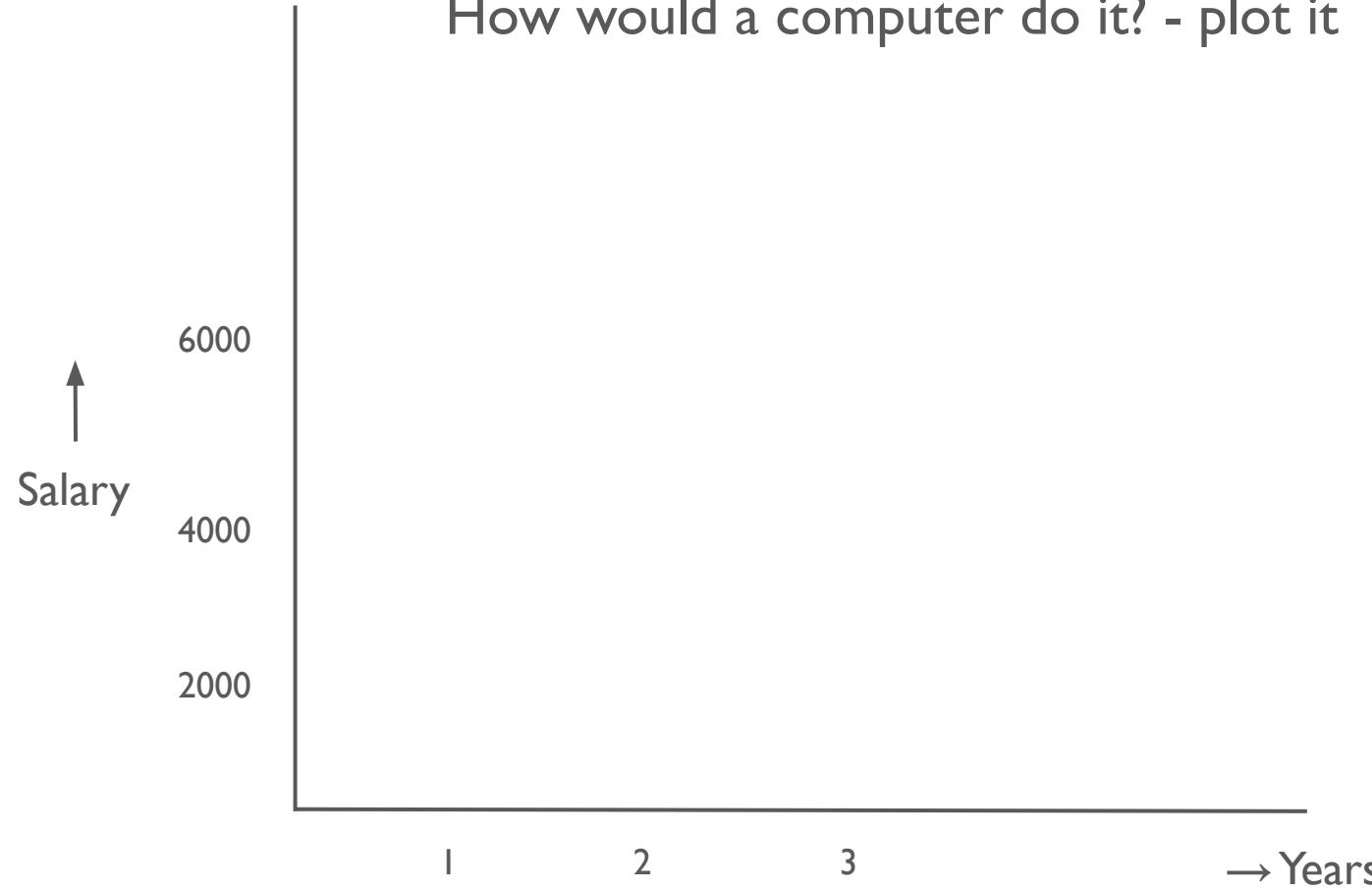
Training

Unknown Salary	
Years of Experience	Salary
4	???

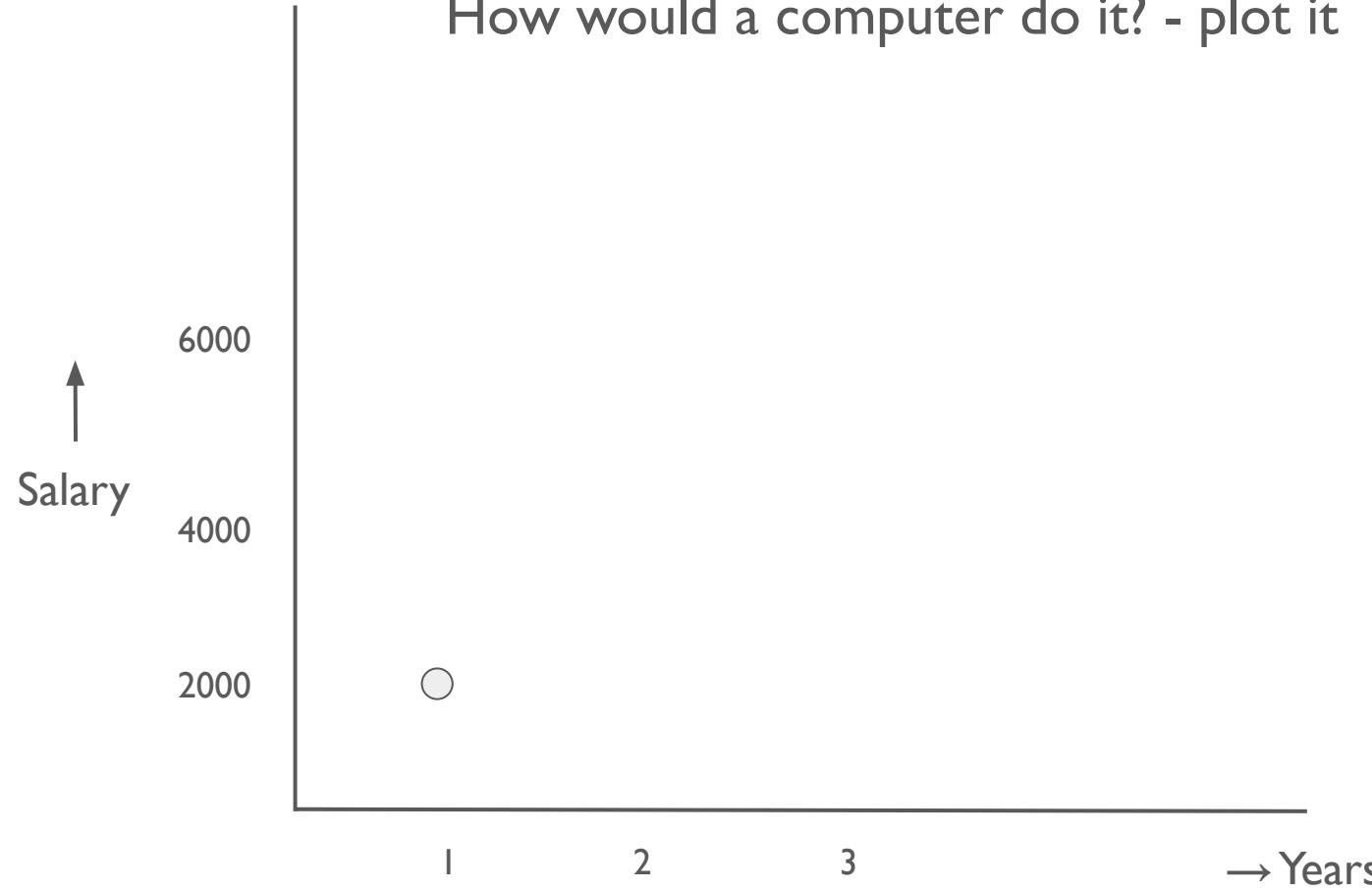
Model

Prediction	
Years of Experience	Salary
4	8000

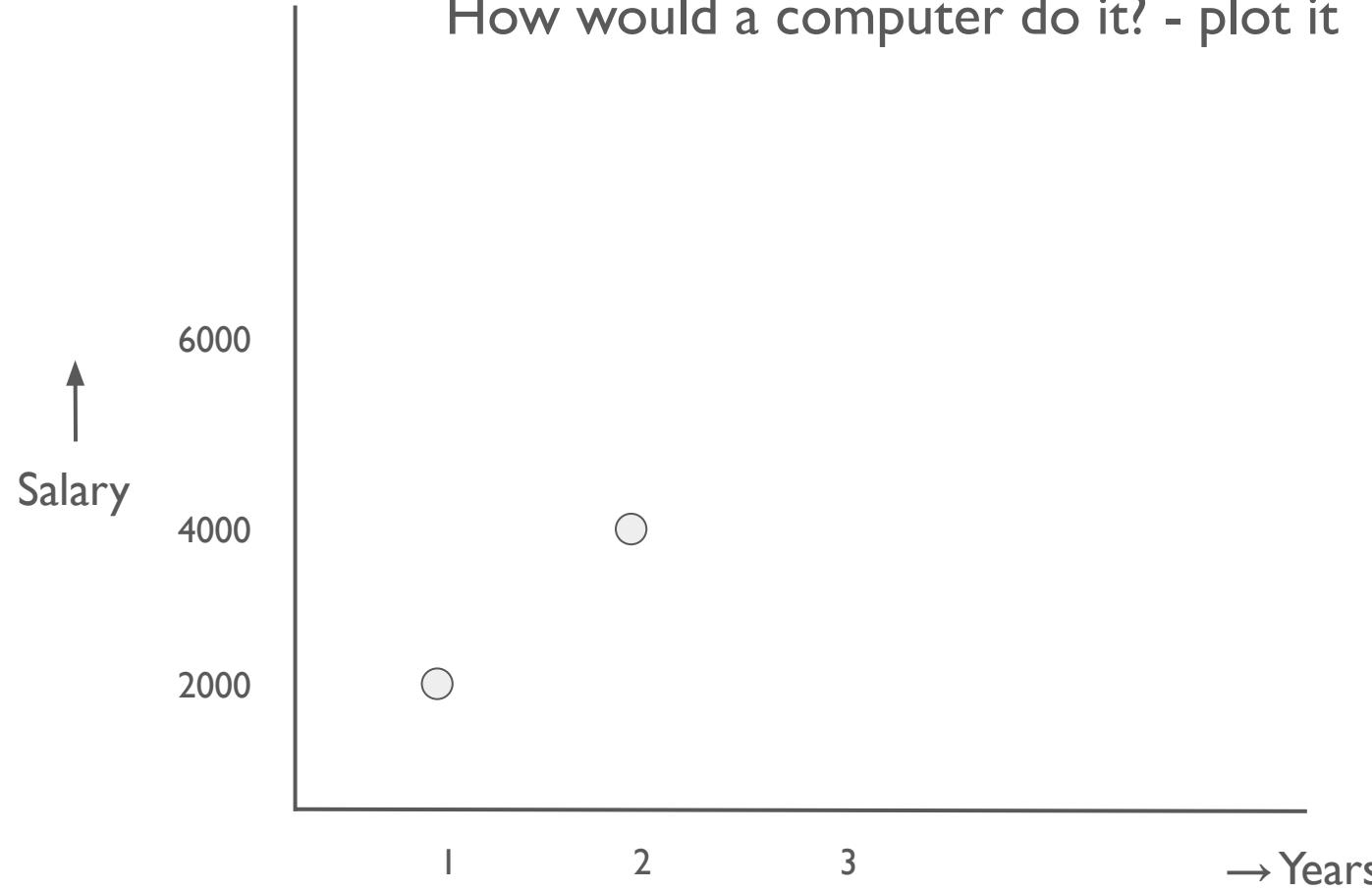
How would a computer do it? - plot it



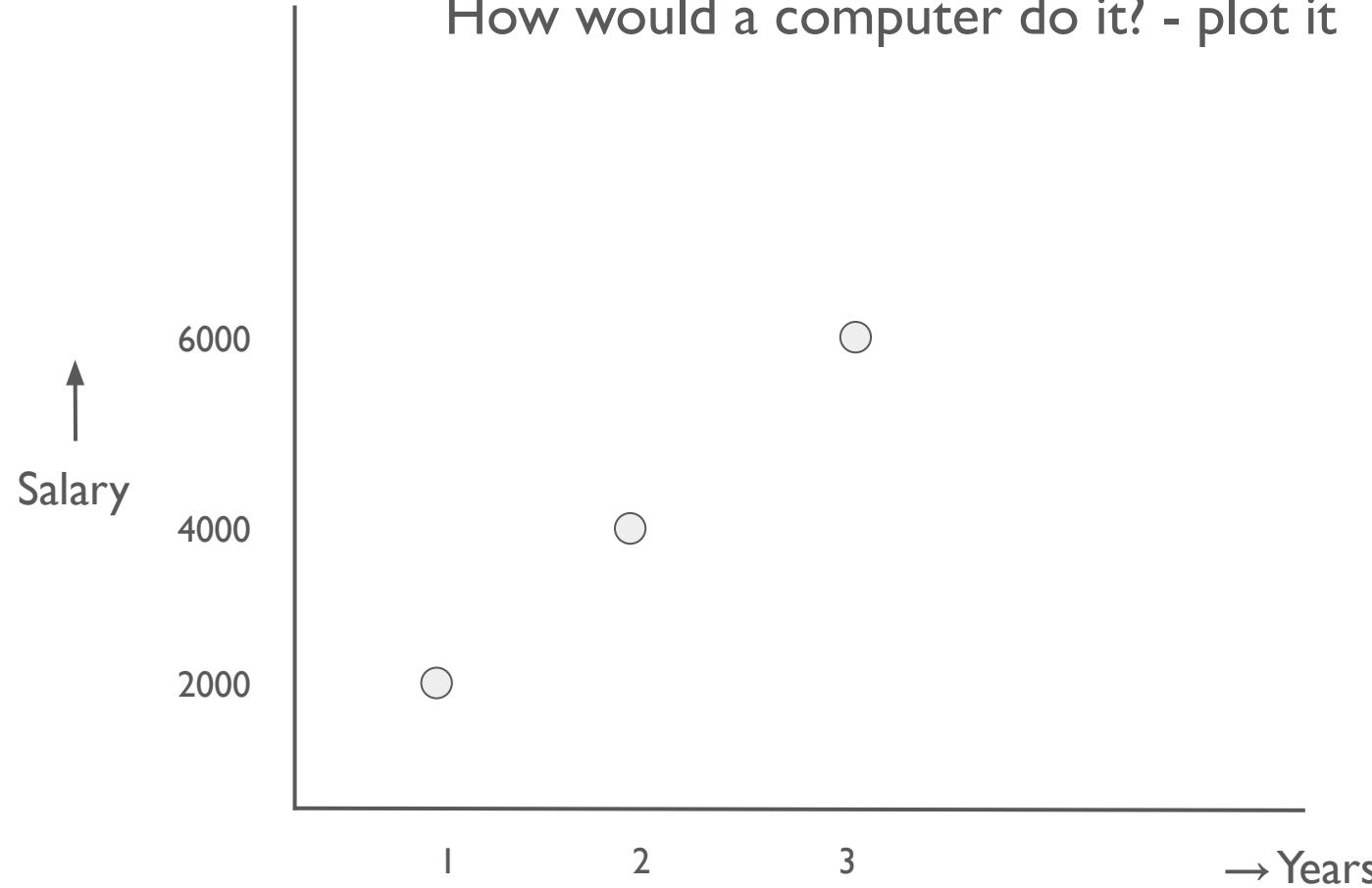
How would a computer do it? - plot it



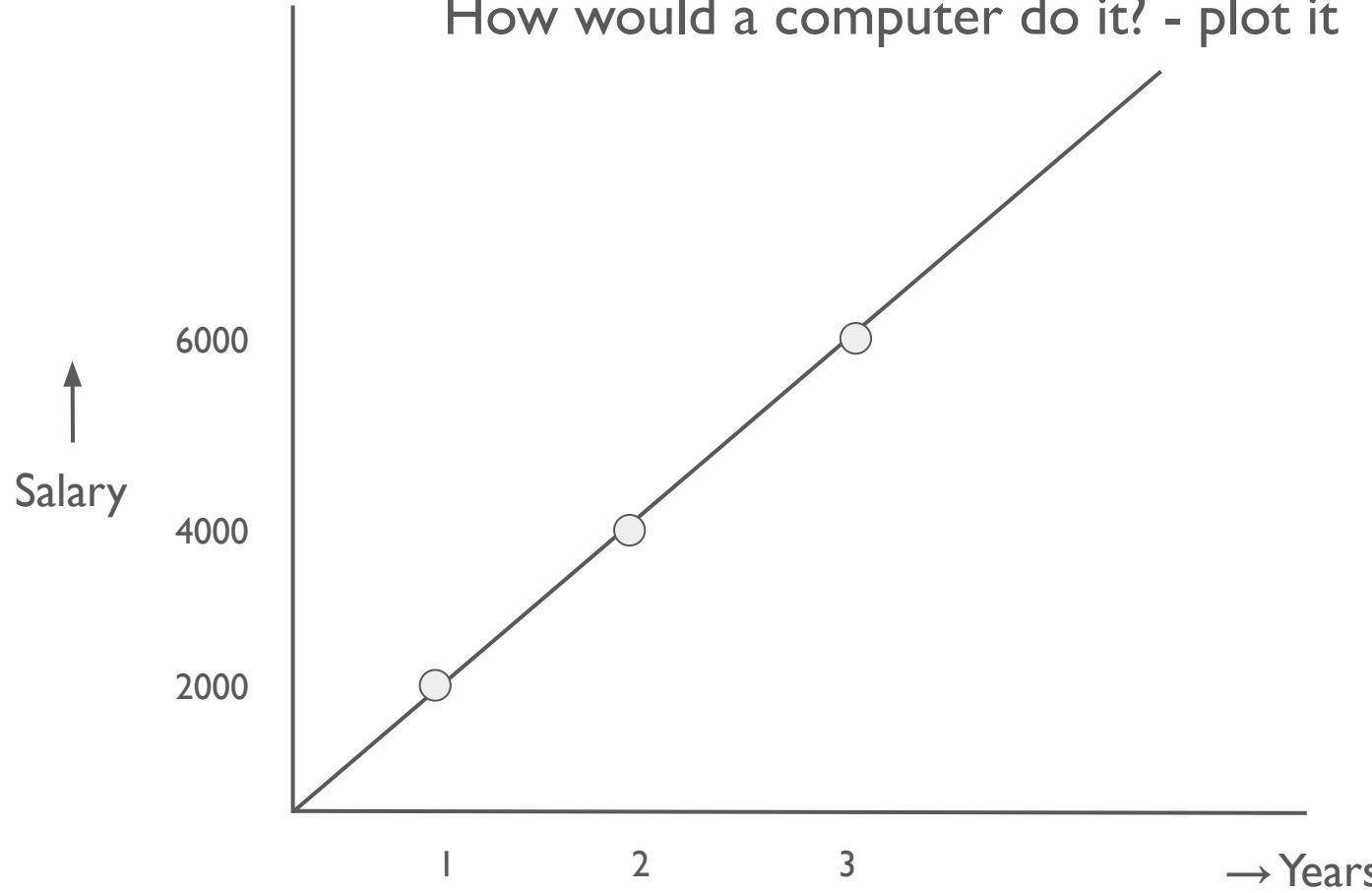
How would a computer do it? - plot it



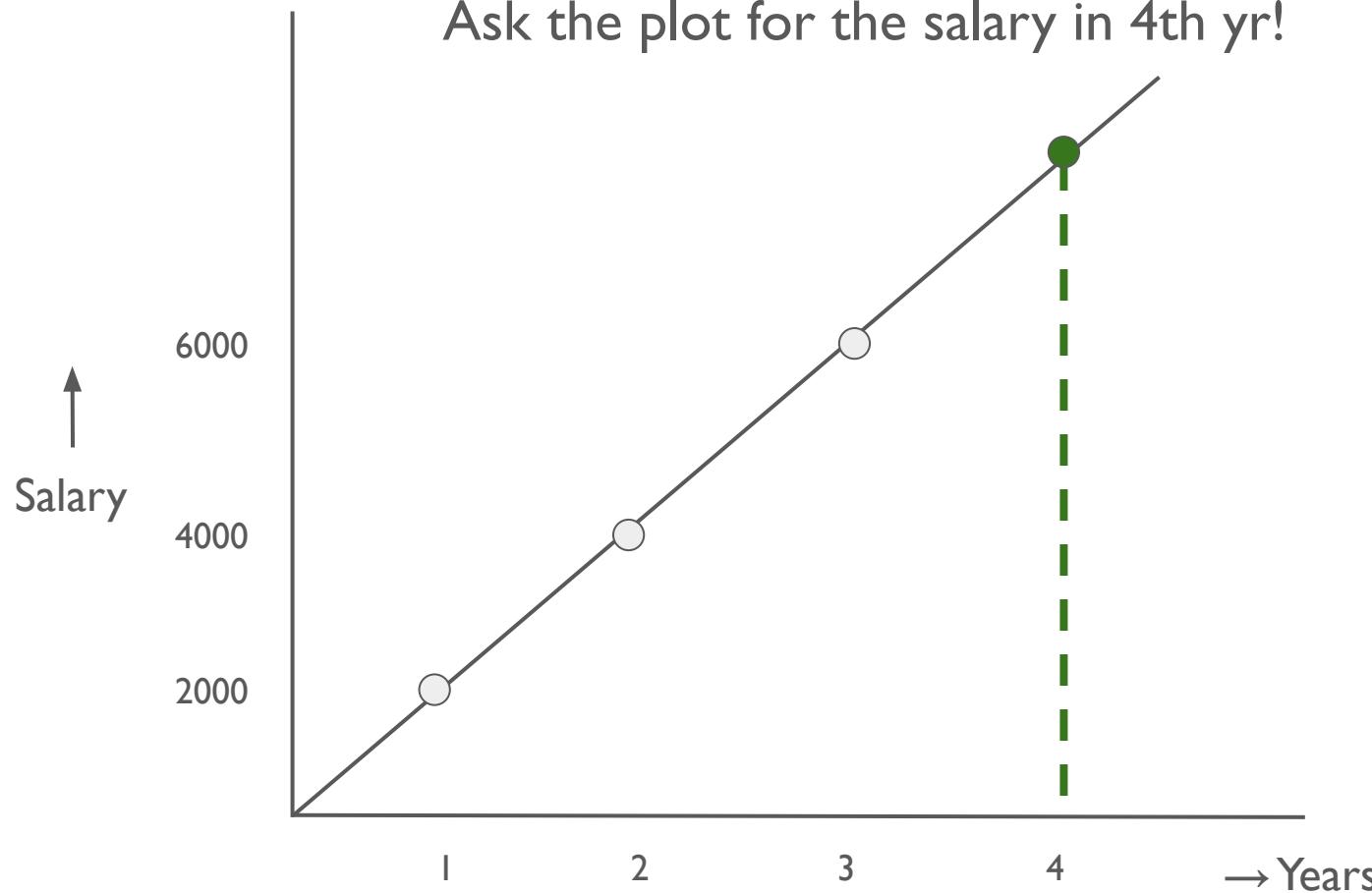
How would a computer do it? - plot it



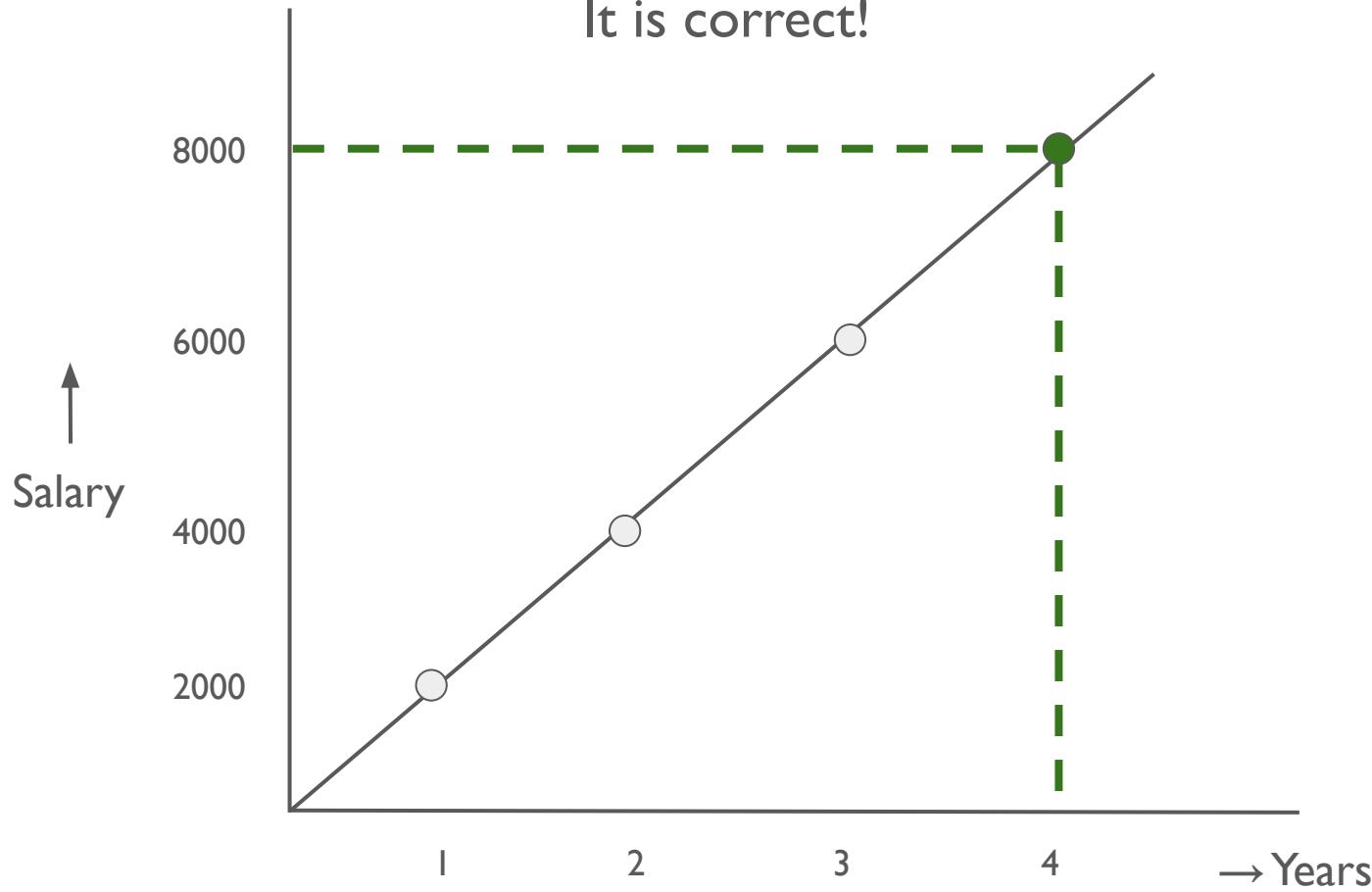
How would a computer do it? - plot it



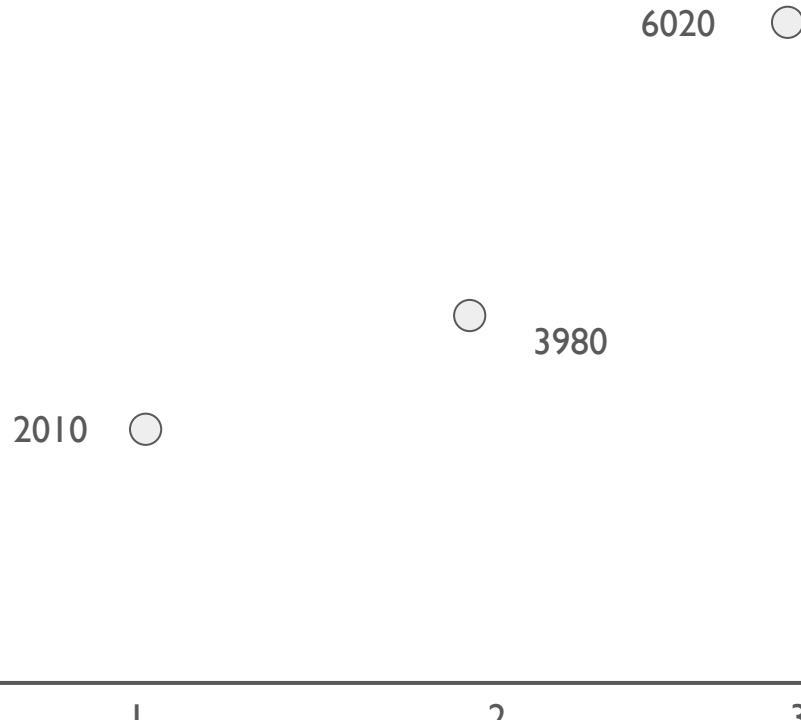
Ask the plot for the salary in 4th yr!



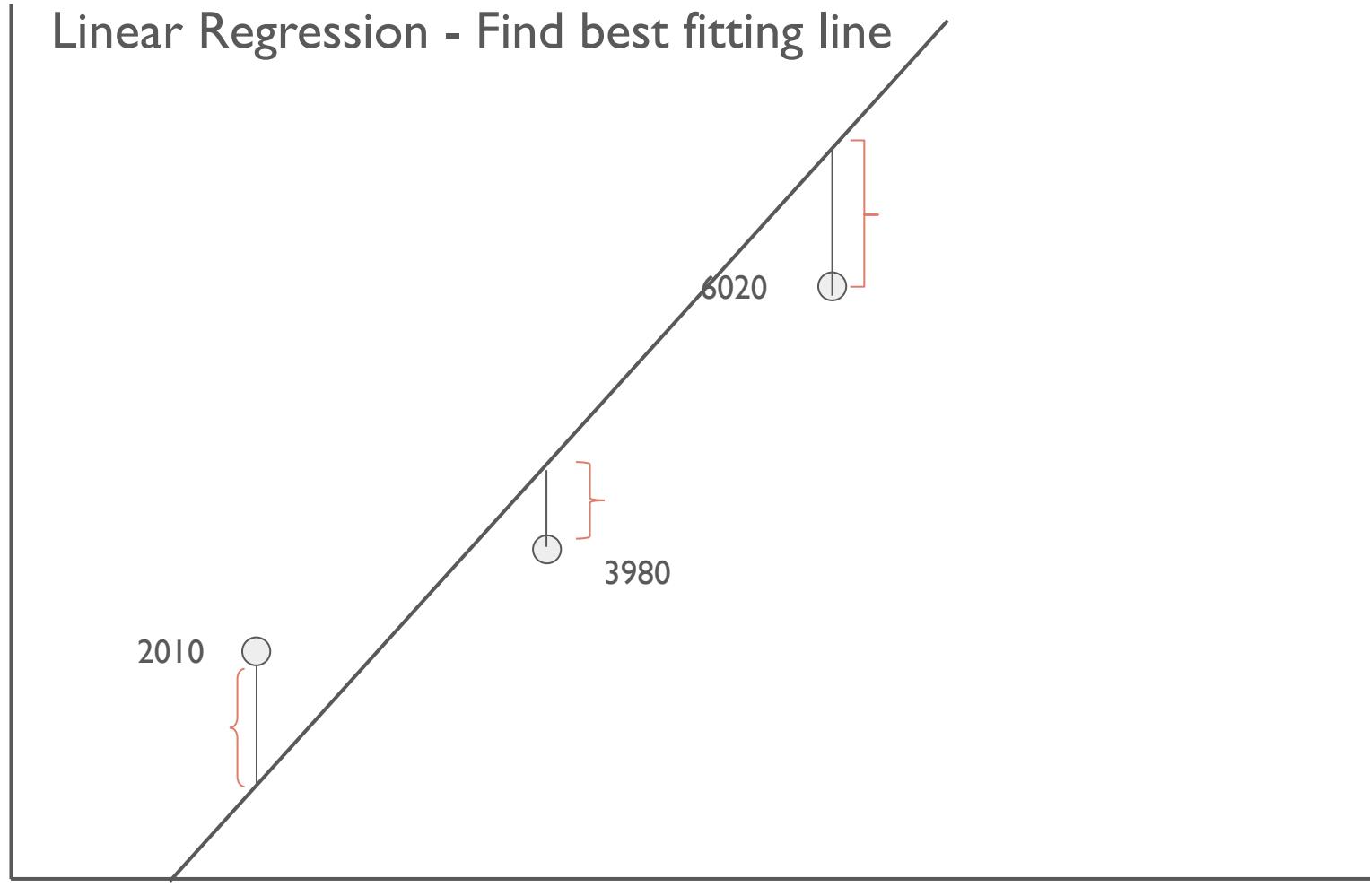
It is correct!



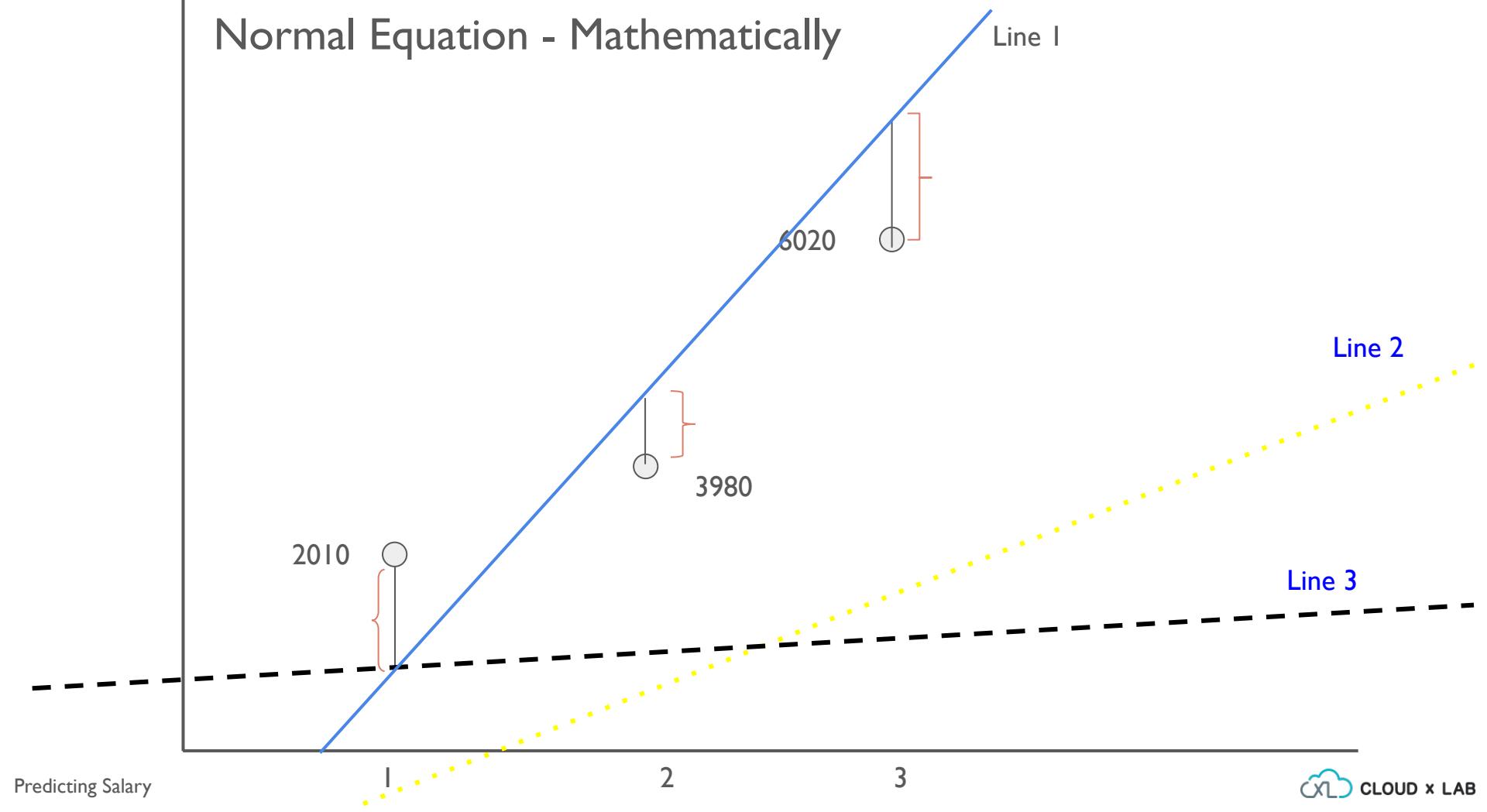
Life isn't perfect! You never get straight line!



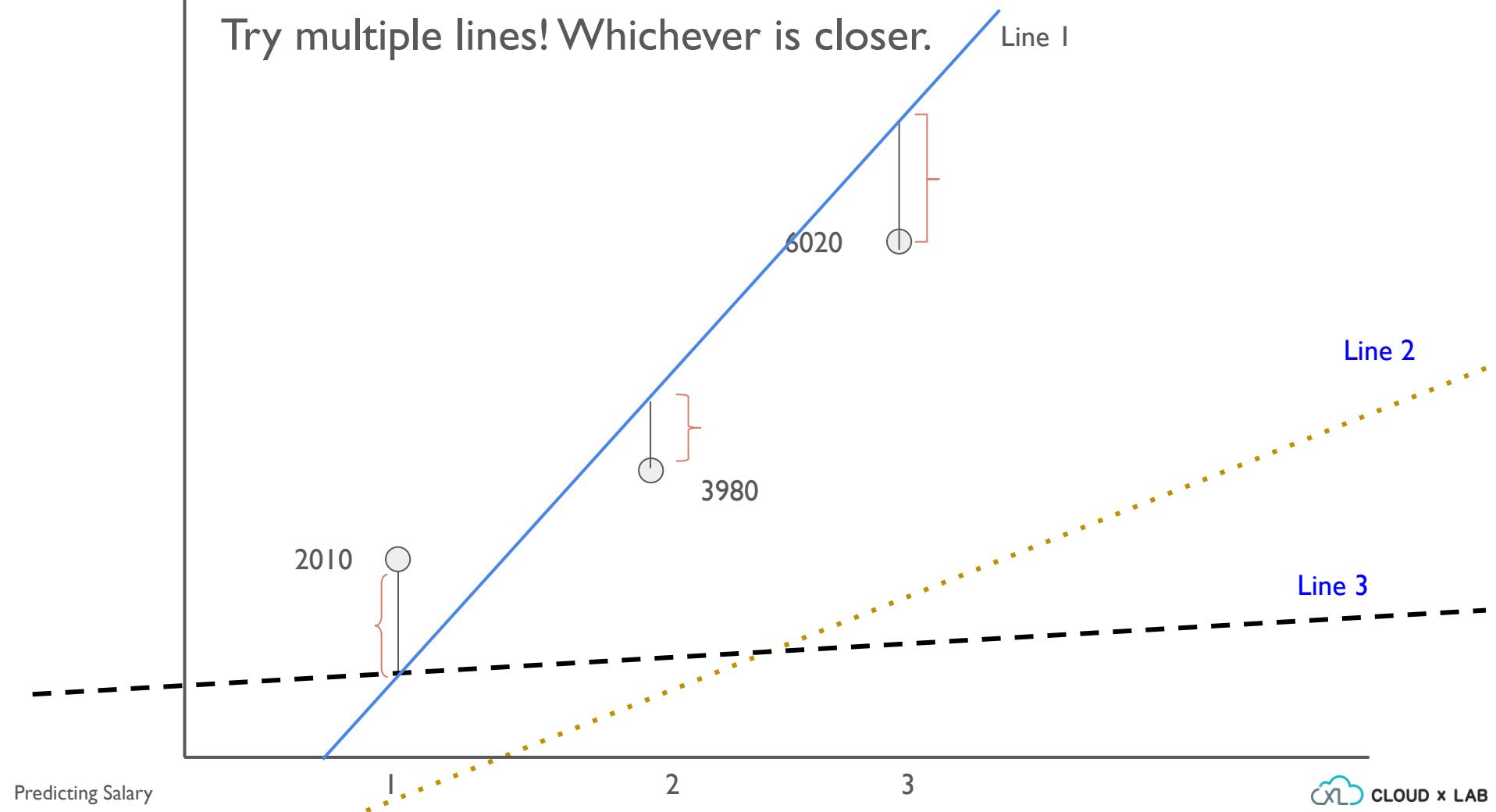
Linear Regression - Find best fitting line



Normal Equation - Mathematically

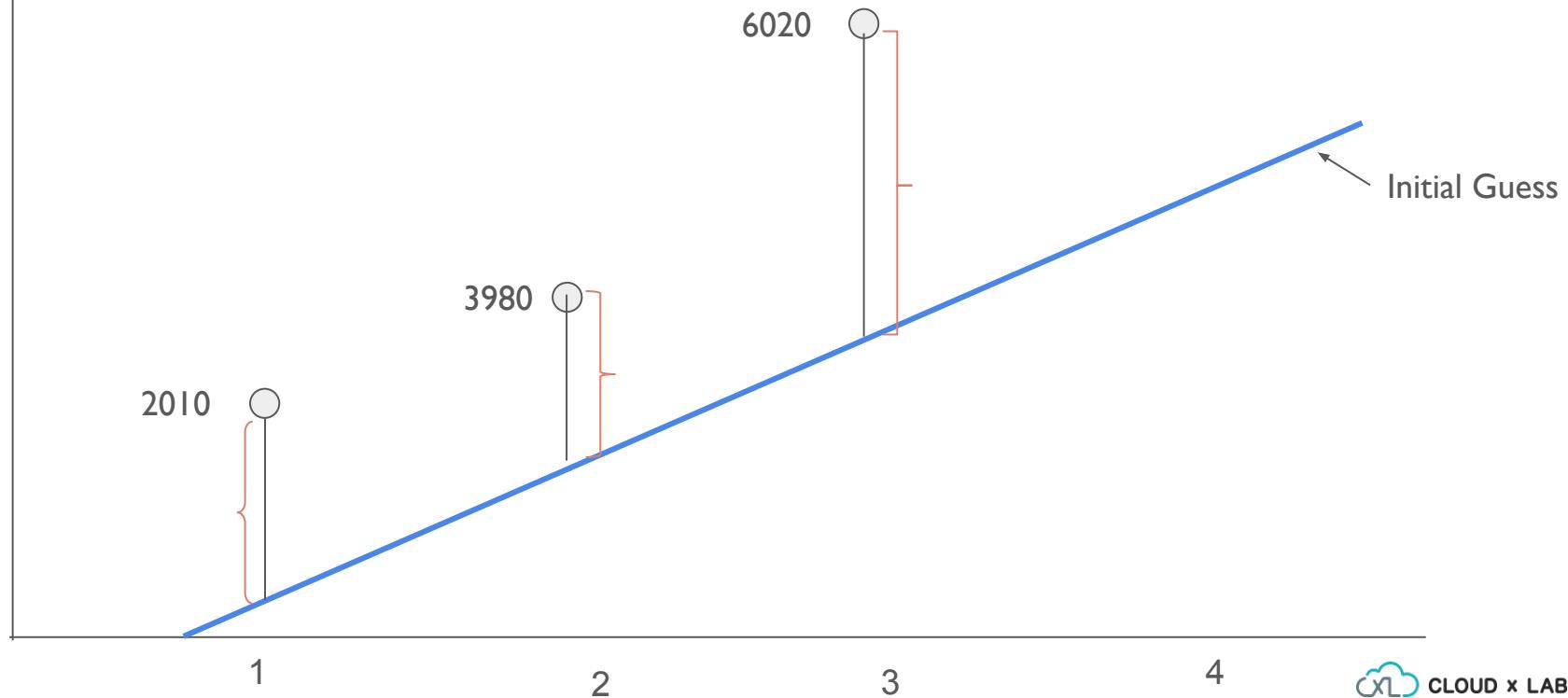


Try multiple lines! Whichever is closer.

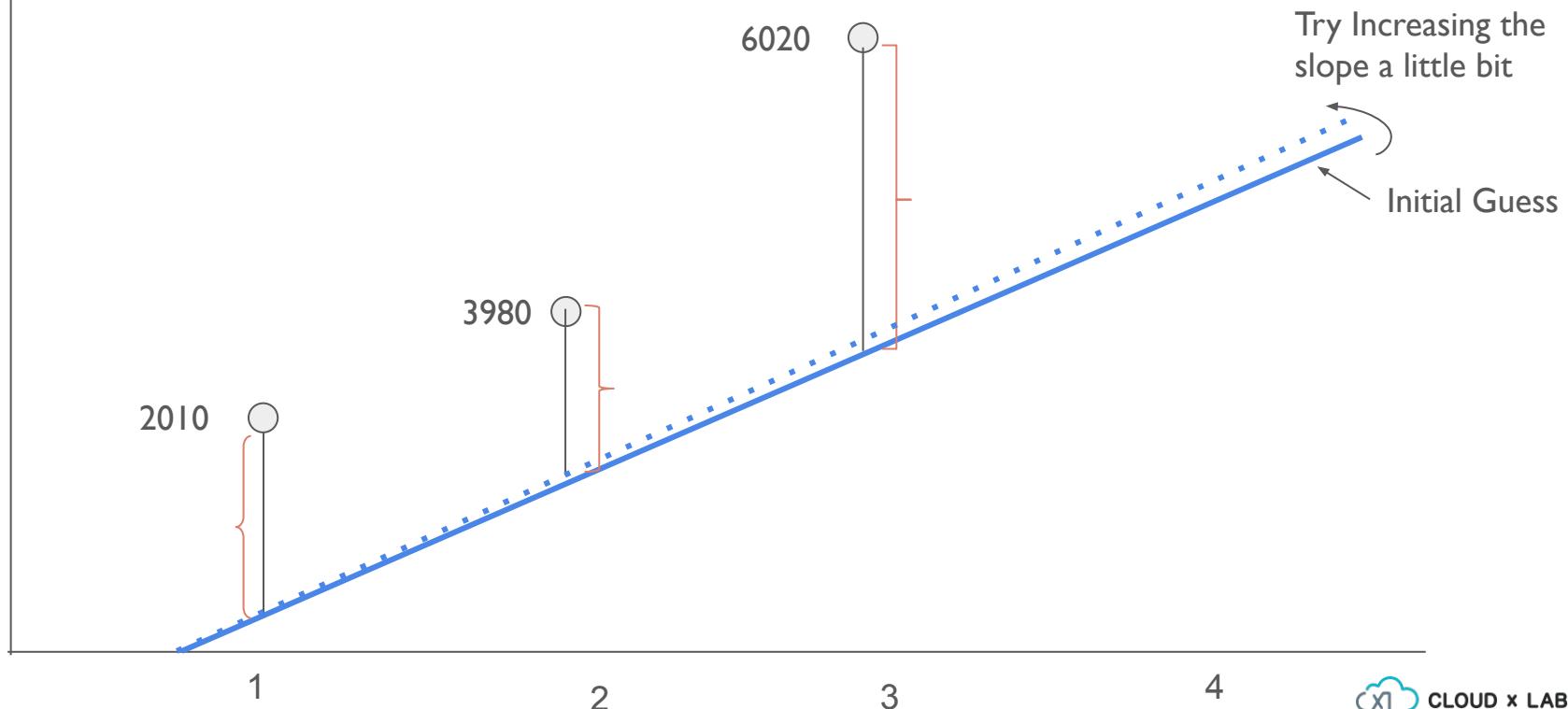


Gradient Descent

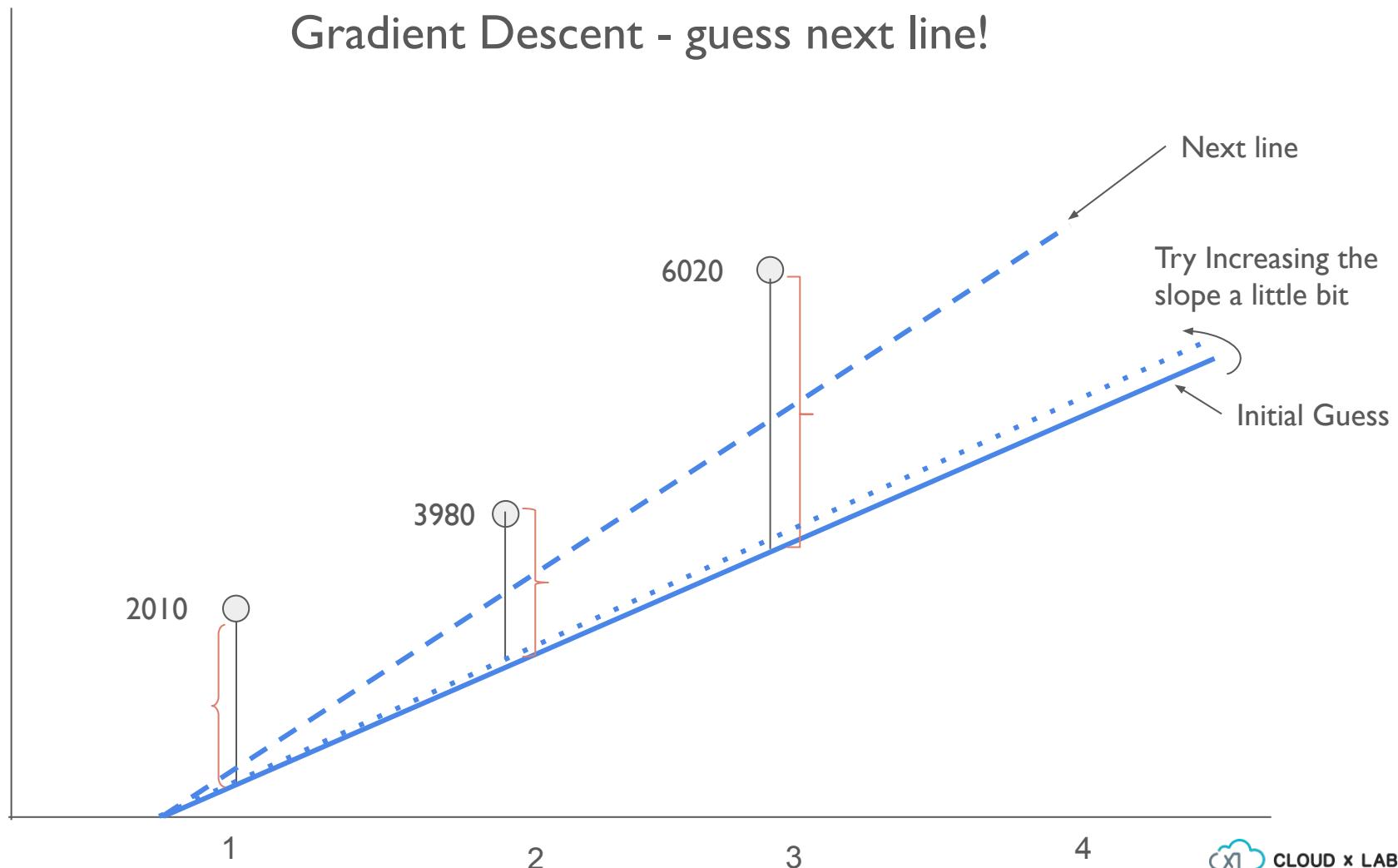
Gradient Descent - A little greedy approach.



