

The Art of Machine Learning

Learn To Make Your Machine Learn

Why Machine Learning?

- Some tasks are too complex to code directly
 - Medical diagnosis assistance to diagnose a particular disease
 - A lot of parameters to handle
 - e.g. body temperature, pain conditions, blood pressure, eyes condition and a lot more
 - To code them explicitly is Impractical, if not impossible
- A practical approach
 - to handle complex task, make an algorithm which can learn the parameters from sample data

What is Machine Learning?

Machine learning is all about building a mathematical model of a sample data in order to make decisions or predictions.



Just tell me how to do ML

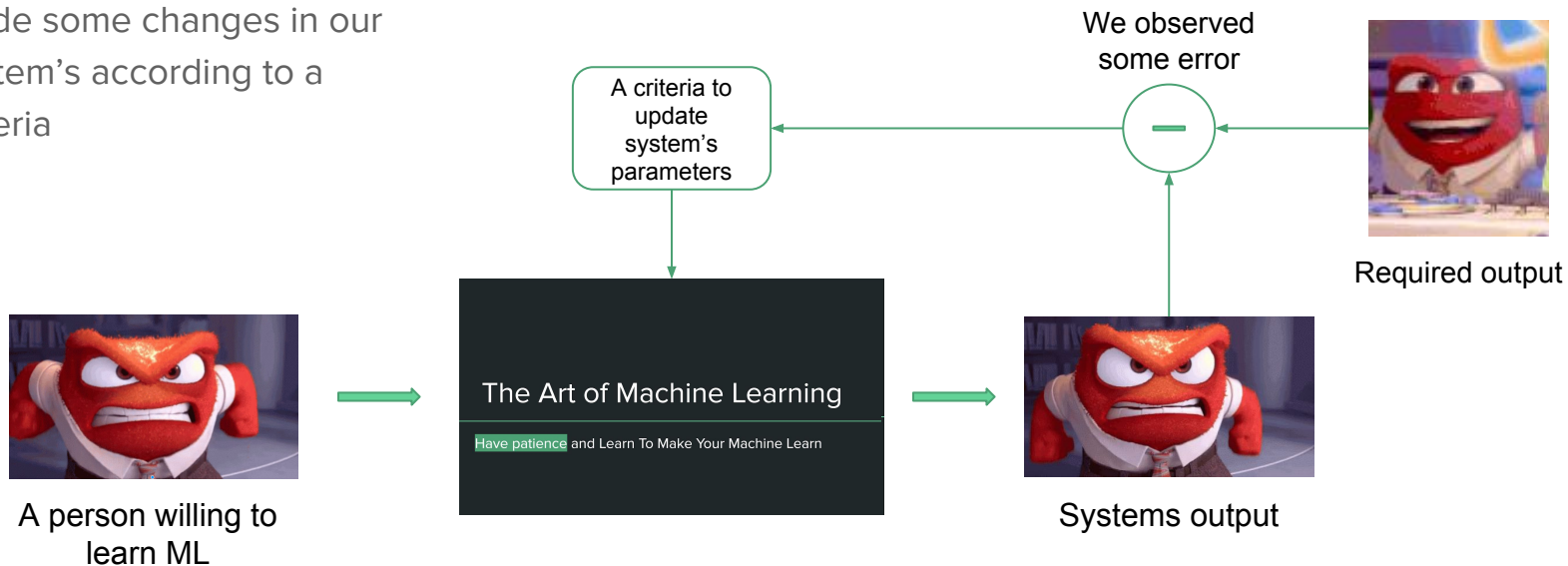
Just to keep the situation under control!

The Art of Machine Learning

Have patience and Learn To Make Your Machine Learn

Wait!

- We defined a system
- We gave it an input
- System gave an output
- We knew the required output
- We found some error by comparison
- Made some changes in our system's according to a criteria



Machine Learning is Easy - Believe me

The most important thing

Patience

- Other important things
 - Don't get scared of complex mathematics
 - It's easy to if you try to find the meaning
 - A lot of implemented work is available (e.g tensorflow, pytorch etc)

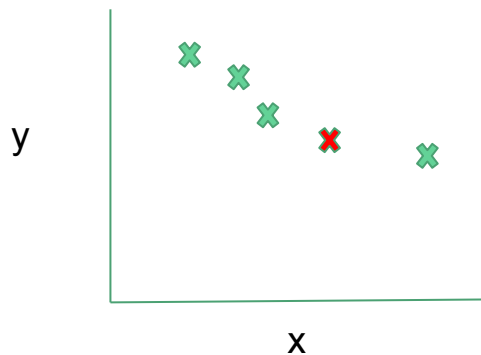
What is the meaning of this equation?

$$ax + b = c$$

Types of Problems in Machine Learning - A Basic understanding

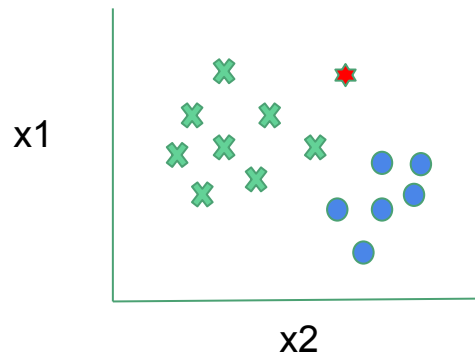
- Prediction

- You have a sample dataset of defined points and you have to predict output of a system for new unseen data point



- Classification

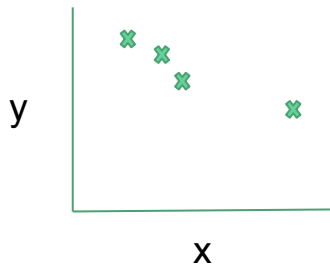
- You have a sample dataset of defined points and their classes, you have to find class of a new data point



Types of Problems in Machine Learning - A Basic understanding

- Prediction

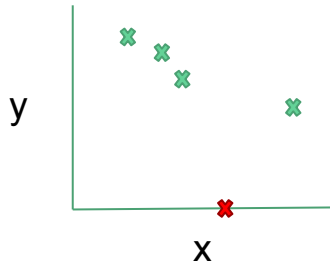
- Let's assume sample data points (x,y)
 - x is input and y is output
 - $(1,10)$ $(2,9)$ $(2.5,8)$ $(8.1,6.8)$



Types of Problems in Machine Learning - A Basic understanding

- Prediction

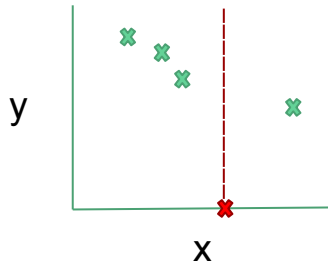
- Let's assume sample data points (x,y)
 - x is input and y is output
 - $(1,10)$ $(2,9)$ $(2.5,8)$ $(8.1,6.8)$
- Now you have a new value for $x = 5.2$
 - $y = ?$



Types of Problems in Machine Learning - A Basic understanding

- Prediction

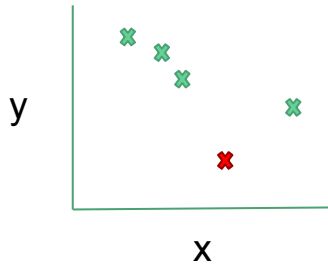
- Let's assume sample data points (x,y)
 - x is input and y is output
 - $(1,10)$ $(2,9)$ $(2.5,8)$ $(8.1,6.8)$
- Now you have a new value for $x = 5.2$
 - $y = ?$



Types of Problems in Machine Learning - A Basic understanding

- Prediction

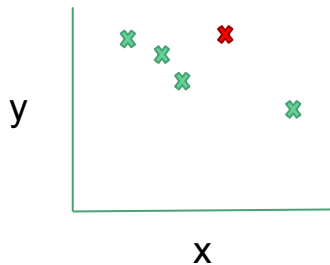
- Let's assume sample data points (x,y)
 - x is input and y is output
 - $(1,10)$ $(2,9)$ $(2.5,8)$ $(8.1,6.8)$
- Now you have a new value for $x = 5.2$
 - $y = ?$



Types of Problems in Machine Learning - A Basic understanding

- Prediction

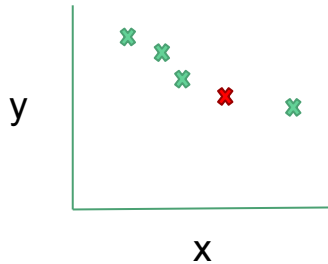
- Let's assume sample data points (x,y)
 - x is input and y is output
 - $(1,10)$ $(2,9)$ $(2.5,8)$ $(8.1,6.8)$
- Now you have a new value for $x = 5.2$
 - $y = ?$