Gradient Descent - find impact of slight change

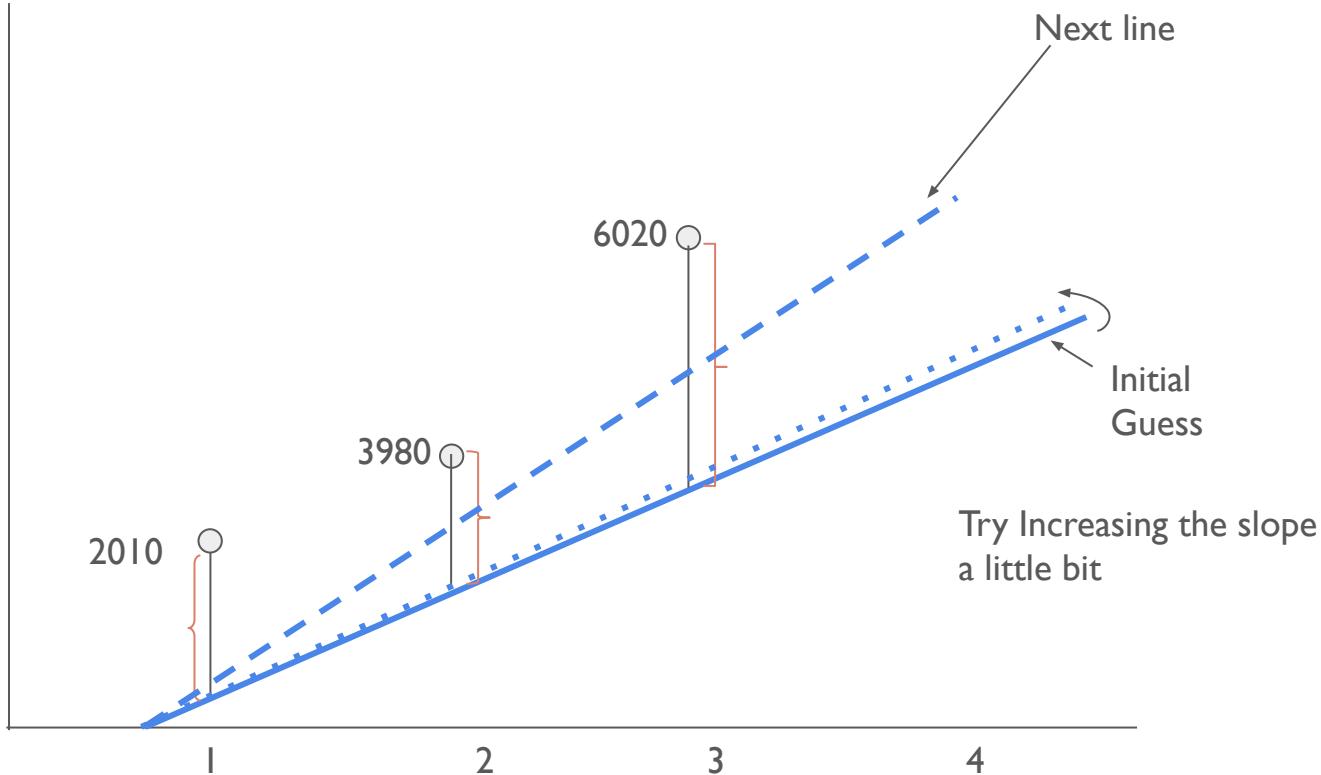


Gradient Descent - guess next line!



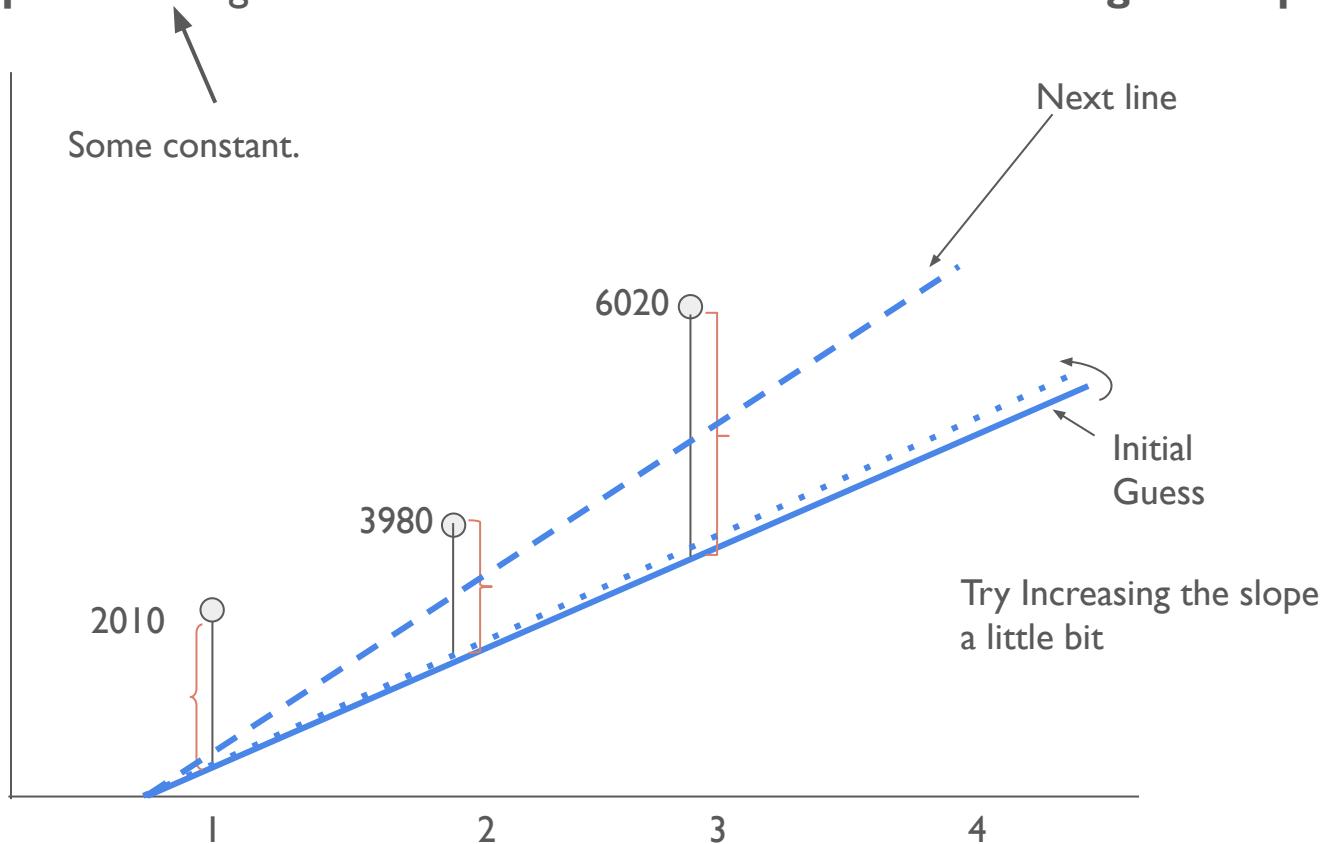
Gradient Descent

Increase in slope is proportional to Rate of decrease of error w.r.t change in slope



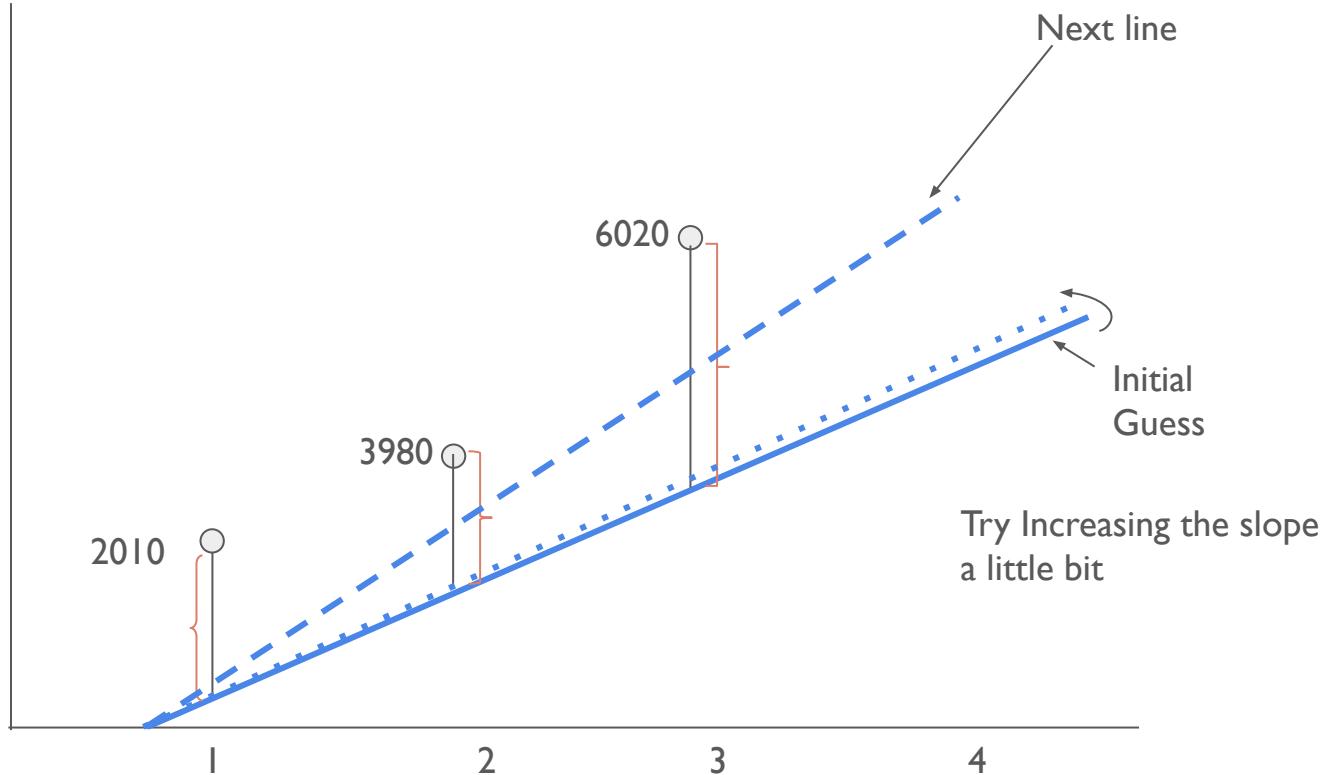
Gradient Descent

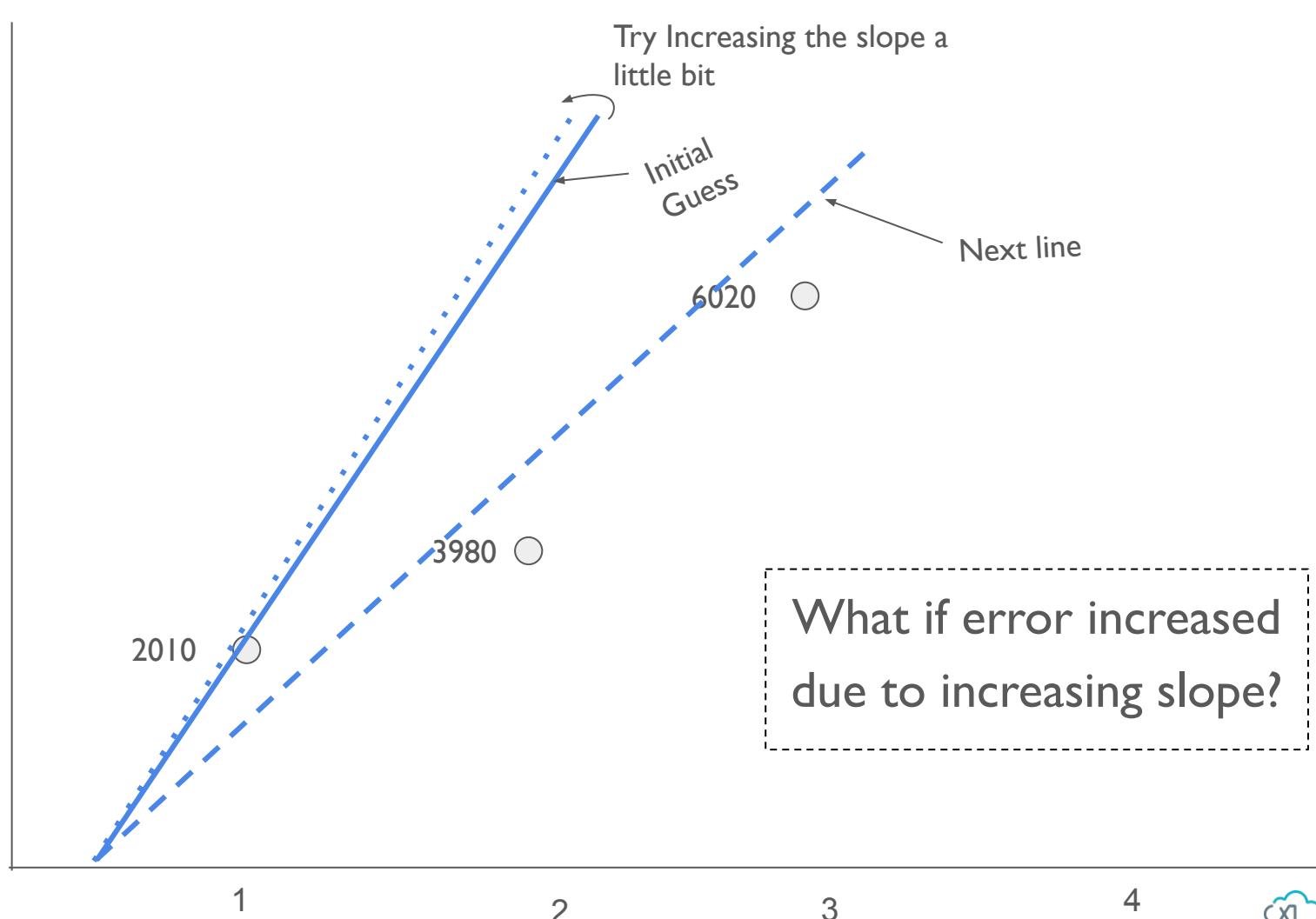
Increase in slope = learning rate * Rate of decrease of error wrt change in slope



Gradient Descent

New Slope = Old Slope + Learning Rate * Rate of decrease of error wrt change in slope

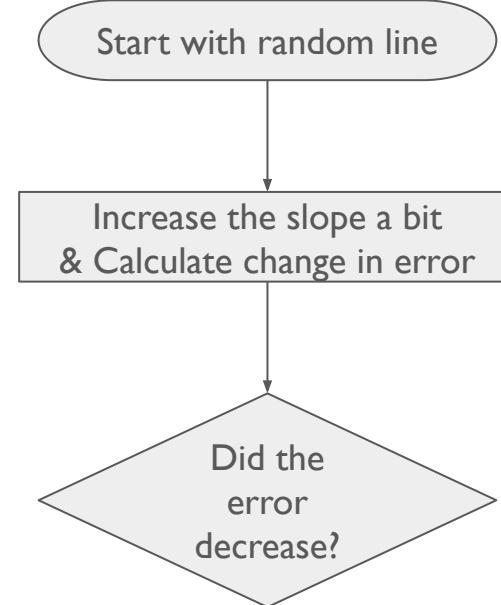




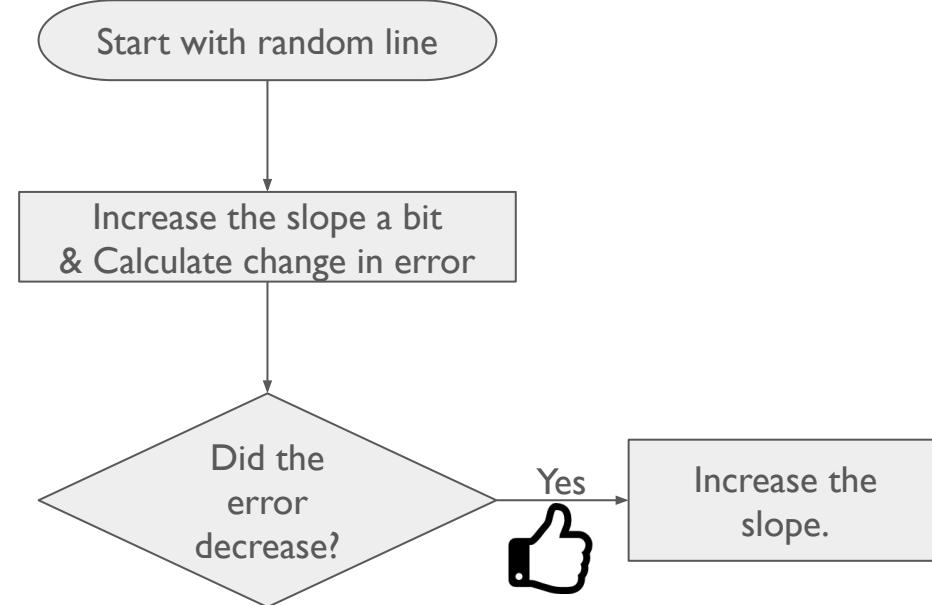
Gradient Descent

Start with random line

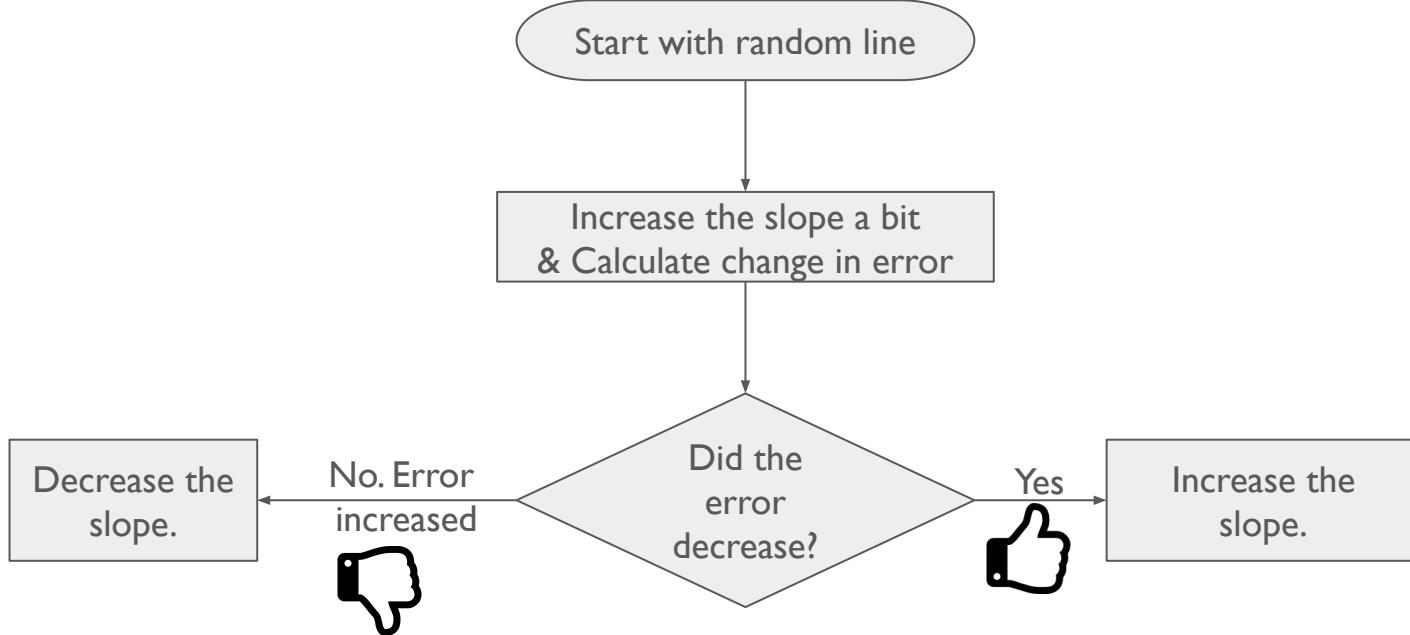
Gradient Descent



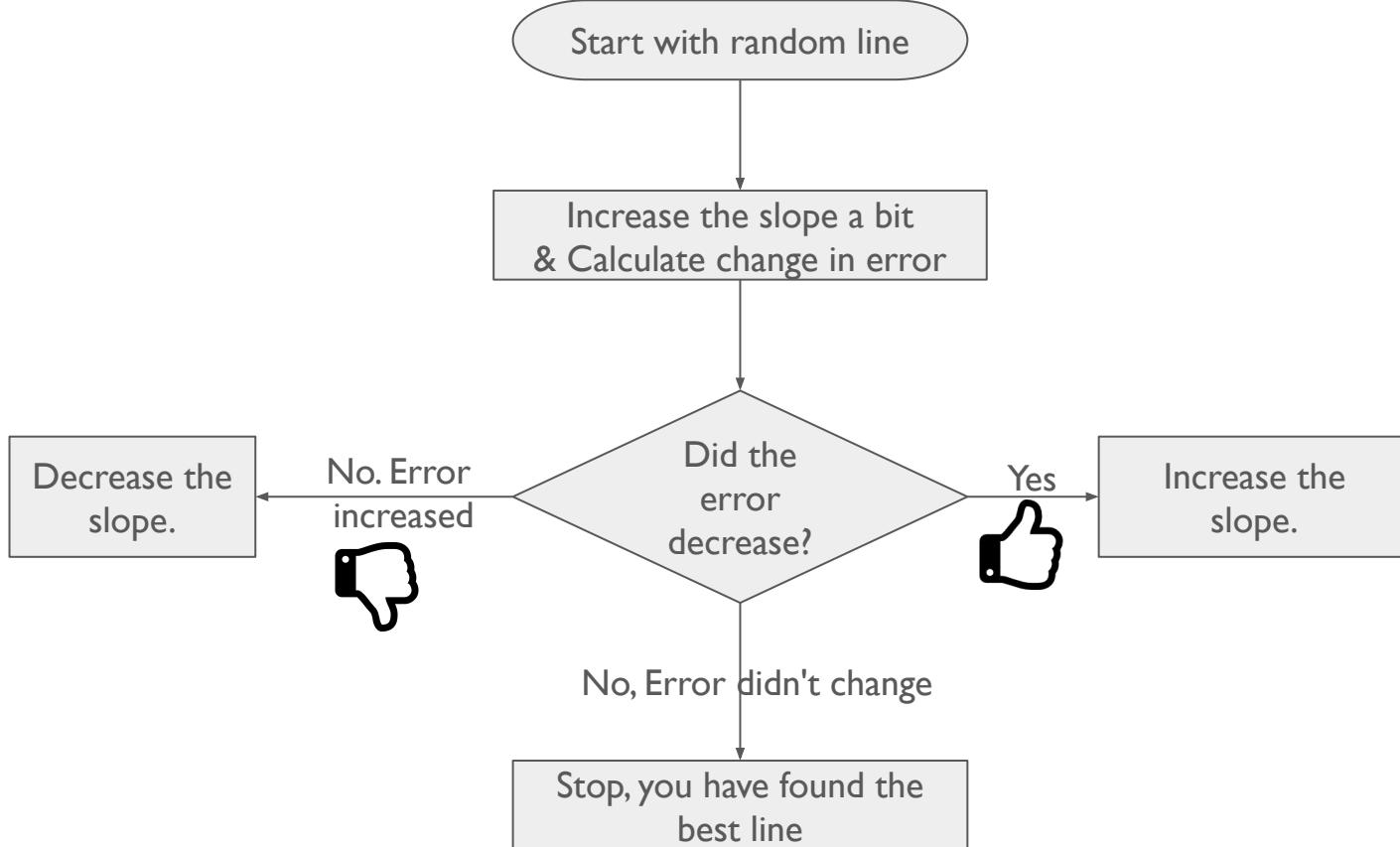
Gradient Descent



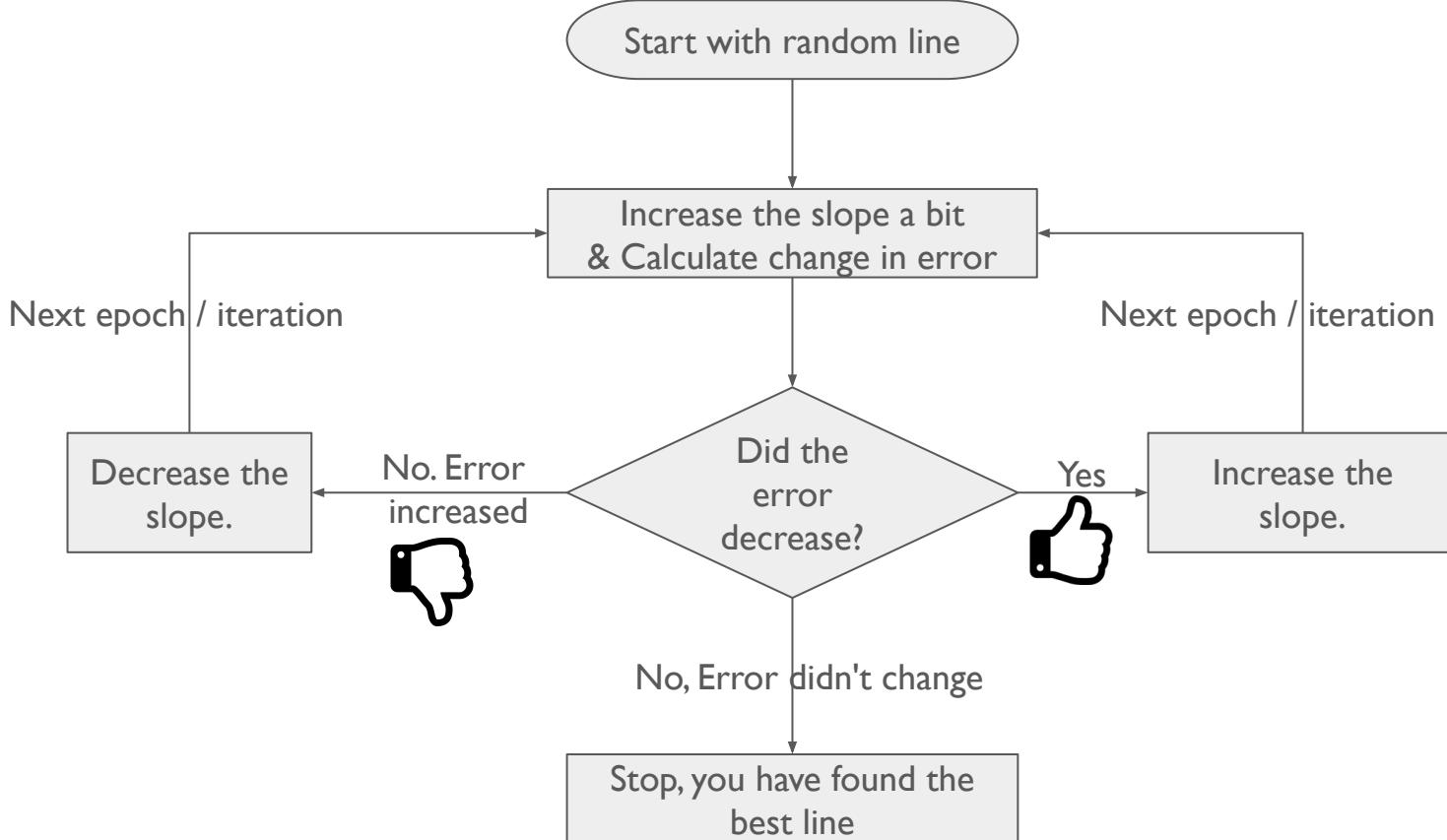
Gradient Descent



Gradient Descent



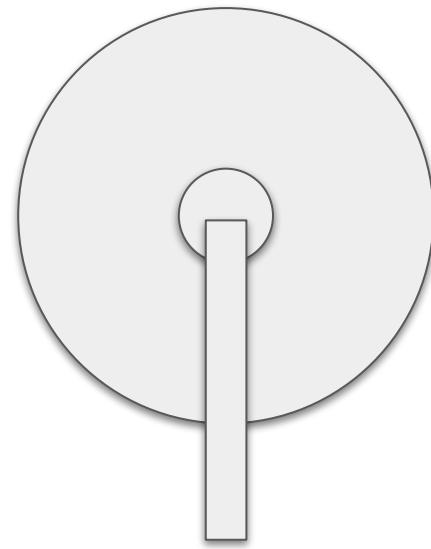
Gradient Descent



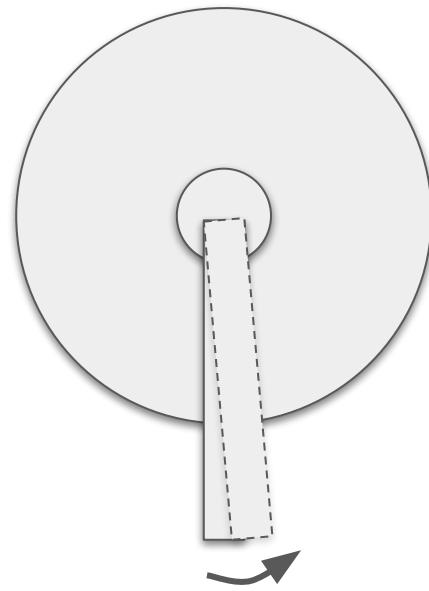
Gradient Descent



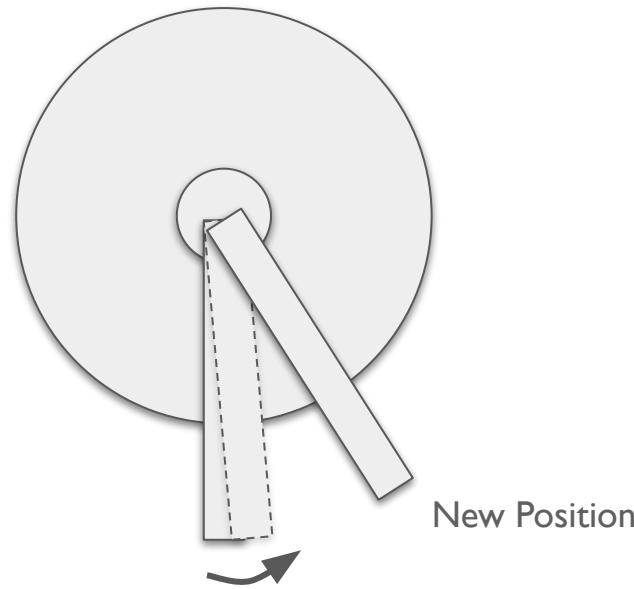
Gradient Descent



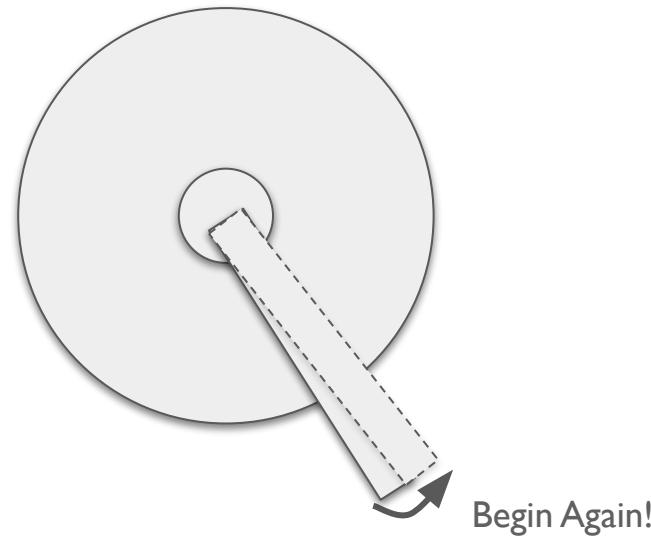
Gradient Descent



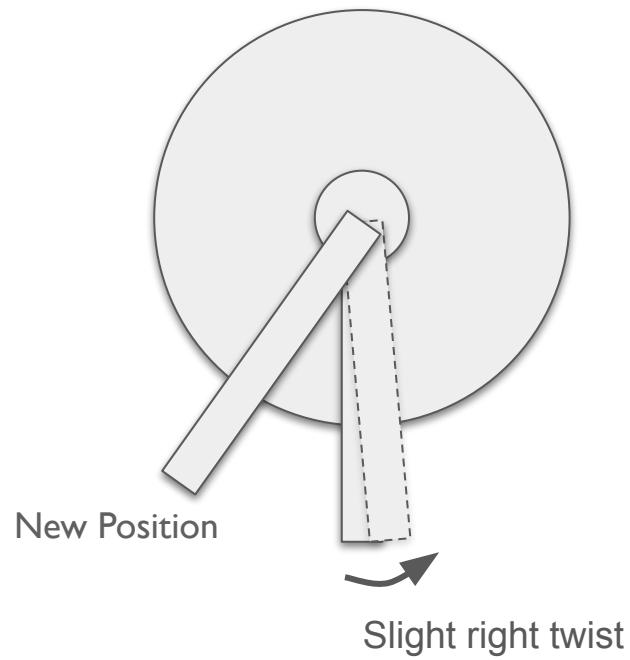
Gradient Descent



Gradient Descent



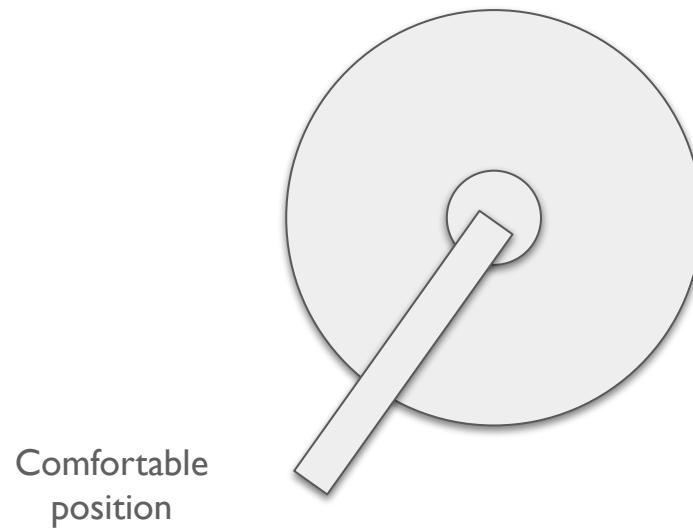
Gradient Descent



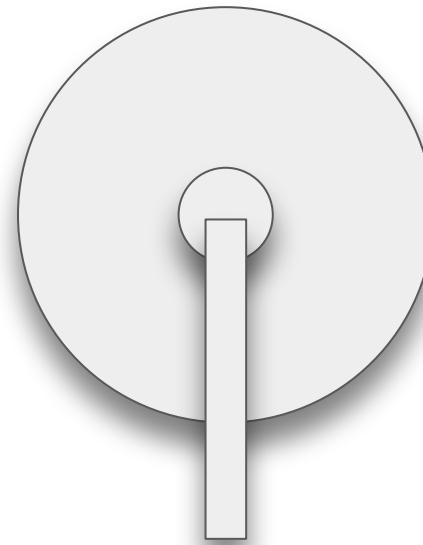
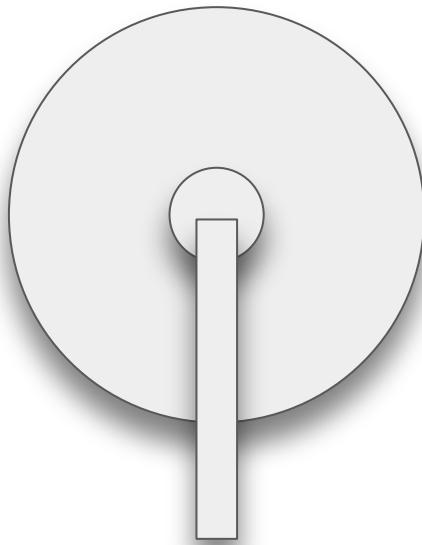
New Position

Slight right twist

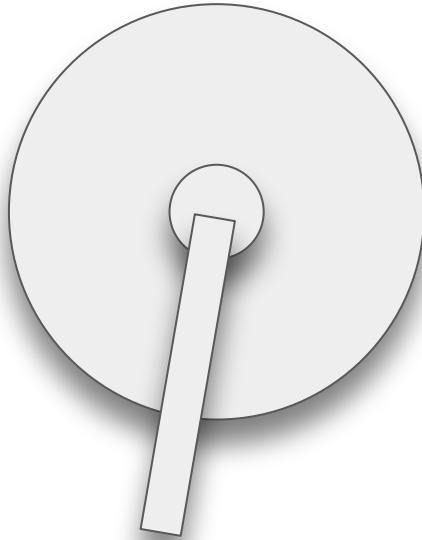
What if we had two knob?

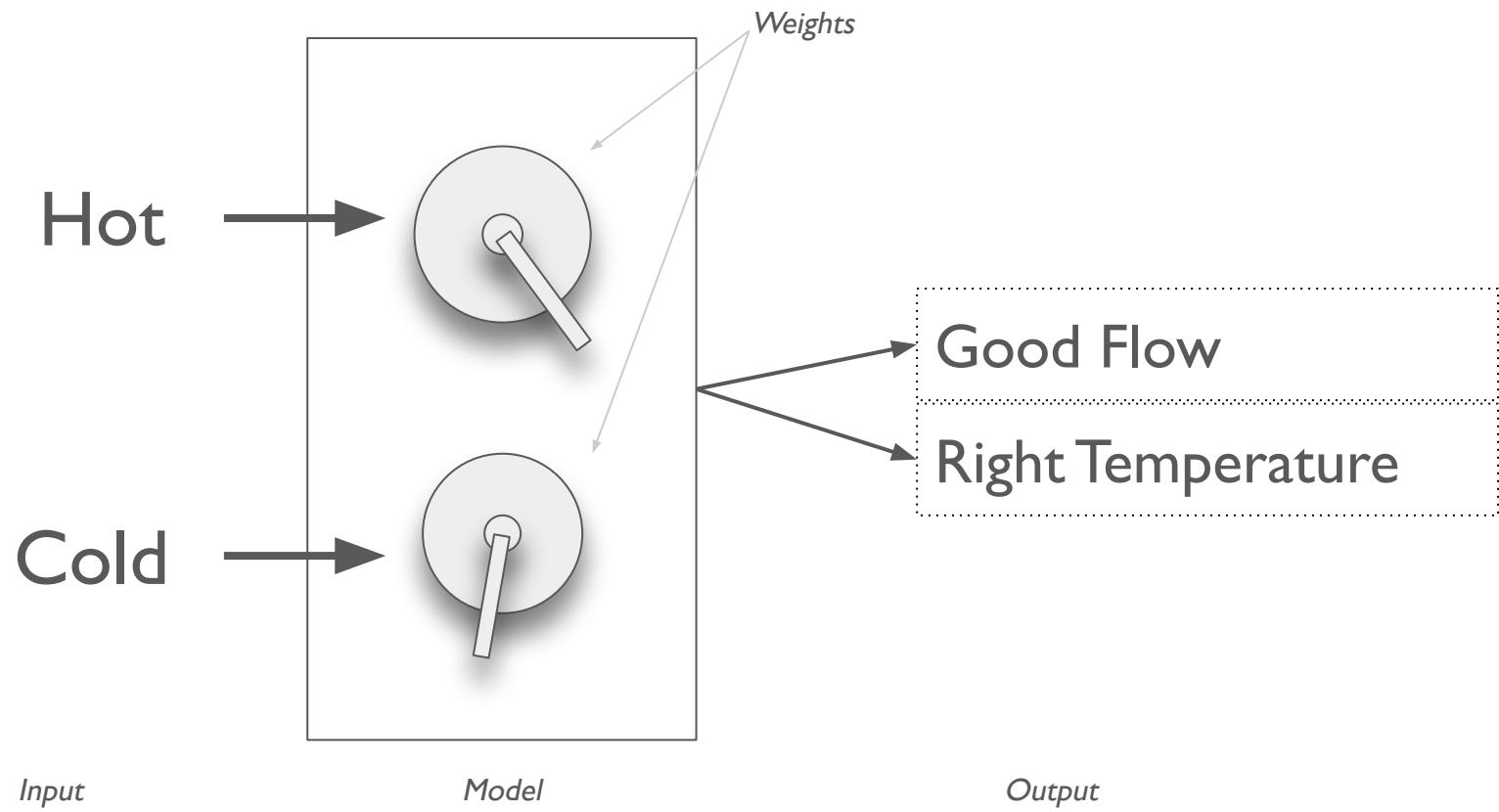


What if we had two knob?

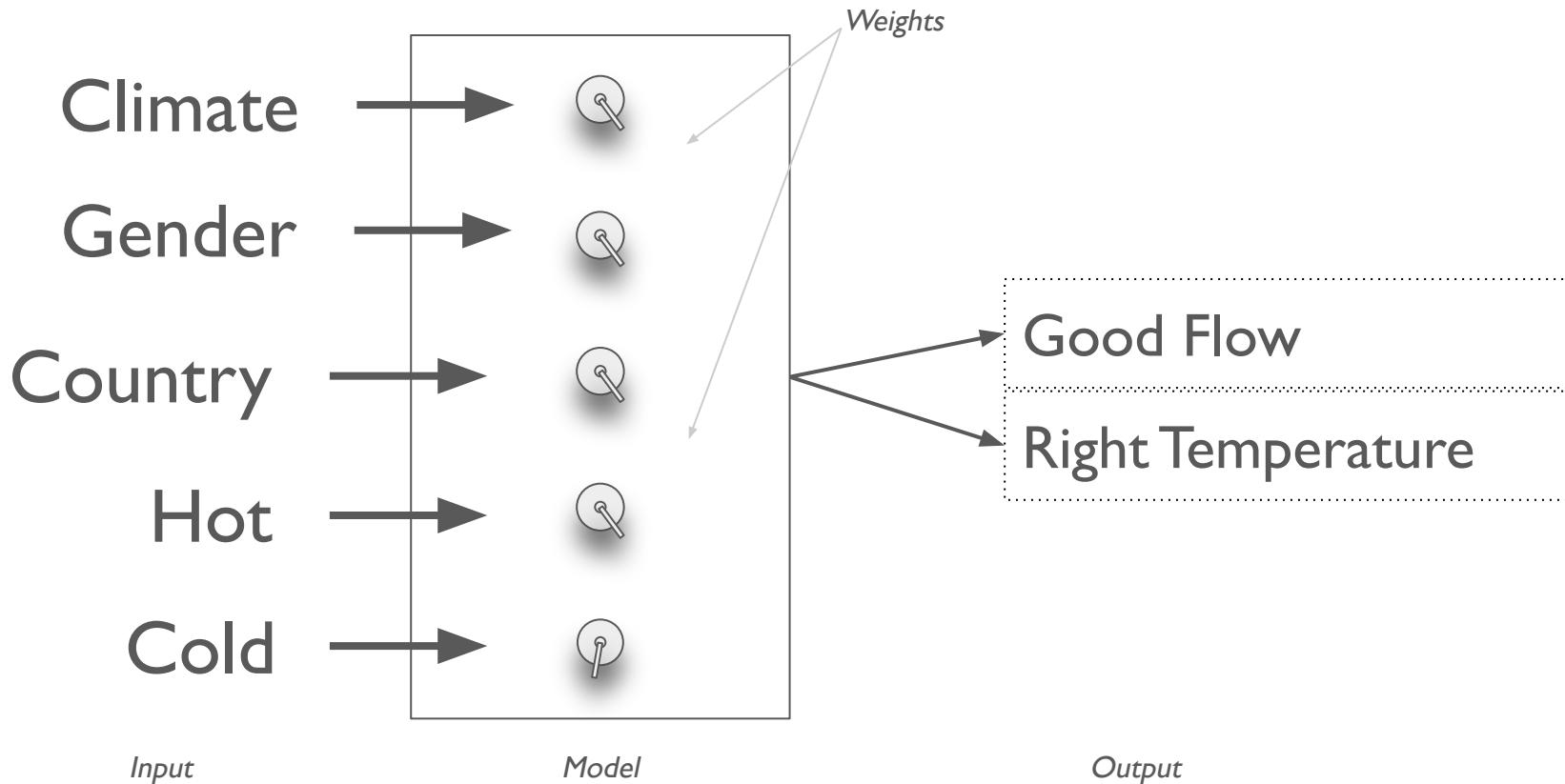


Gradient Descent





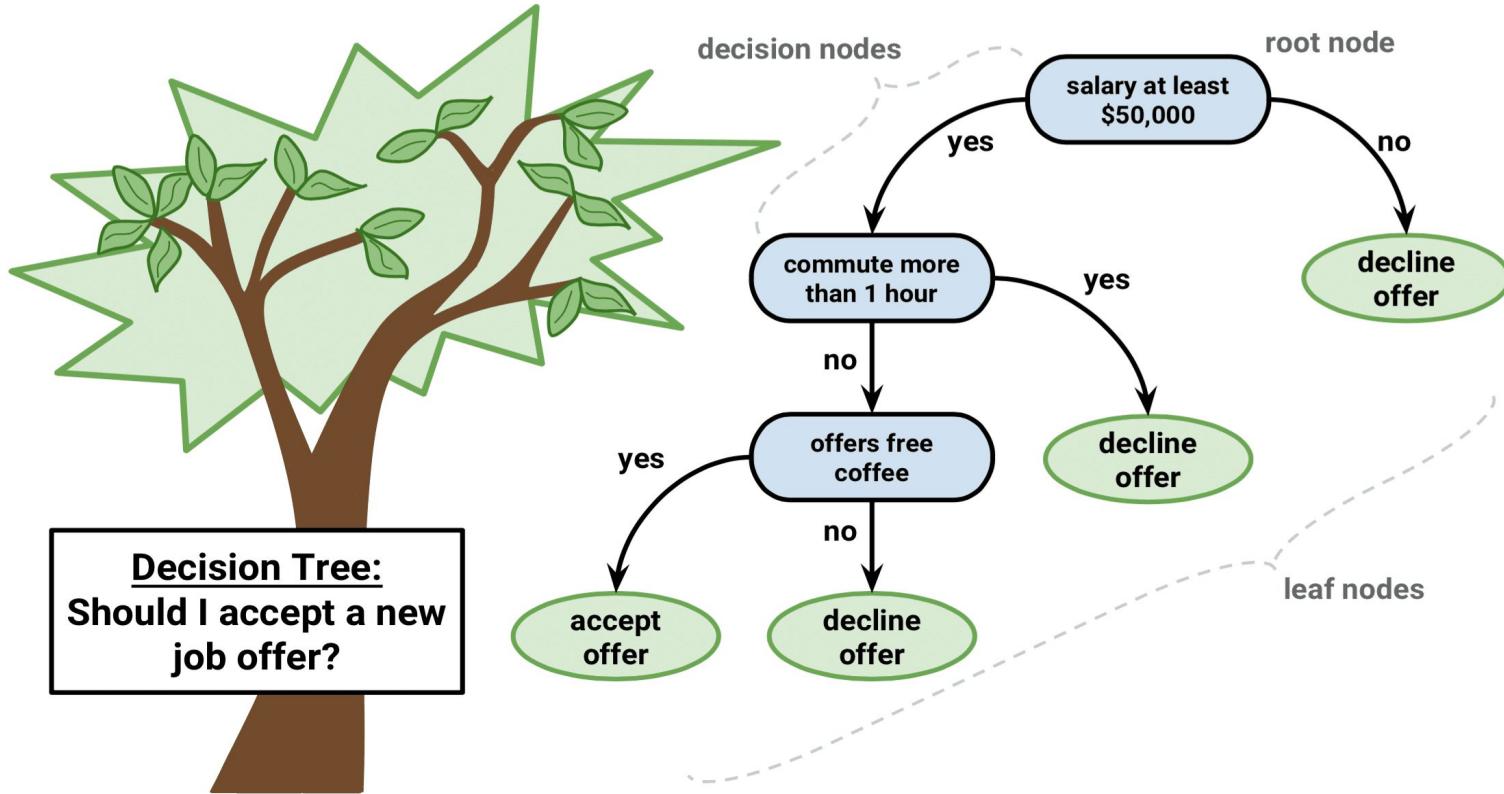
What if we had more input features?