Types of Problems in Machine Learning - A Basic understanding

- Prediction

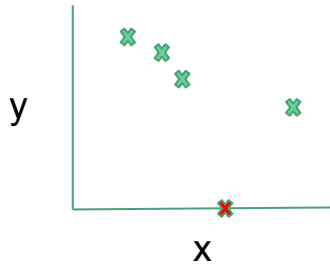
- Let's assume sample data points (x,y)
 - x is input and y is output
 - $(1,10)$ $(2,9)$ $(2.5,8)$ $(8.1,6.8)$
- Now you have a new value for $x = 5.2$
 - $y = ?$



Types of Problems in Machine Learning - A Basic understanding

- Prediction

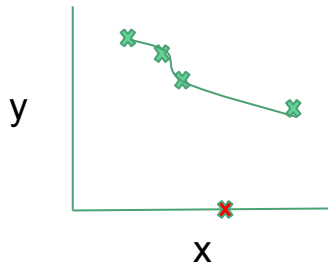
- Let's assume sample data points (x,y)
 - x is input and y is output
 - $(1,10)$ $(2,9)$ $(2.5,8)$ $(8.1,6.8)$
- Now you have a new value for $x = 5.2$
 - $y = ?$



Types of Problems in Machine Learning - A Basic understanding

- Prediction

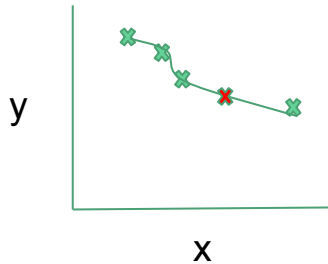
- Let's assume sample data points (x,y)
 - x is input and y is output
 - $(1,10)$ $(2,9)$ $(2.5,8)$ $(8.1,6.8)$
- Now you have a new value for $x = 5.2$
 - $y = ?$



Types of Problems in Machine Learning - A Basic understanding

- Prediction

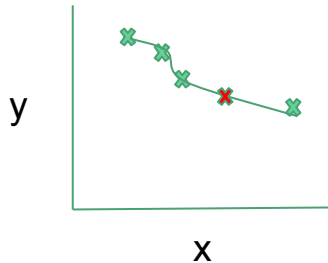
- Let's assume sample data points (x,y)
 - x is input and y is output
 - $(1,10)$ $(2,9)$ $(2.5,8)$ $(8.1,6.8)$
- Now you have a new value for $x = 5.2$
 - $y = ?$



Types of Problems in Machine Learning - A Basic understanding

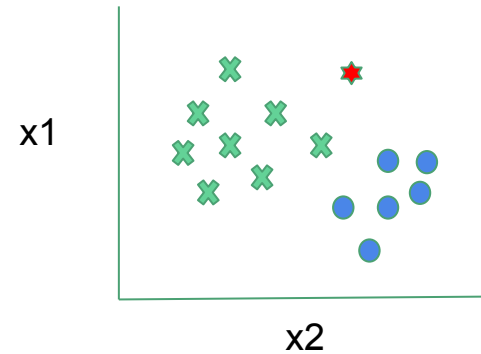
● Prediction

- Let's assume sample data points (x,y)
 - x is input and y is output
 - (1,10) (2,9) (2.5,8) (8.1,6.8)
- Now you have a new value for $x = 5.2$
 - $y = ?$



● Classification

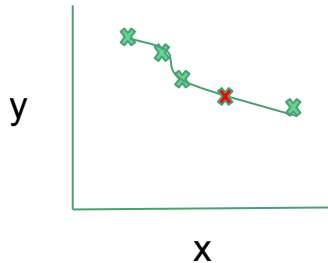
- You have a sample dataset of defined points and their classes, you have to find class of a new data point



Types of Problems in Machine Learning - A Basic understanding

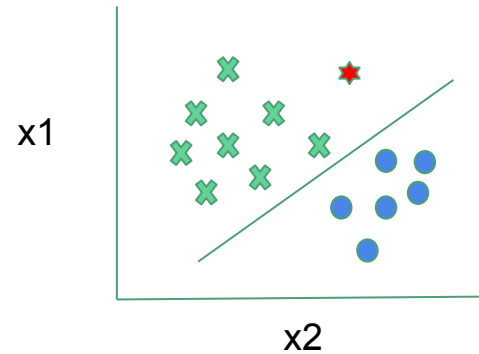
● Prediction

- Let's assume sample data points (x,y)
 - x is input and y is output
 - (1,10) (2,9) (2.5,8) (8.1,6.8)
- Now you have a new value for $x = 5.2$
 - $y = ?$



● Classification

- You have a sample dataset of defined points and their classes, you have to find class of a new data point