Decision Trees

Decision Tree

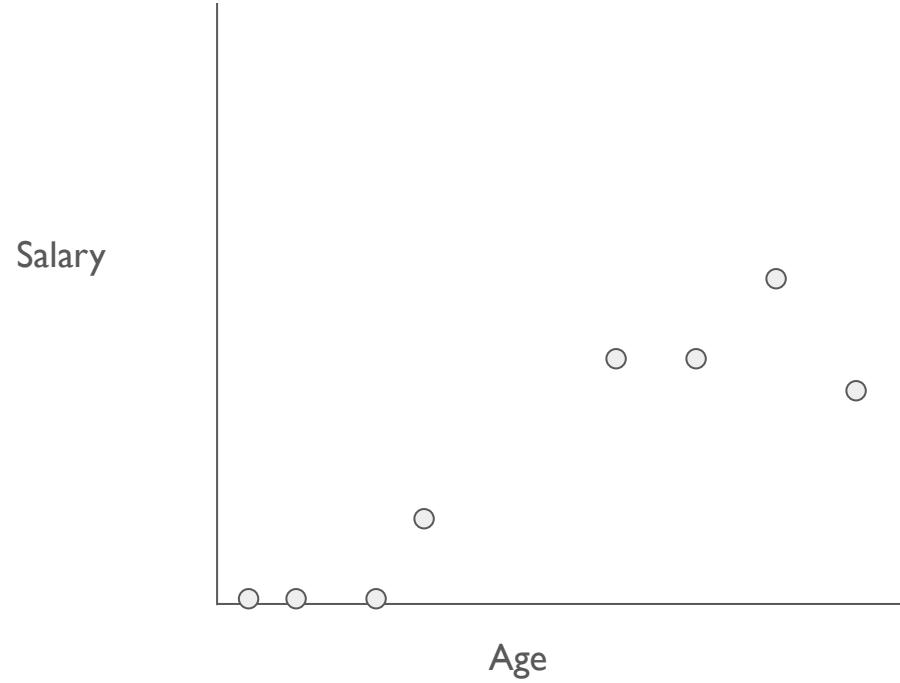


Decision Tree

Age	Salary
10	0
5	0
2	0
13	5
25	20
30	20
35	25
40	18
age	?

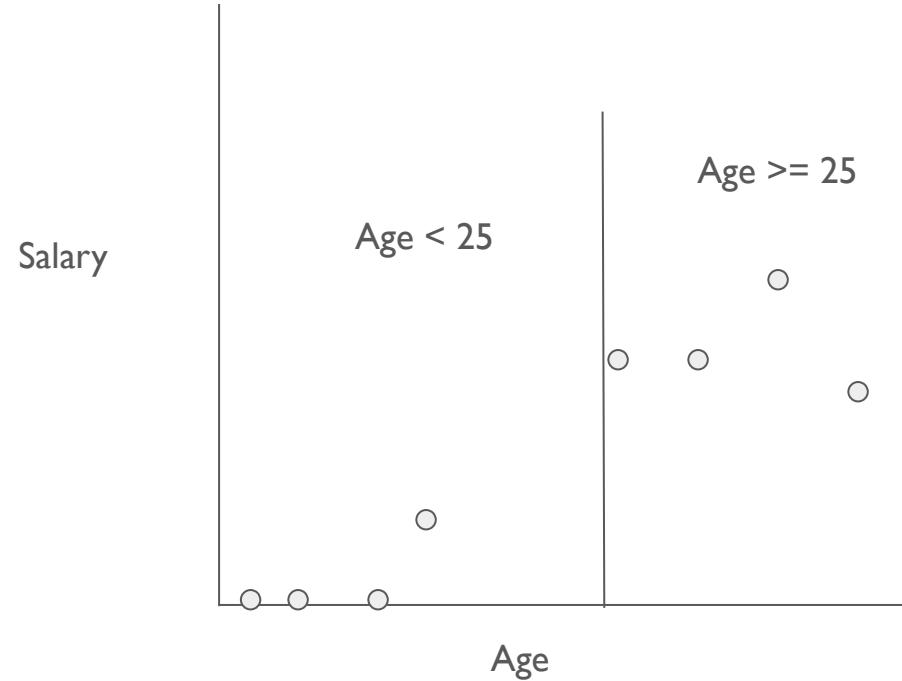
Decision Tree

Age	Salary
10	0
5	0
2	0
13	5
25	20
30	20
35	25
40	18
age	?



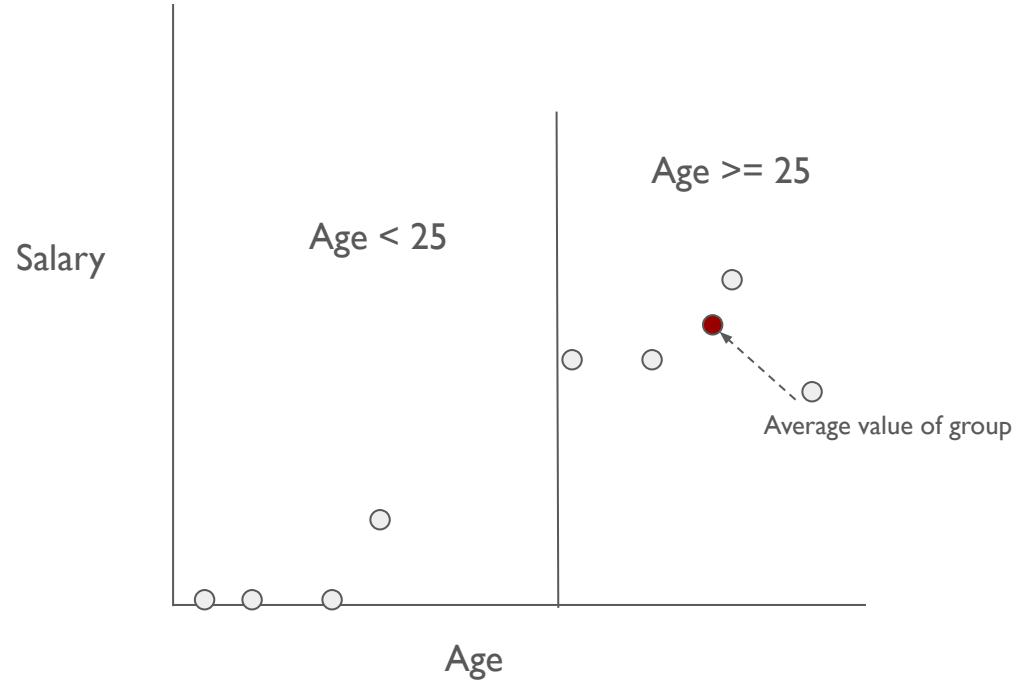
Decision Tree

Age	Salary
10	0
5	0
2	0
13	5
25	20
30	20
35	25
40	18
age	?



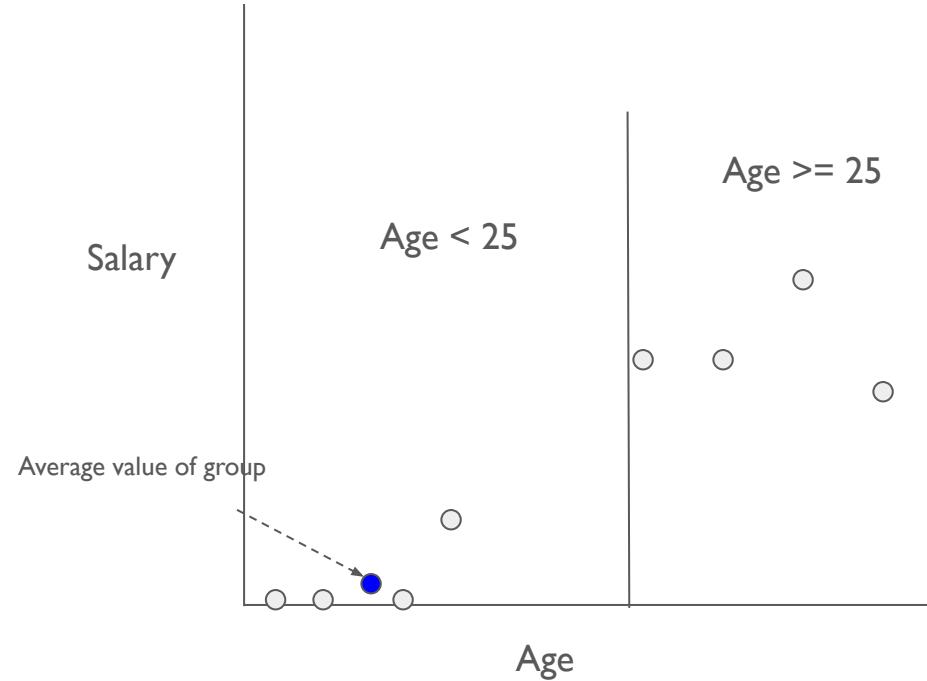
Decision Tree

Age	Salary
10	0
5	0
2	0
13	5
25	20
30	20
35	25
40	18
age	?

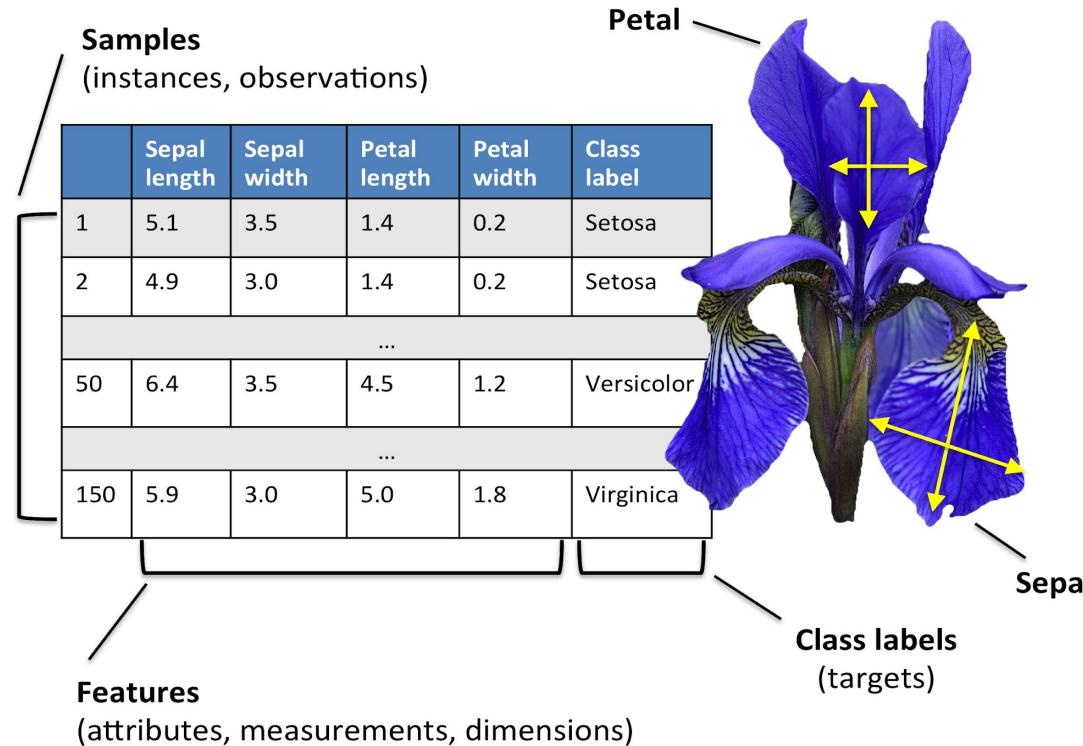


Decision Tree

Age	Salary
10	0
5	0
2	0
13	5
25	20
30	20
35	25
40	18
age	?

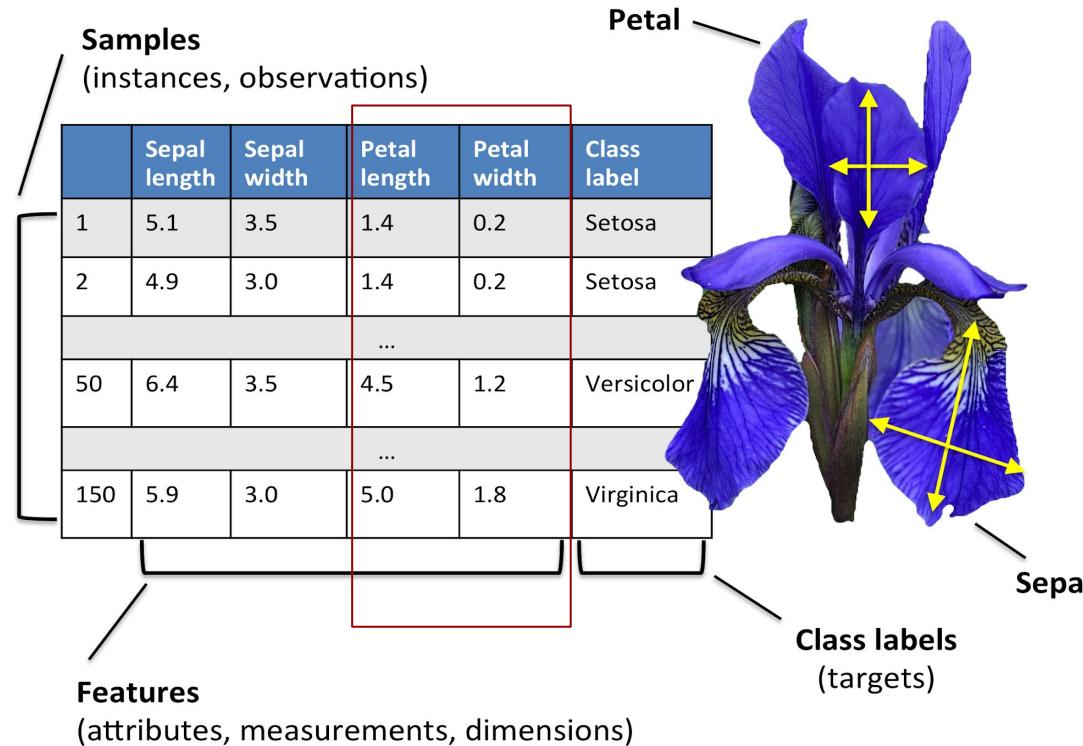


IRIS dataset



We'll only use two features i.e. petal length and petal width.

IRIS dataset

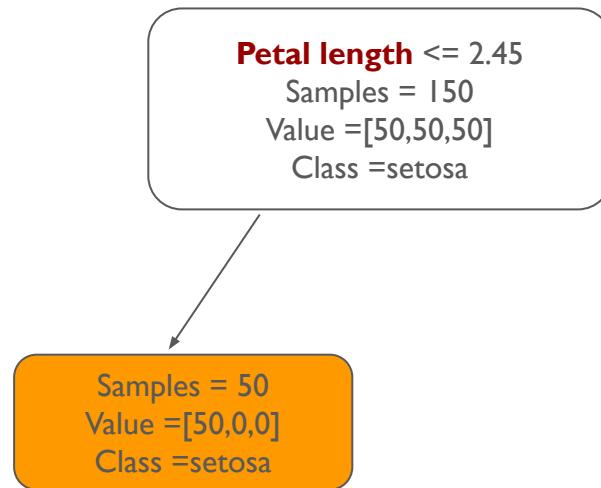


We'll only use two features i.e. petal length and petal width.

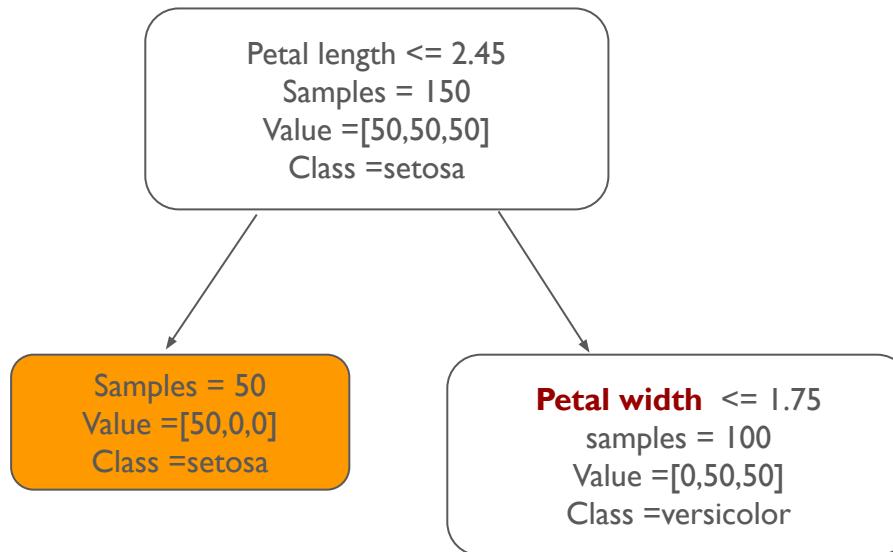
Decision Tree - IRIS dataset - Model

Petal length <= 2.45
Samples = 150
Value =[50,50,50]
Class =setosa

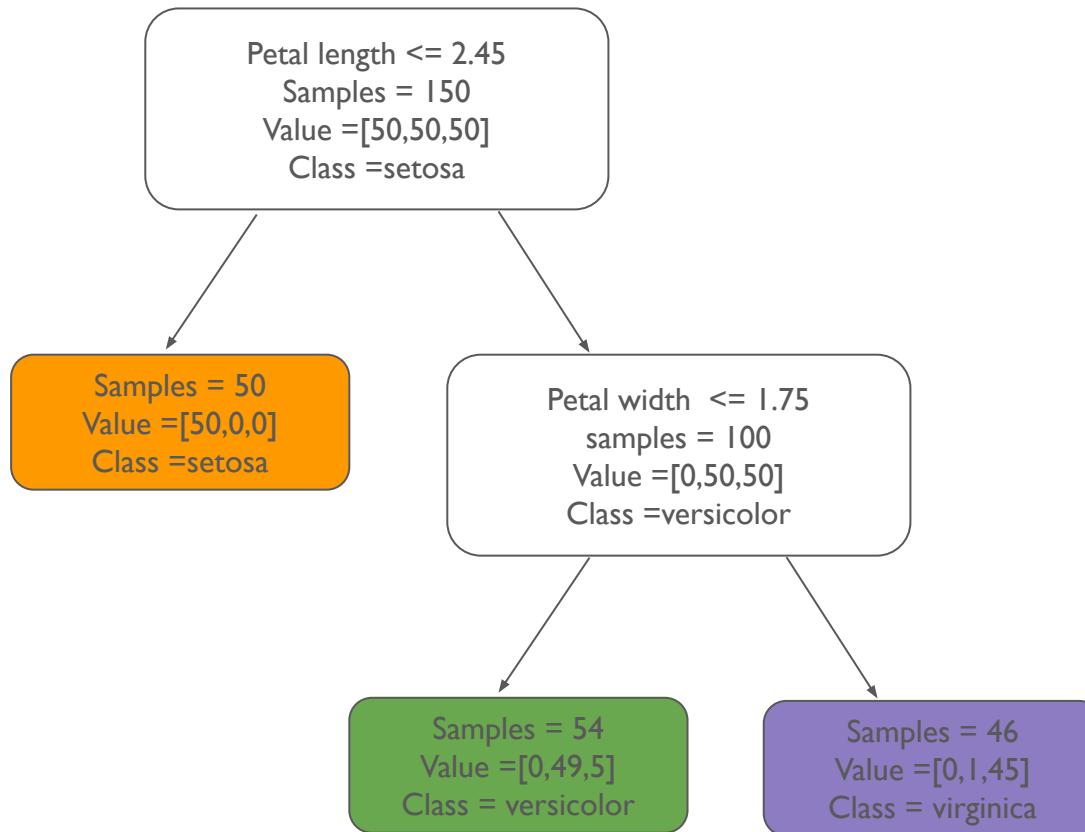
Decision Tree - IRIS dataset - Model



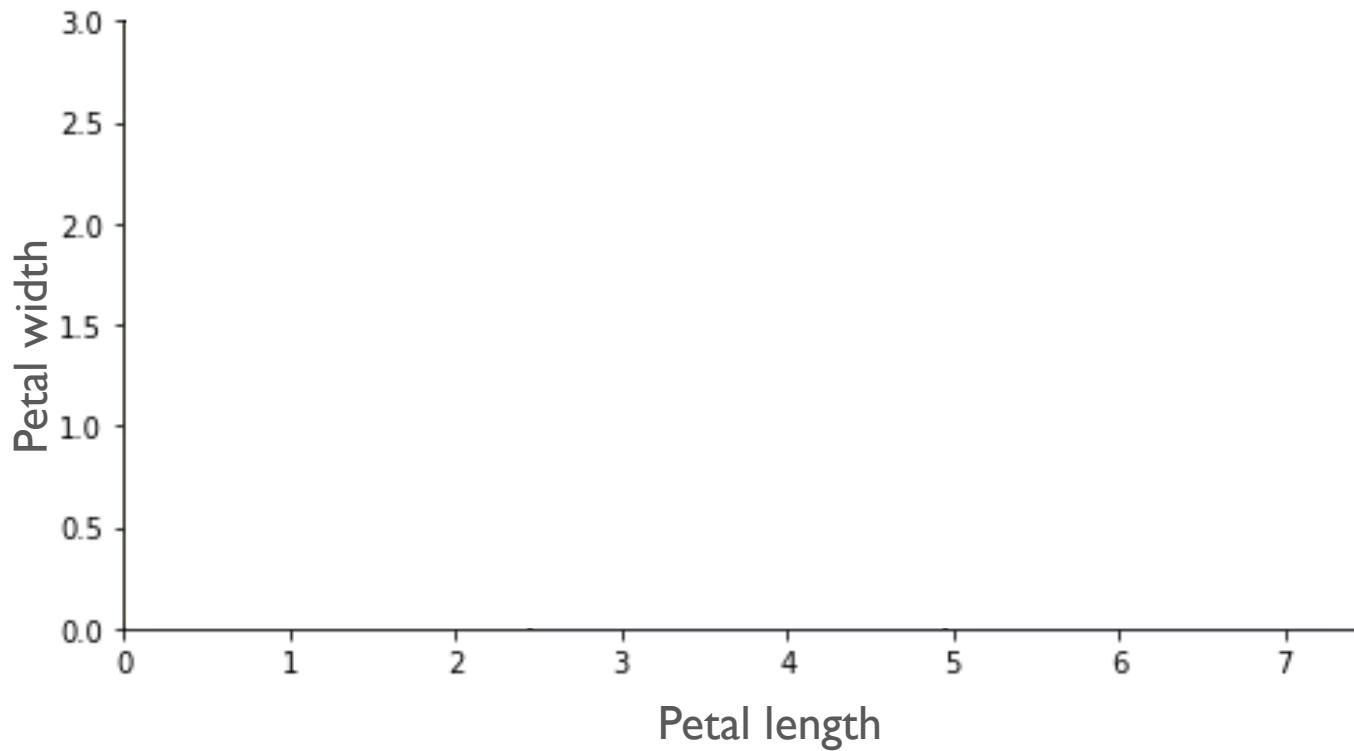
Decision Tree - IRIS dataset - Model



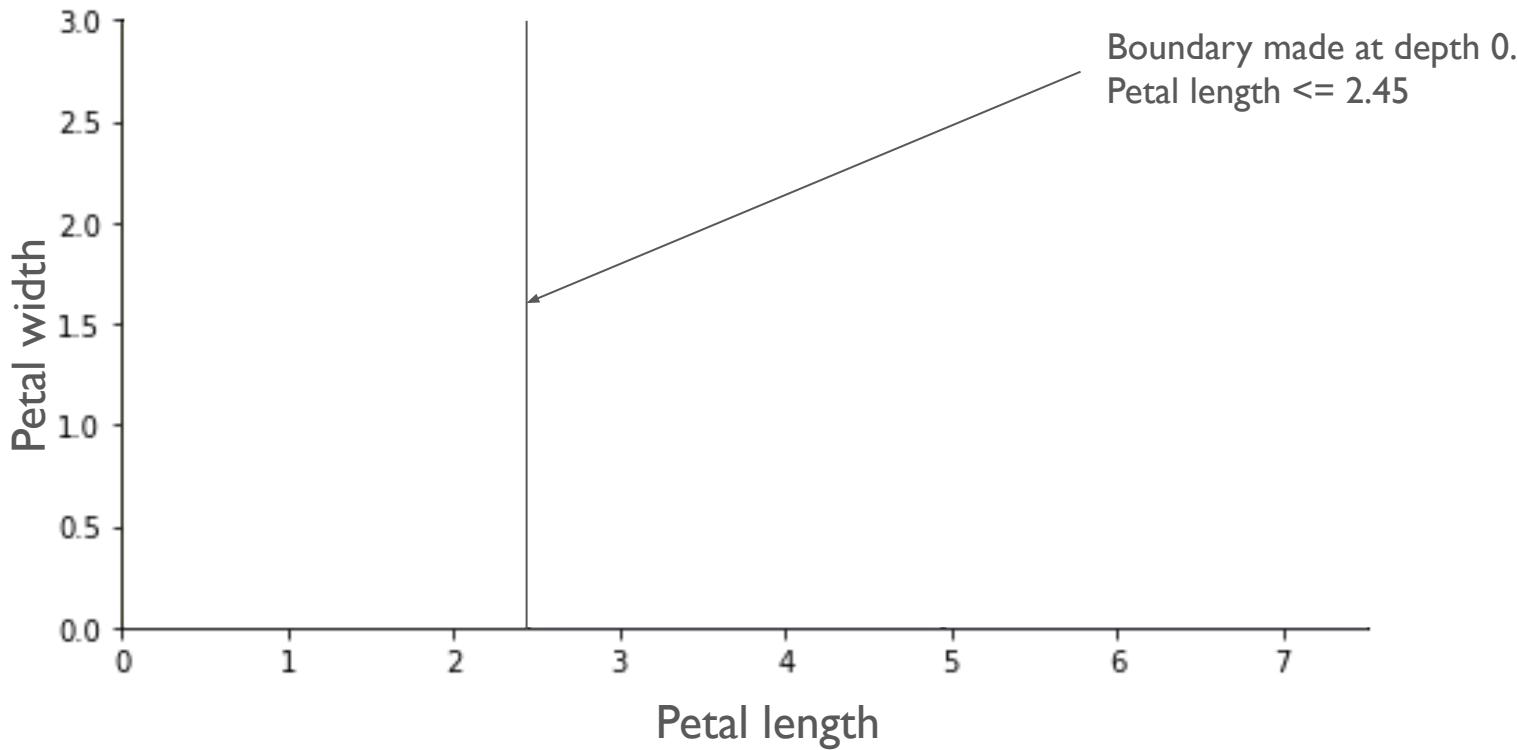
Decision Tree - IRIS dataset - Model



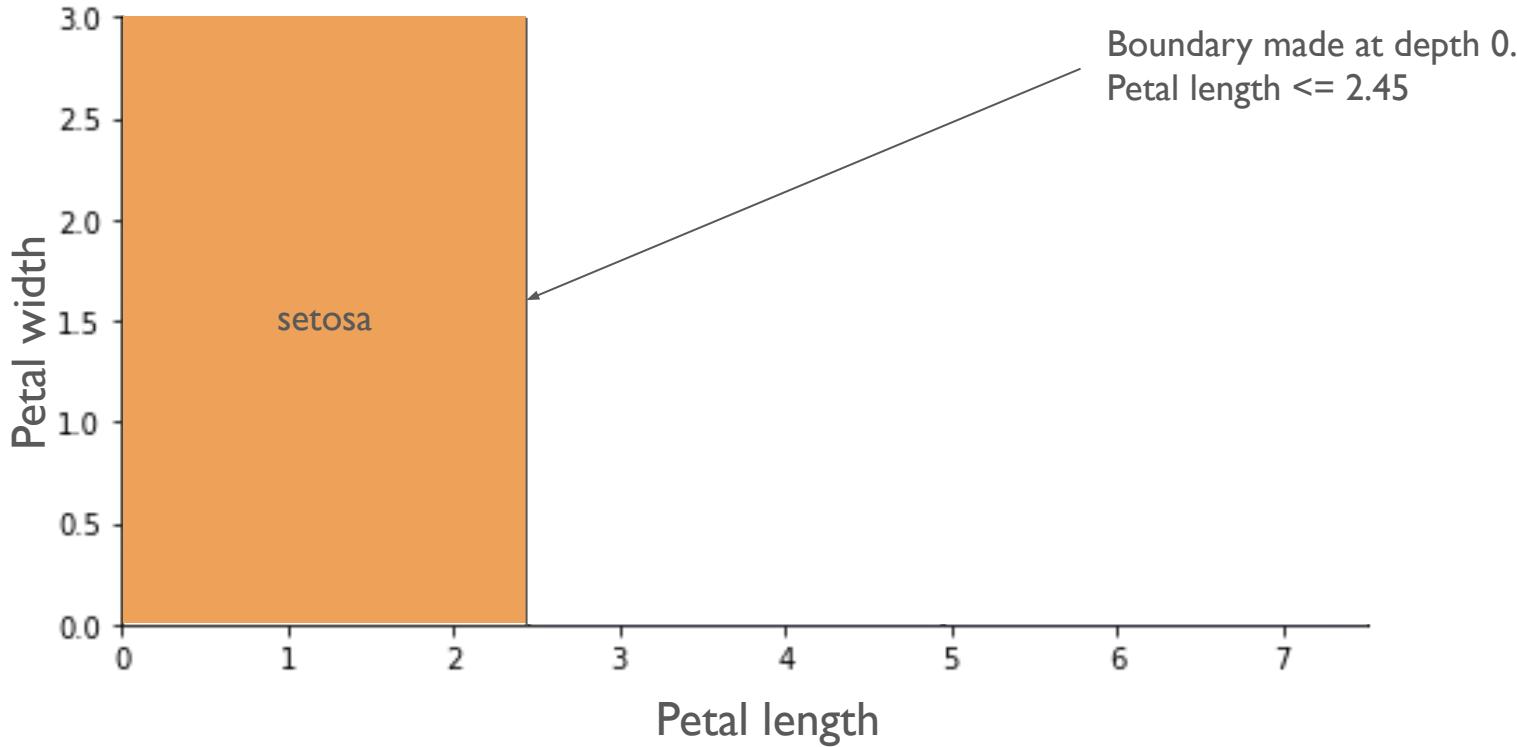
Decision Tree - Decision Boundaries



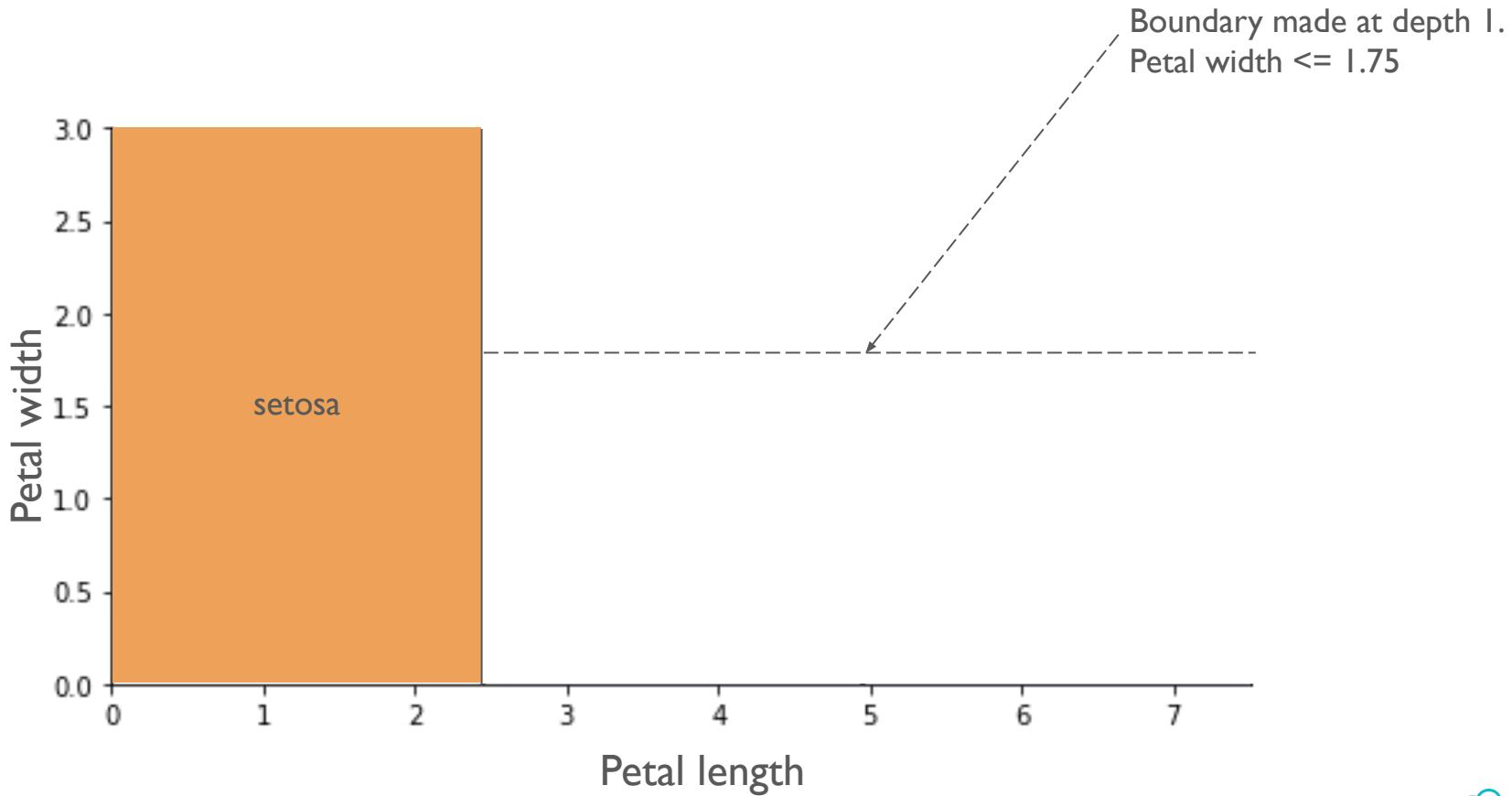
Decision Tree - Decision Boundaries



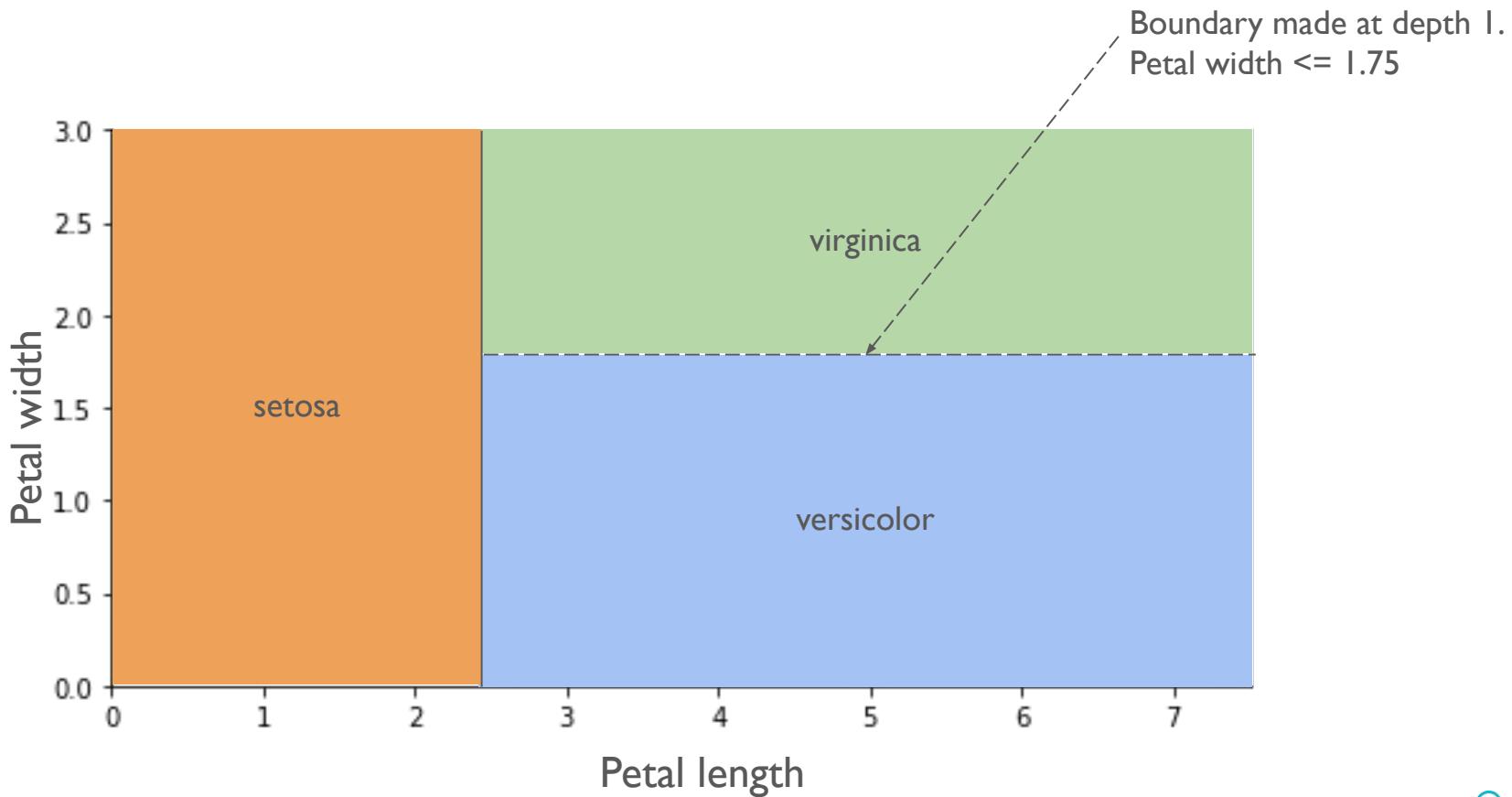
Decision Tree - Decision Boundaries



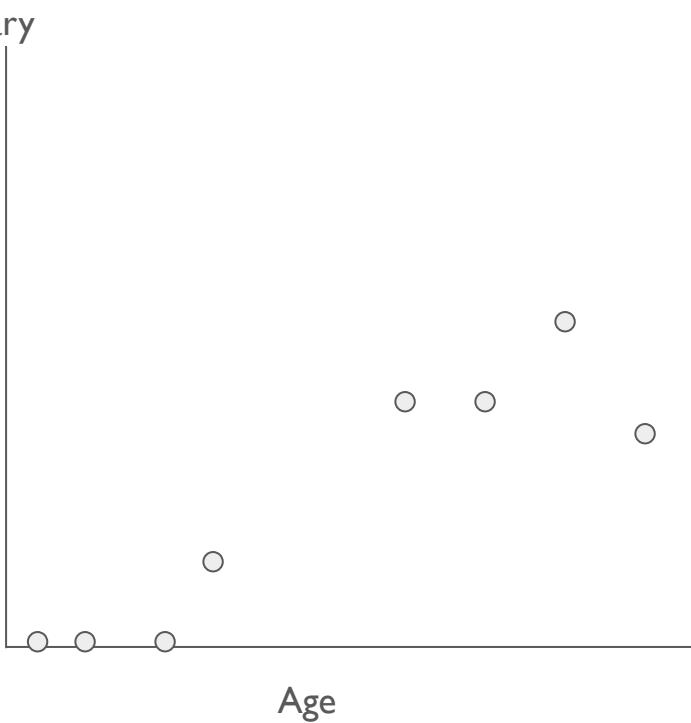
Decision Tree - Decision Boundaries



Decision Tree - Decision Boundaries



Decision Tree - how does it work?

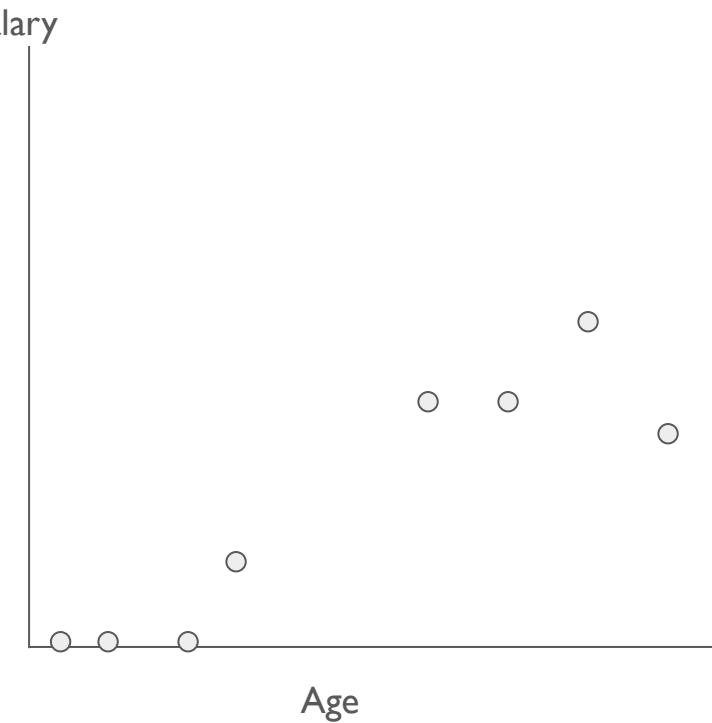


Decision Tree - how does it work?

It tries to find a boundary that separates similar instances.

We go over all the features.

I. Pick a feature

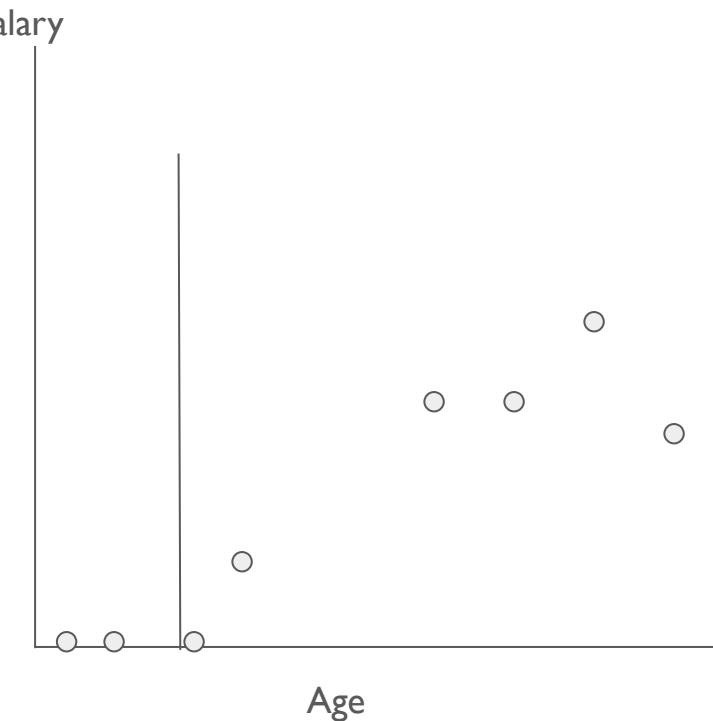


Decision Tree - how does it work?

It tries to find a boundary that separates similar instances.

We go over all the features.

1. Pick a feature
2. Pick some boundary value

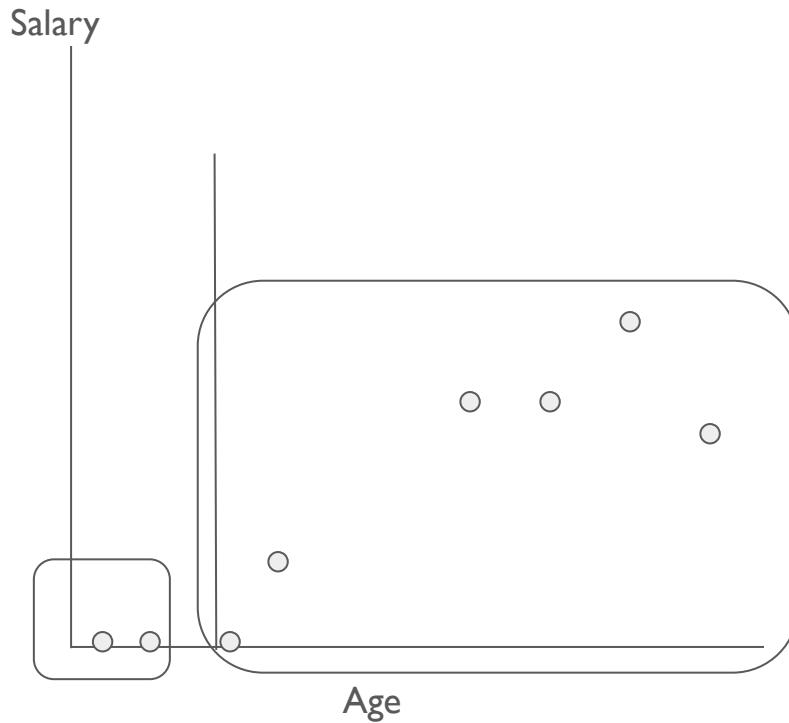


Decision Tree - how does it work?

It tries to find a boundary that separates similar instances.

We go over all the features.

1. Pick a feature
2. Pick some boundary value
3. Find the similarity in instances on same side

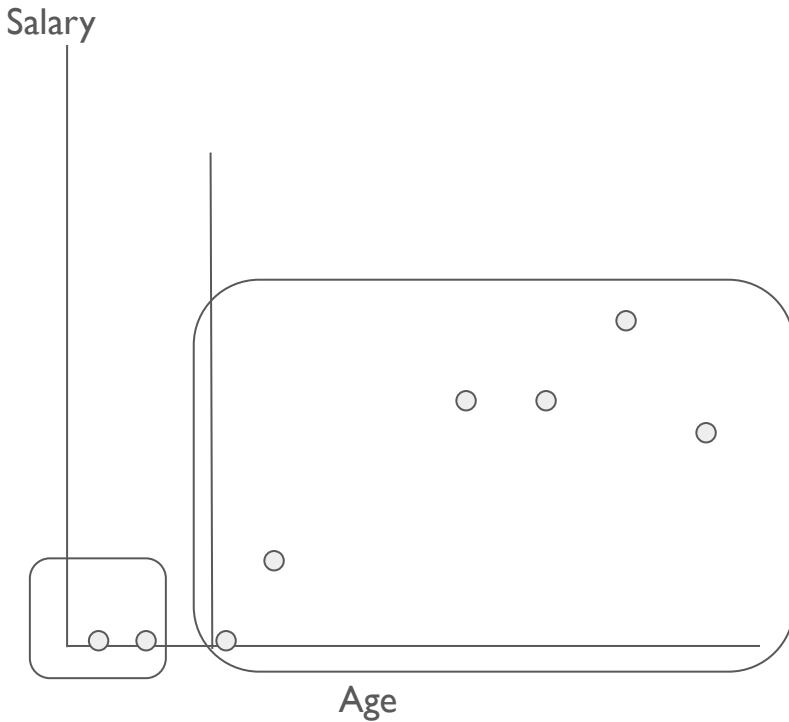


Decision Tree - how does it work?

It tries to find a boundary that separates similar instances.

We go over all the features.

1. Pick a feature
2. Pick some boundary value
3. Find the similarity in instances on same side
4. If the similarity is better, keep it.
5. Go to step 1

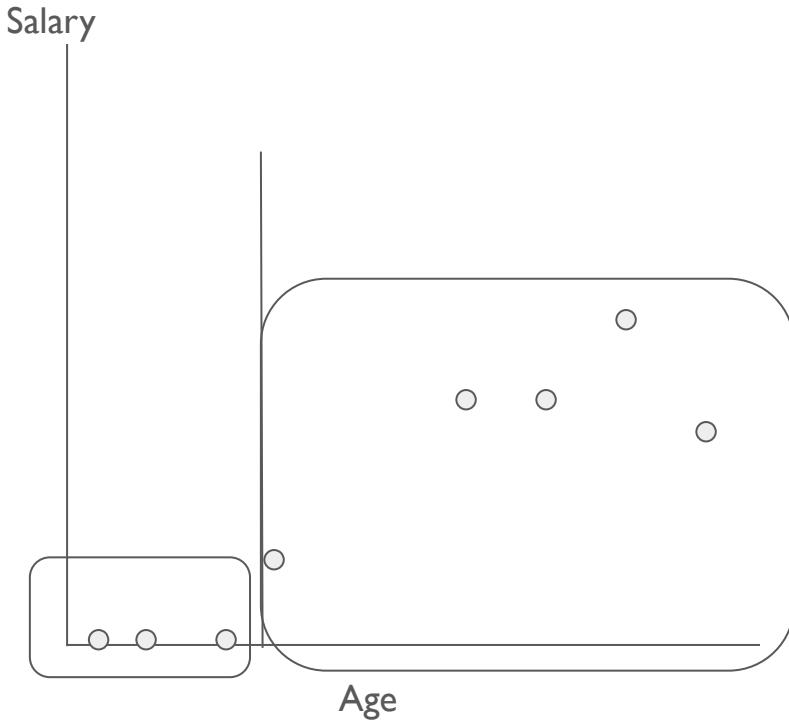


Decision Tree - how does it work?

It tries to find a boundary that separates similar instances.

We go over all the features.

1. Pick a feature
2. Pick some boundary value
3. Find the similarity in instances on same side
4. If the similarity is better, keep it.
5. Go to step 1

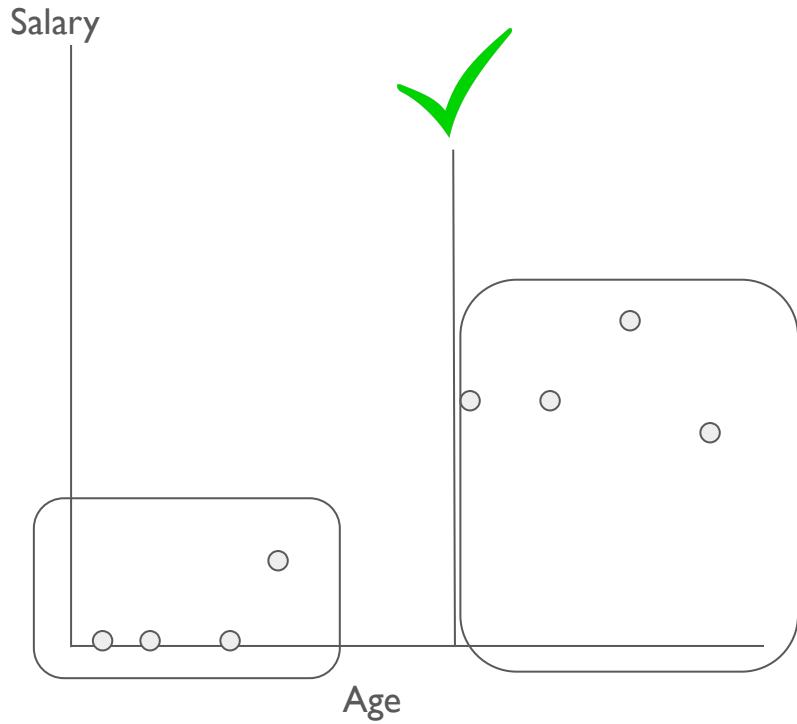


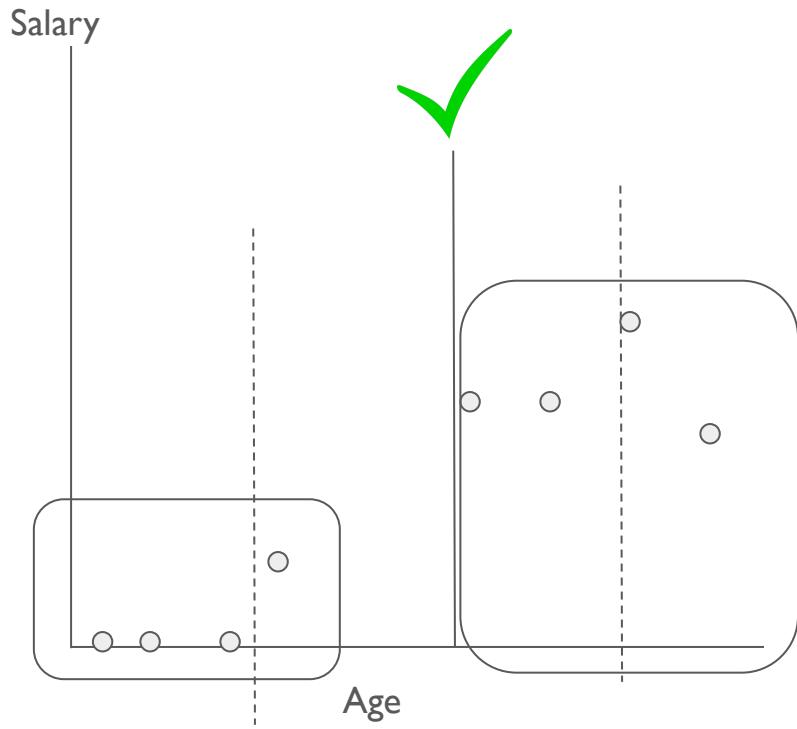
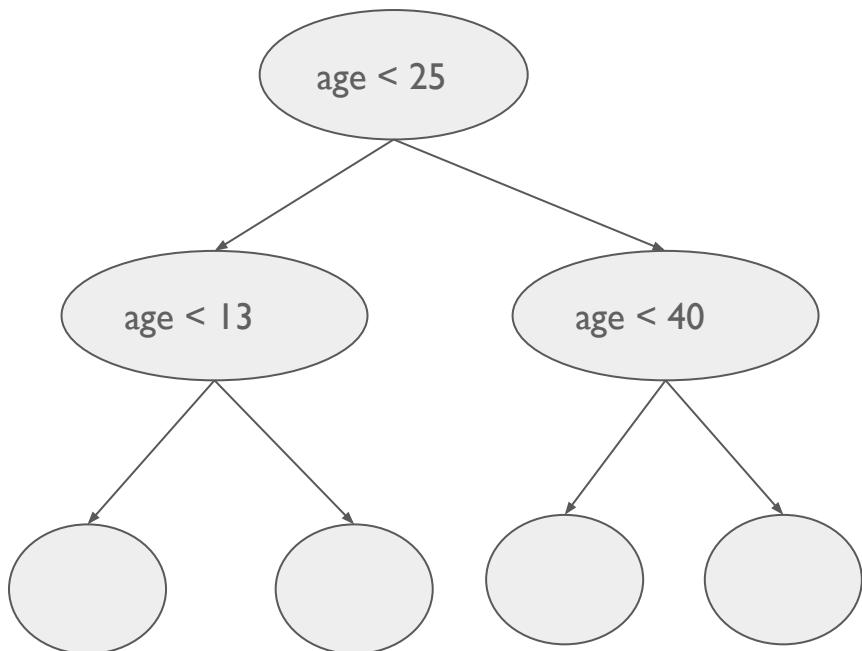
Decision Tree - how does it work?

It tries to find a boundary that separates similar instances.

We go over all the features.

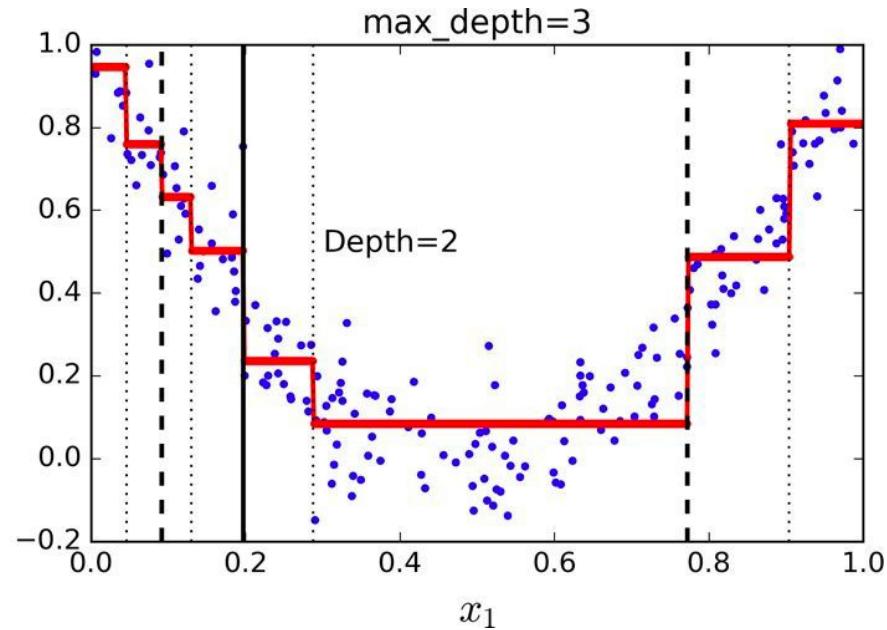
1. Pick a feature
2. Pick some boundary value
3. Find the similarity in instances on same side
4. If the similarity is better, keep it.
5. Go to step 1





Decision Tree - Complex Models?

Decision Trees can come up with really complex models.



Decision Tree - Final Word

I. Easy to Visualize

Decision Tree - Final Word

- I. Easy to Visualize
2. Easy to Interpret & explain

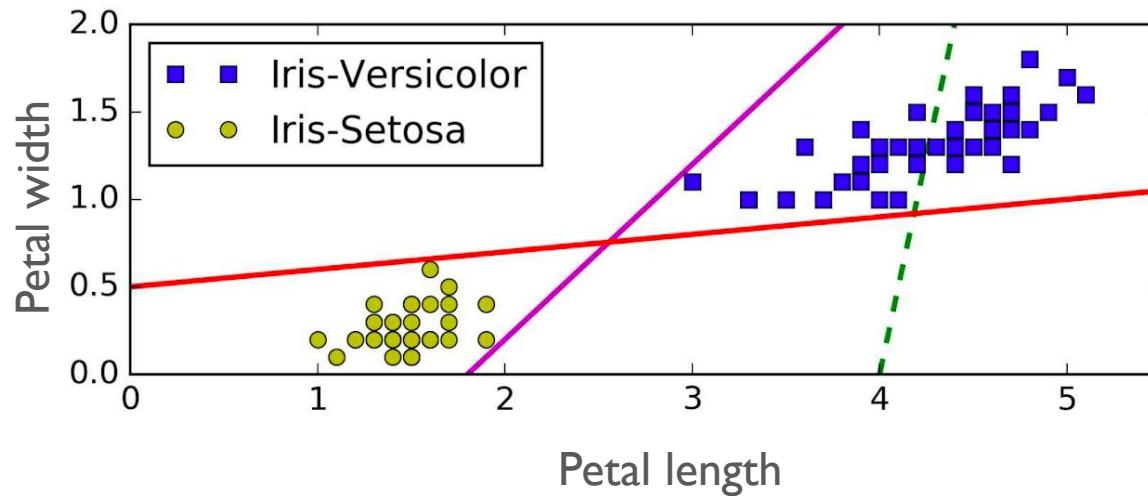
Decision Tree - Final Word

1. Easy to Visualize
2. Easy to Interpret & explain
3. If not designed well, ends up overfitting

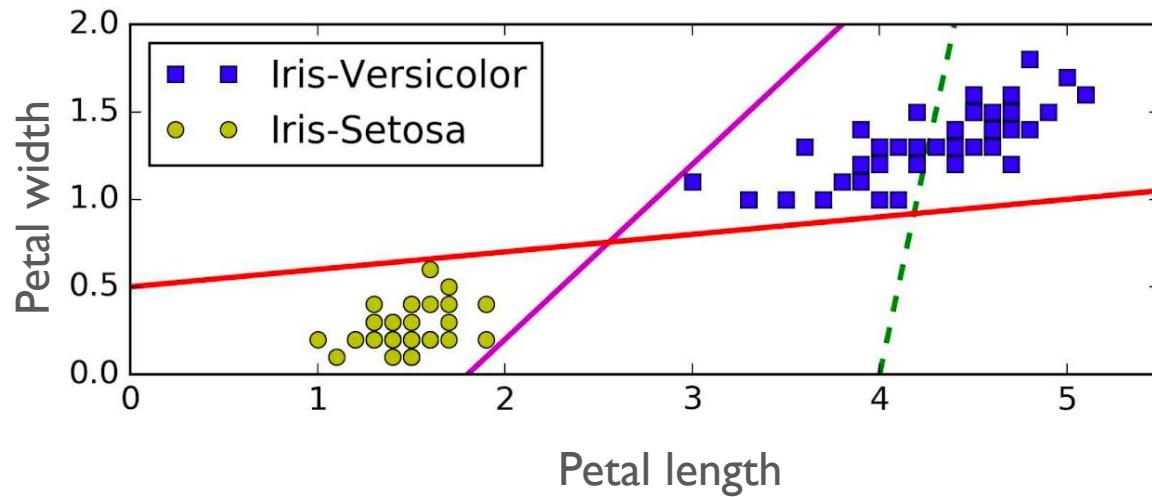
SVM

What is Linear Classification?

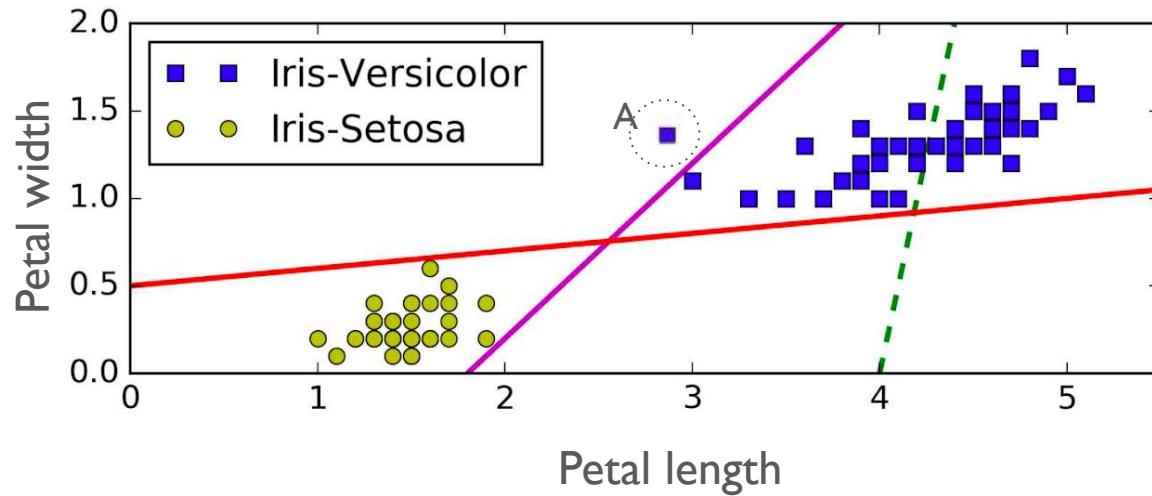
The two classes can be separated easily with a ‘straight’ line



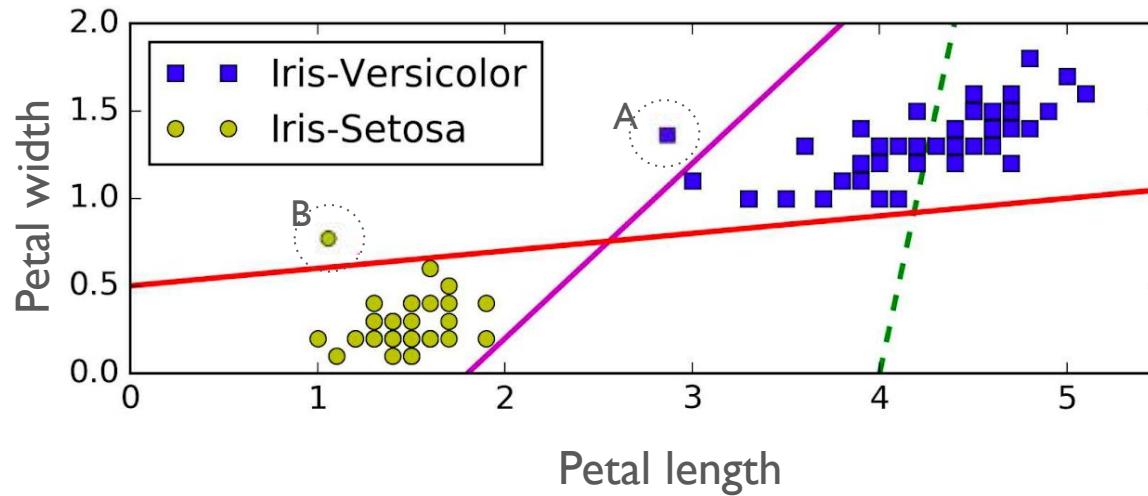
Is Linear Classification good enough?



Is Linear Classification good enough?

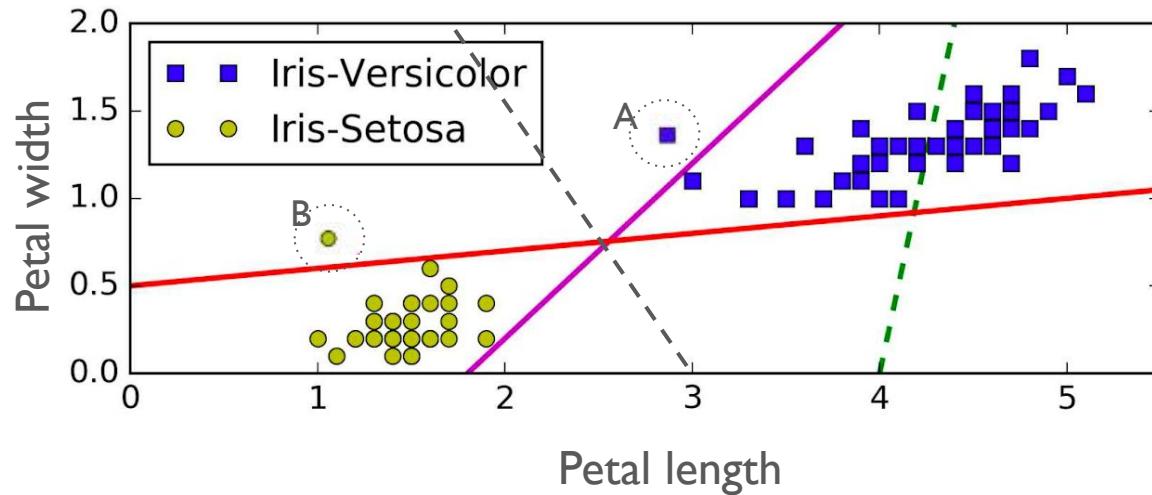


Is Linear Classification good enough?

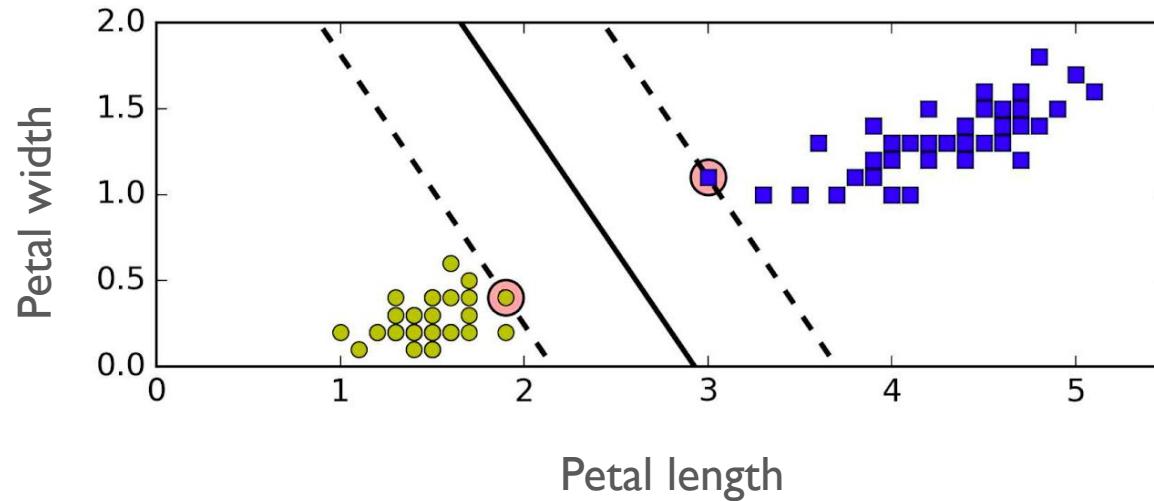


Is Linear Classification good enough?

How about this line?

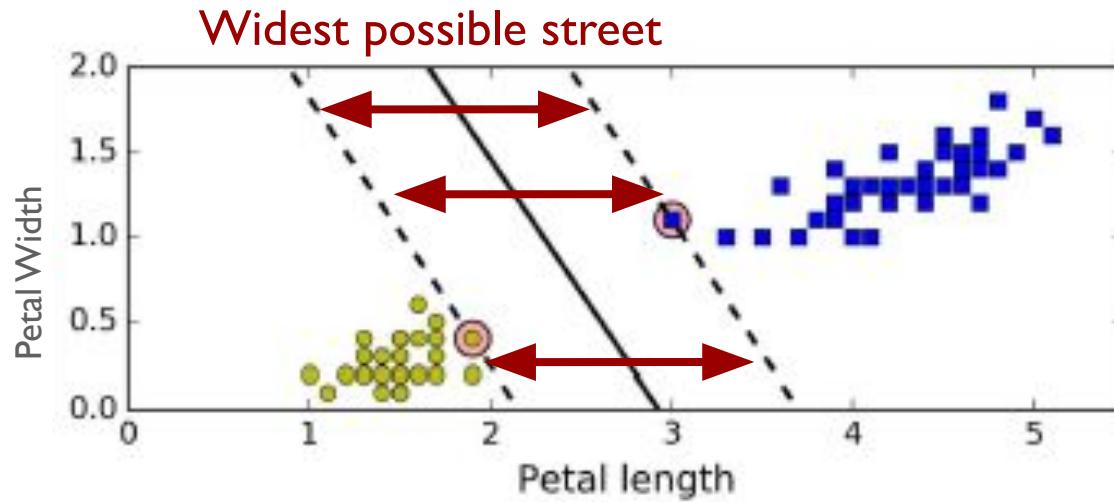


What is Linear Classification?

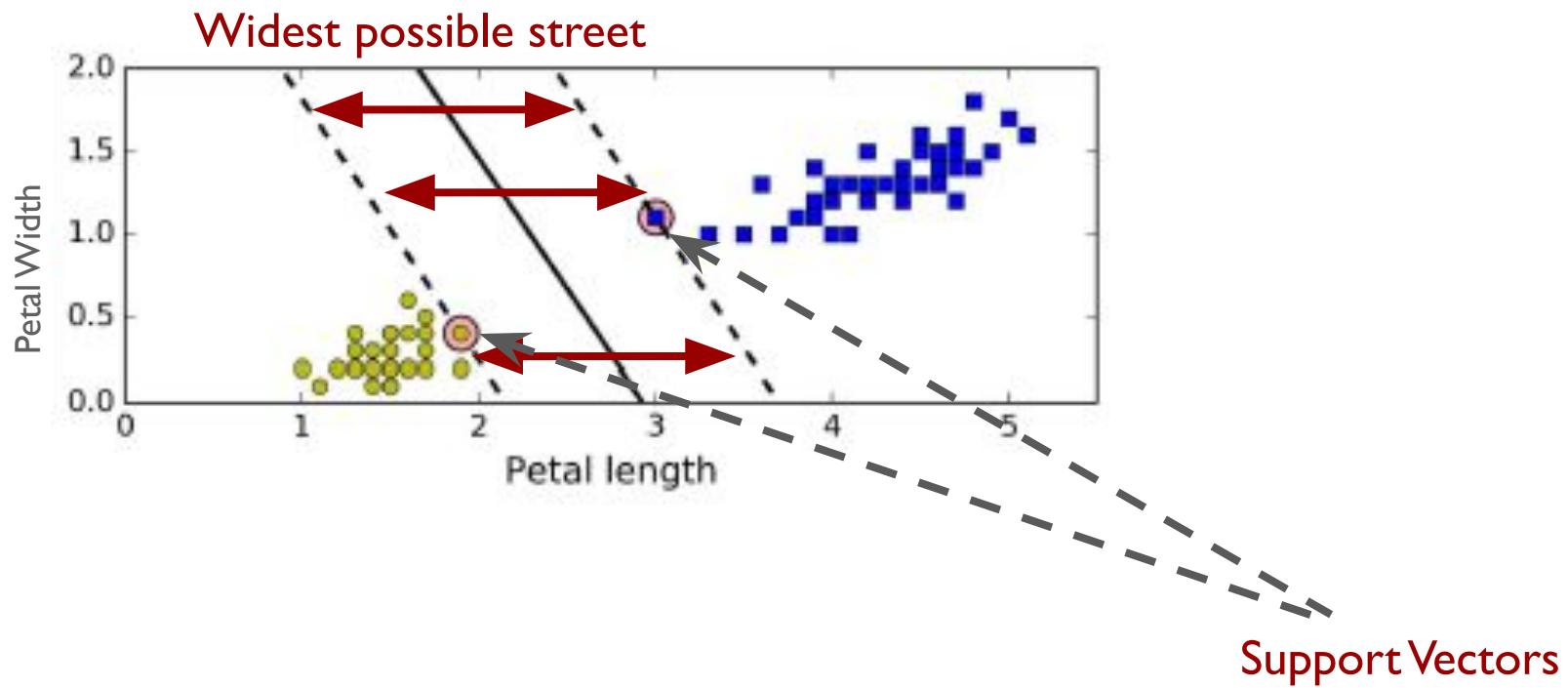


‘Straight’ is the keyword. It means linear classification.

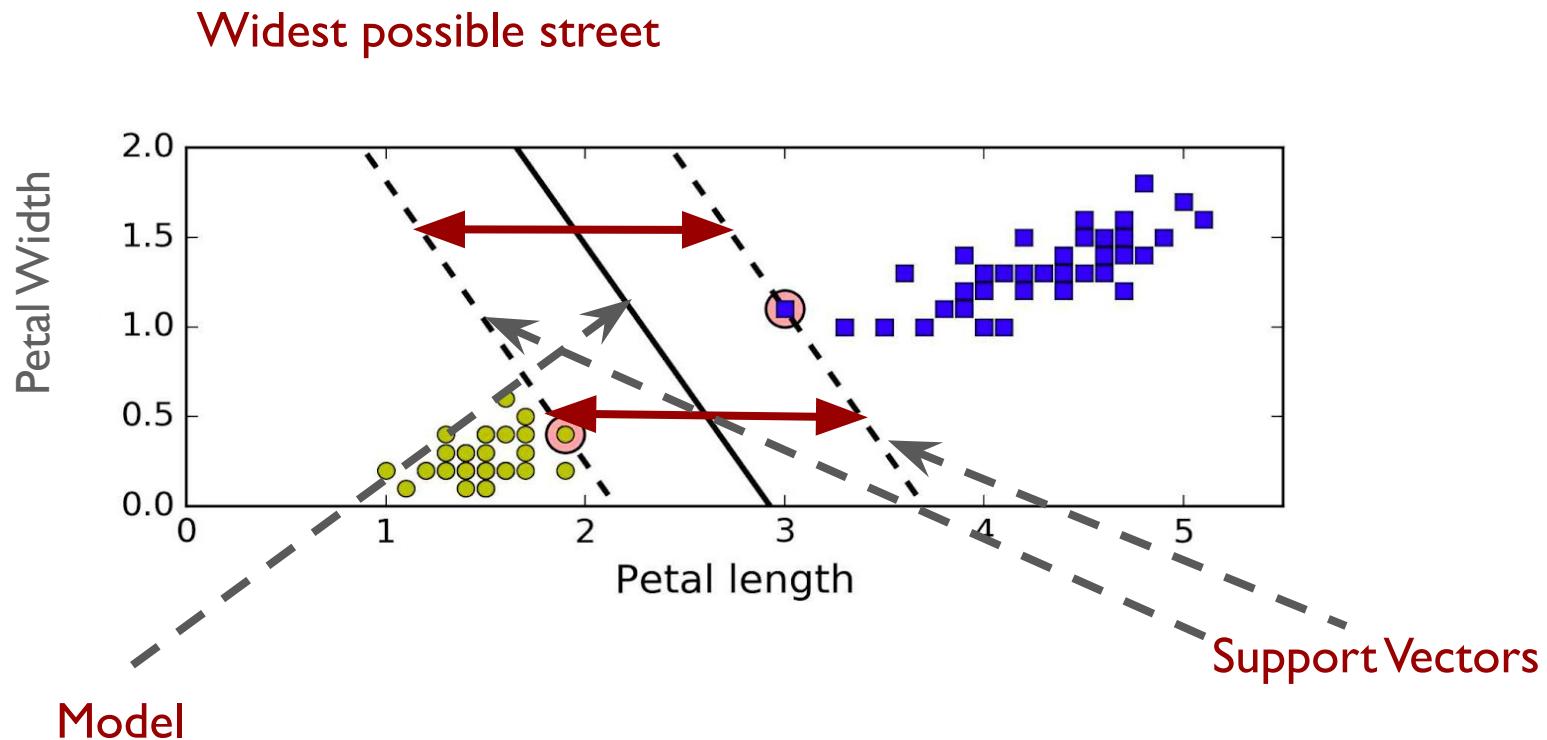
What is Linear Classification?



What is Linear Classification?



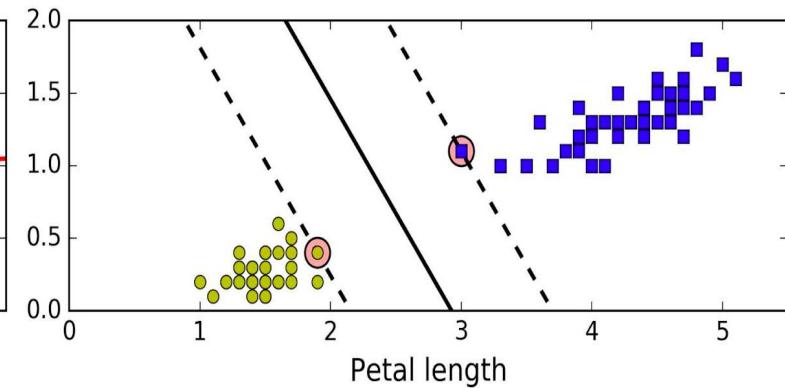
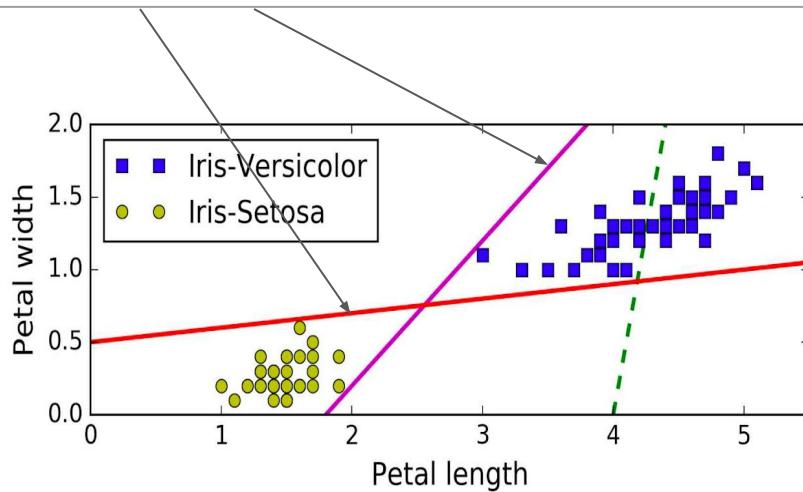
What is Linear Classification?



Linear SVM Classification - Large Margin

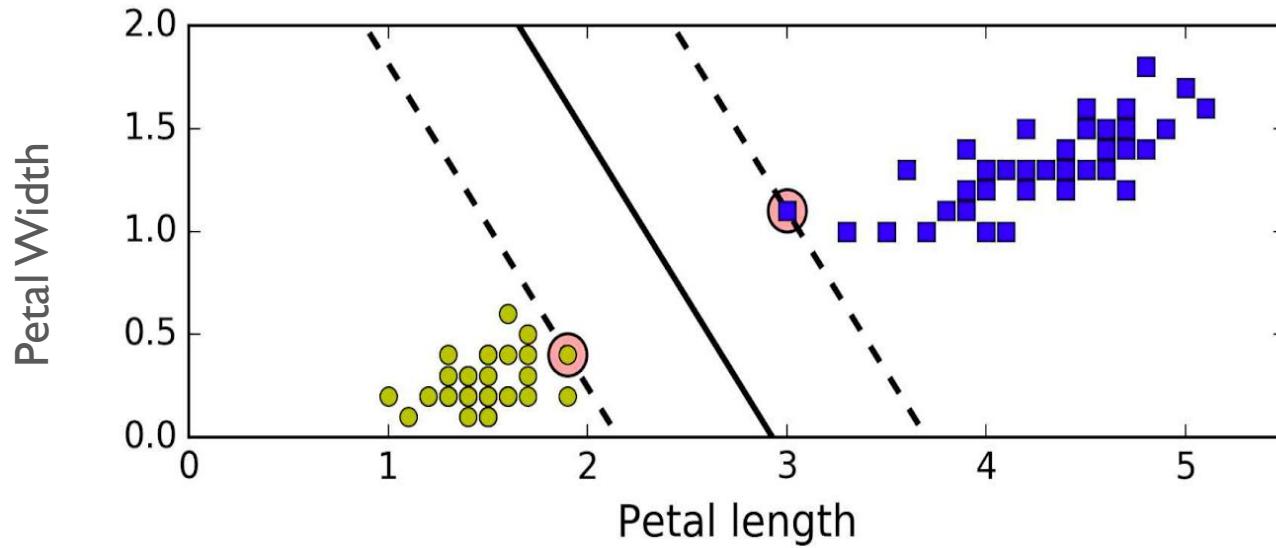
Violet and red decision boundaries are very close to the instances - bad model

Decision Boundary as far away from training instances - good model



Large Margin Classification

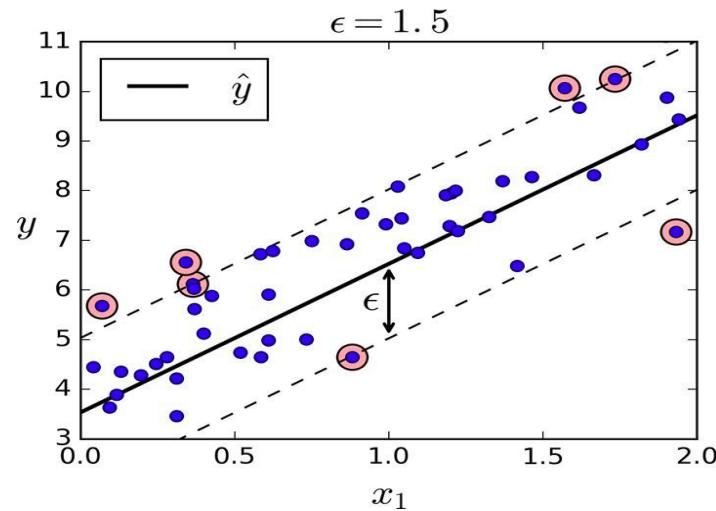
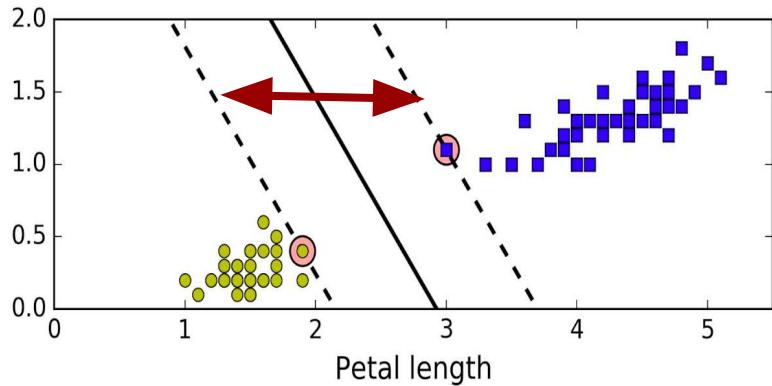
How does SVM work?



SVM Regression

SVM Classifier	SVM Regression
Find the largest possible street between the two classes limiting margin violations	Fit as many instances as possible on the street while limiting margin violations

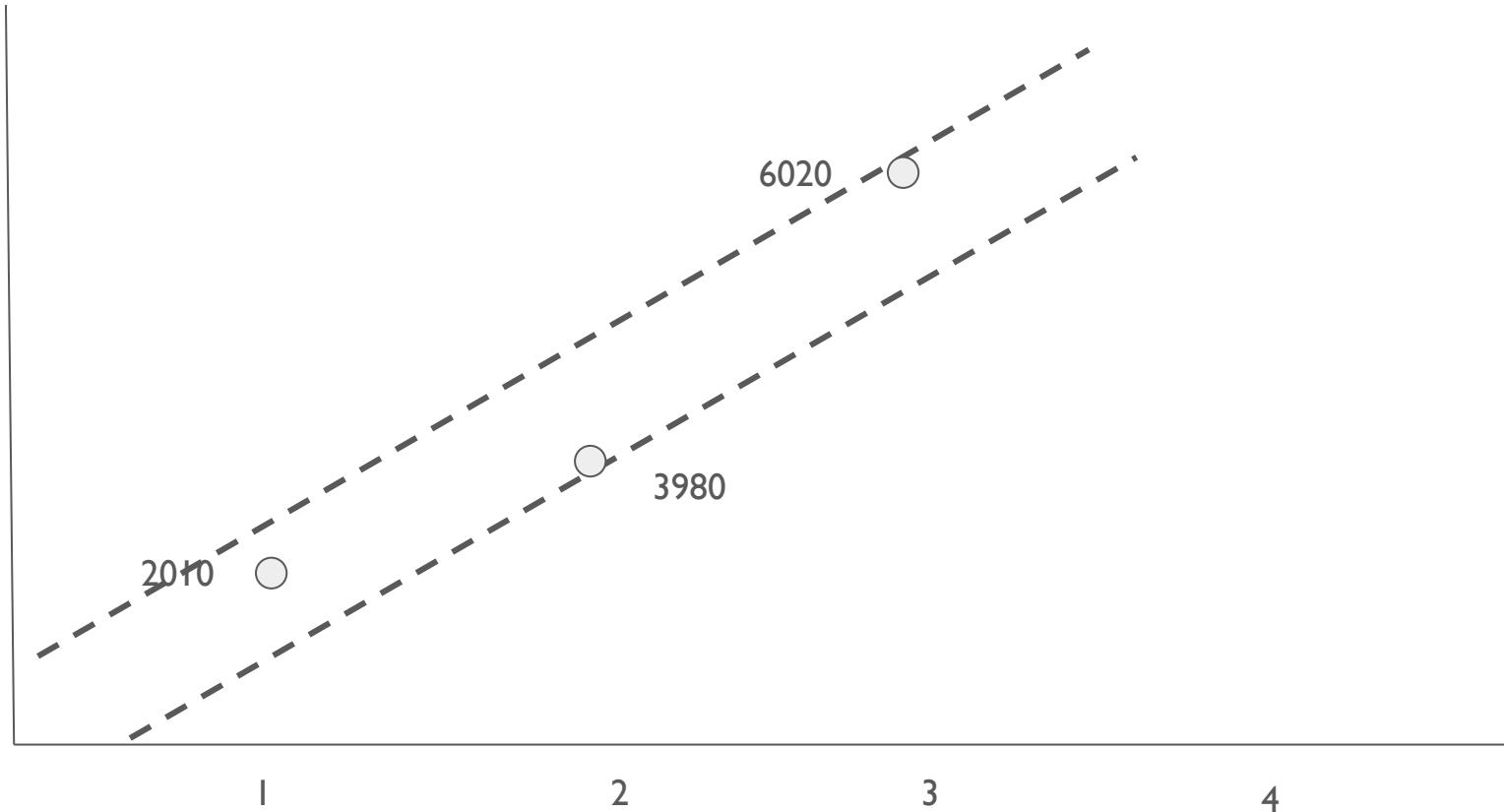
Widest possible street



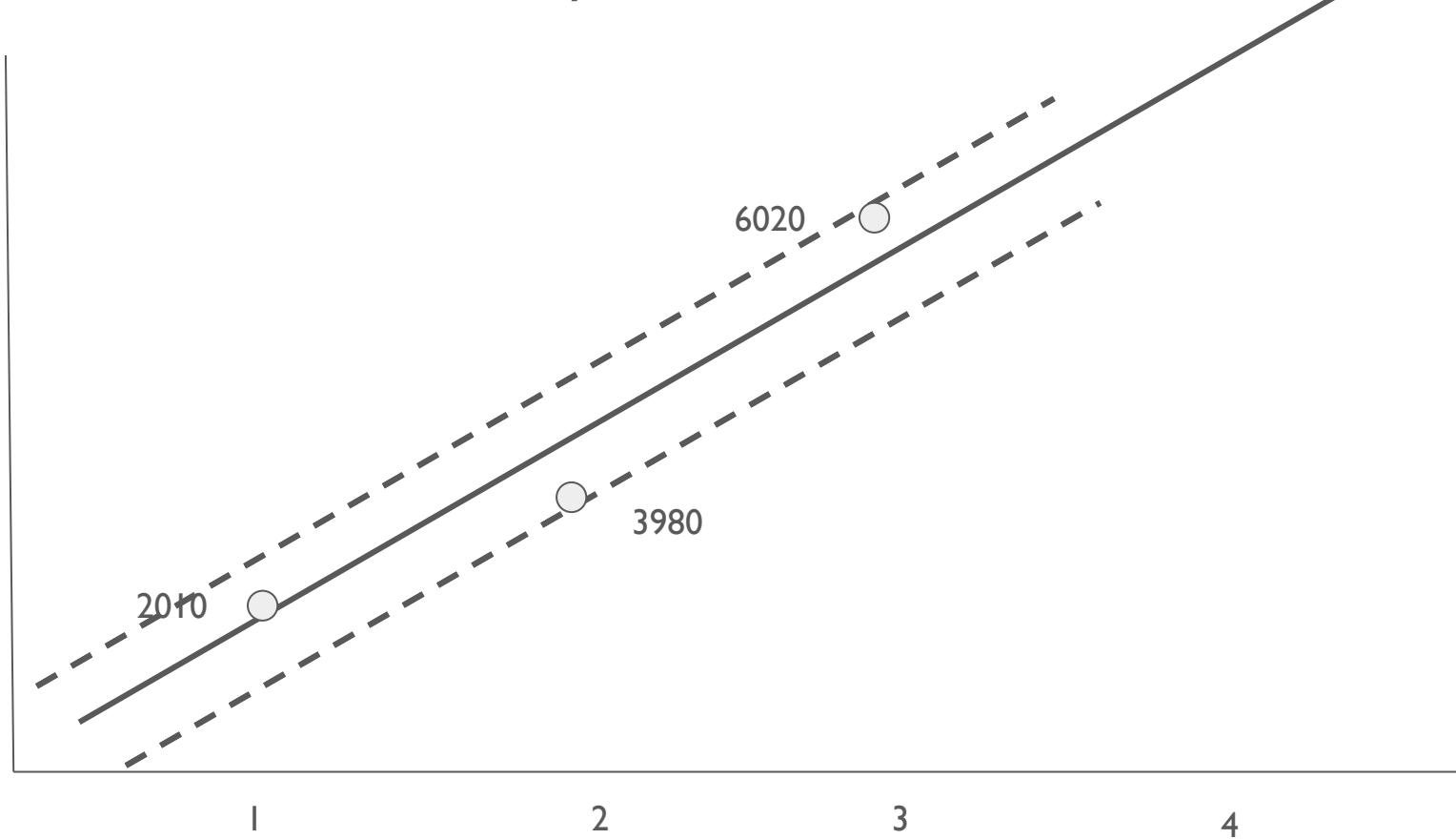
Can we solve it with SVM regression?