Basic steps to build ML system

- Collect and analyze data for ML problem (training data)
- Make a mathematical model of the problem
- Define an error function
- Define a criteria to update system parameters to minimize error

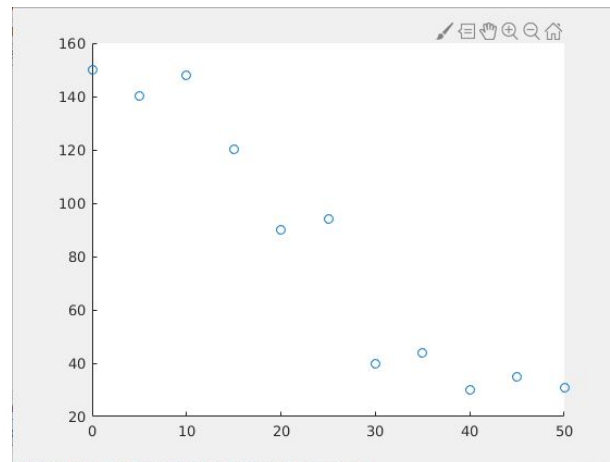
A simple prediction example

- Predict no. of coffee cups sold on random day

no. of cups sold y	avg temp of the day x
150	0
140	5
148	10
120	15
90	20
94	25
40	30
44	35
30	40
35	45
31	50