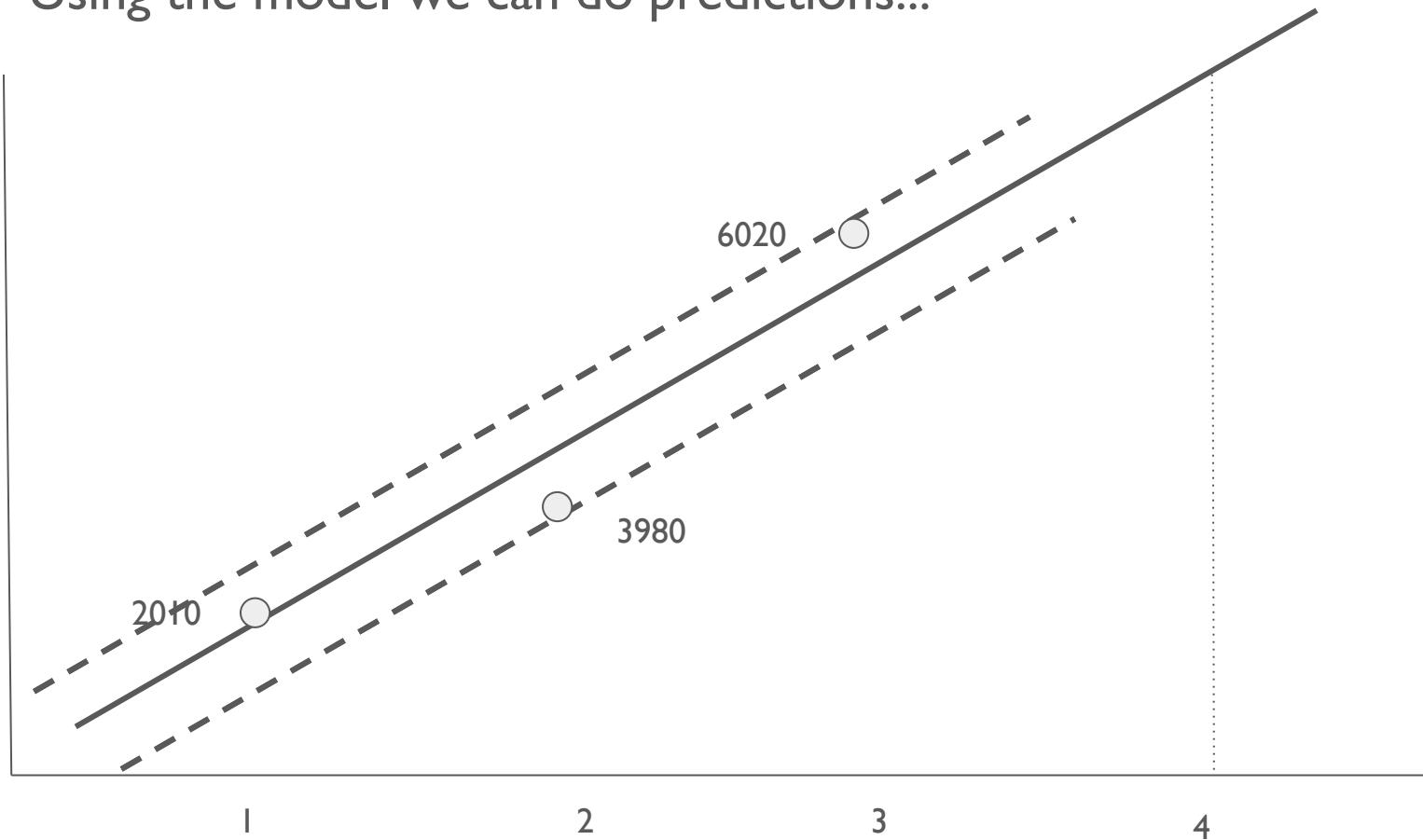
Find narrowest street that can fit the data points



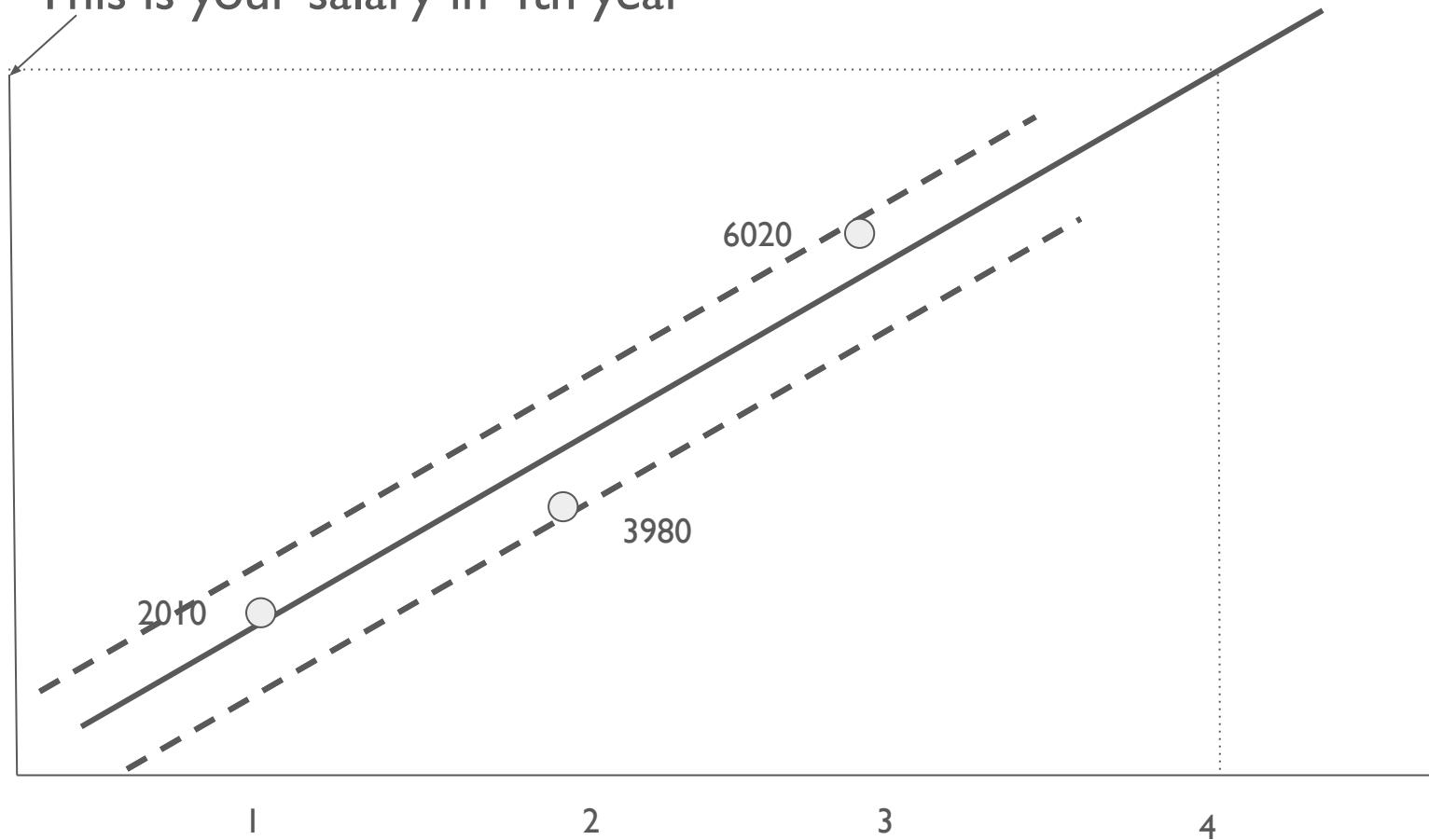
Central line of street is your model!



Using the model we can do predictions...



This is your salary in 4th year



SVM

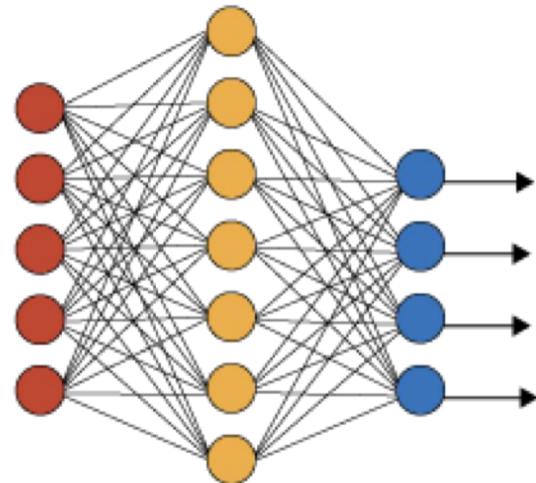
- Very powerful and versatile model
- Capable of performing
 - Classification
 - Regression
 - Outlier detection
- Before neural networks, they were best performing
- Well suited for small or medium sized datasets

Deep Learning - Artificial Neural Network(ANN)



Computing systems inspired by the biological neural networks that constitute animal brains.

Simple Neural Network

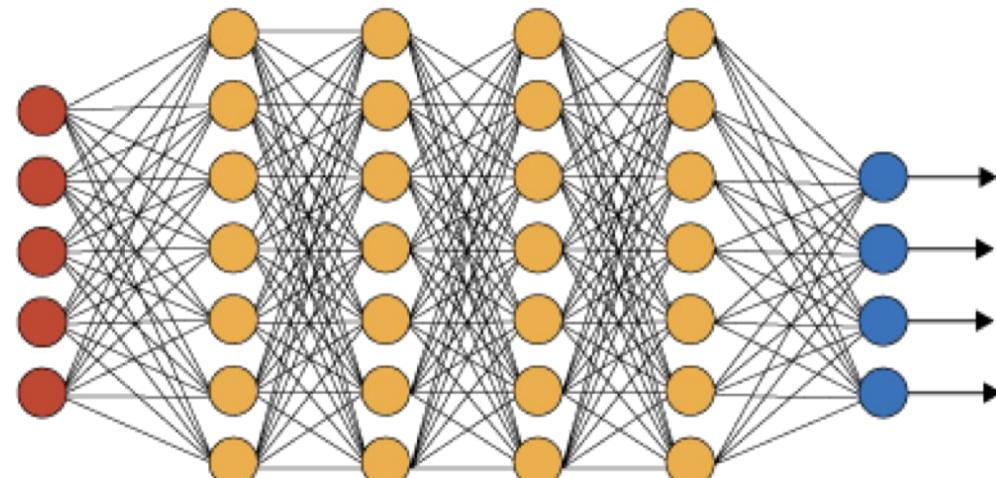


● Input Layer

● Hidden Layer

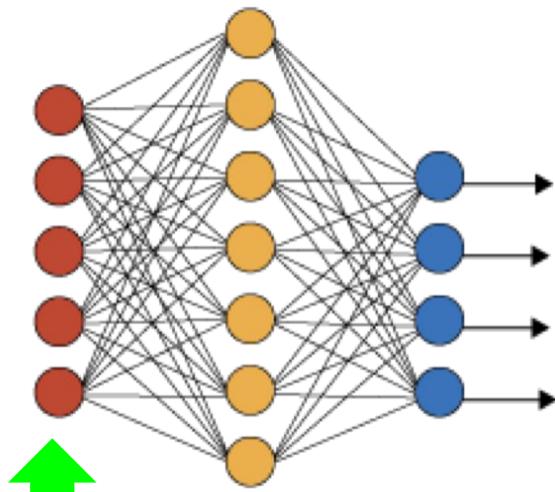
● Output Layer

Deep Learning Neural Network



Multiple layers of neurons

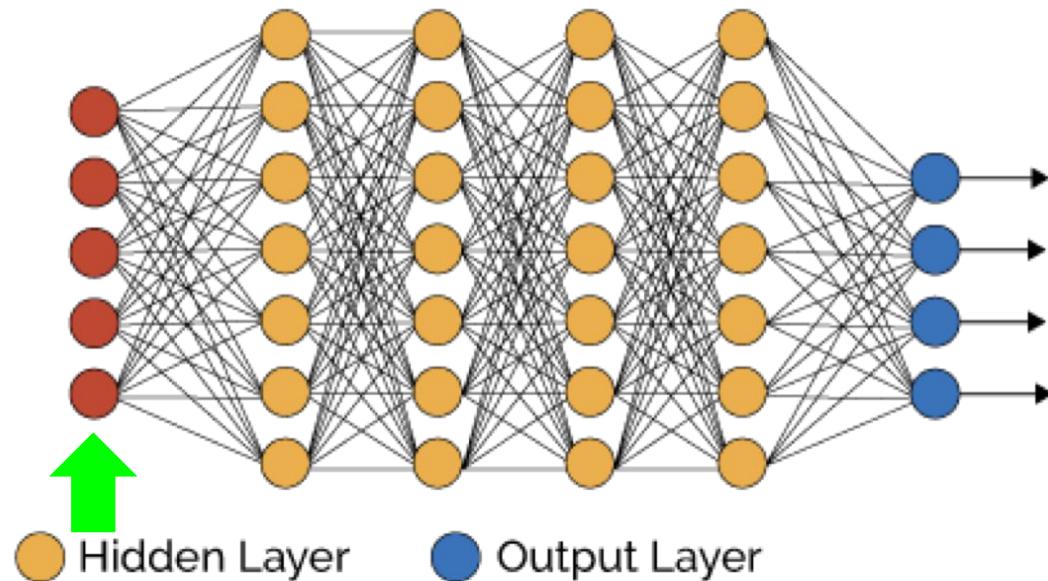
Simple Neural Network



● Input Layer



Deep Learning Neural Network

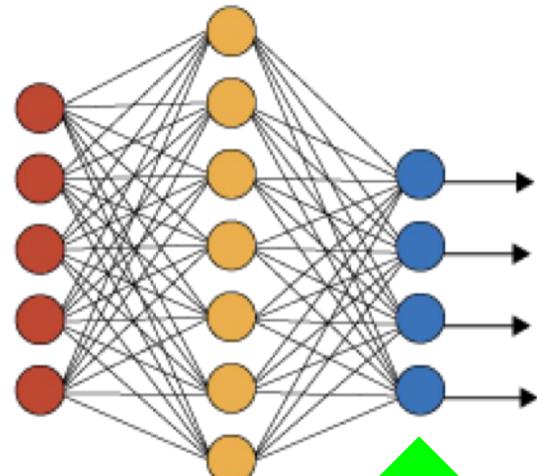


● Hidden Layer

● Output Layer

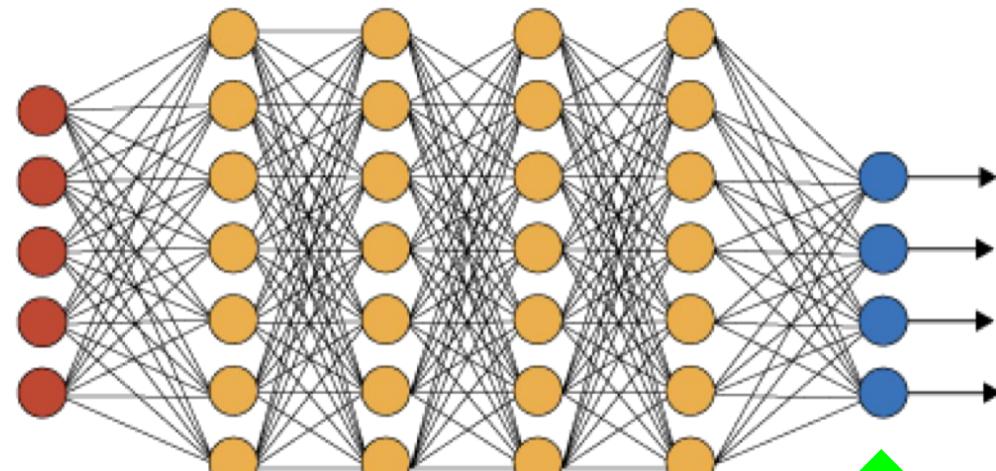
Multiple layers of neurons

Simple Neural Network



● Input Layer

Deep Learning Neural Network

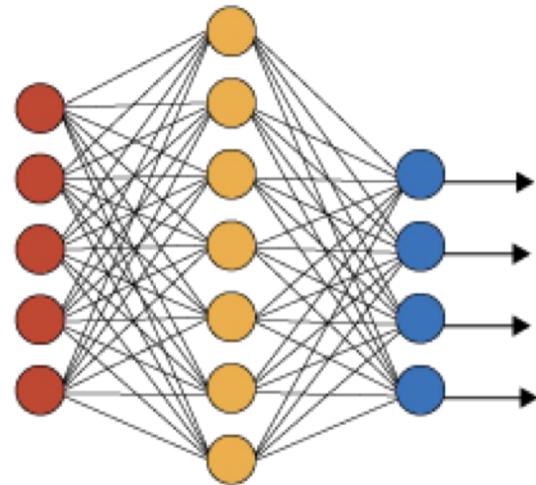


● Hidden Layer

● Output Layer

Multiple layers of neurons

Simple Neural Network

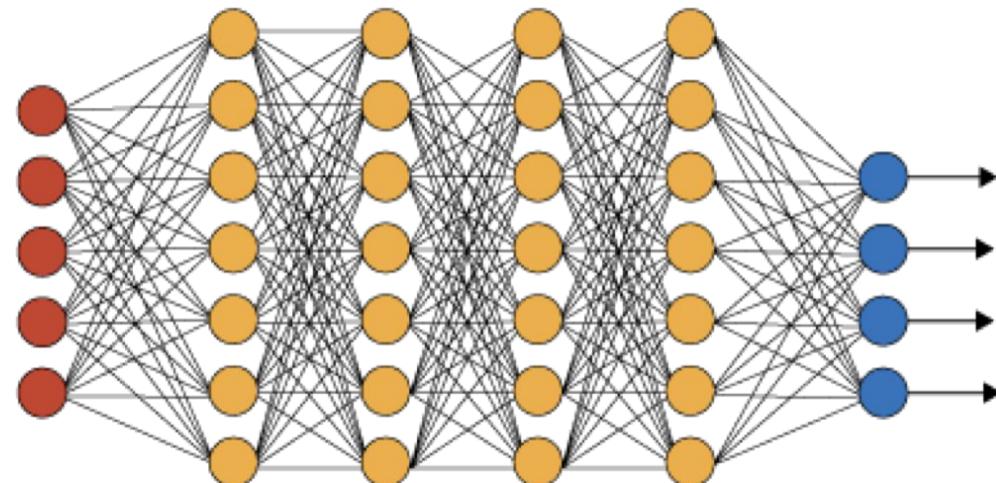


● Input Layer

● Hidden Layer

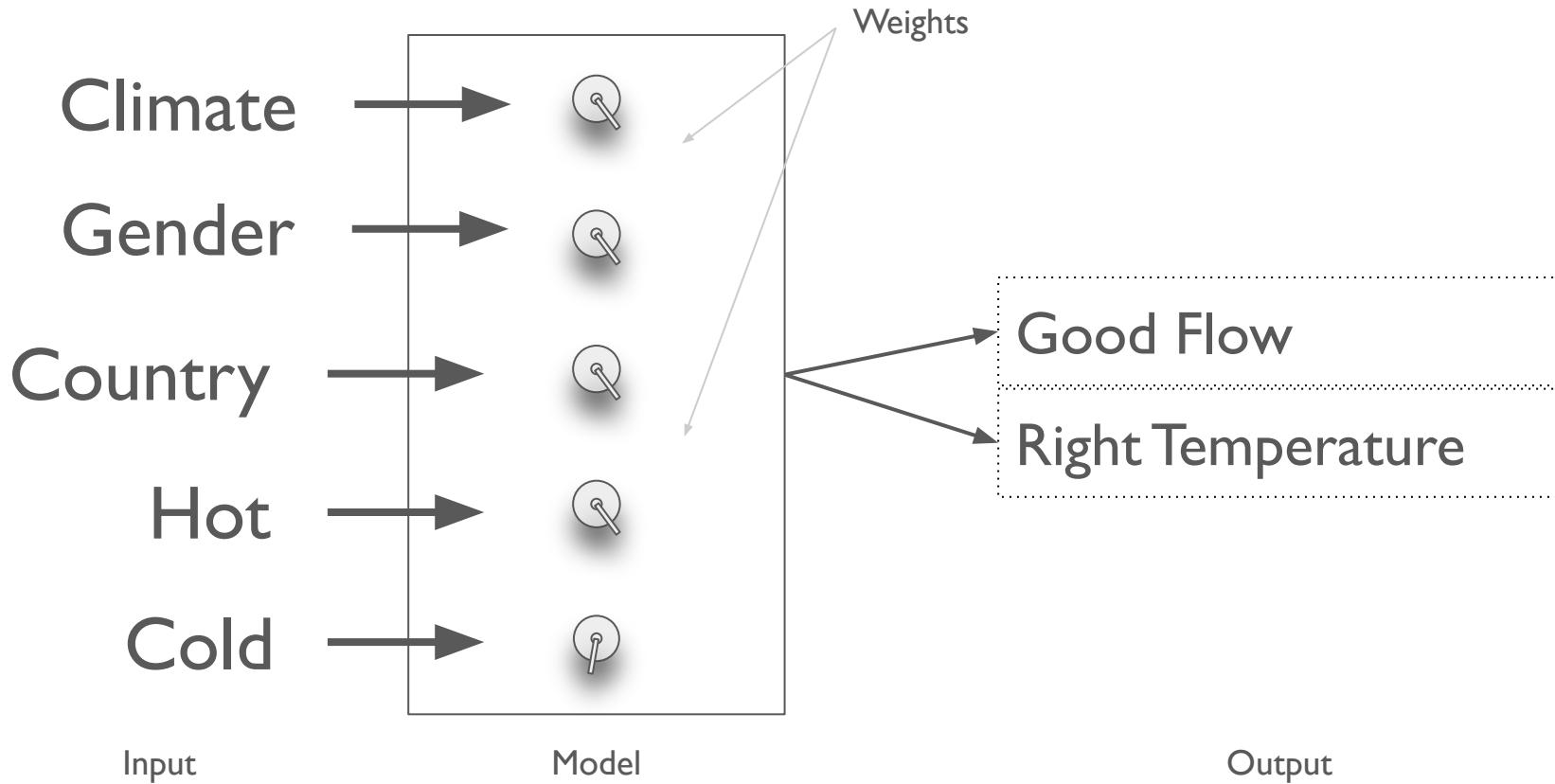
● Output Layer

Deep Learning Neural Network

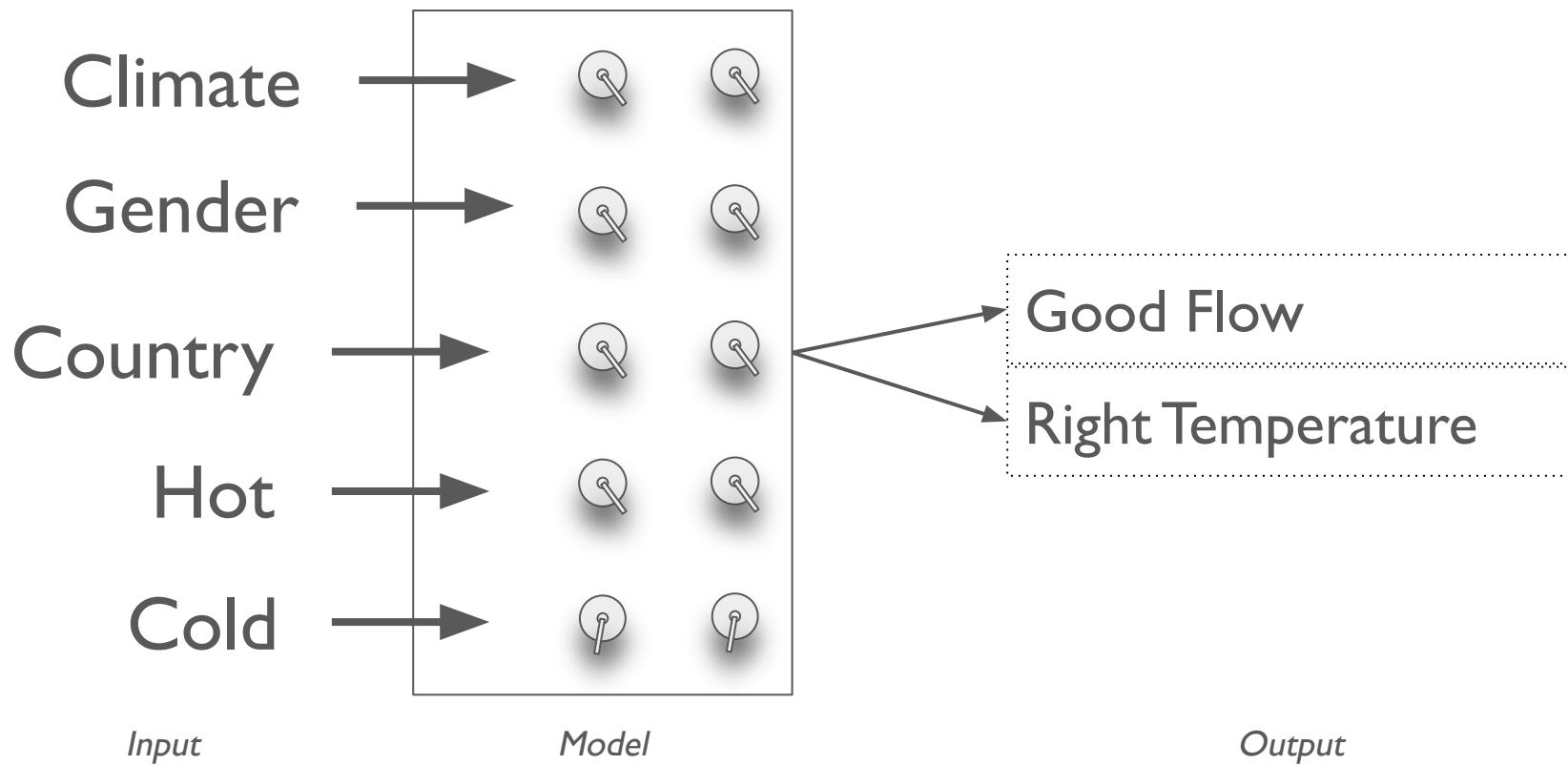


Multiple layers of neurons

What if we had more input features?

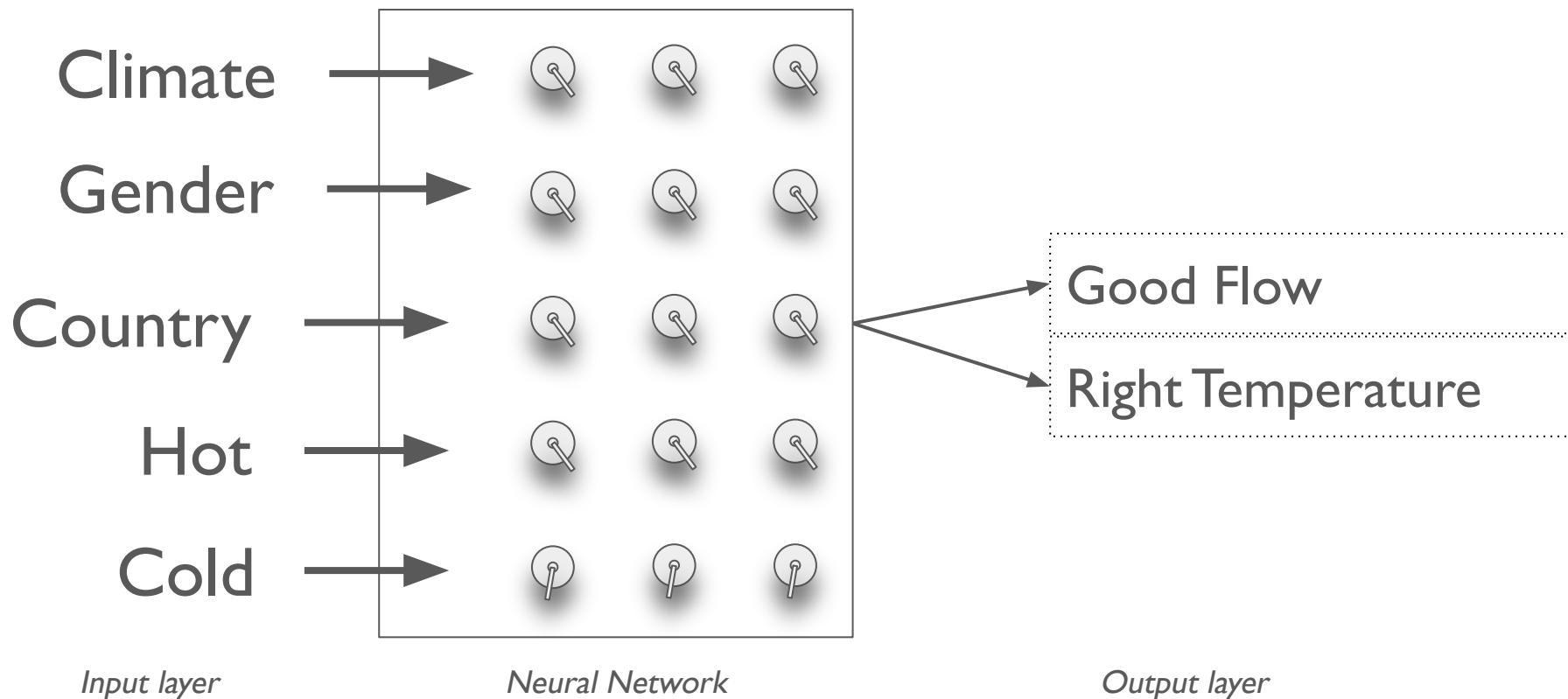


We could improve the model further...

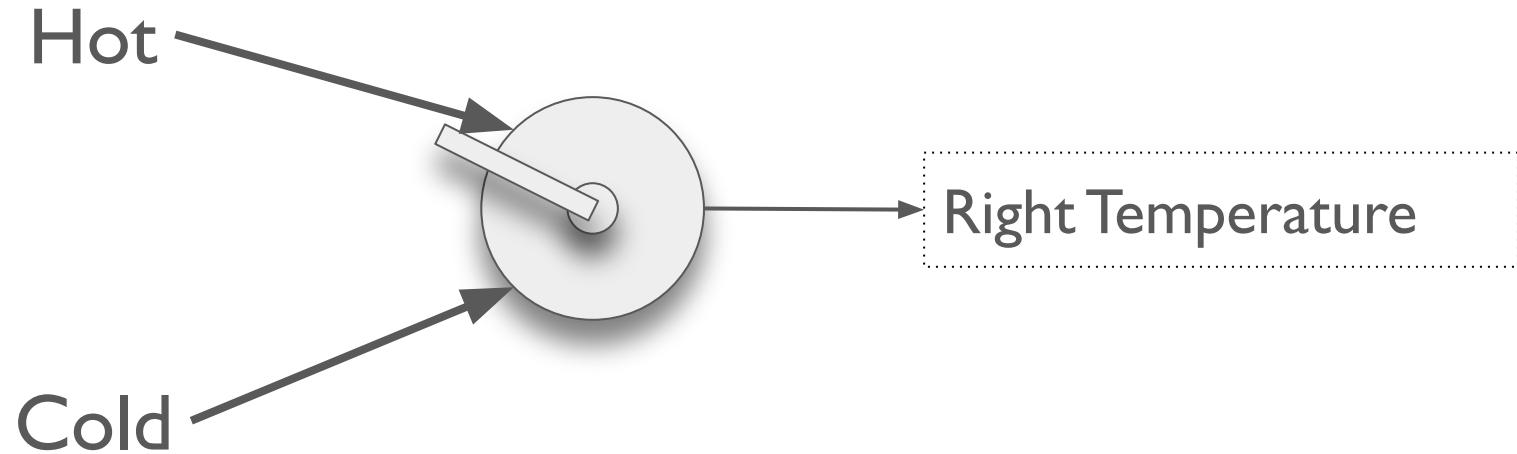


And further....

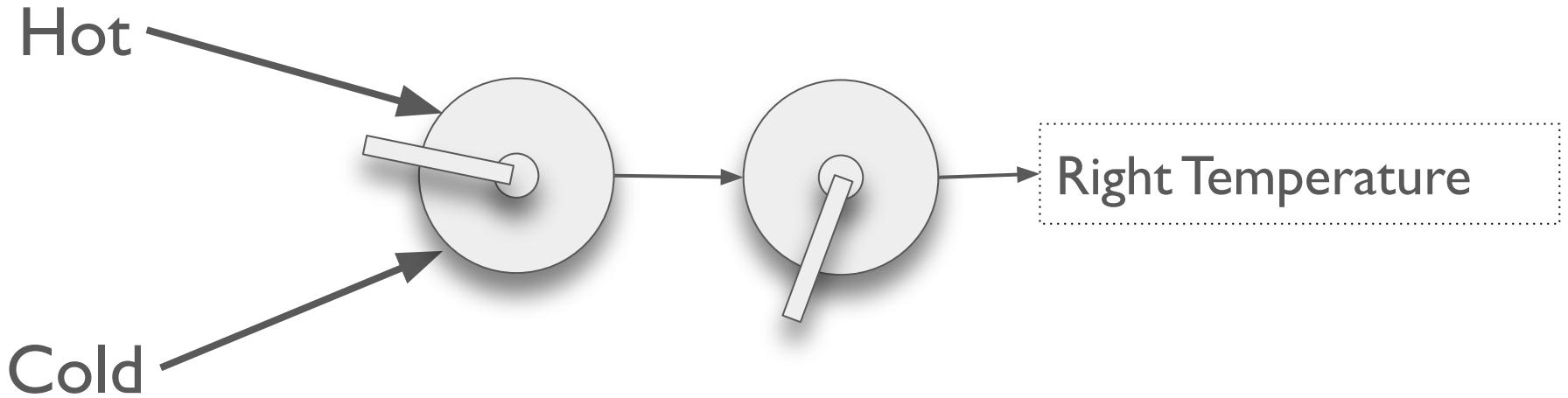
Deep Neural Network...



Training Neural Networks - Single Neuron

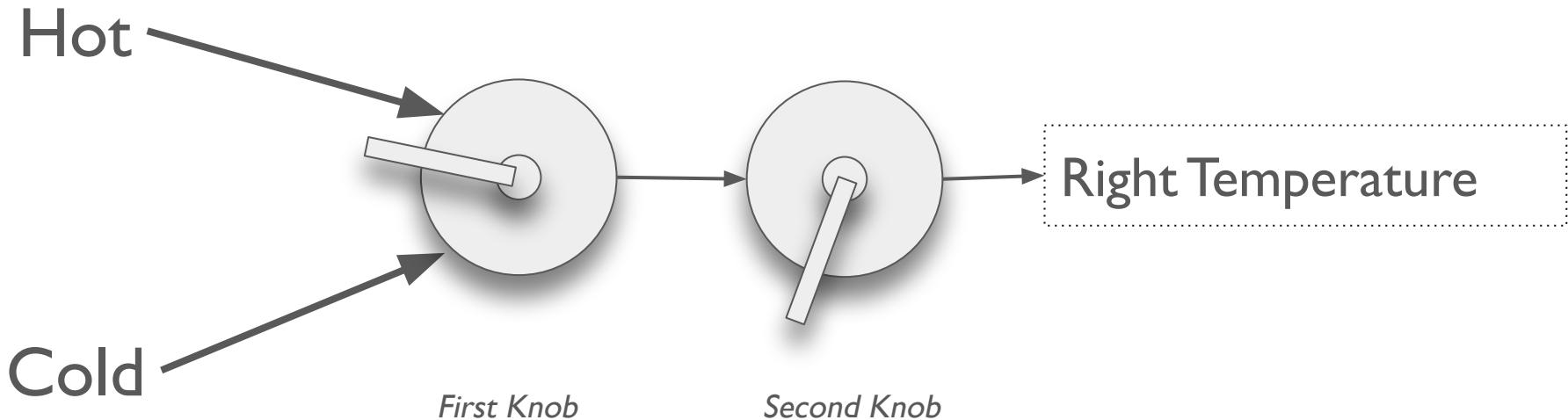


Training Neural Networks - How?



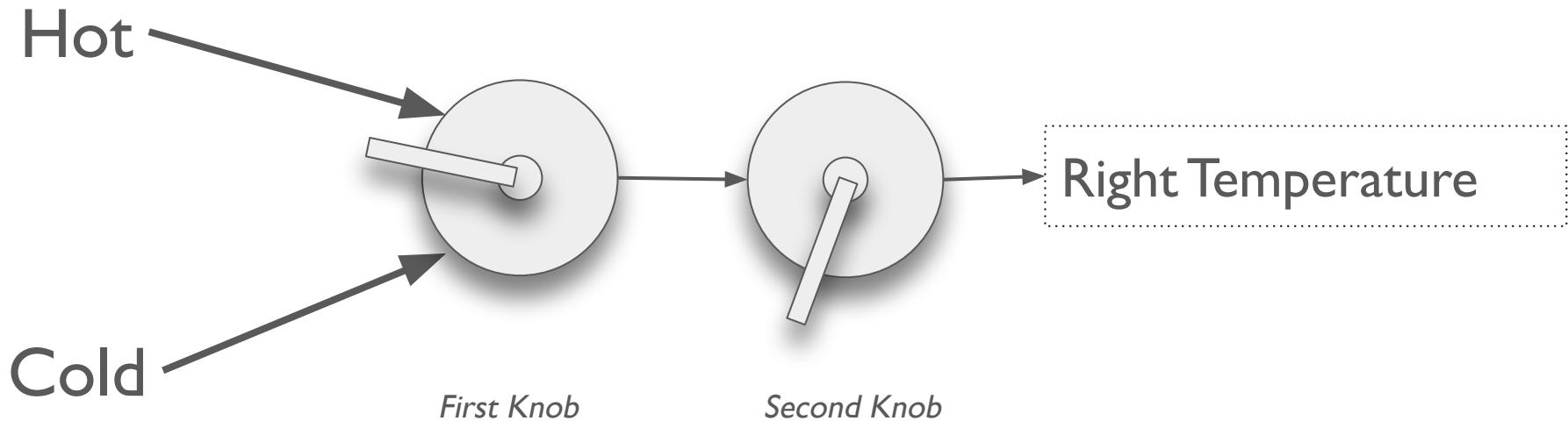
Training Neural Networks - How?

Backpropagation - Two neuron



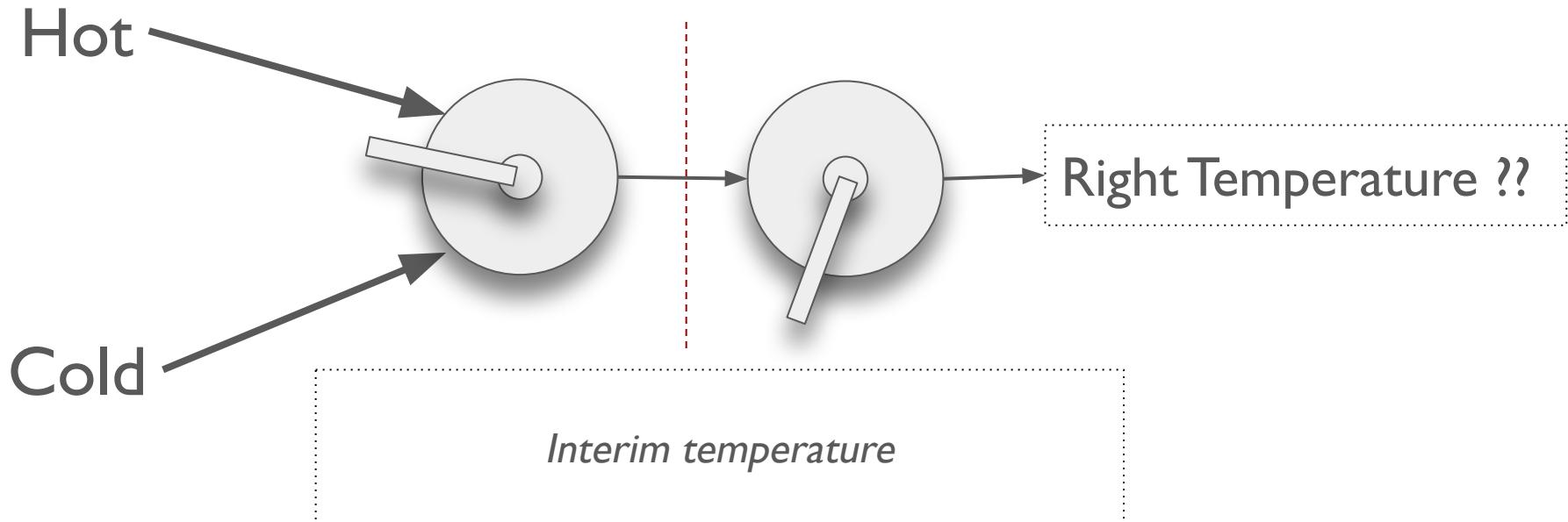
Training Neural Networks - Backpropagation

I. Initialisation



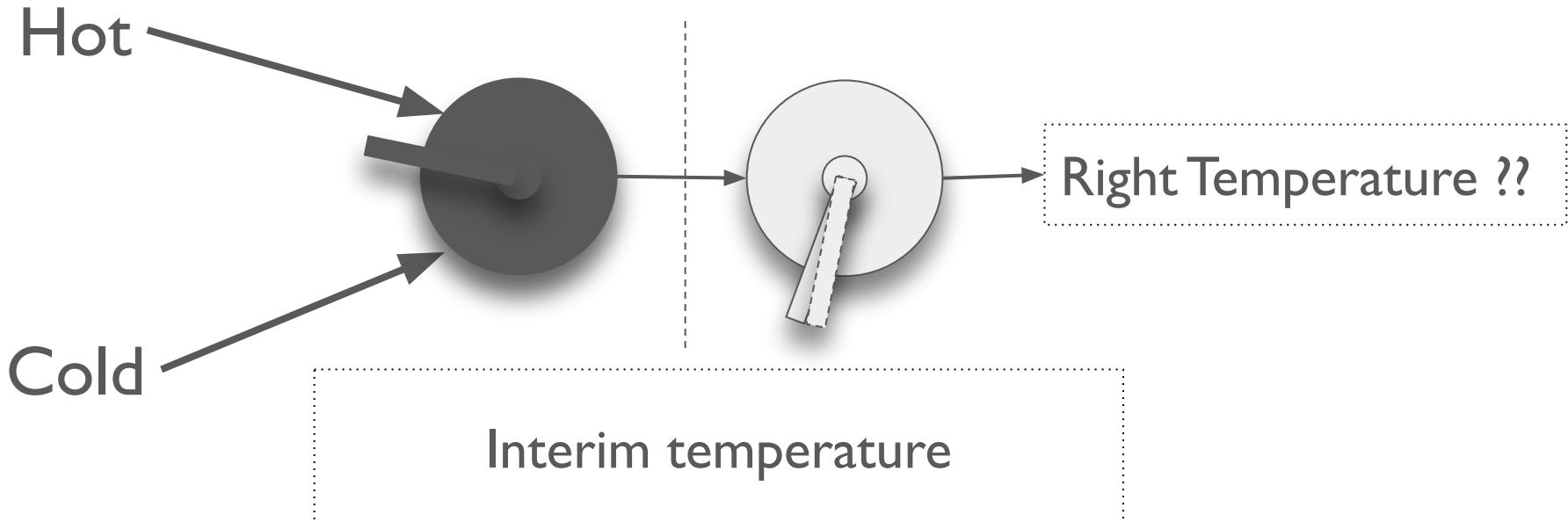
Training Neural Networks - Backpropagation

2. Forward Pass



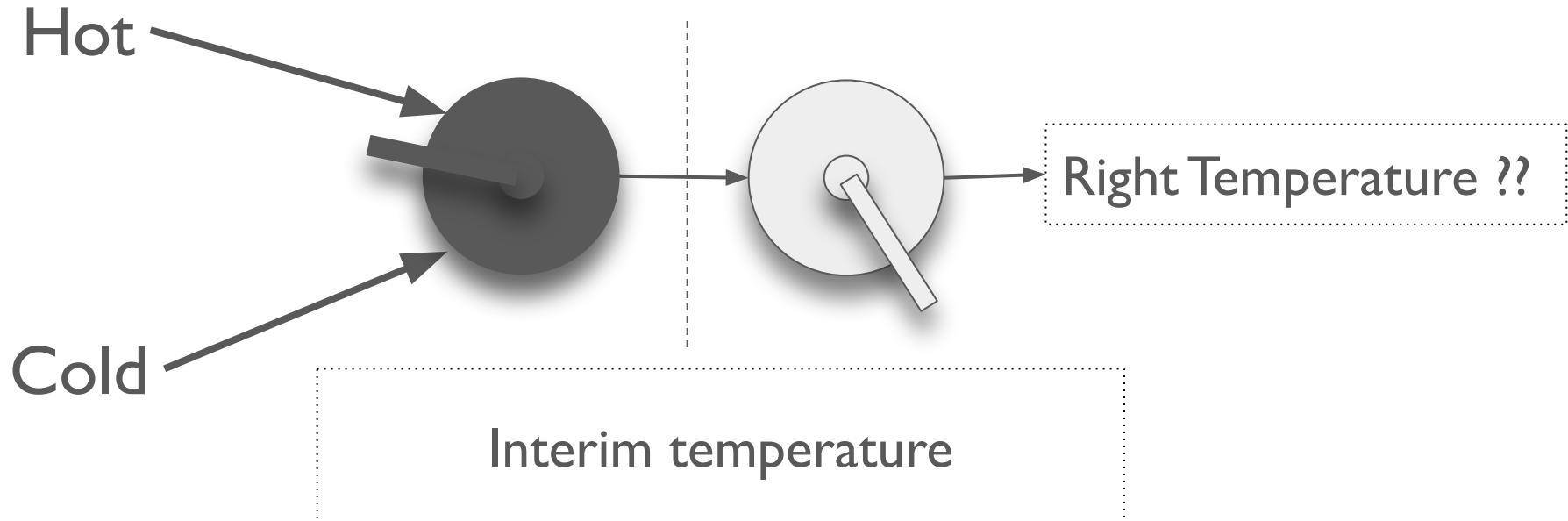
Training Neural Networks - Backpropagation

3. Reverse Pass - Tweak Second knob.



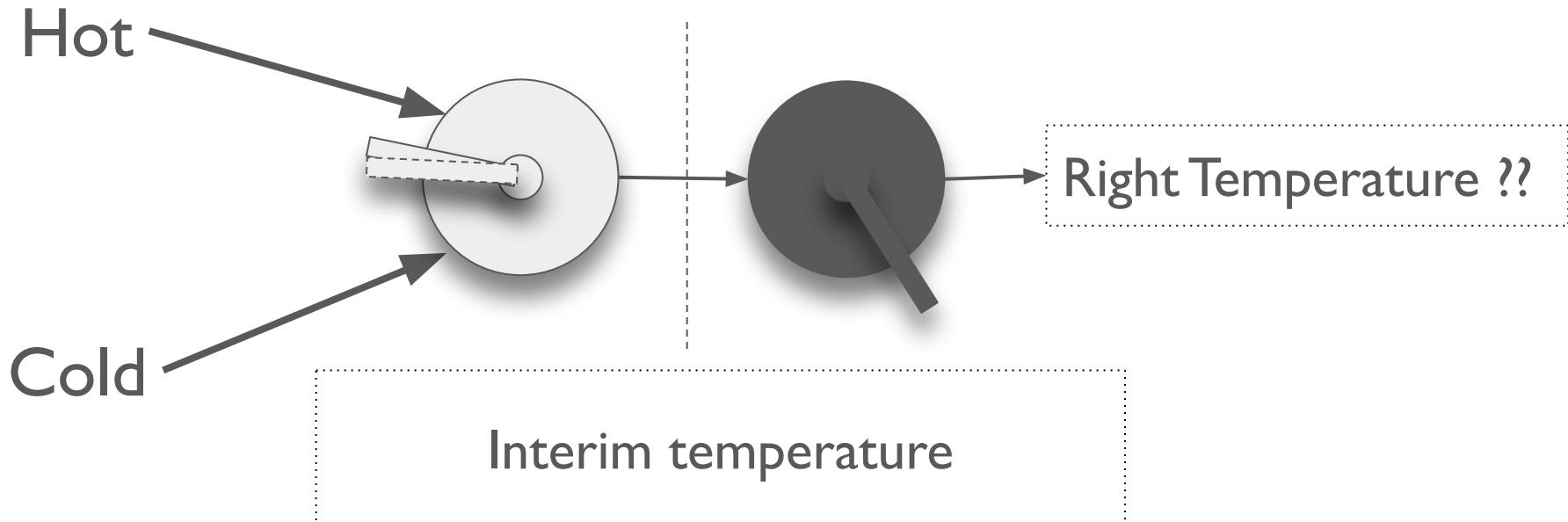
Training Neural Networks - Backpropagation

3. Reverse Pass - Tweak first knob.



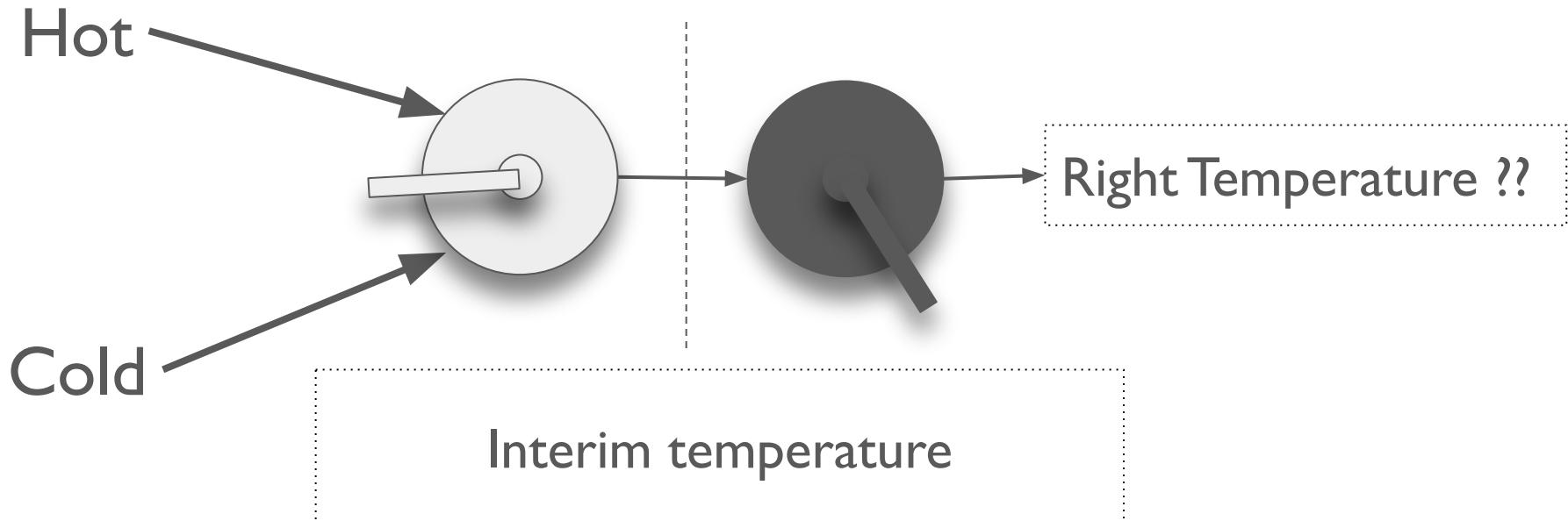
Training Neural Networks - Backpropagation

4. Reverse Pass - Tweak first knob.



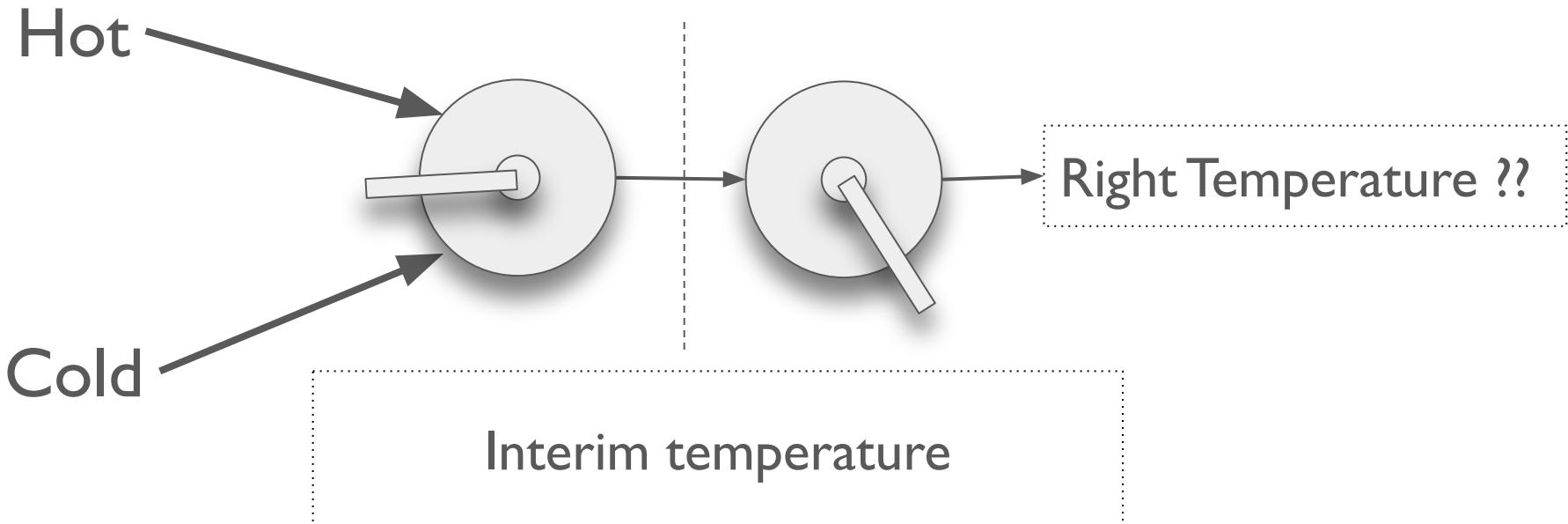
Training Neural Networks - Backpropagation

4. Reverse Pass - Tweak first knob.



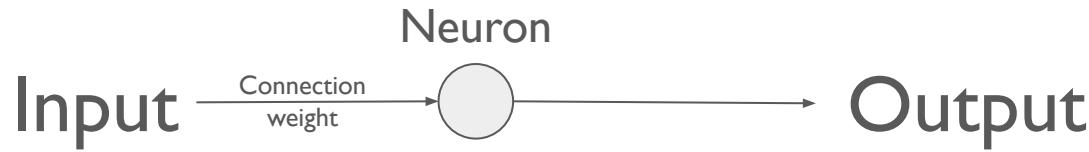
Training Neural Networks - Backpropagation

Back to I. Forward Pass - Take next instance

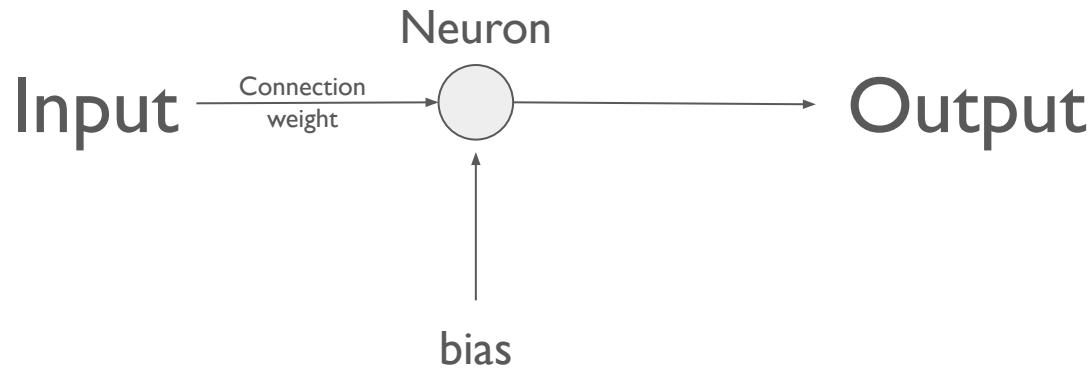


Artificial Neuron - Let's get real!