Step1: Analyze Data

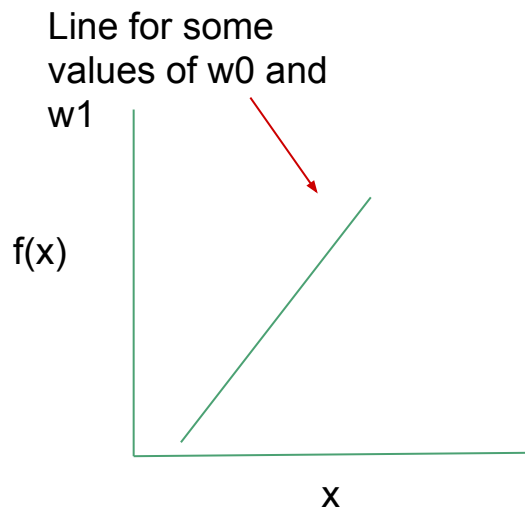
No. of cups
sold



Avg temperature of the day

A simple prediction example

- Predict no. of coffee cups sold on random day



Step2: Make a mathematical model

Let's fit a linear line in the data trend

$$ax + b = c$$

$$(b - c) + (a)x = 0$$

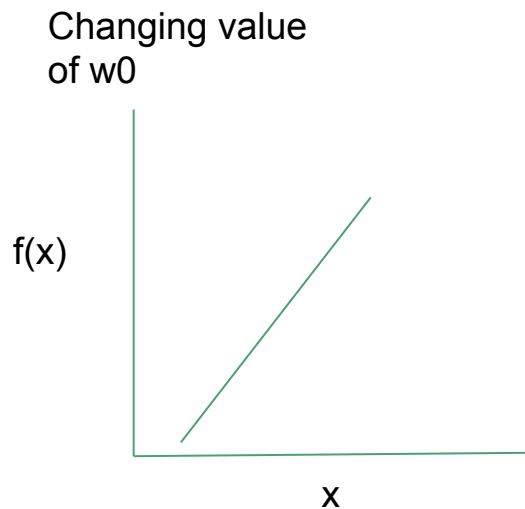
$$(w_0) + (w_1)x = 0$$

These are called systems parameters

$$f(x) = (w_0) + (w_1)x$$

A simple prediction example

- Predict no. of coffee cups sold on random day



Step2: Make a mathematical model

Let's fit a linear line in the data trend

$$ax + b = c$$
$$(b - c) + (a)x = 0$$

$$(w_0) + (w_1)x = 0$$

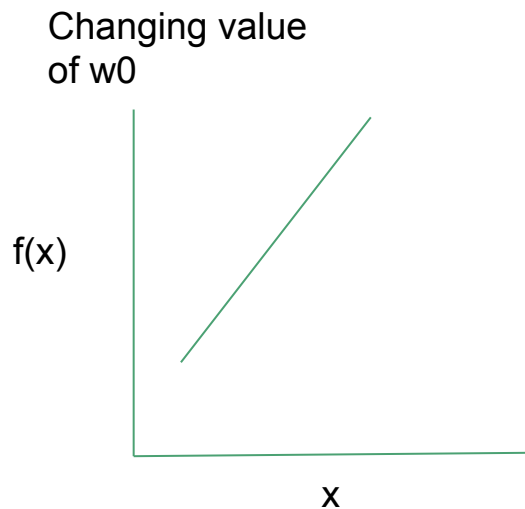


These are called systems
parameters

$$f(x) = (w_0) + (w_1)x$$

A simple prediction example

- Predict no. of coffee cups sold on random day



Step2: Make a mathematical model

Let's fit a linear line in the data trend

$$ax + b = c$$
$$(b - c) + (a)x = 0$$

$$(w_0) + (w_1)x = 0$$

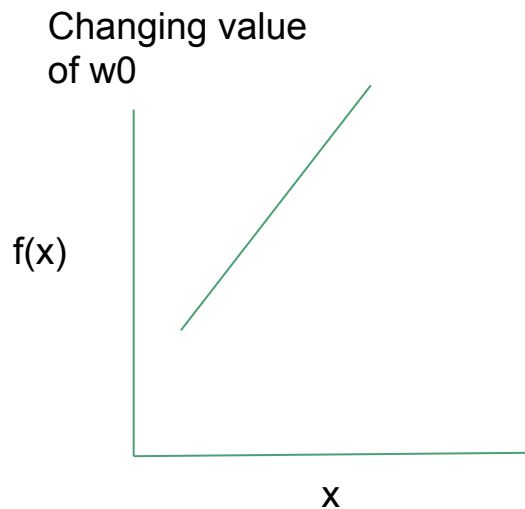


These are called systems
parameters

$$f(x) = (w_0) + (w_1)x$$

A simple prediction example

- Predict no. of coffee cups sold on random day



Step2: Make a mathematical model

Let's fit a linear line in the data trend

$$ax + b = c$$
$$(b - c) + (a)x = 0$$

$$(w_0) + (w_1)x = 0$$

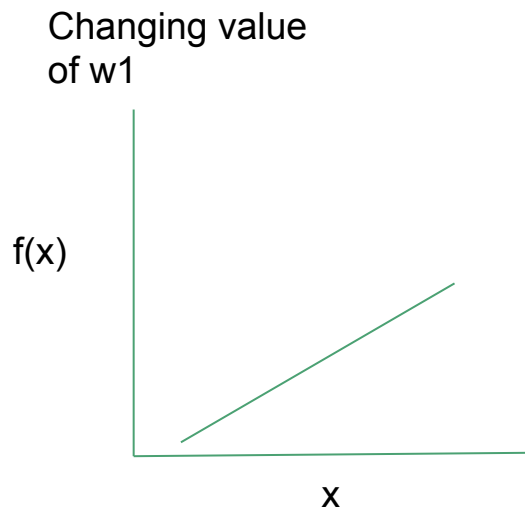


These are called systems
parameters

$$f(x) = (w_0) + (w_1)x$$

A simple prediction example

- Predict no. of coffee cups sold on random day



Step2: Make a mathematical model

Let's fit a linear line in the data trend

$$ax + b = c$$
$$(b - c) + (a)x = 0$$

$$(w_0) + (w_1)x = 0$$