Artificial Neuron - Let's get real!

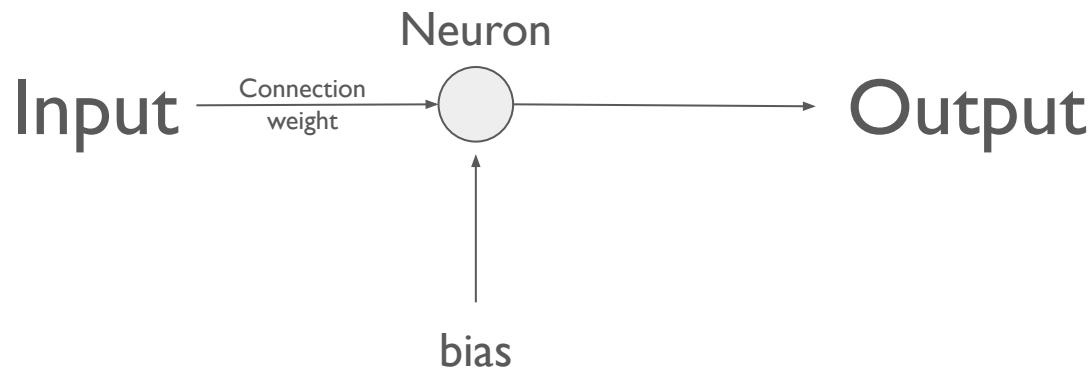


Artificial Neuron - Let's get real!



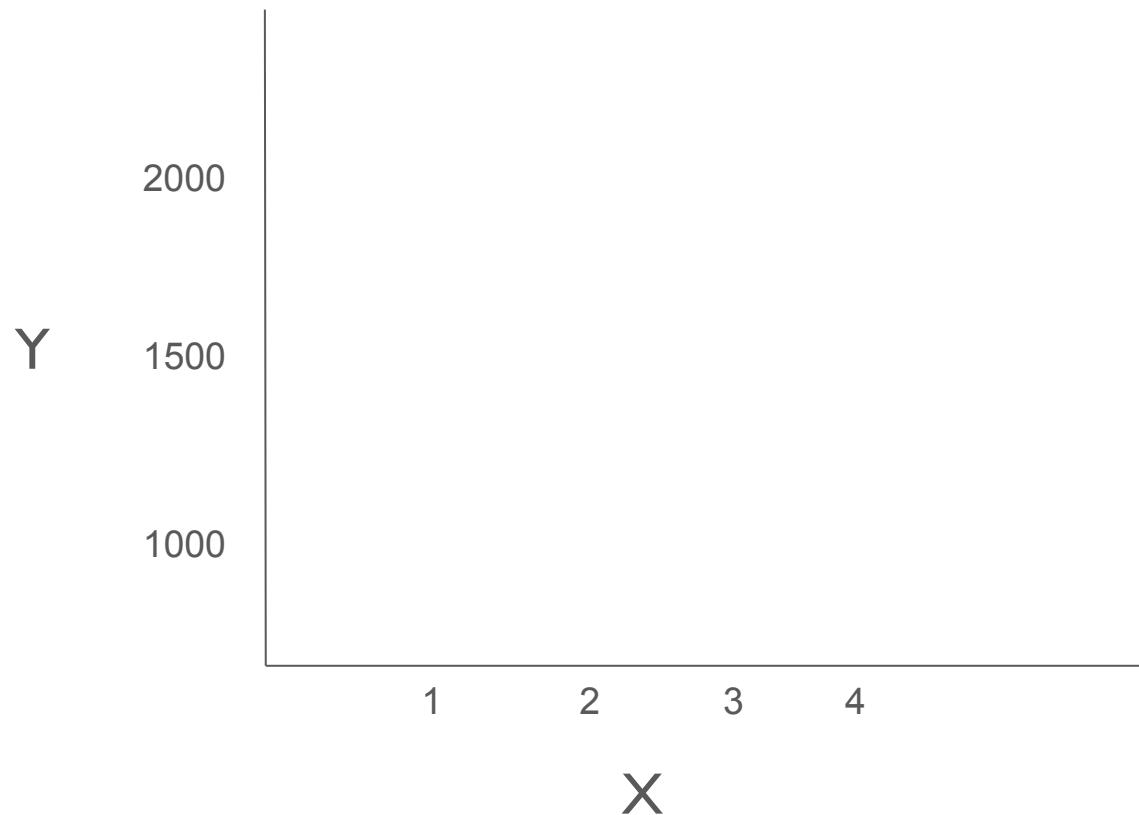
Artificial Neuron - Let's get real!

$$\text{Output} = \text{Input} * \text{weight} + \text{bias}$$

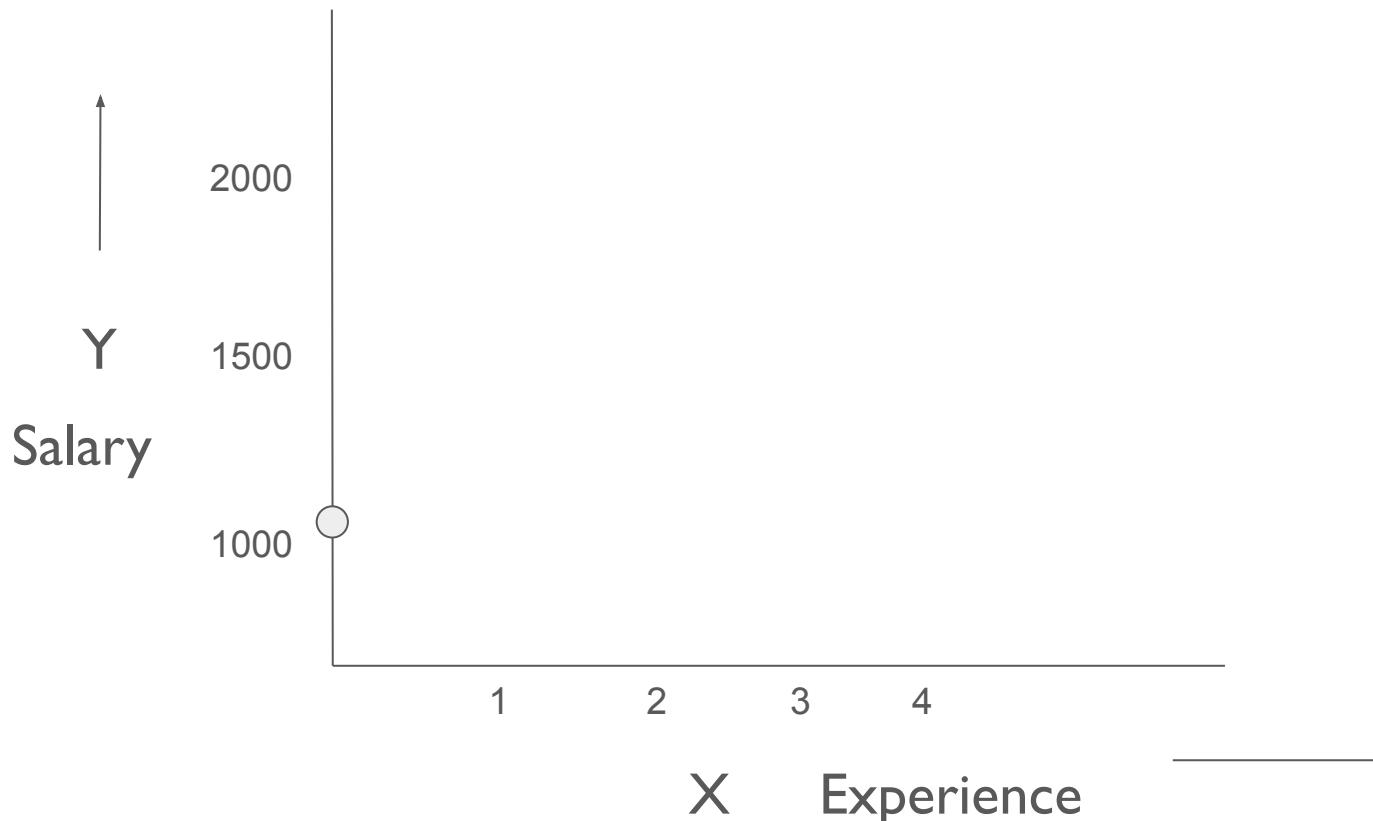


Why do we need both weight and bias?

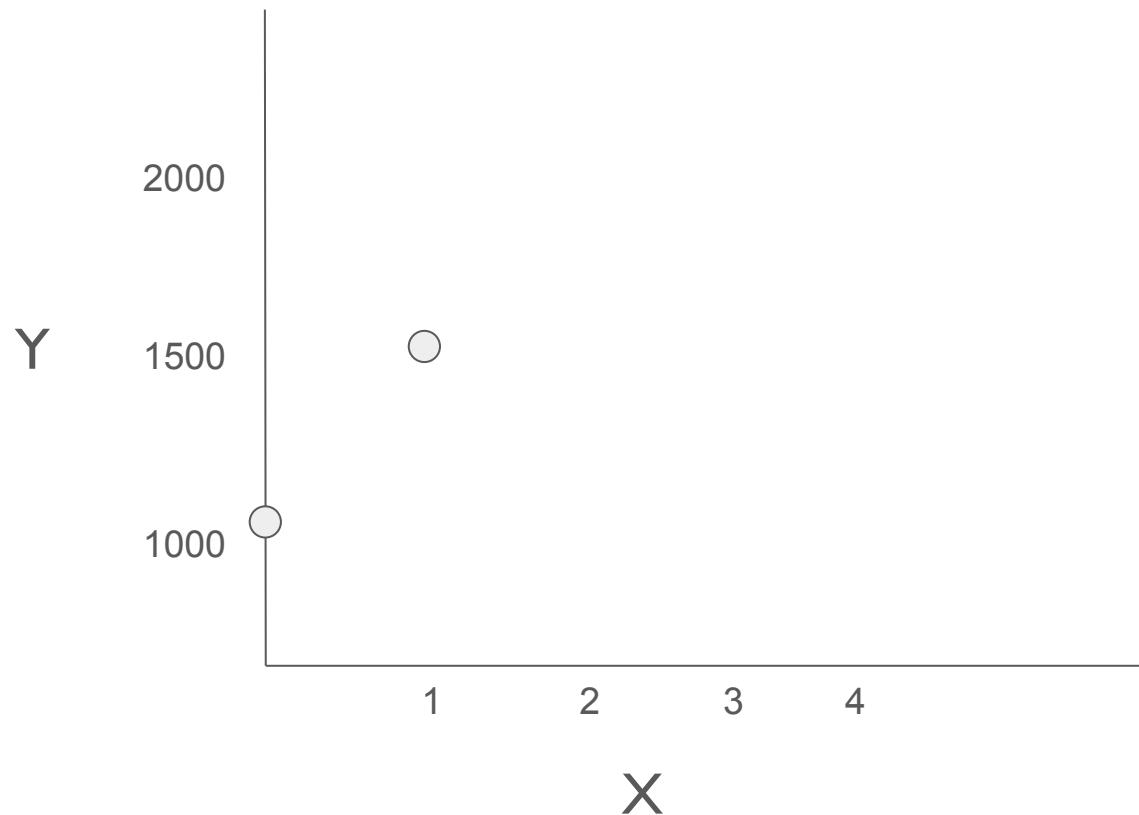
Predicting Salary



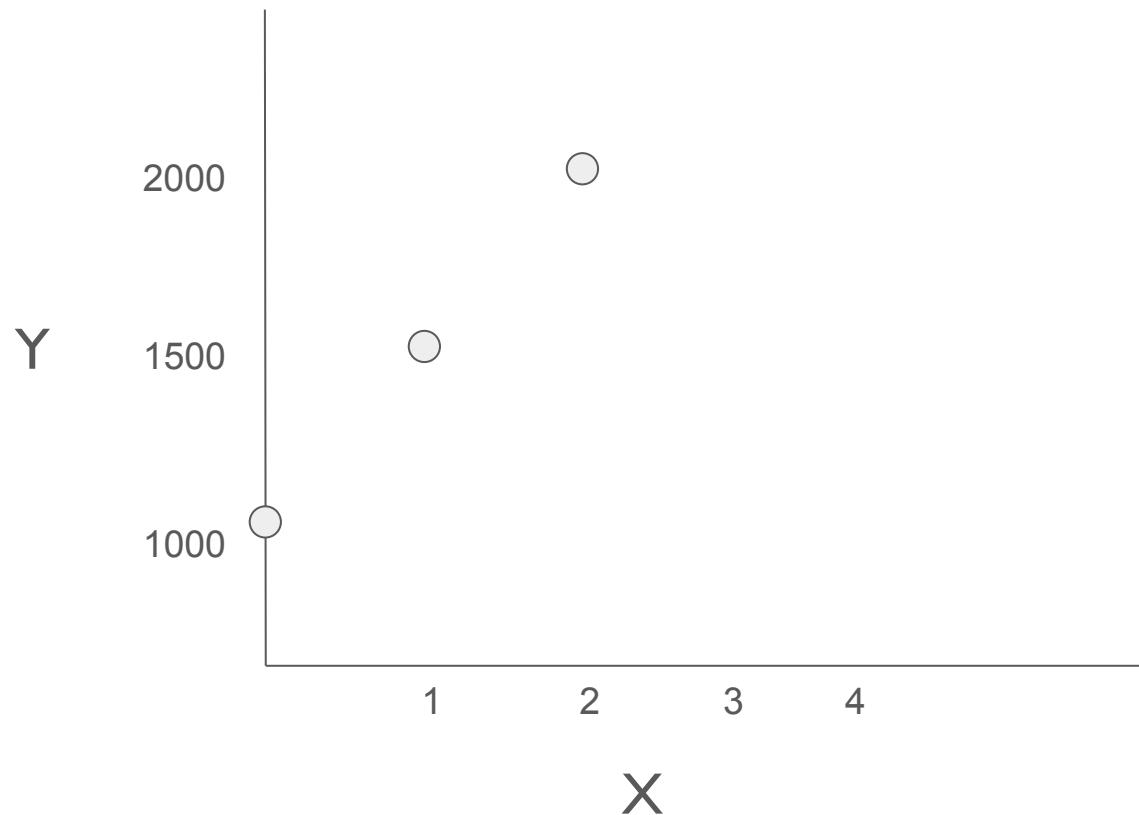
Predicting Salary



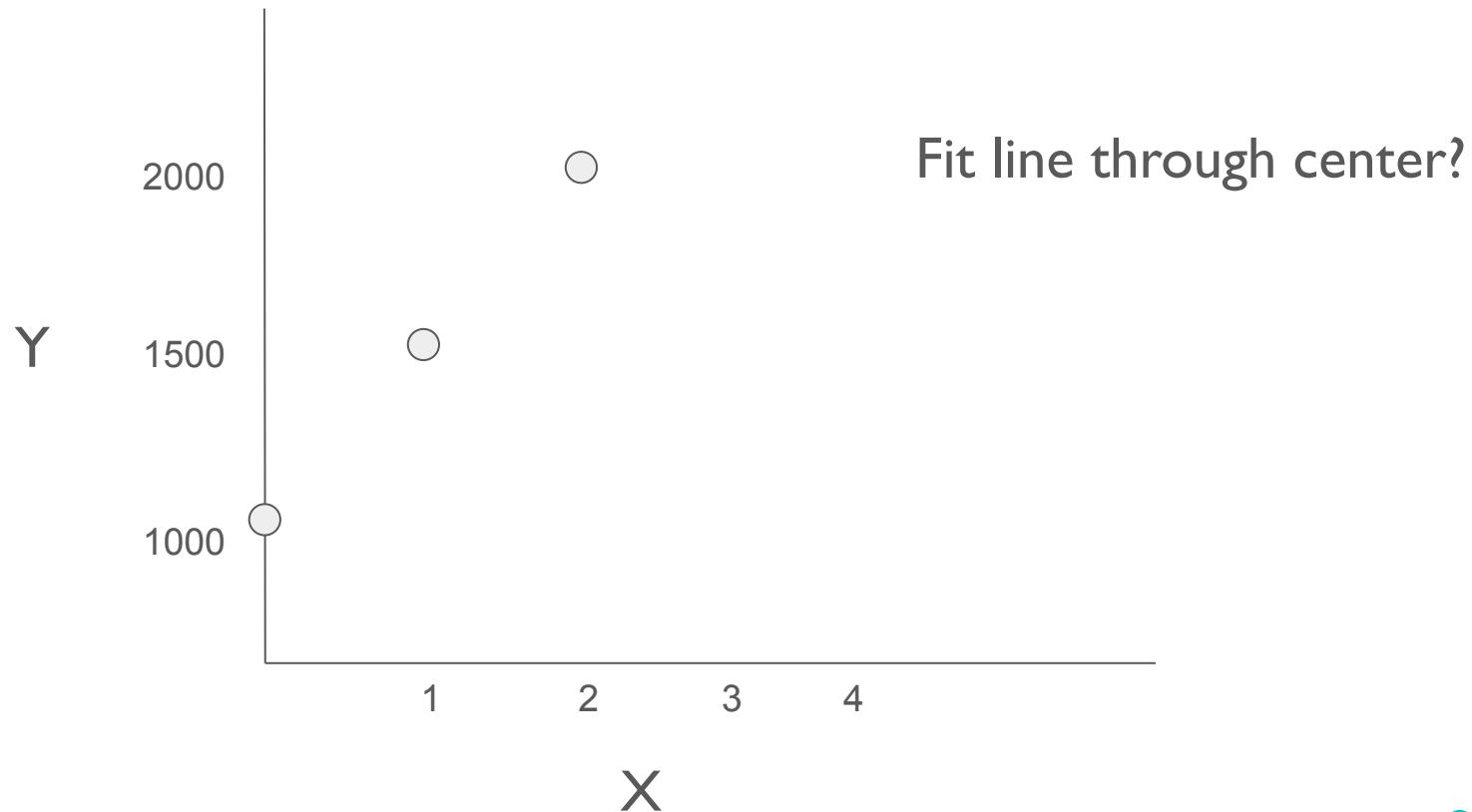
Predicting Salary



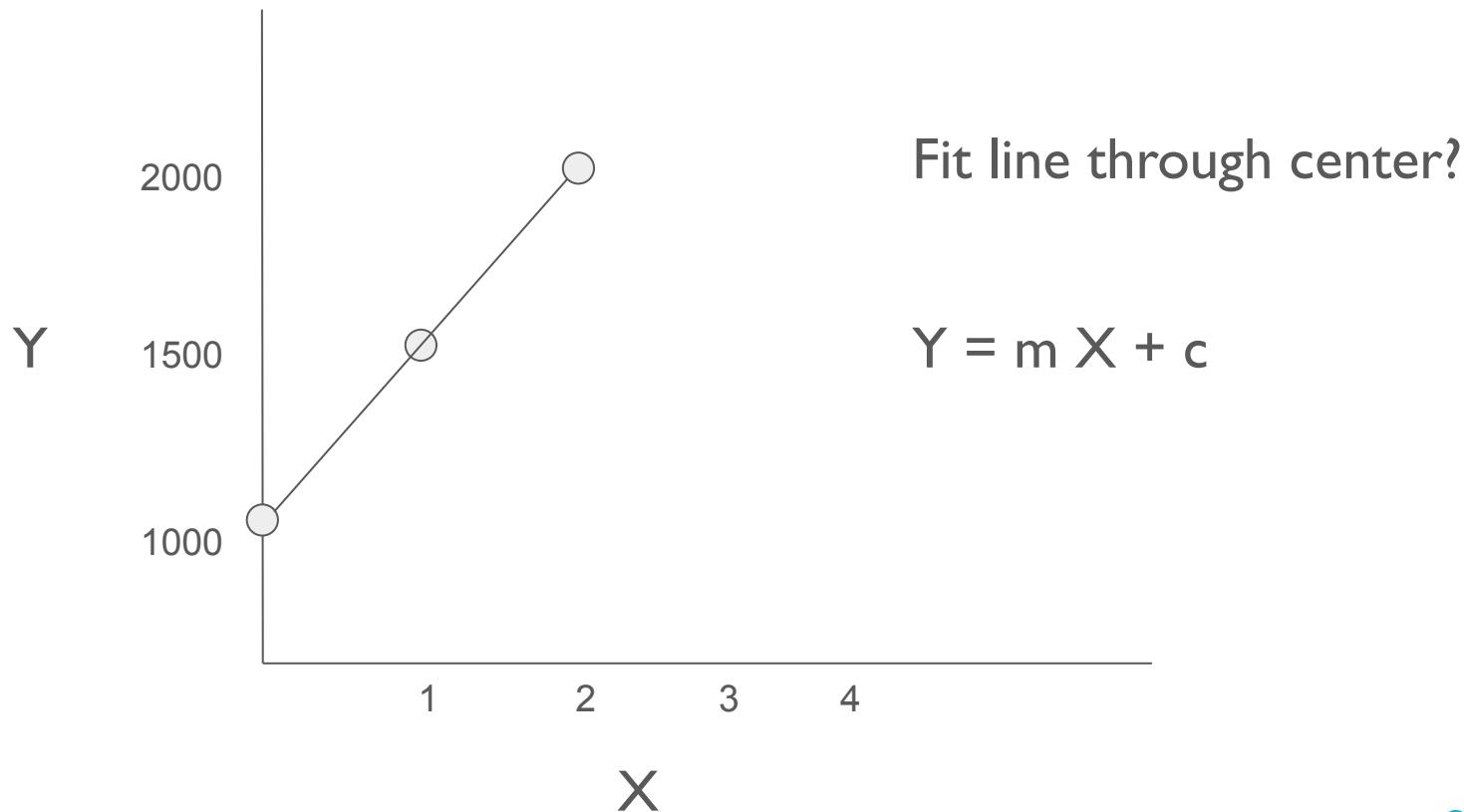
Predicting Salary



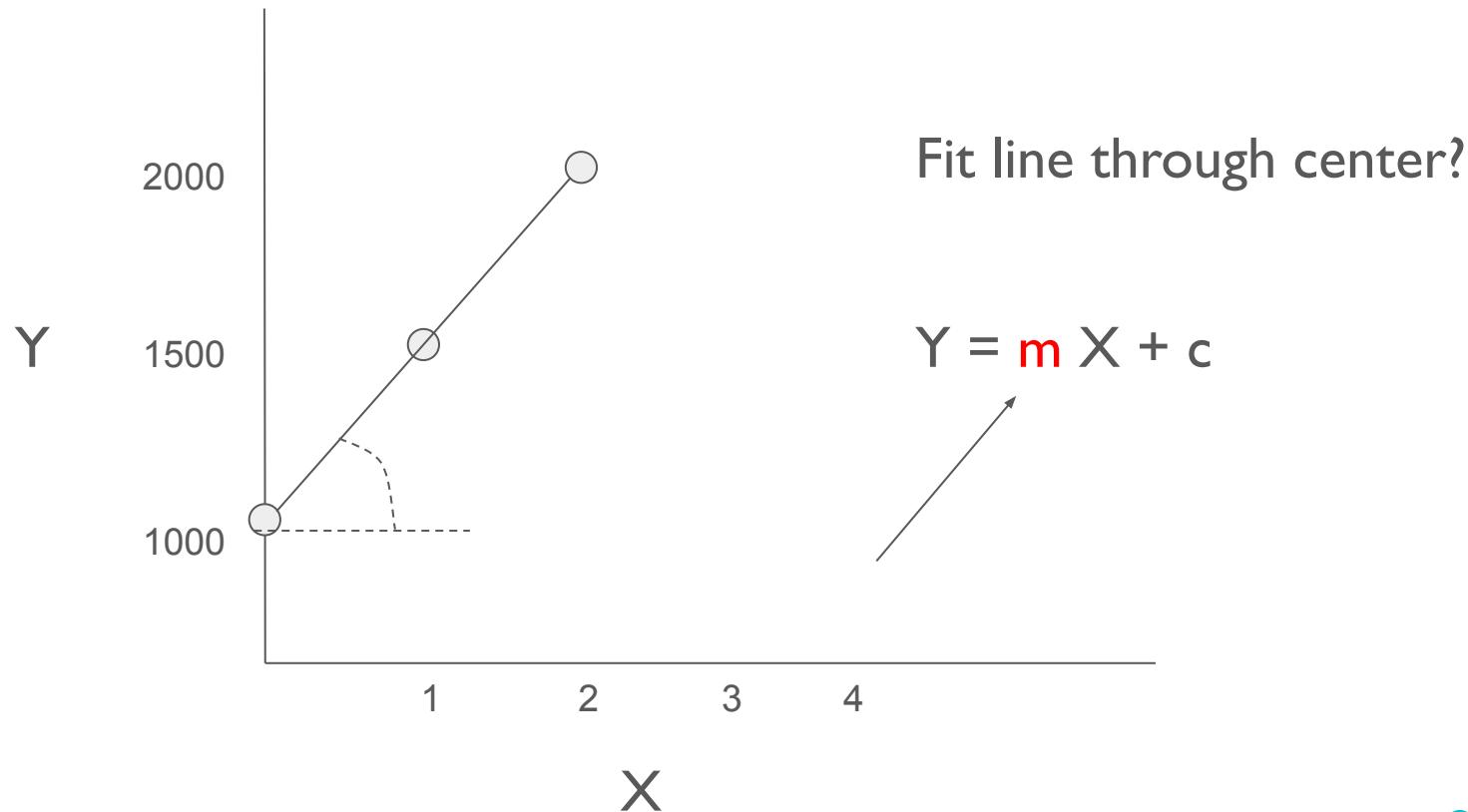
Predicting Salary



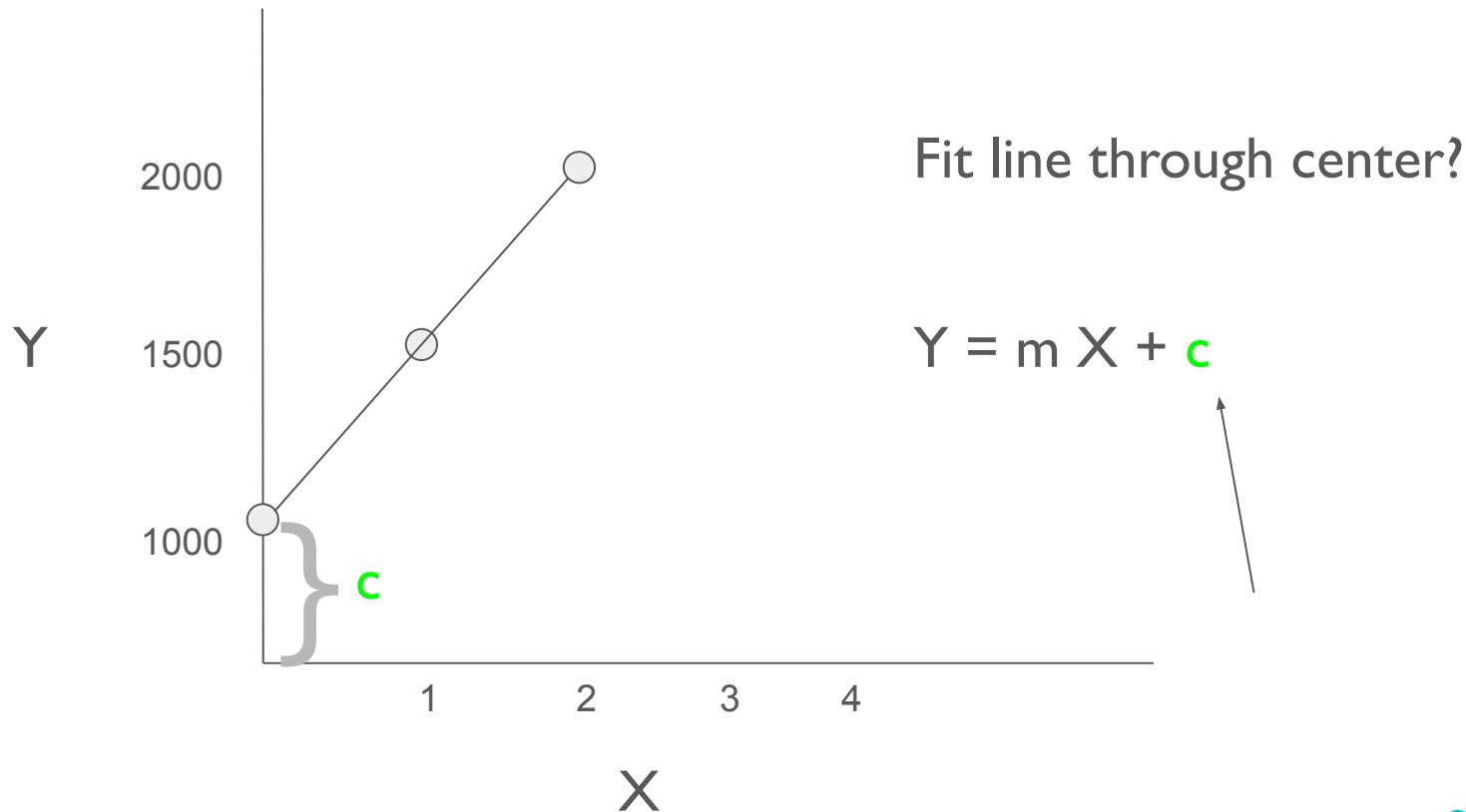
Predicting Salary



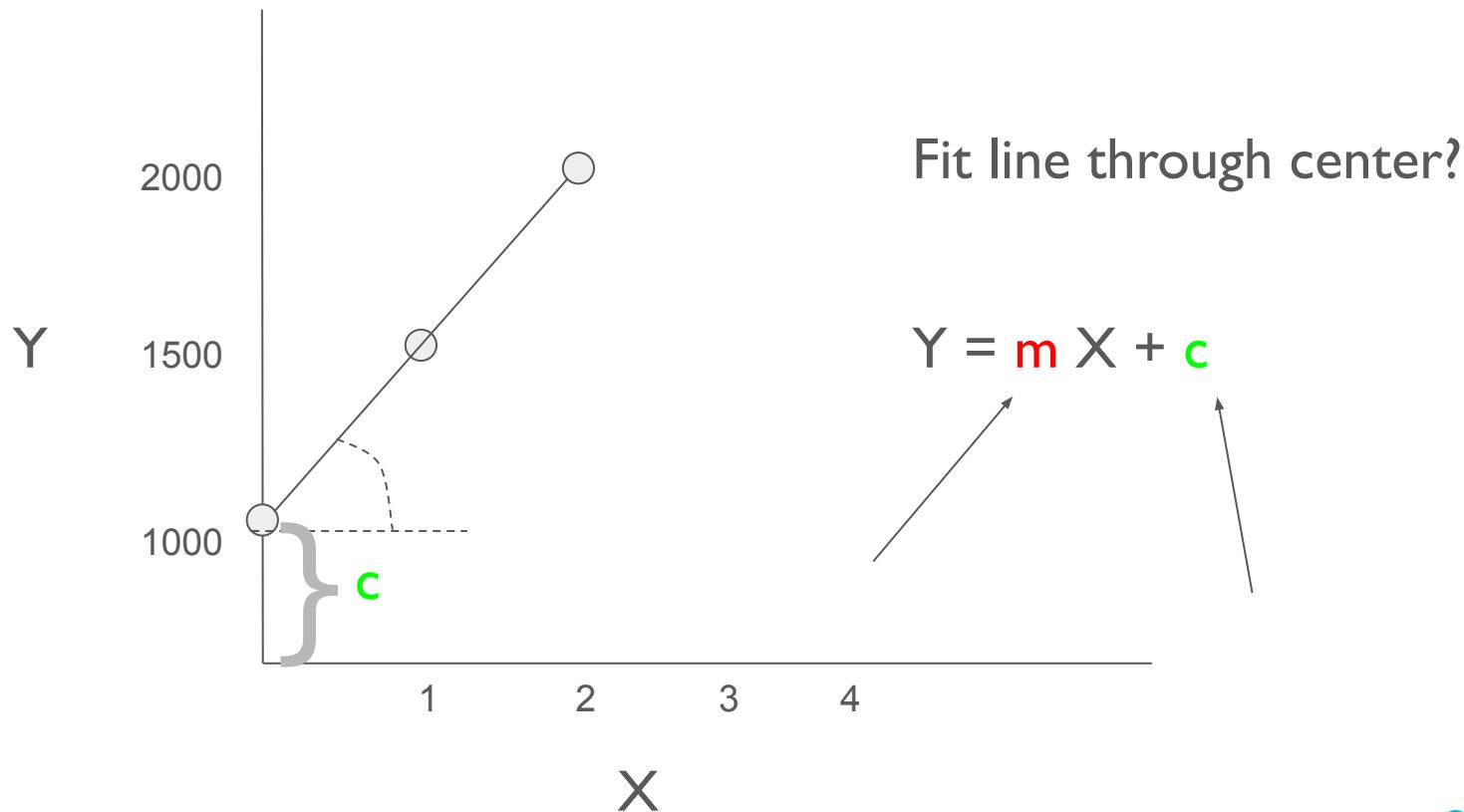
Predicting Salary



Predicting Salary

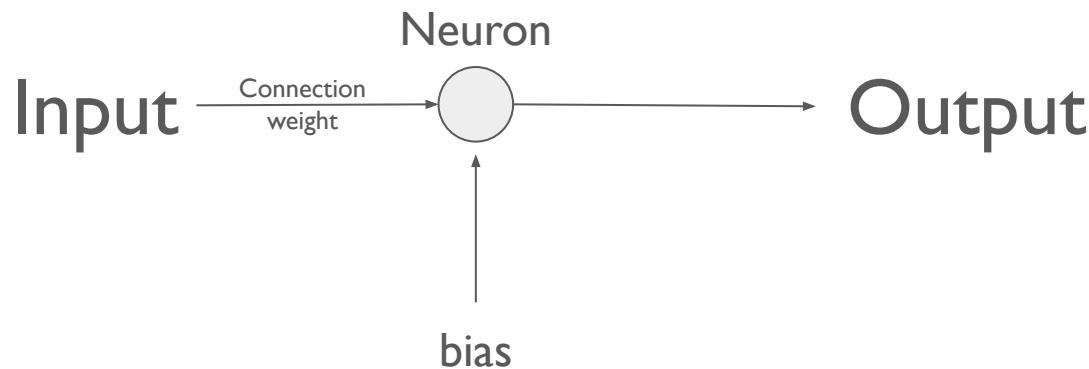


Predicting Salary

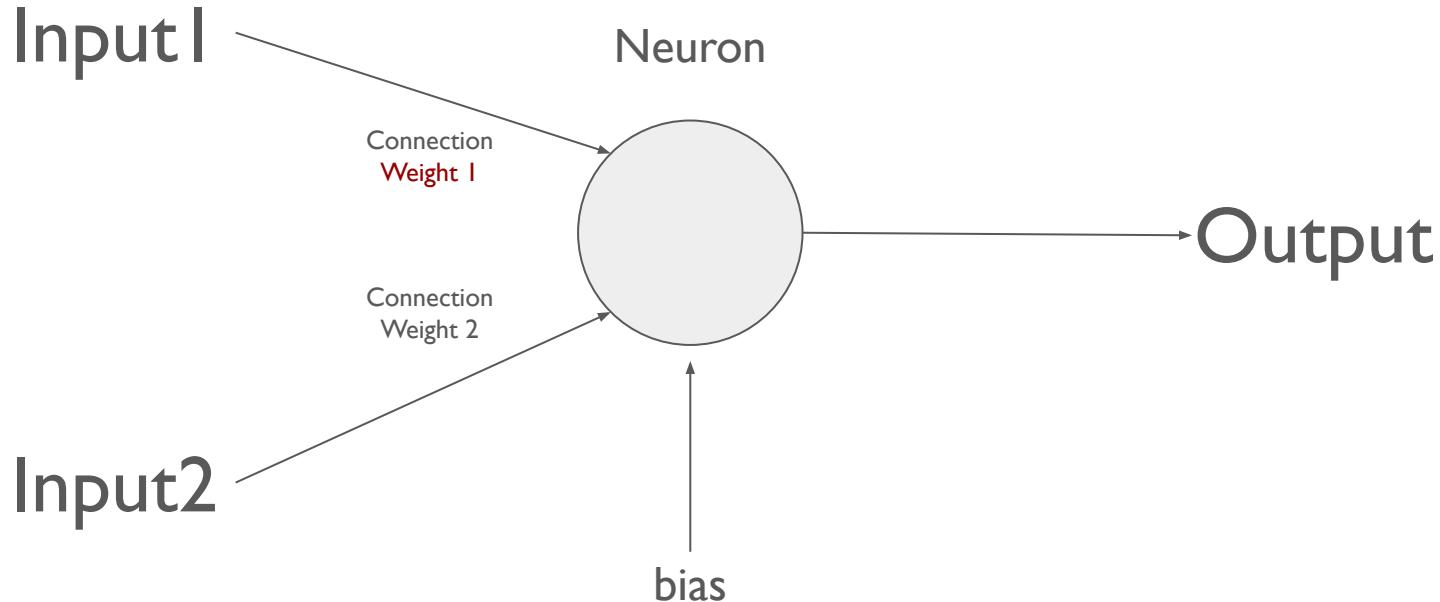


Artificial Neuron - Let's get real!

$\text{Output} = \text{Input} * \text{weight} + \text{bias}$

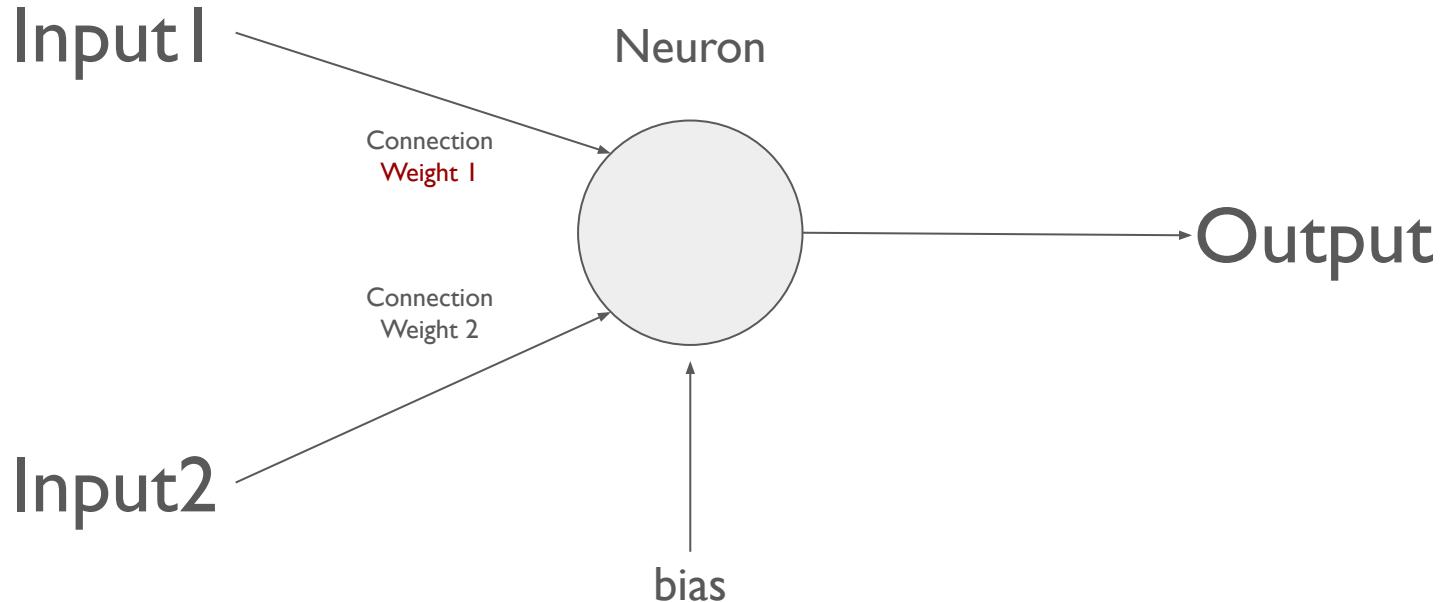


Artificial Neuron - Let's get real!