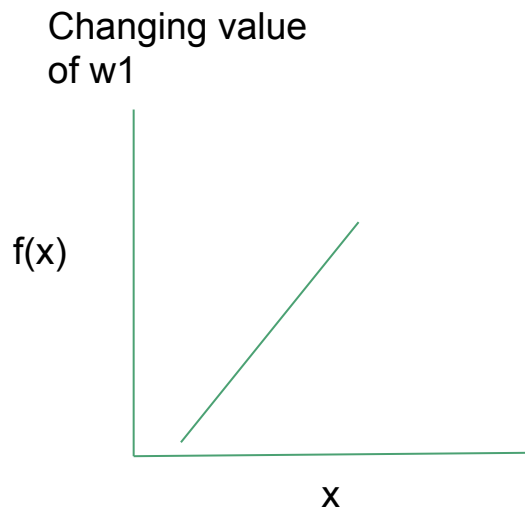


These are called systems
parameters

$$f(x) = (w_0) + (w_1)x$$

A simple prediction example

- Predict no. of coffee cups sold on random day



Step2: Make a mathematical model

Let's fit a linear line in the data trend

$$ax + b = c$$
$$(b - c) + (a)x = 0$$

$$(w_0) + (w_1)x = 0$$

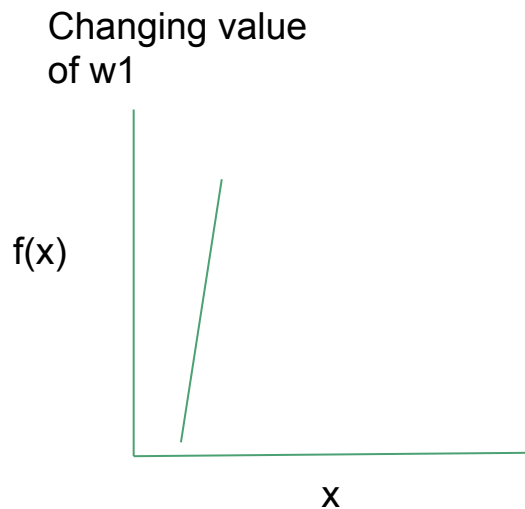


These are called systems
parameters

$$f(x) = (w_0) + (w_1)x$$

A simple prediction example

- Predict no. of coffee cups sold on random day



Step2: Make a mathematical model

Let's fit a linear line in the data trend

$$ax + b = c$$
$$(b - c) + (a)x = 0$$

$$(w_0) + (w_1)x = 0$$



These are called systems
parameters

$$f(x) = (w_0) + (w_1)x$$

A simple prediction example

- Predict no. of coffee cups sold on random day

Step3: find an error function

Let's take

$$h = f(x) = (w_0) + (w_1)x$$

Error function

$$E = \sum (|y - h|) \quad \text{For each point in given data}$$

What we want?

$$\operatorname{argmin}_w \sum (|y - h|) = ?$$

A simple prediction example

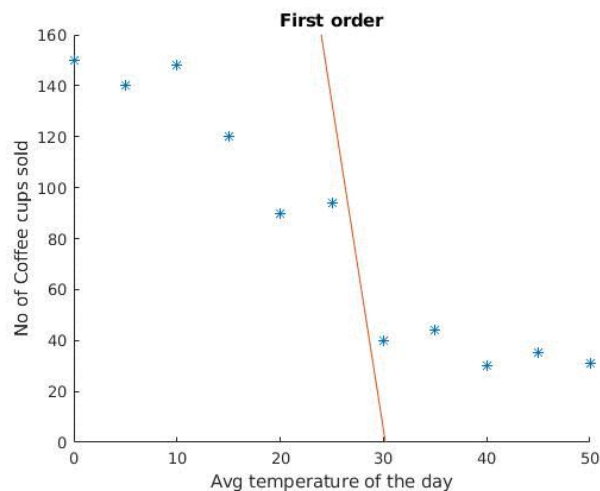
- Predict no. of coffee cups sold on random day

Step4: find a criteria to update w

What will happen to the value of error when we change w_0 and w_1 ?

A simple prediction example

- Predict no. of coffee cups sold on random day



Step4: find a criteria to update w

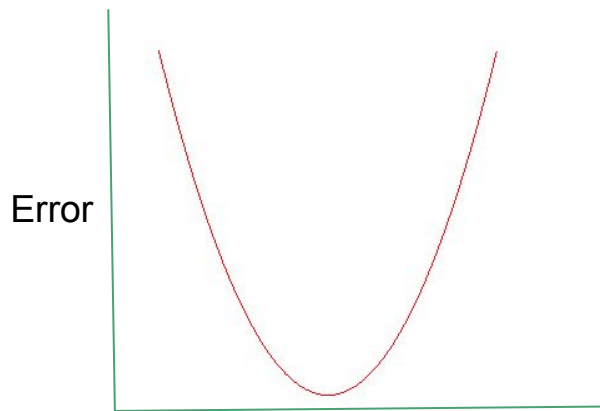
What will happen to the value of error when we change w_0 and w_1 ?

A simple prediction example

- Predict no. of coffee cups sold on random day

Step4: find a criteria to update w

What will happen to the value of error when we change w_0 and w_1 ?