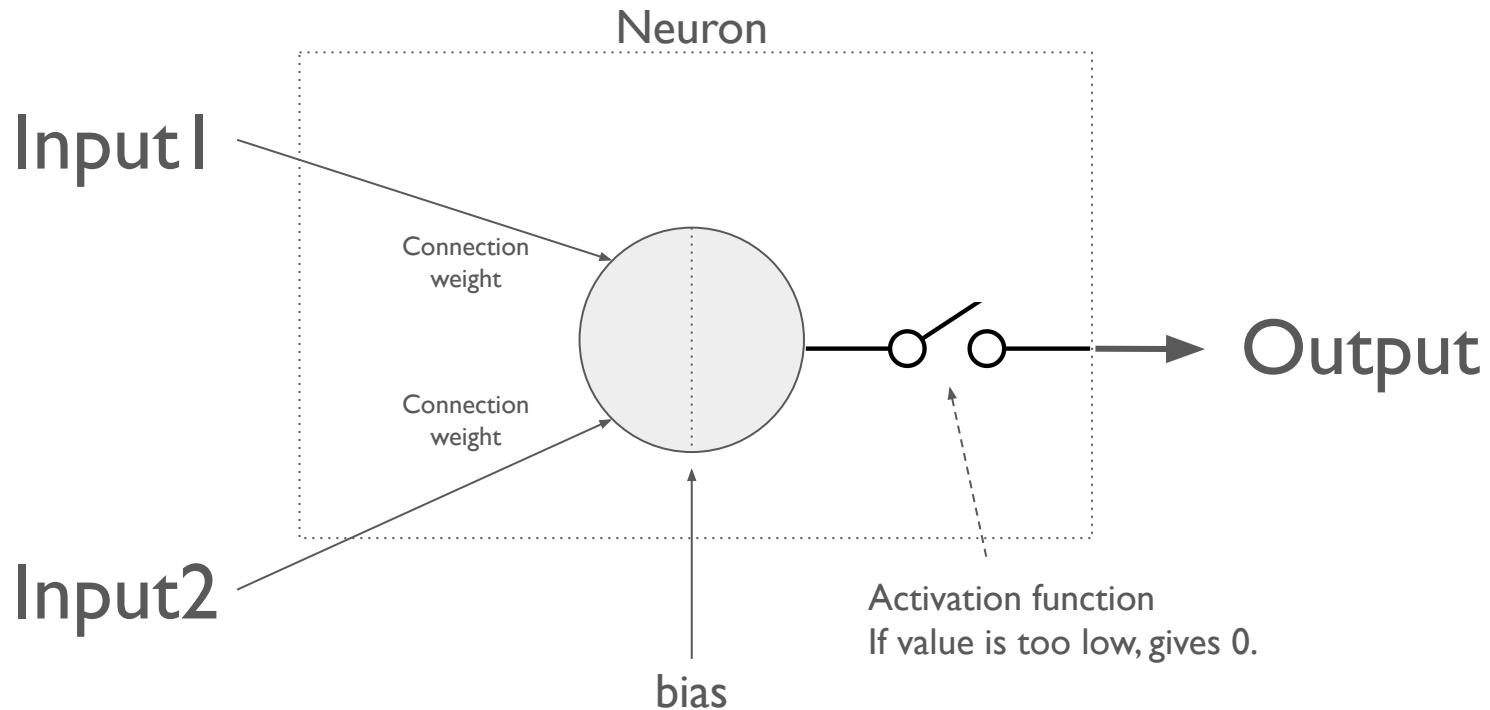


Artificial Neuron - Let's get real!

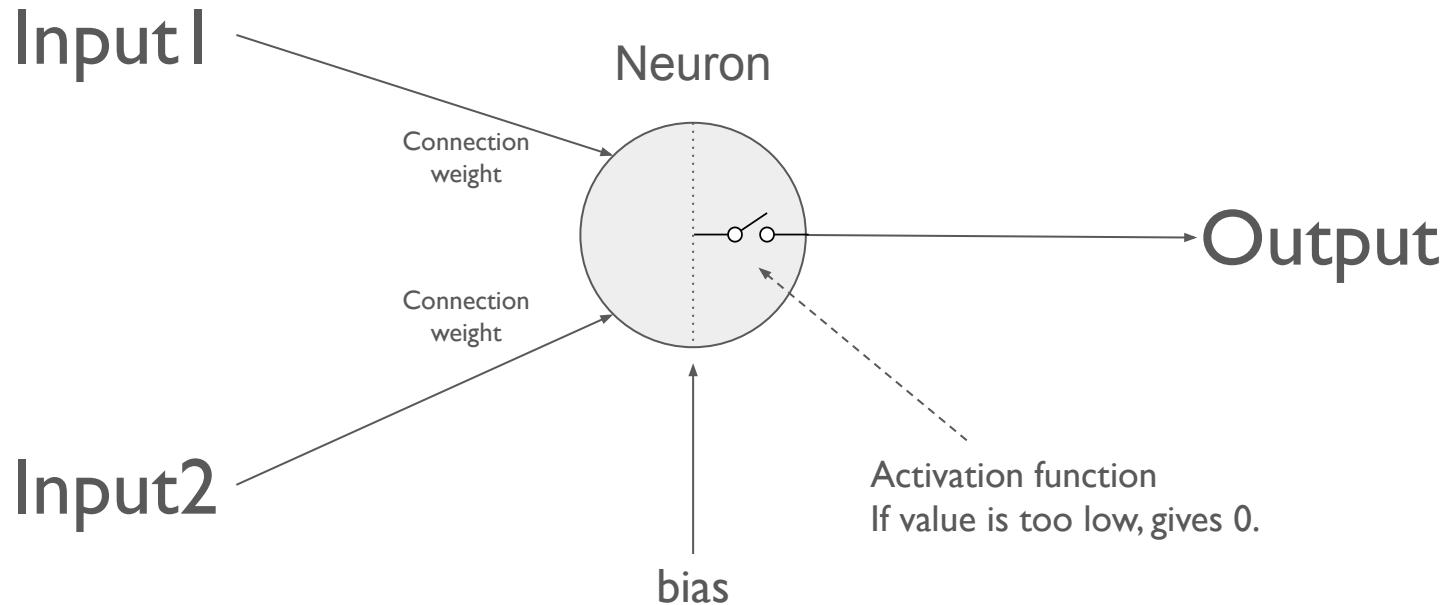
$$\text{Output} = \text{Input1} * \text{weight1} + \text{input2} * \text{weight2} + \text{bias}$$



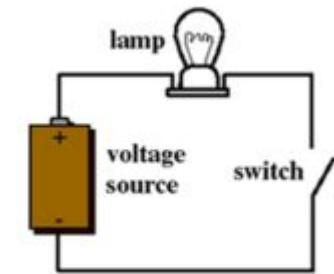
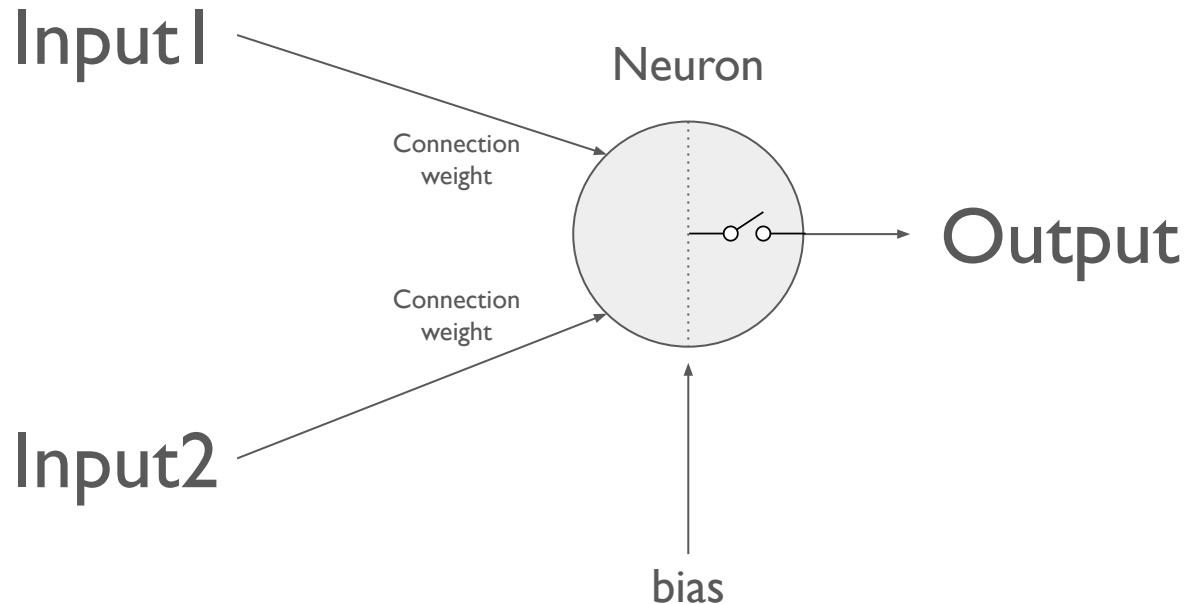
Artificial Neuron - Let's get real!



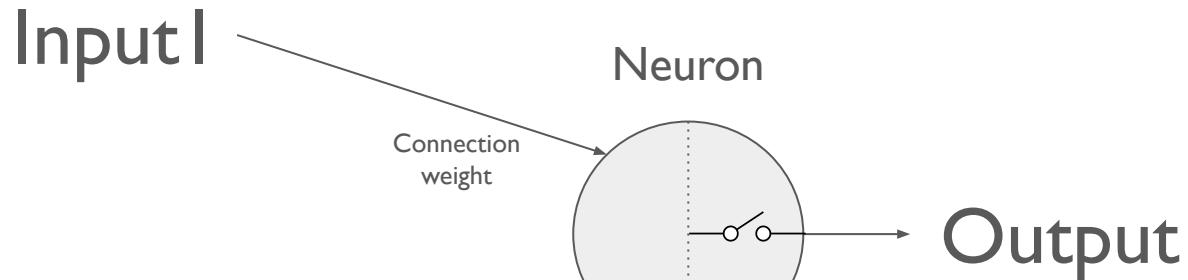
Artificial Neuron - Let's get real!



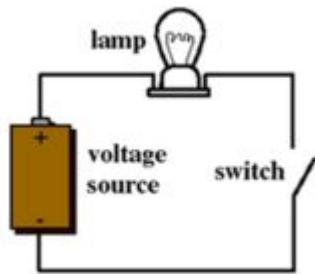
Artificial Neuron - To make it like a electric circuit



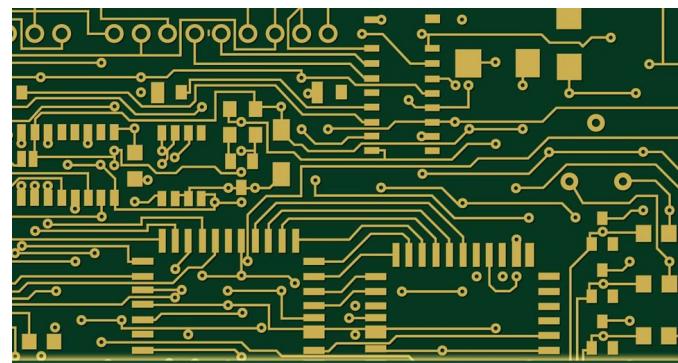
Artificial Neuron - To make it like a electric circuit



Input2



Output



Electric Circuits

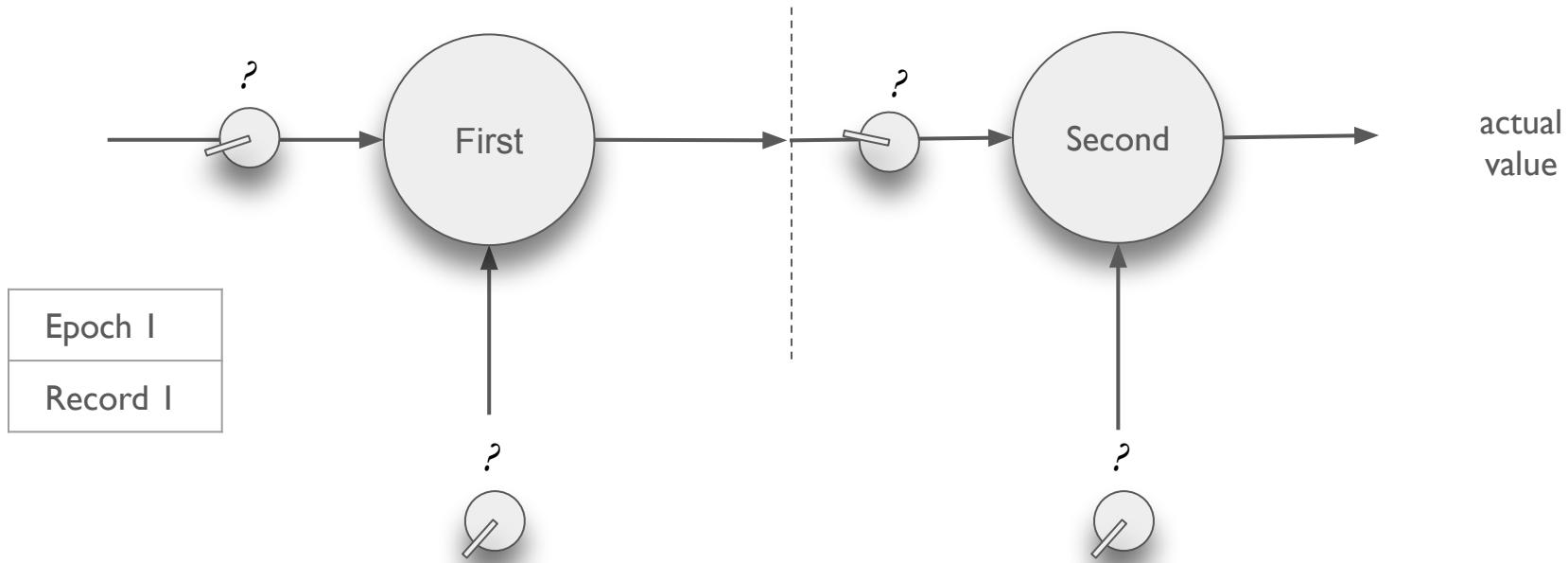
Backpropagation with example

Training a network based on years of work experience as input and Salary as output

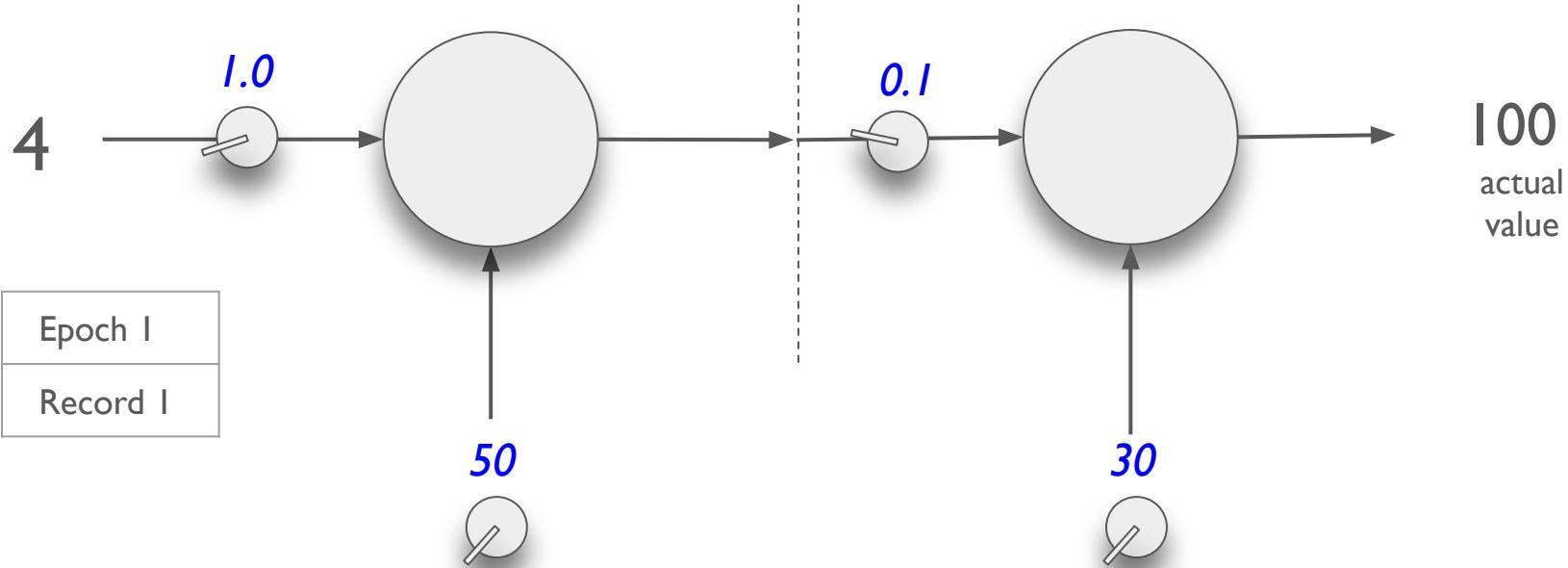
EMP	Work Ex	Salary
emp1	4	100
emp2	5	130
emp3	2	30
emp4	5	140

Training a network based on years of work experience as input and Salary as output

EMP	Work Ex	Salary
emp1	4	100
emp2	5	130
emp3	2	30
emp4	5	140

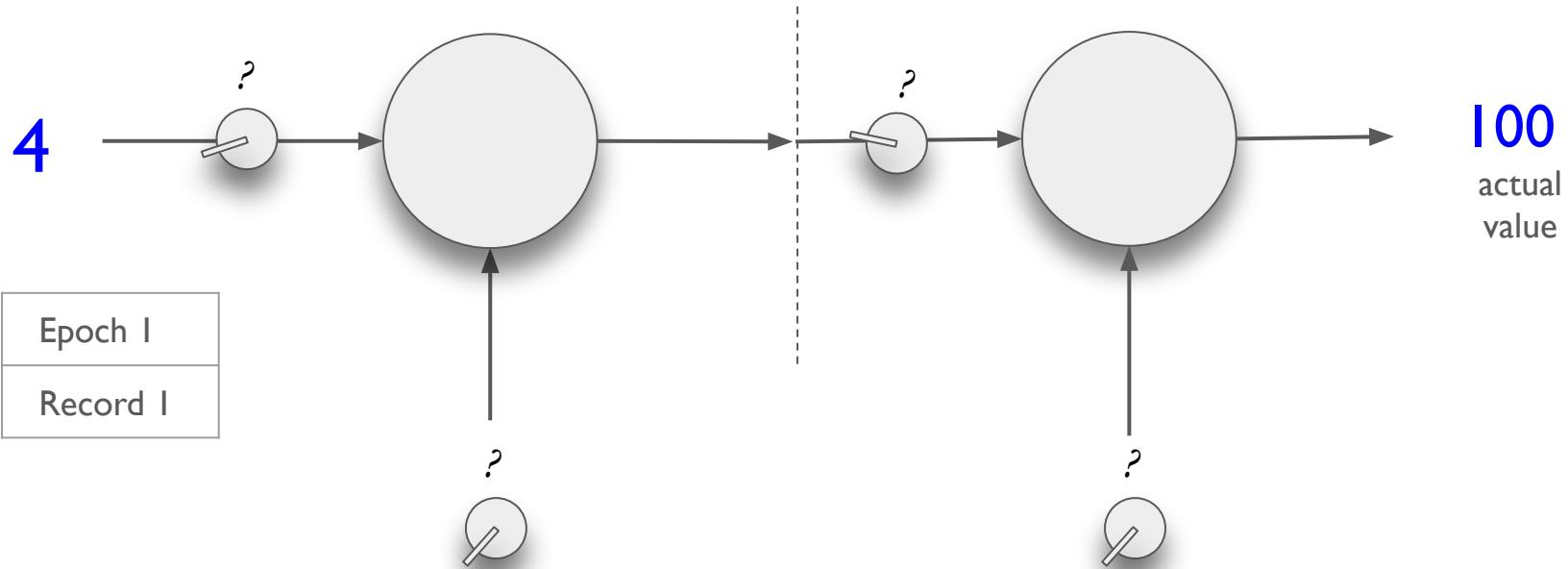


EMP	Work Ex	Salary
empl	4	100
emp2	5	130
emp3	2	30
emp4	5	140



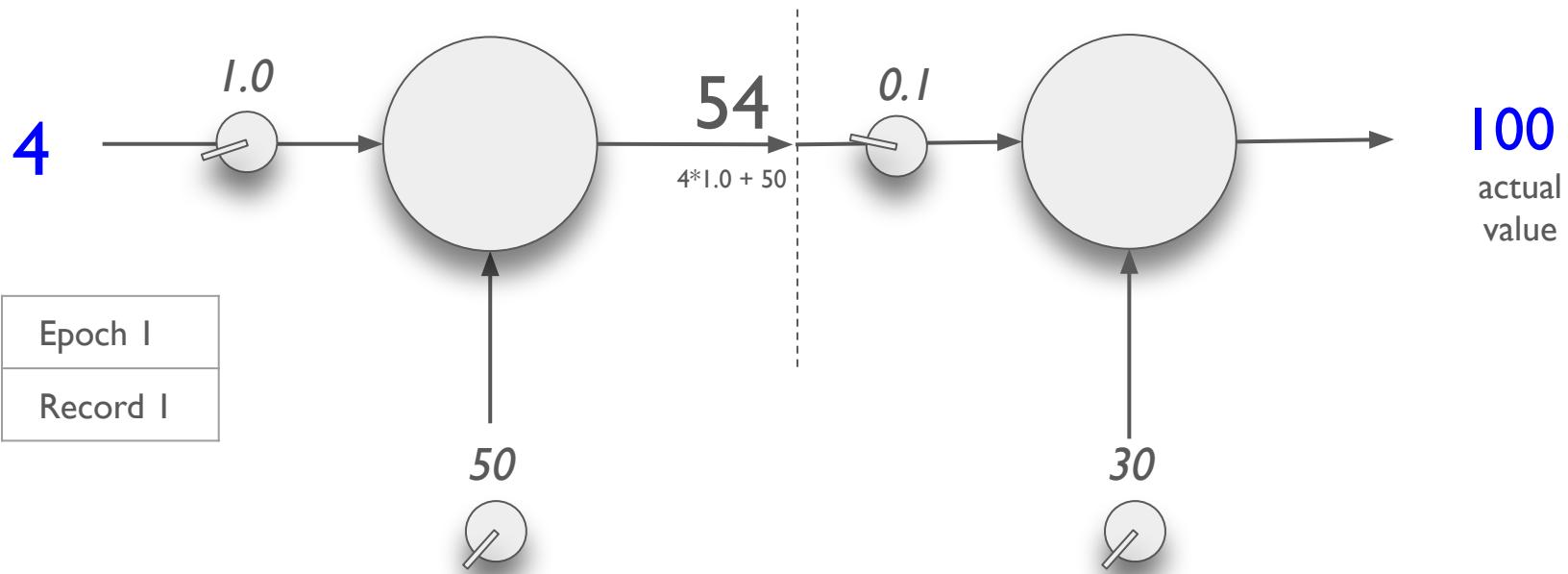
EMP	Work Ex	Salary
empl	4	100
emp2	5	130
emp3	2	30
emp4	5	140

Two Neurons - 4 knobs.
Take first record...



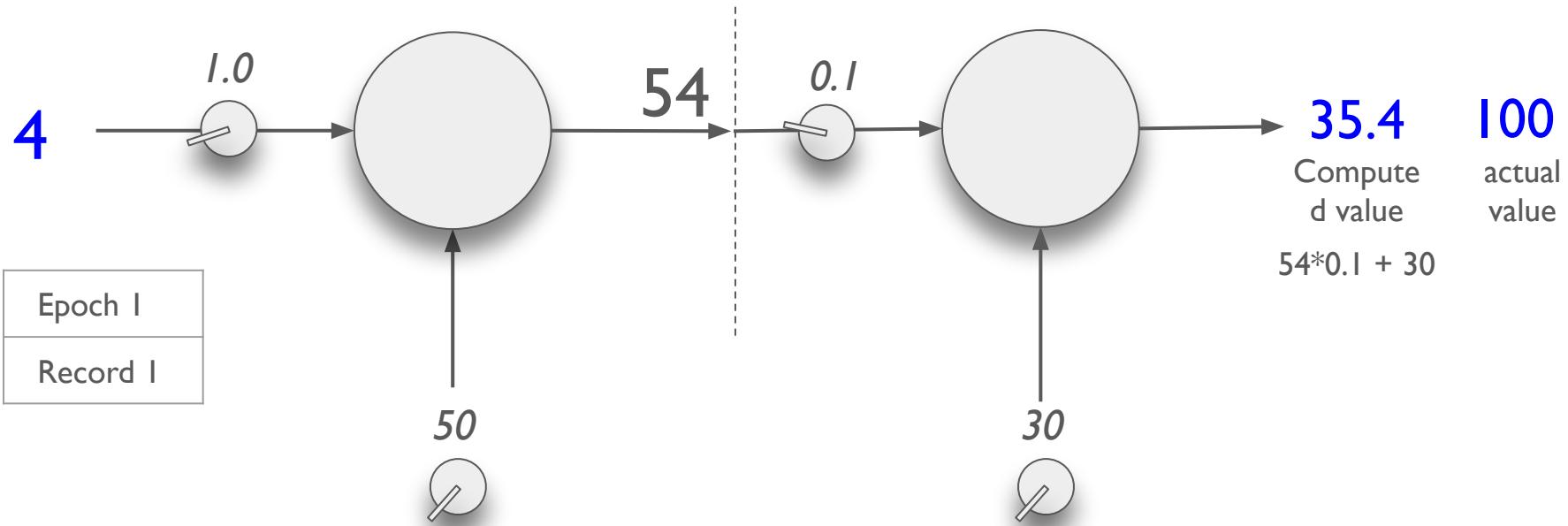
EMP	Work Ex	Salary
emp1	4	100
emp2	5	130
emp3	2	30
emp4	5	140

← Do the forward pass



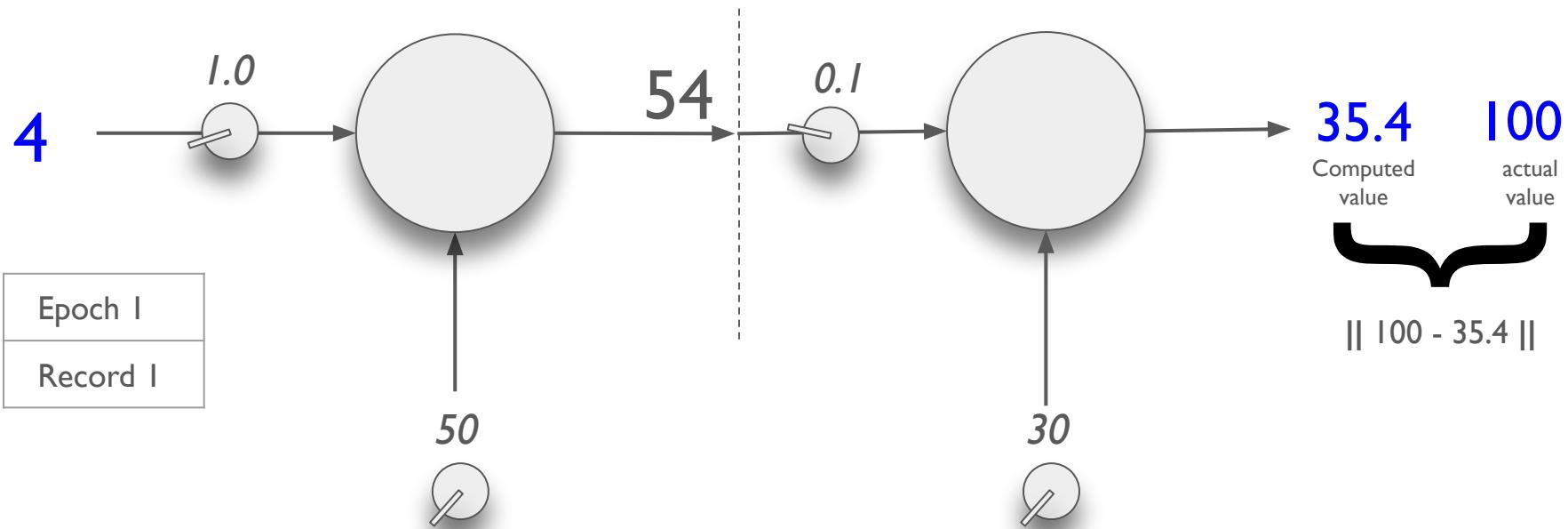
EMP	Work Ex	Salary
empl	4	100
emp2	5	130
emp3	2	30
emp4	5	140

← Do the forward pass



EMP	Work Ex	Salary
emp1	4	100
emp2	5	130
emp3	2	30
emp4	5	140

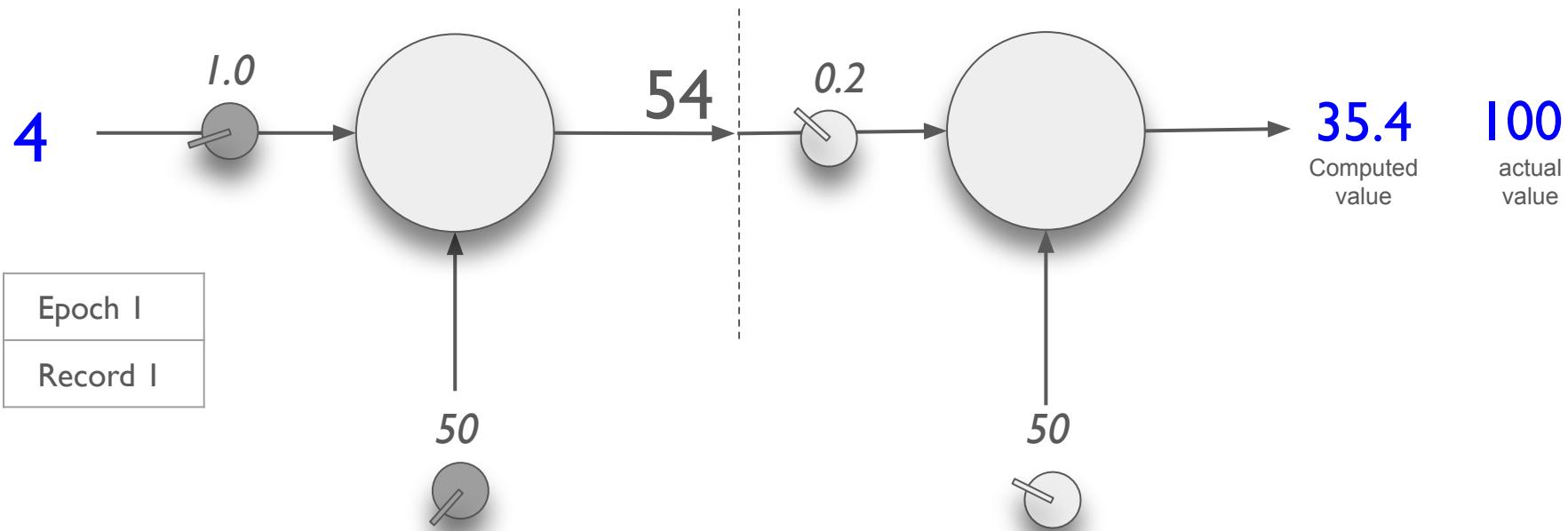
← Compute the error



EMP	Work Ex	Salary
empl	4	100
emp2	5	130
emp3	2	30
emp4	5	140



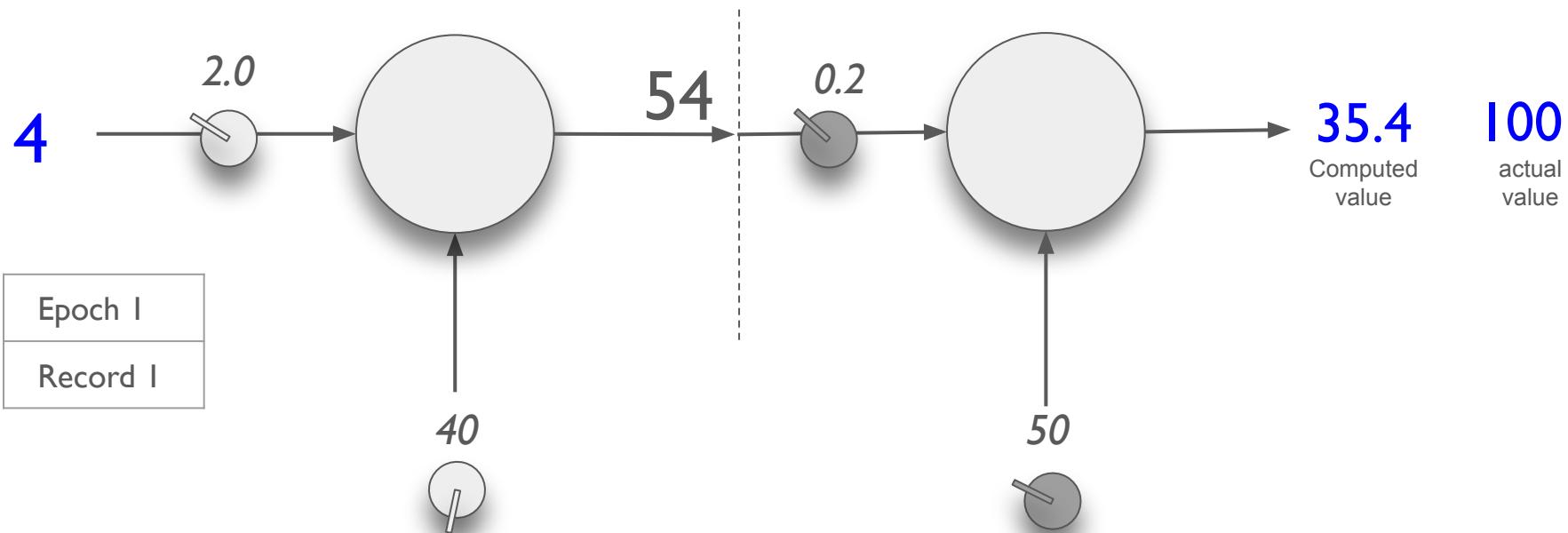
Back Pass: Tweak the knob of the second neuron



EMP	Work Ex	Salary
empl	4	100
emp2	5	130
emp3	2	30
emp4	5	140

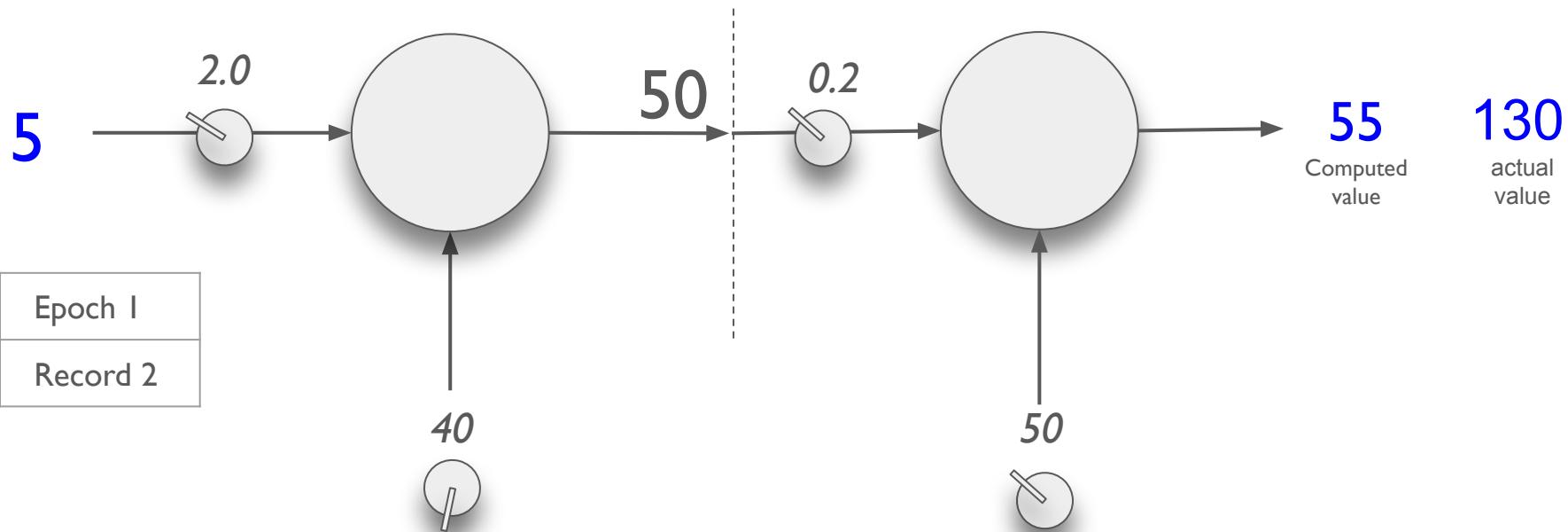


Back Pass: Freeze the second knob, tweak the knob of the first neuron



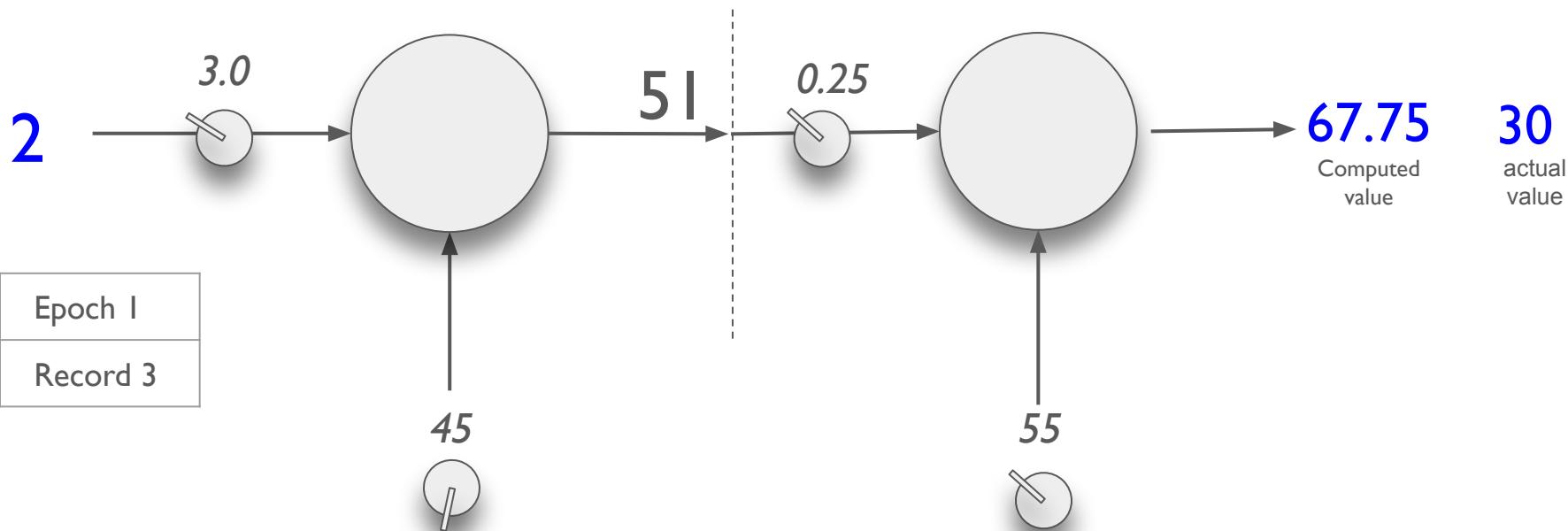
EMP	Work Ex	Salary
empl	4	100
emp2	5	130
emp3	2	30
emp4	5	140

Next Record. Repeat the process,
Tweak the weights...



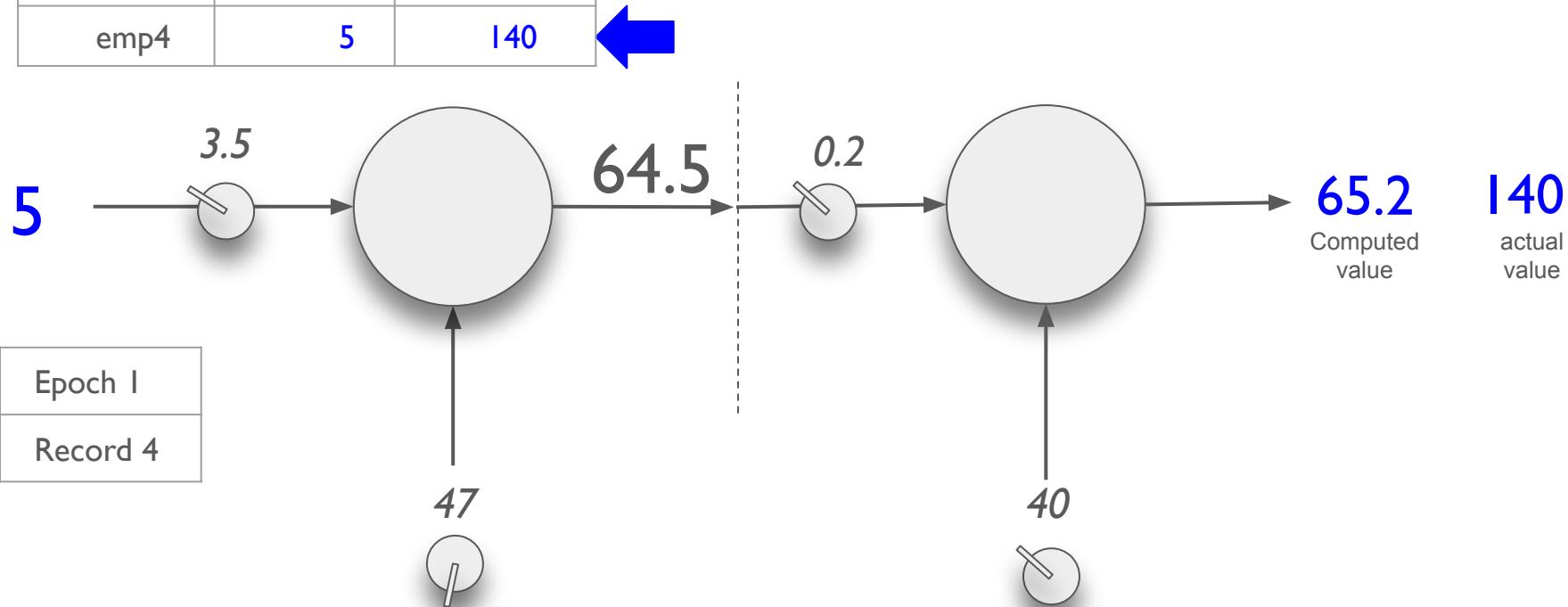
EMP	Work Ex	Salary
empl	4	100
emp2	5	130
emp3	2	30
emp4	5	140

Next Record. Repeat the process,
Tweak the weights...



EMP	Work Ex	Salary
empl	4	100
emp2	5	130
emp3	2	30
emp4	5	140

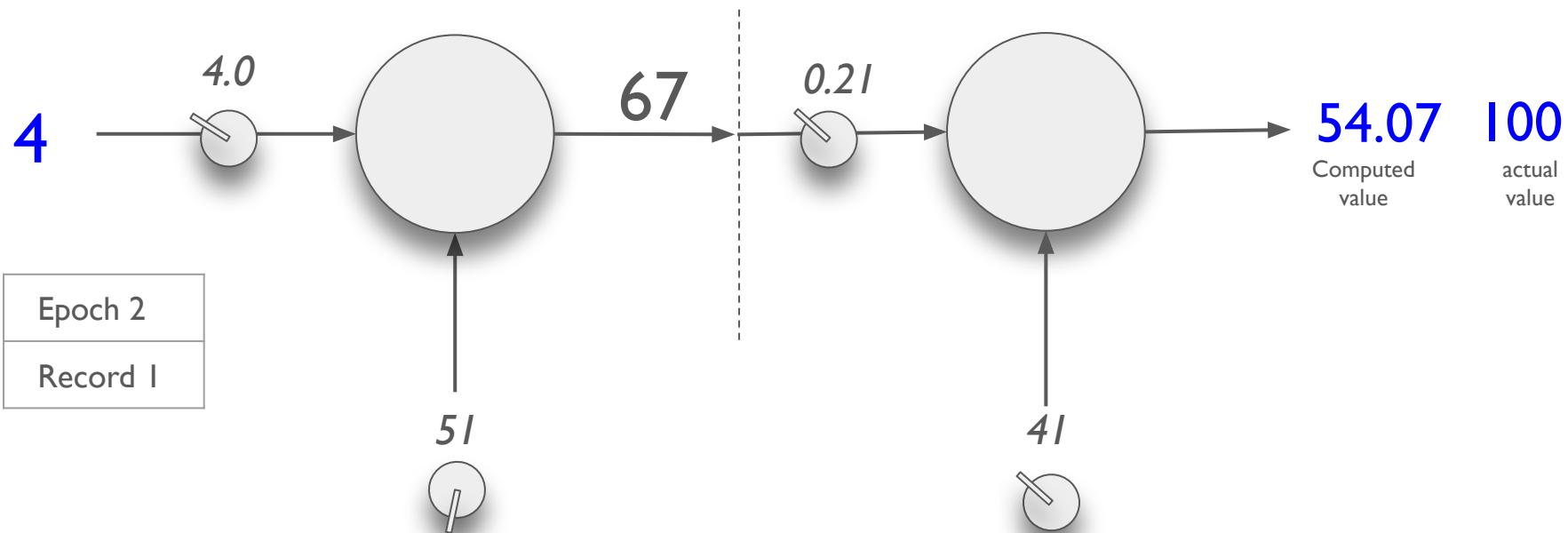
Next Record. Repeat the process,
Tweak the weights...



EMP	Work Ex	Salary
empl	4	100
emp2	5	130
emp3	2	30
emp4	5	140



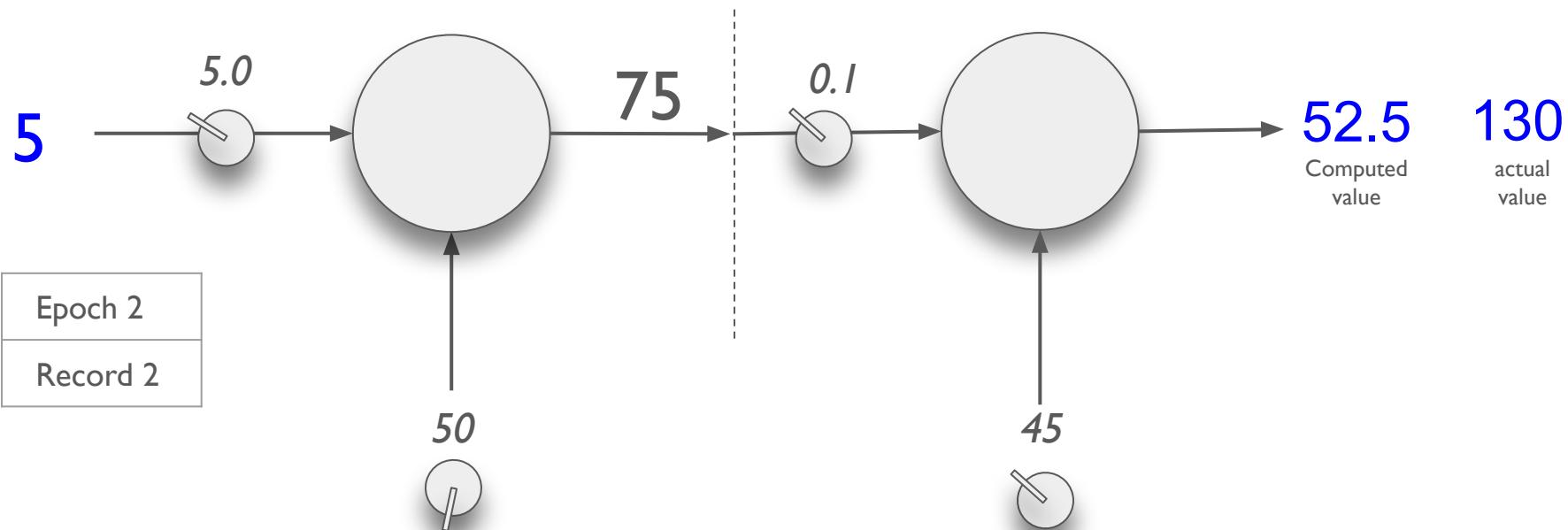
Next Epoch. Begin again from the first record ...