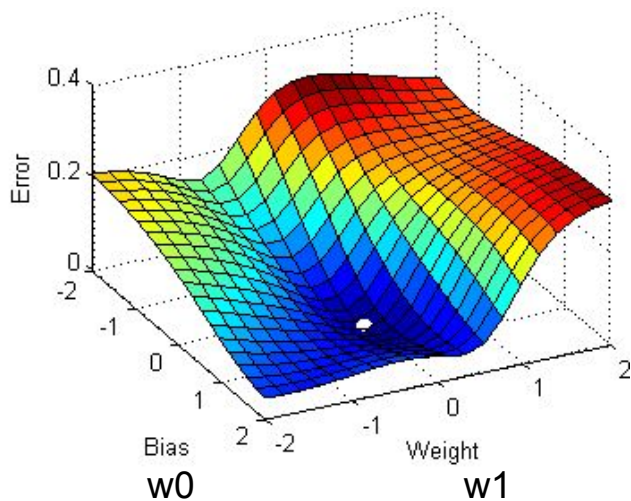


A simple prediction example

- Predict no. of coffee cups sold on random day

Step4: find a criteria to update w

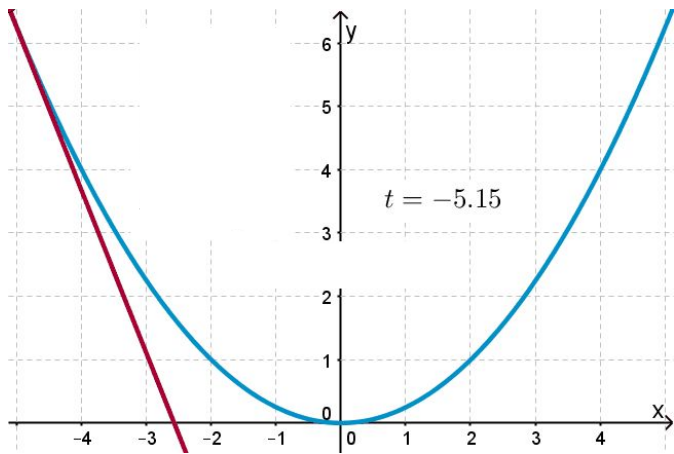
What will happen to the value of error when we change w_0 and w_1 ?



A simple prediction example

- Predict no. of coffee cups sold on random day

Step4: find a criteria to update w



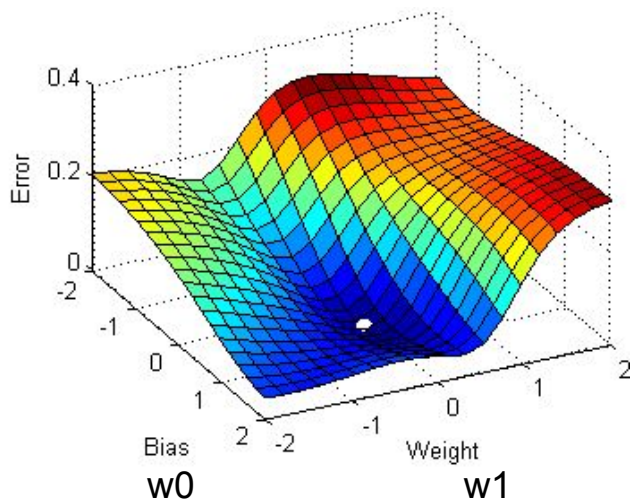
What is derivative?

A simple prediction example

- Predict no. of coffee cups sold on random day

Step4: find a criteria to update w

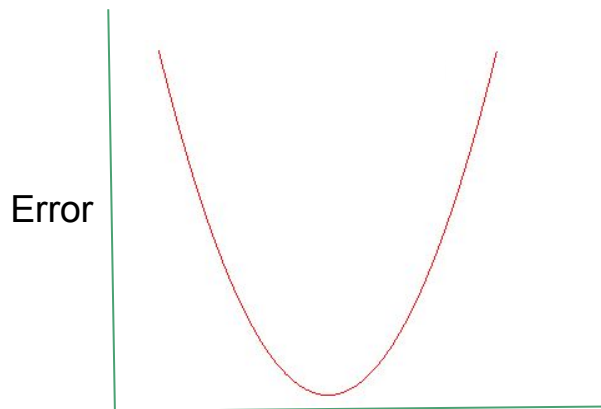
What will happen if we update the weight in negative direction on derivative of this error function?