EMP	Work Ex	Salary
emp1	4	100
emp2	5	130
emp3	2	30
emp4	5	140

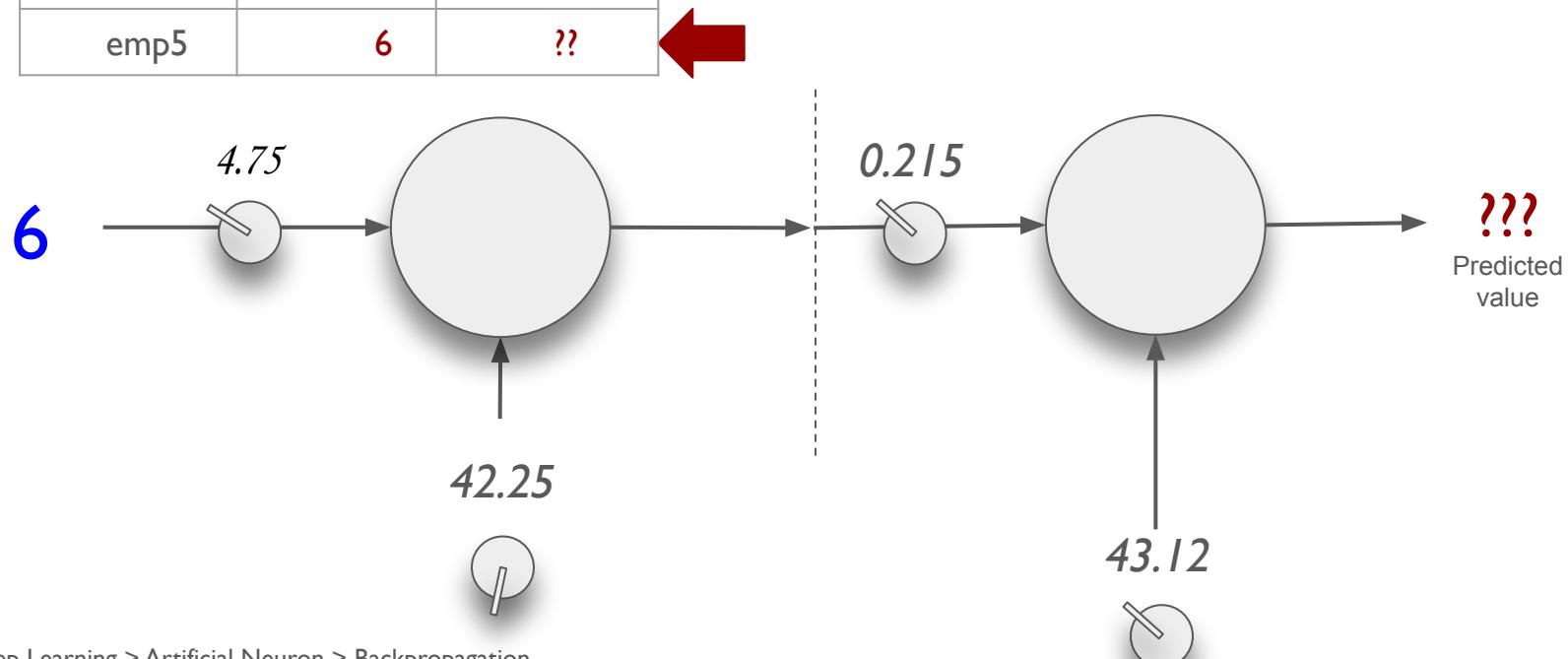


Epoch #2 , Second Record, and so on.



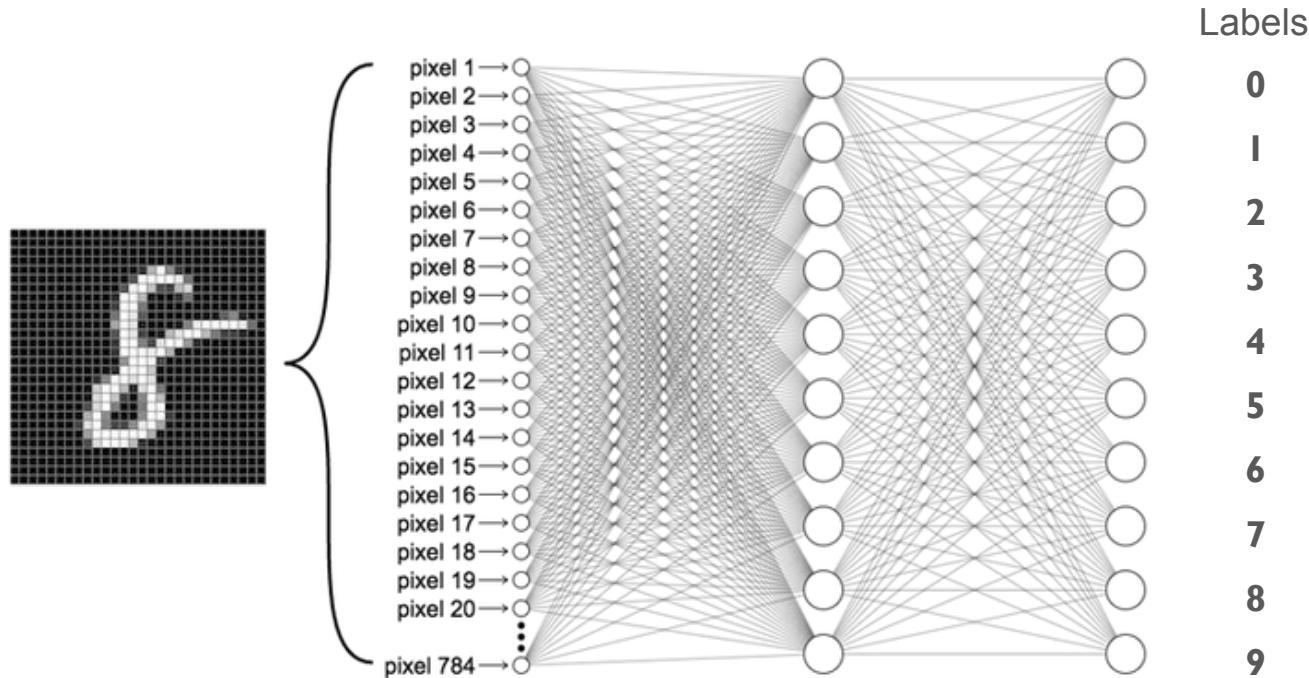
EMP	Work Ex	Salary
empl	4	100
emp2	5	130
emp3	2	30
emp4	5	140
emp5	6	??

Once it has been trained, it is ready to do the predictions



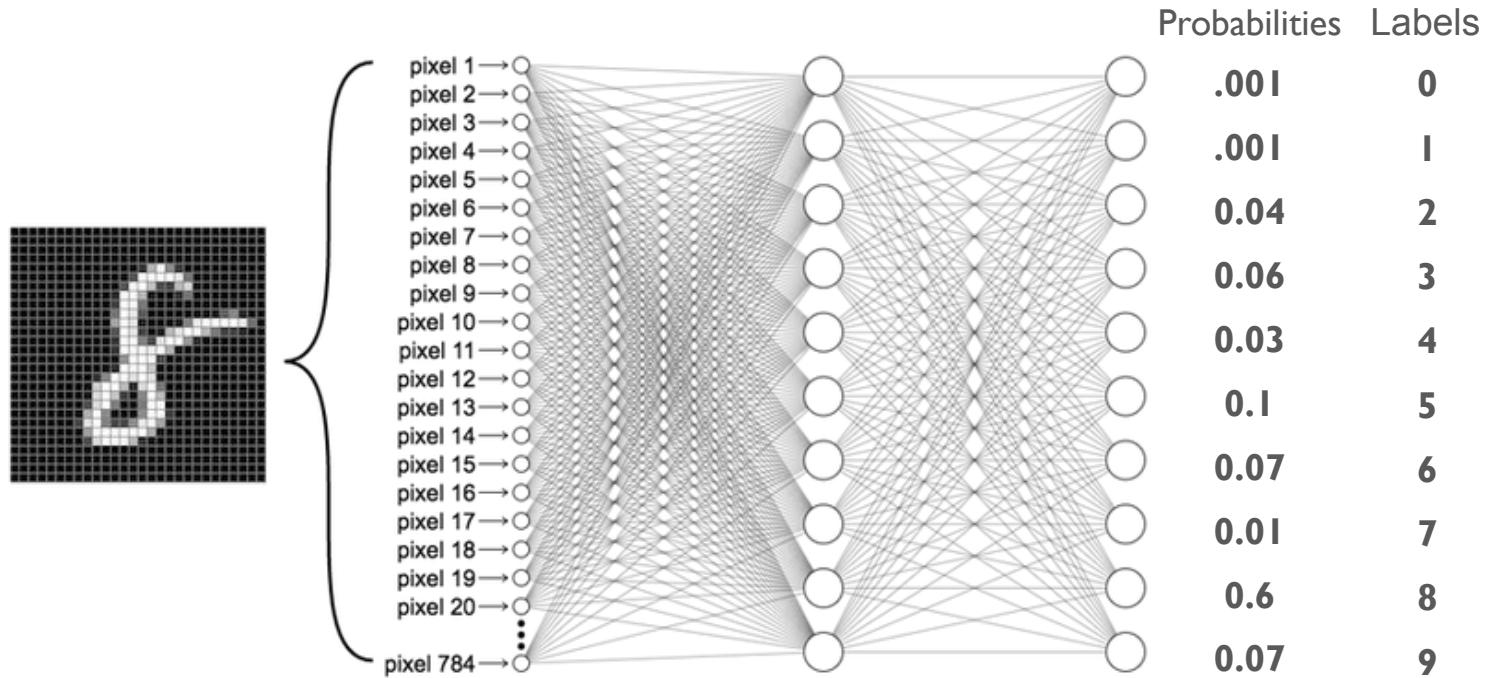
ANN or Fully connected Neural Network

MNIST - classifying a handwritten image into 10 labels.



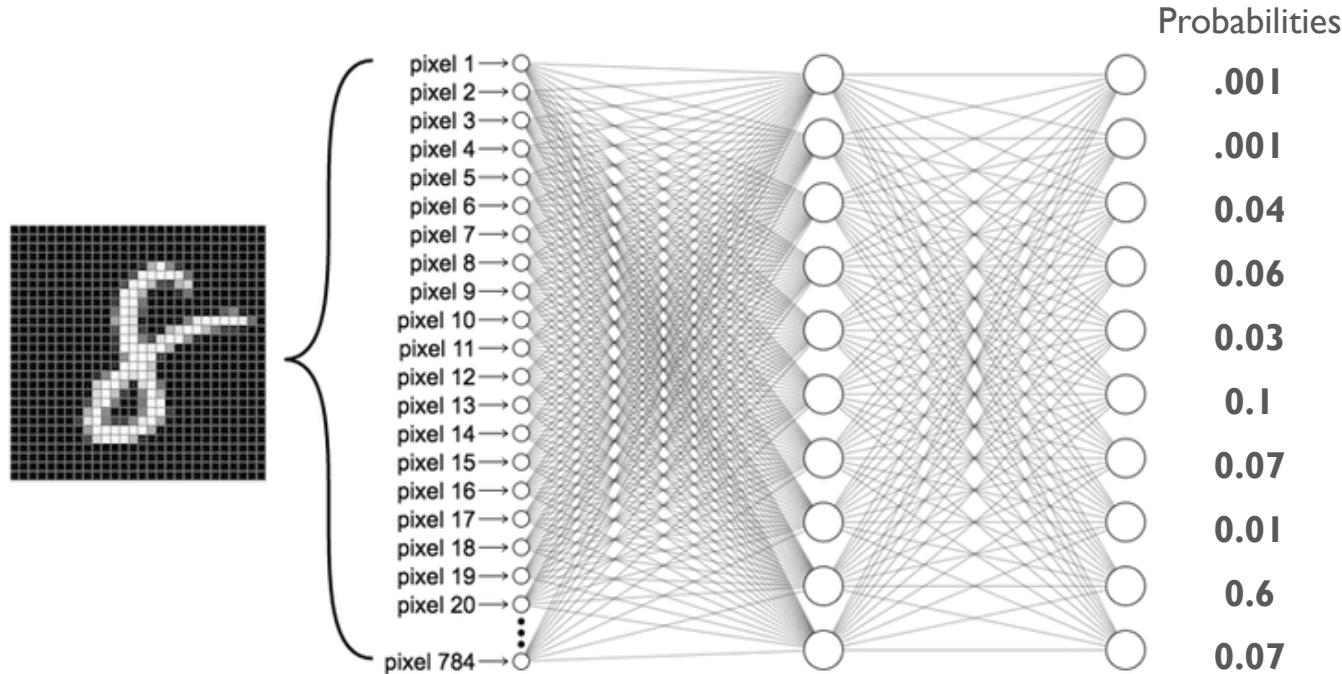
Training Neural Network - MNIST

It would basically compute probabilities of various digits



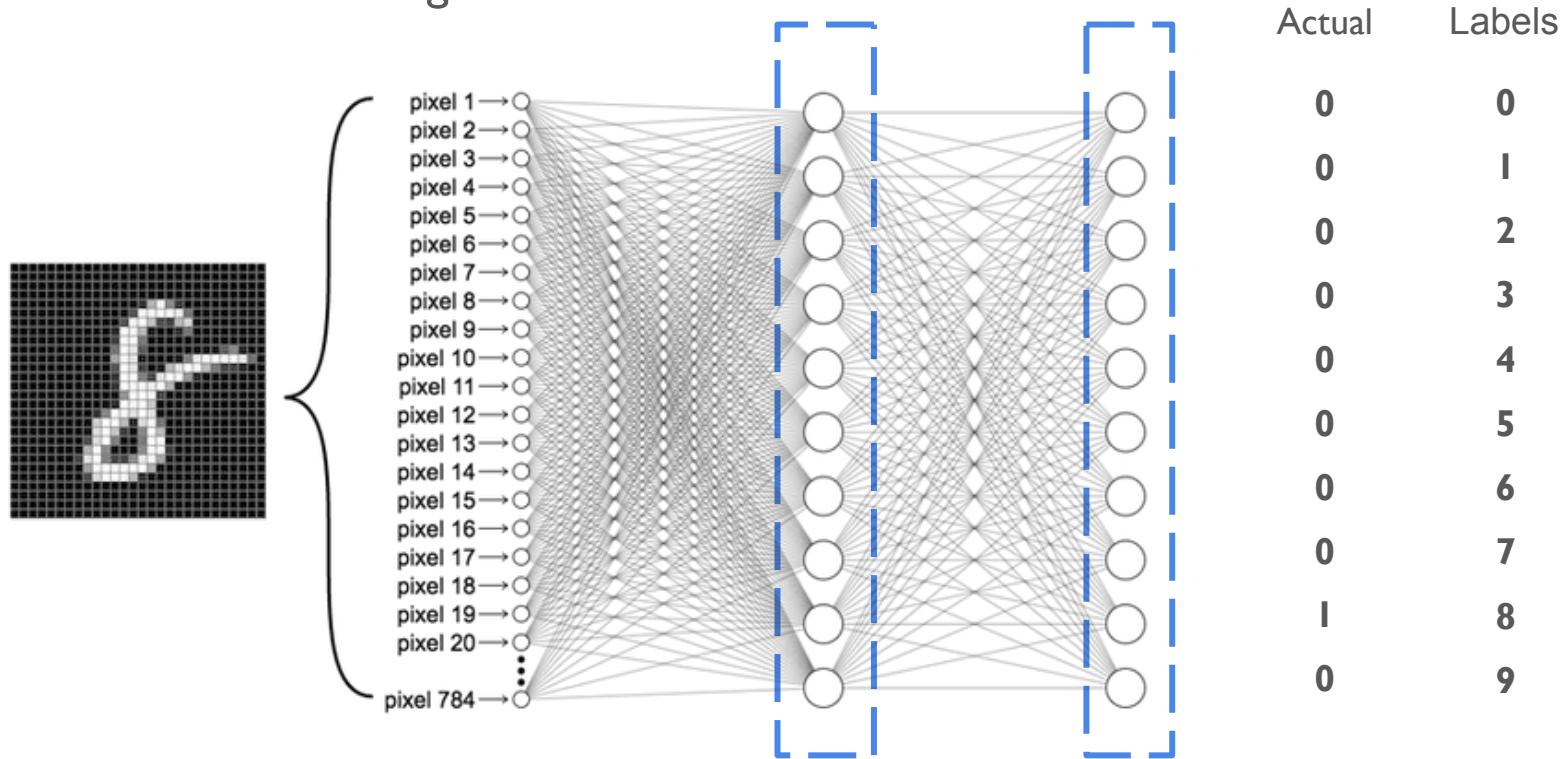
Training Neural Network - MNIST

How to train it?



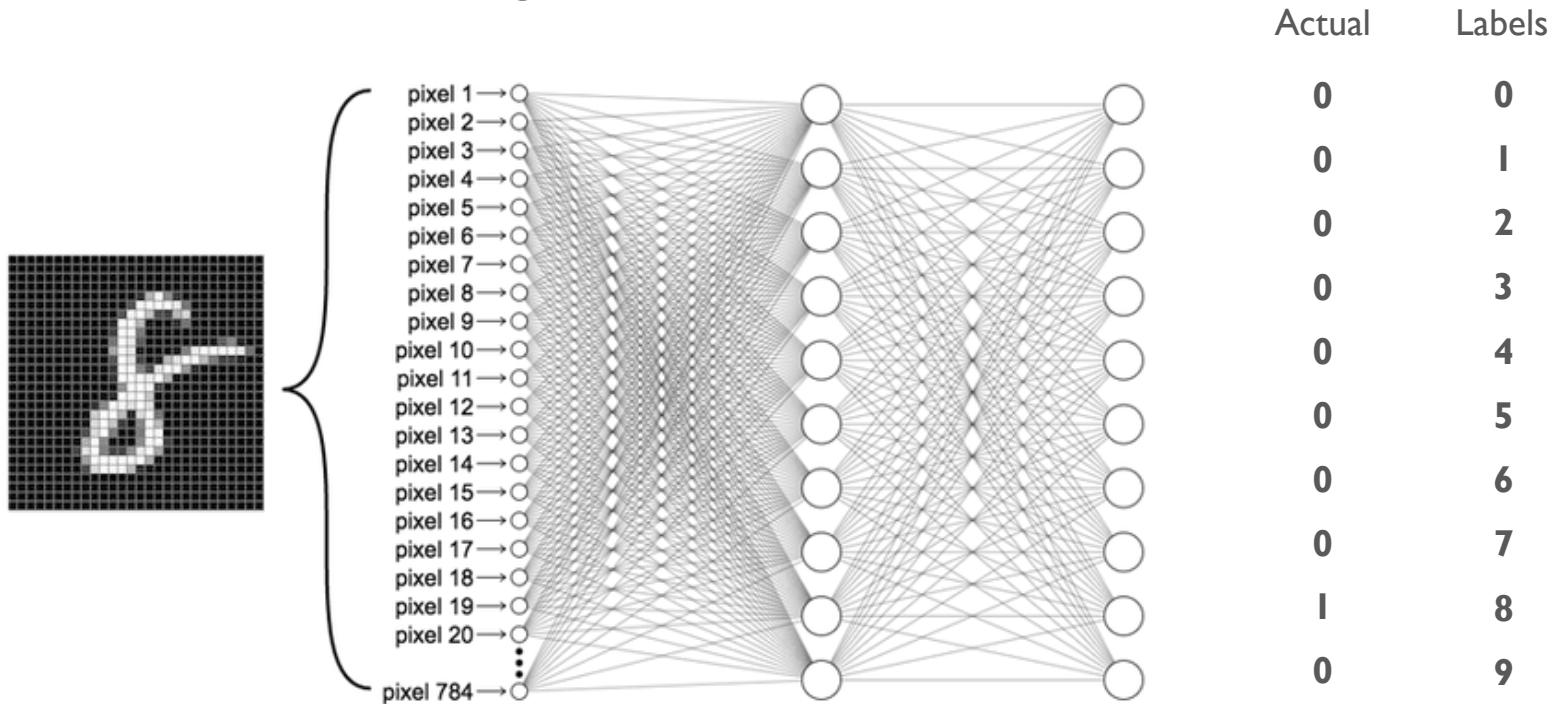
ANN or Fully connected Neural Network

Step 1 - Initialize the weights



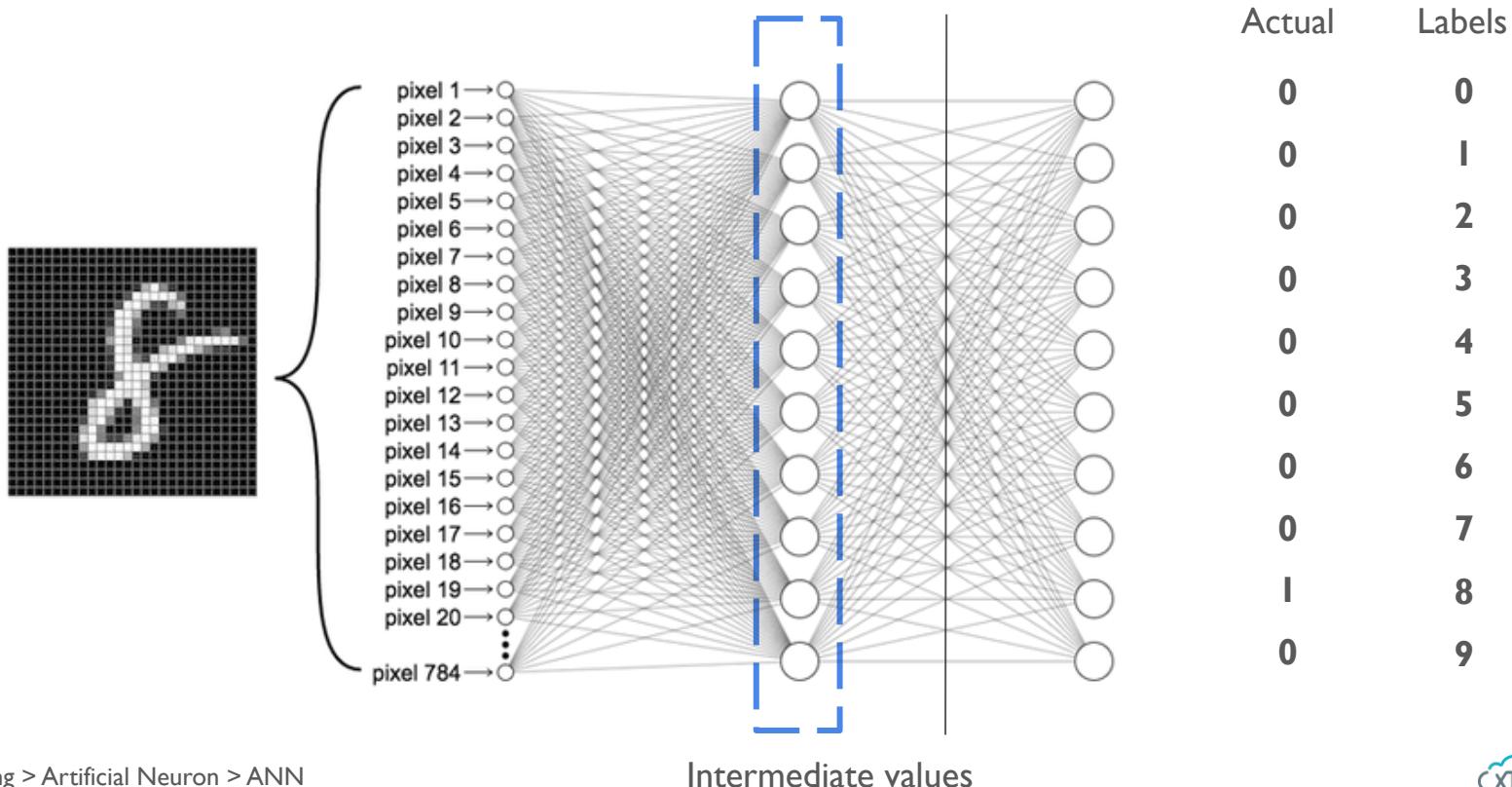
Training Neural Network - MNIST

Take an instance from training data.



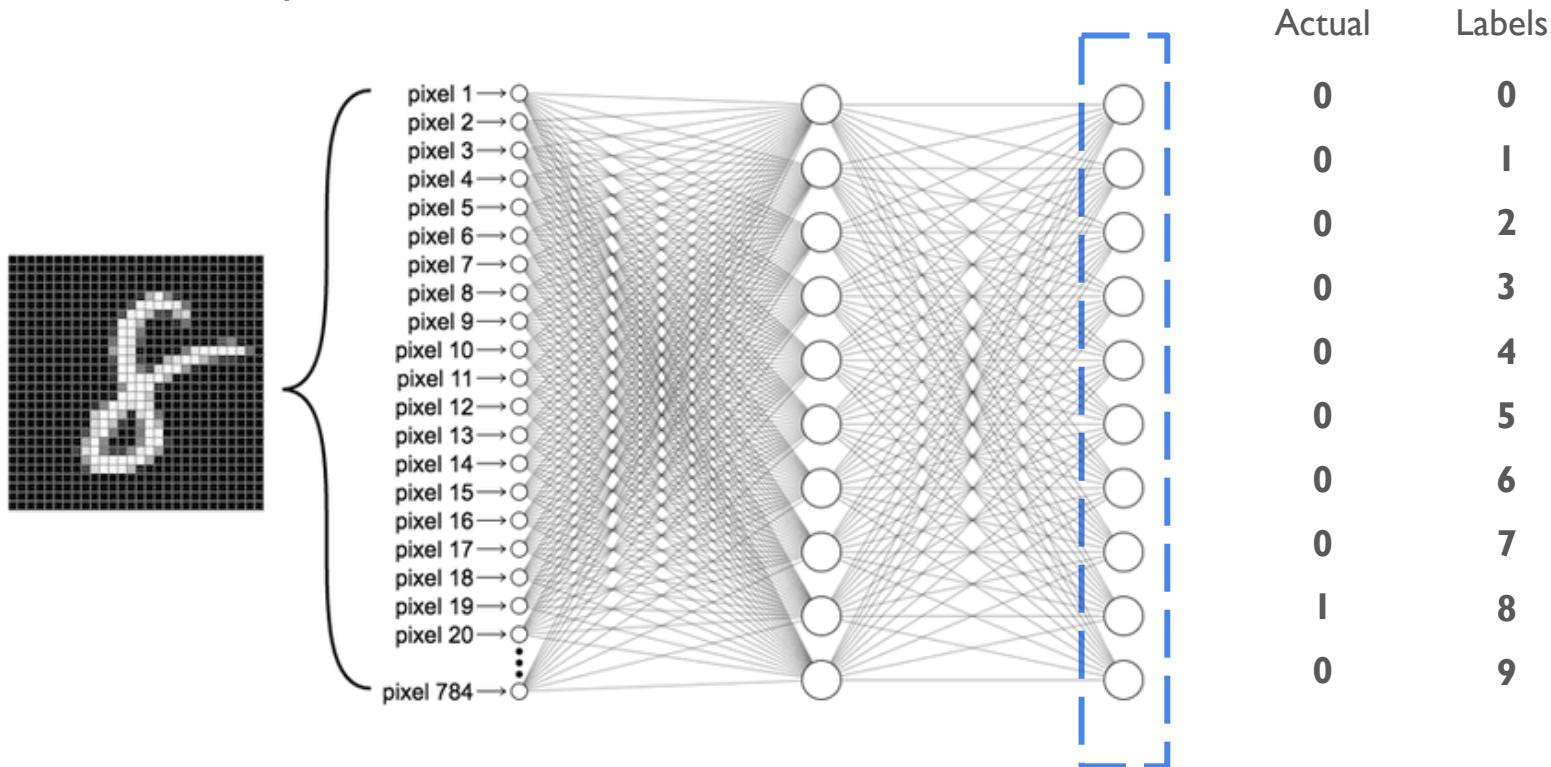
ANN or Fully connected Neural Network

Step 2 - Forward pass



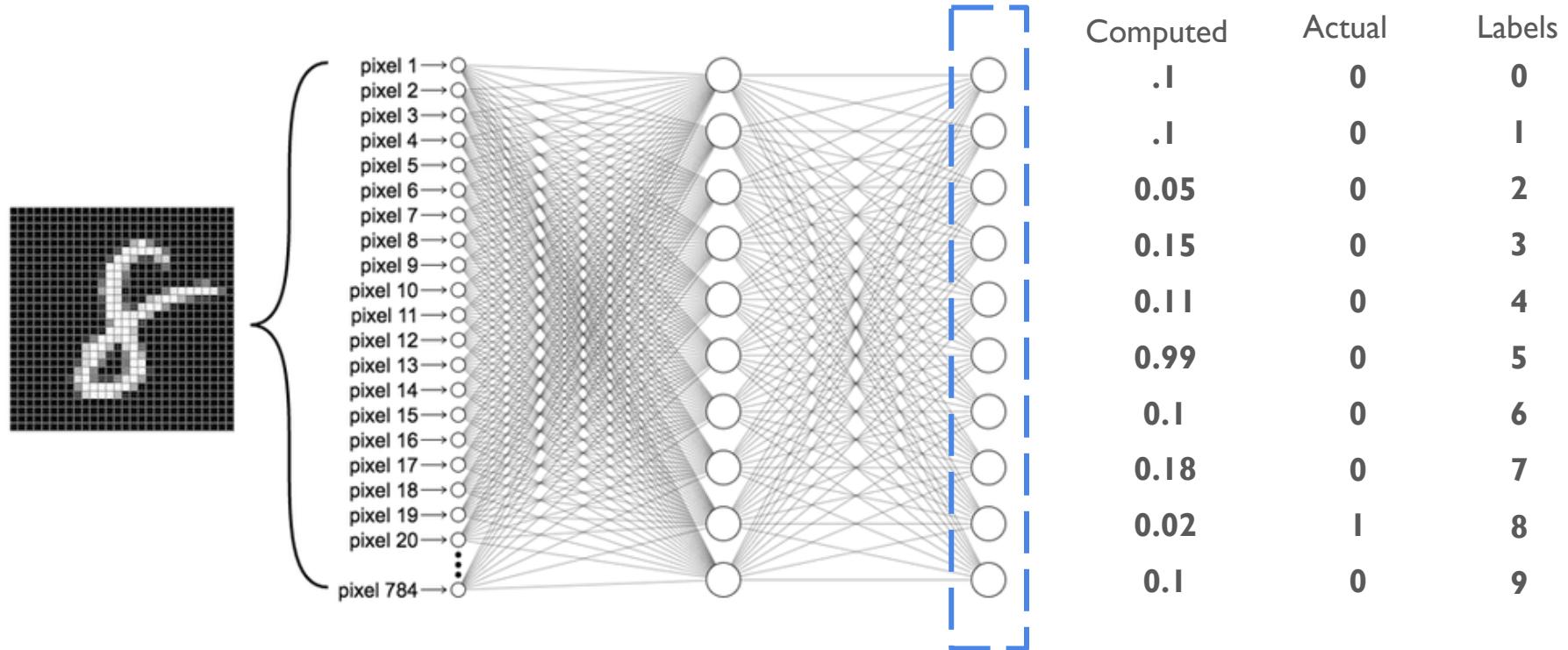
ANN or Fully connected Neural Network

Step 2 - Forward pass



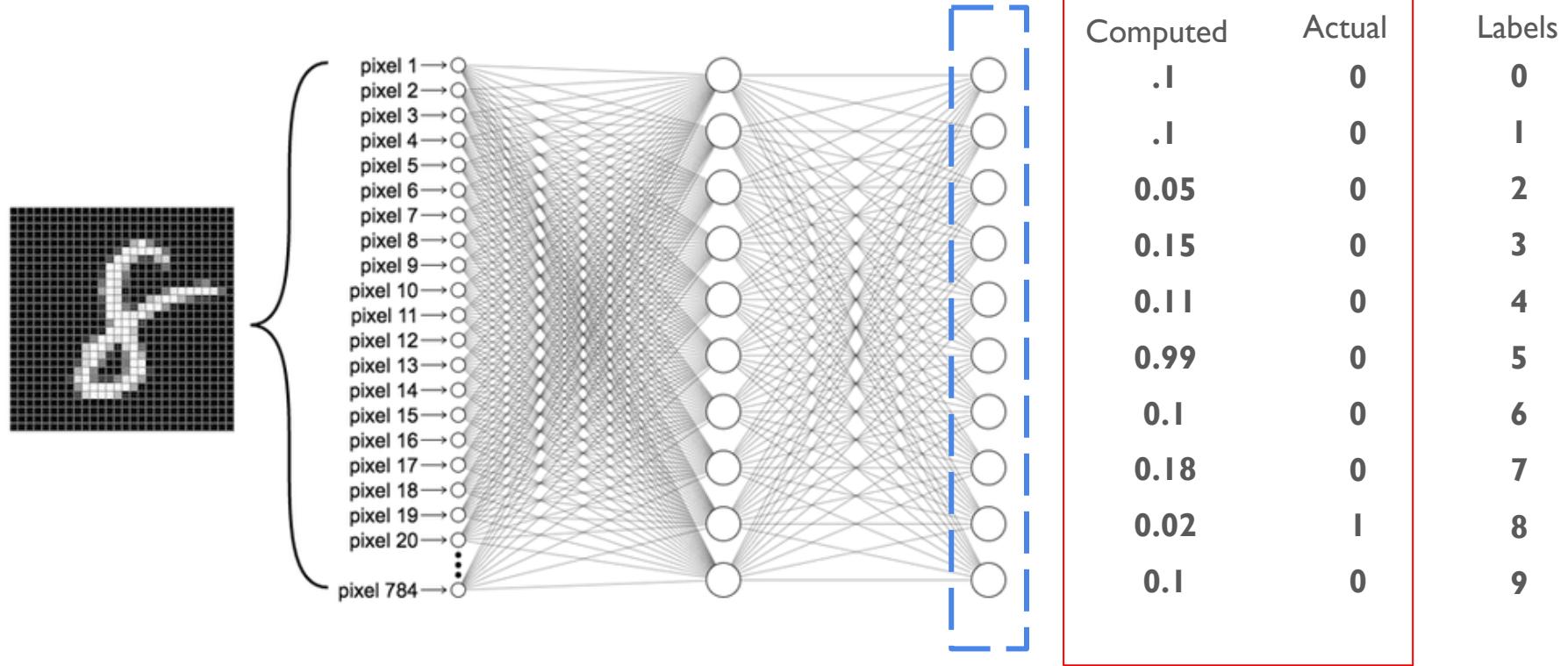
ANN or Fully connected Neural Network

Step 2 - Forward pass



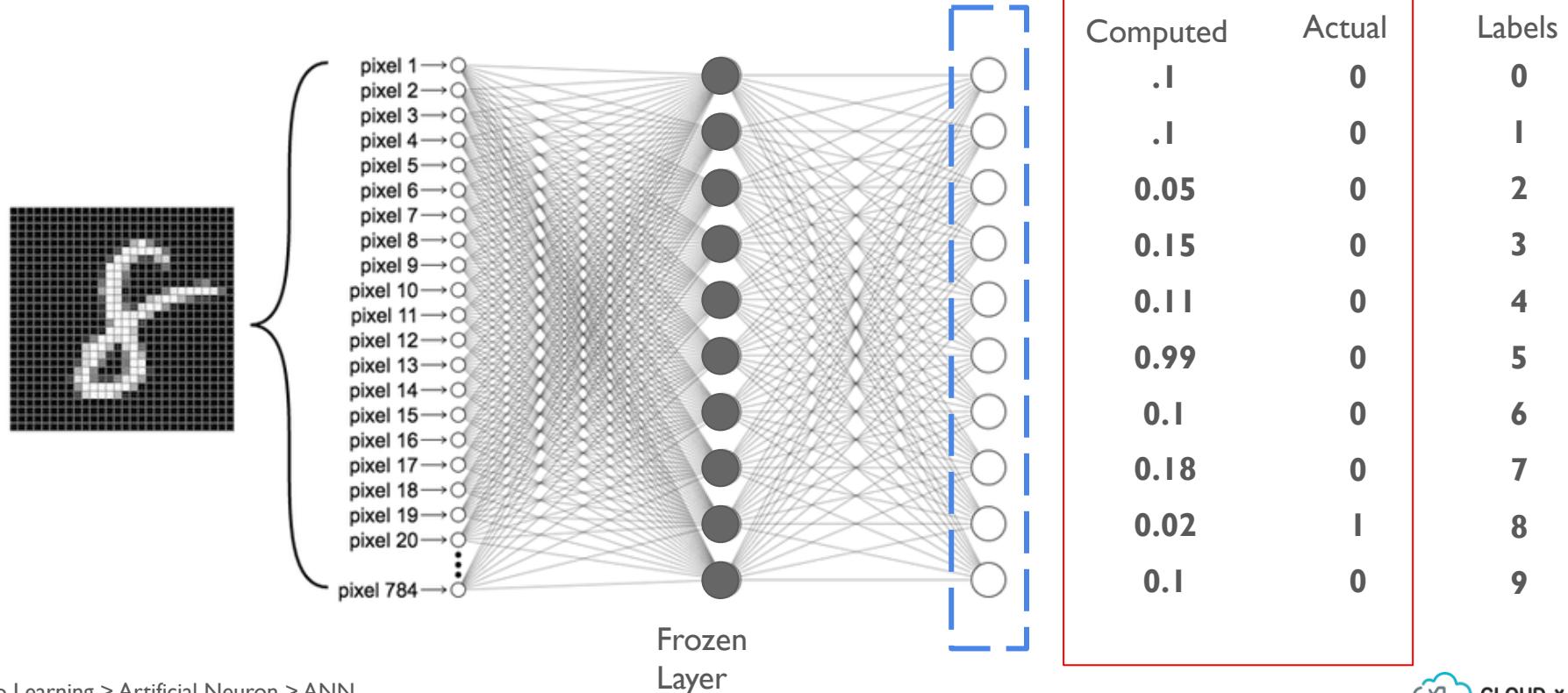
ANN or Fully connected Neural Network

Step 2 - Forward pass



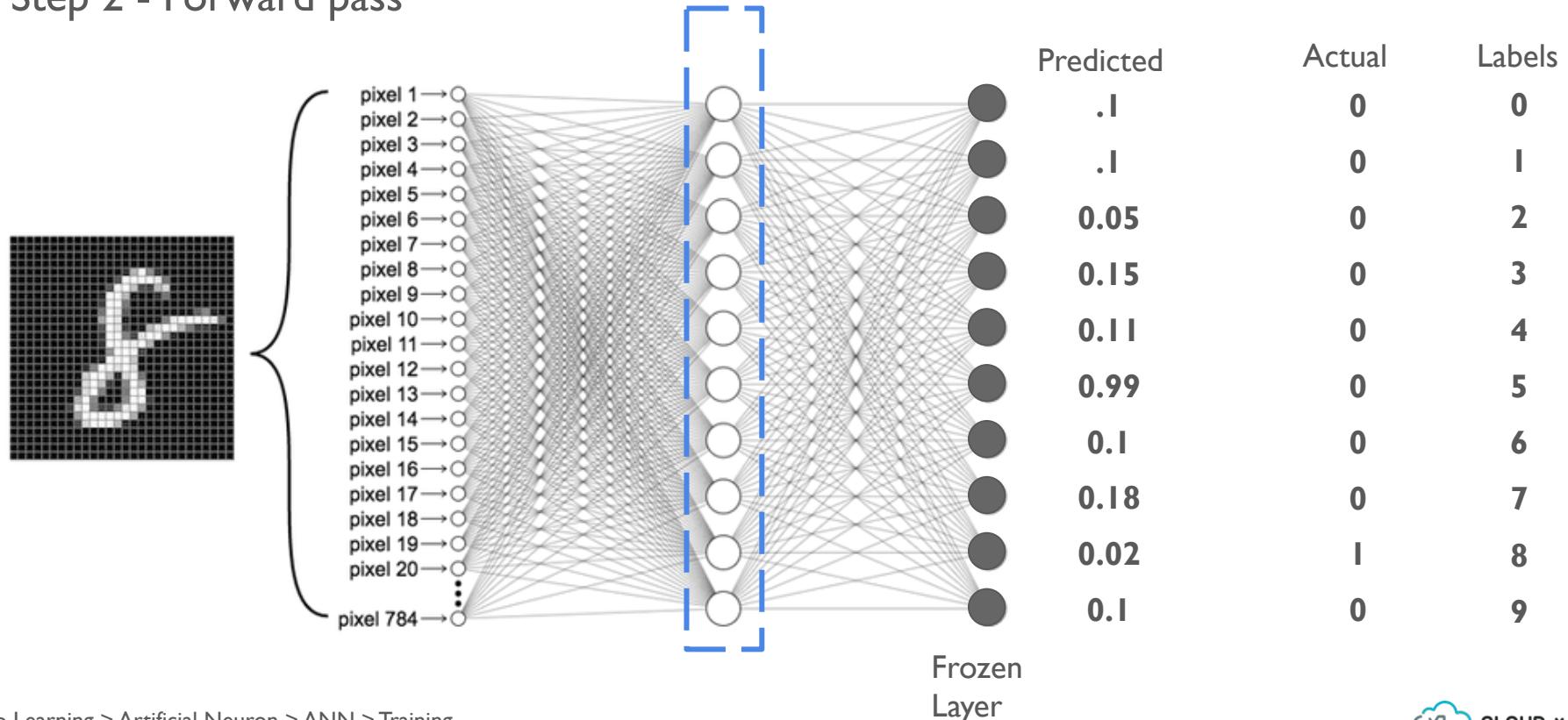
ANN or Fully connected Neural Network

Step 3 - Back Pass - Second Layer



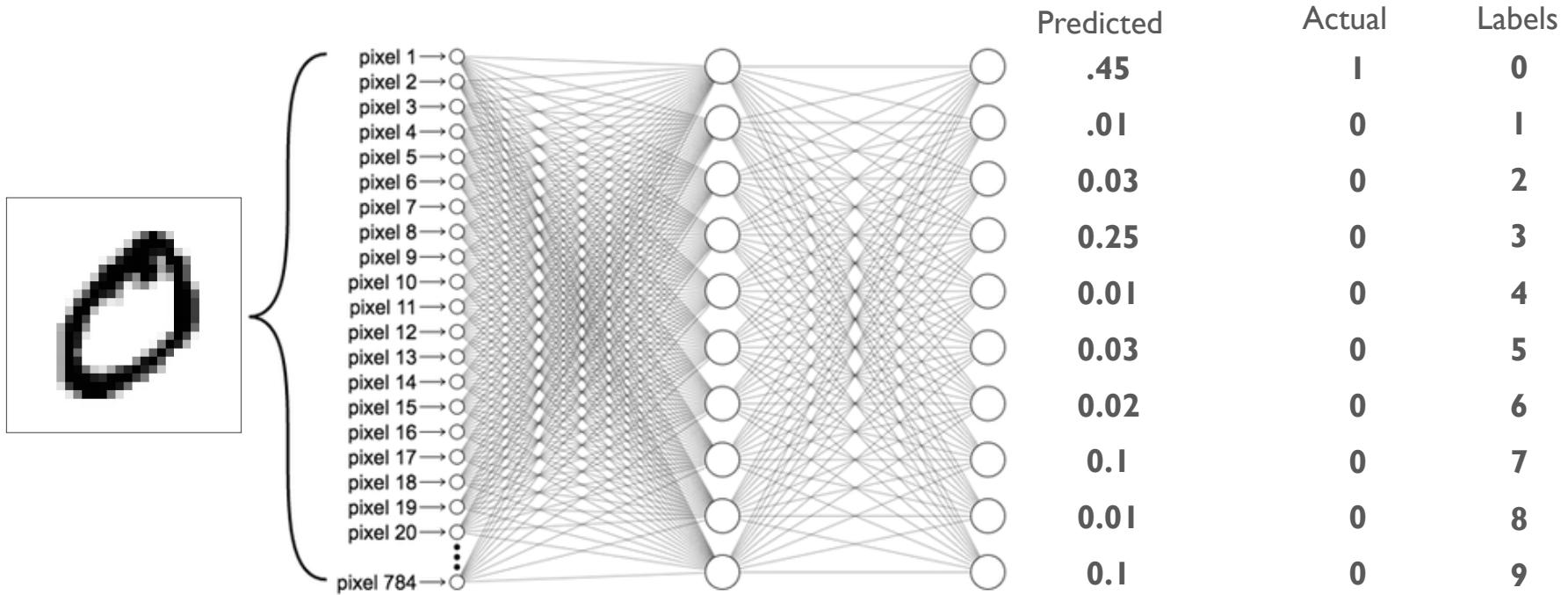
Training Neural Network - MNIST

Step 2 - Forward pass



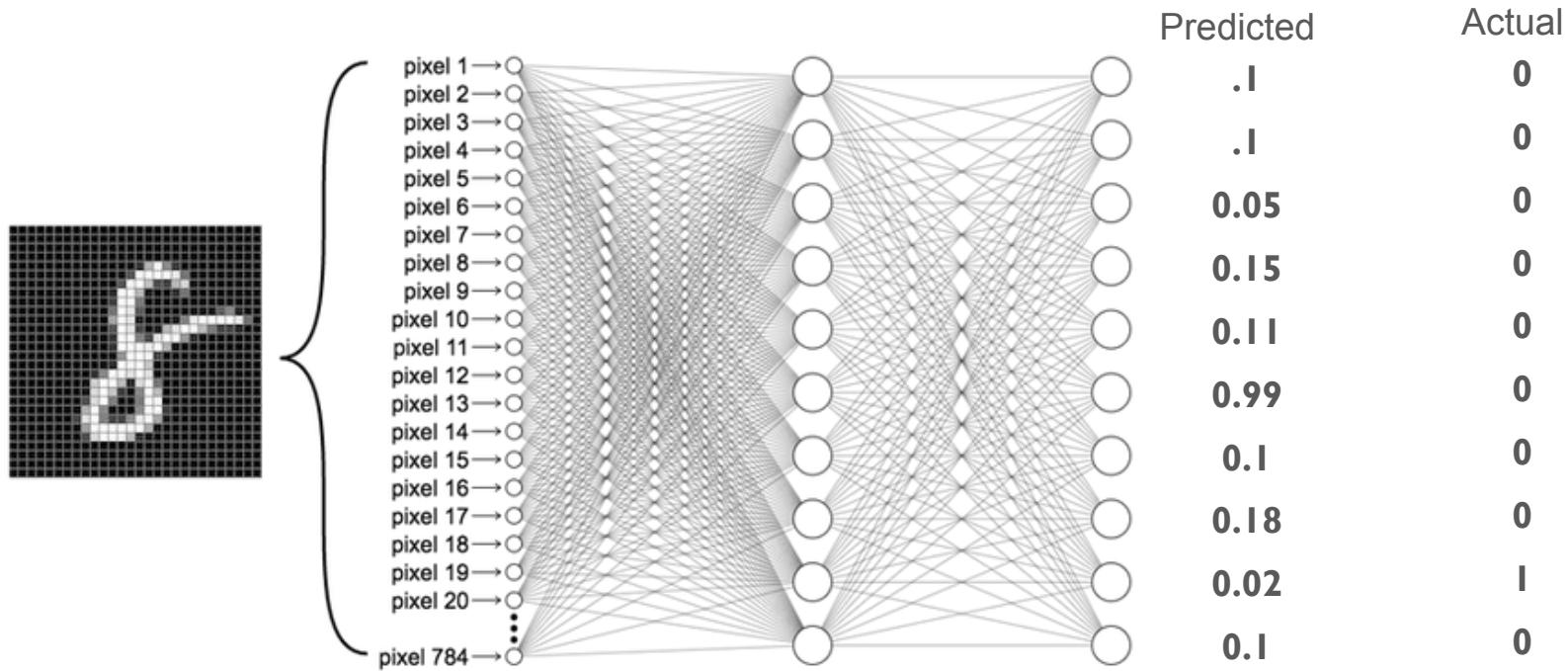
Training Neural Network - MNIST

Pick next instance



Training Neural Network - MNIST

Next Iteration or epoch



Summary

- Linear Regression

- Linear Regression
- Decision Trees

- Linear Regression
- Decision Trees
- SVM

- Linear Regression
- Decision Trees
- SVM
- Neural Networks

Questions?

<https://discuss.cloudxlab.com>

reachus@cloudxlab.com