A simple prediction example

- Predict no. of coffee cups sold on random day

Step4: find a criteria to update w



What will happen if we update the weight in negative direction on derivative of this error function?

updating weights

$$w_0 = w_0 - \alpha \frac{d(|y - h|)}{dw_0}$$

$$w_1 = w_1 - \alpha \frac{d(|y - h|)}{dw_1}$$

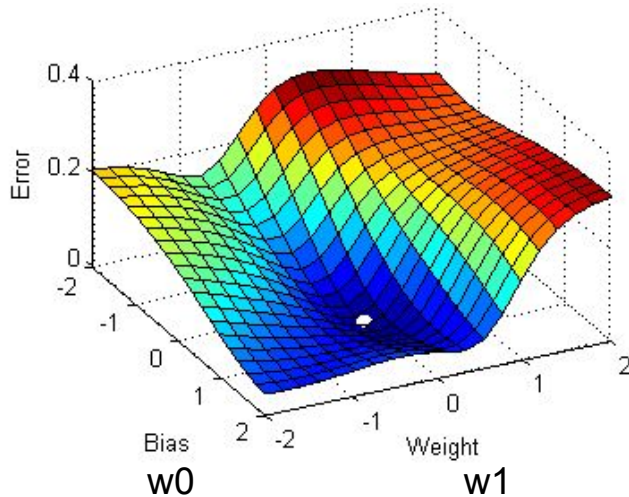
learning rate

A simple prediction example

- Predict no. of coffee cups sold on random day

Step4: find a criteria to update w

What will happen if we update the weight in negative direction on derivative of this error function?



updating weights

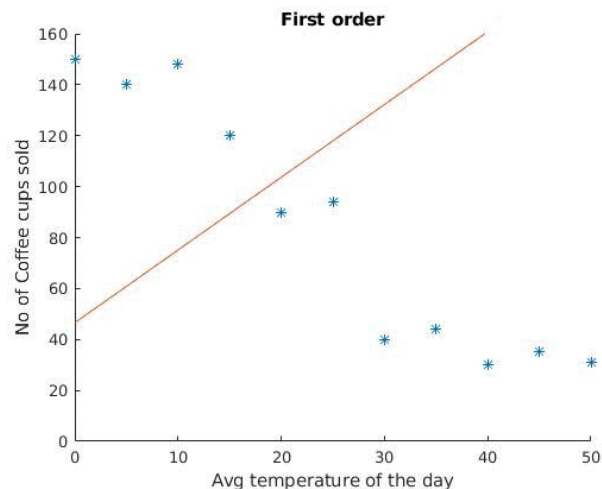
$$w0 = w0 - \alpha \frac{d(|y - h|)}{dw0}$$

$$w1 = w1 - \alpha \frac{d(|y - h|)}{dw1}$$

learning rate

A simple prediction example

- Predict no. of coffee cups sold on random day



Step4: find a criteria to update w

What will happen if we update the weight in negative direction on derivative of this error function?

updating weights

$$w_0 = w_0 + \alpha \frac{d(|y - h|)}{dw_0}$$

$$w_1 = w_1 + \alpha \frac{d(|y - h|)}{dw_1}$$

← dw_1

learning rate

Learning/Training steps

1- randomly initialize w_0 and w_1

2- compute error between h (hypothesis) and y (ground truth)

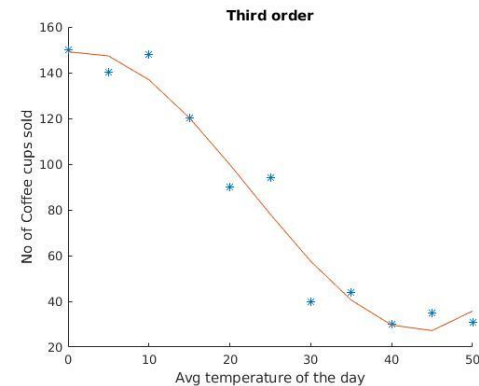
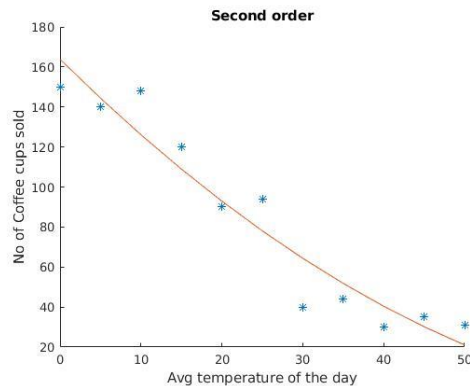
3- update weights

Repeat steps 2 and 3 until error becomes very small

- Higher order polynomial fitting

$$W_0 + W_1x + W_2x^2 = 0$$

$$W_0 + W_1x + W_2x^2 + W_3x^3 = 0$$



- Higher dimensional problems

- e.g. no. of coffee cups sold depends on,
 - X_1 = temperature of the day
 - X_2 = no. of people passing nearby

$$w_0 + w_{11}x_1 + w_{21}x_2 + w_{12}x_1^2 + w_{22}x_2^2 = 0$$

Thank you!