CSCI 1470/2470
Spring 2022

Intro to Machine Learning

Deep Learning

Ritambhara Singh

January 26, 2024

Friday

ChatGPT prompt “minimalist landscape painting of a deep underwater scene with a blue tang fish in the bottom right corner”

What do you hope to learn/be able to do by the end of this course?

~220 Responses! ☺

Understand what deep learning is and what can we do with it

To be able to build my own projects in ML/DL

Get to know the details of different deep learning models and be able to implement them

Read the methods section of papers in my field and understand it

Good for career

I hope to be able to use deep learning to solve complex tasks by the end of the semester.

be able to explain deep learning concepts to someone without a computer science background

Academic research

Tensorflow! How to apply deep learning to many tasks

Make AI tools for animation, help figure out how artists can get compensated for having their work sampled by AI mod

Learn more about the various fields in which DL models overtook human abilities (and how)

I would like to learn this type of data processing for work in medical research

What do you hope to learn/be able to do by the end of this course?

~220 Responses! ☺

Major Themes

(1)Apply deep learning to real-world problems in various types of domains (NLP, Image, Biology) and/or current research

(2)Enhance understanding of concepts and mathematical background

(3)Develop (or improve) implementation skills

(4)Think critically about applying deep learning models and ethical considerations

What do you hope to learn/be able to do by the end of this course?

have a life

200k ml job pls

I want to make friends

Easy A

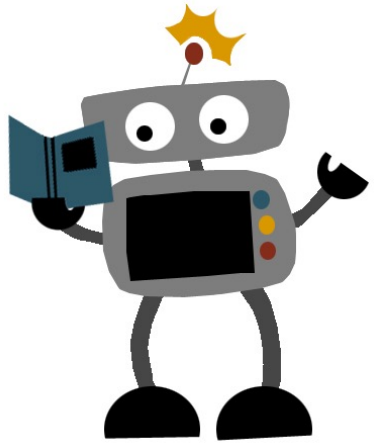
put something on my resume

Deep learning

Earth told me to learn deeply

Learn how to cook

Recap: What is Machine Learning?



Input: X

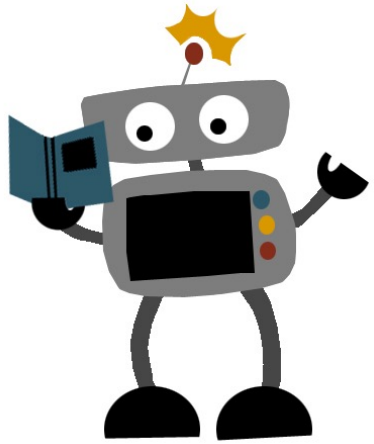
Output: Y

”Cooking?”

Function: f

f(X)→Y

Recap: What is Machine Learning?



Supervised Learning

Input: X

Learned function: f

Output: Y

"Cooking?"

$f(X) \rightarrow Y$

Today’s goal - Learn about some basic concepts of machine learning

(1) How do we represent input/output?

(2) Learning the function f

(3) Training a machine learning model

(4) Learning good models

How do we represent input/output?

Machines work with numbers!

How can we represent input image as numbers?

Input: X

"Model"

Function: f

f(X)→Y

Output: Y

"Cooking?"

How can we represent output labels as numbers?

How do we represent input/output?

Classification

Binary classification

Input: X

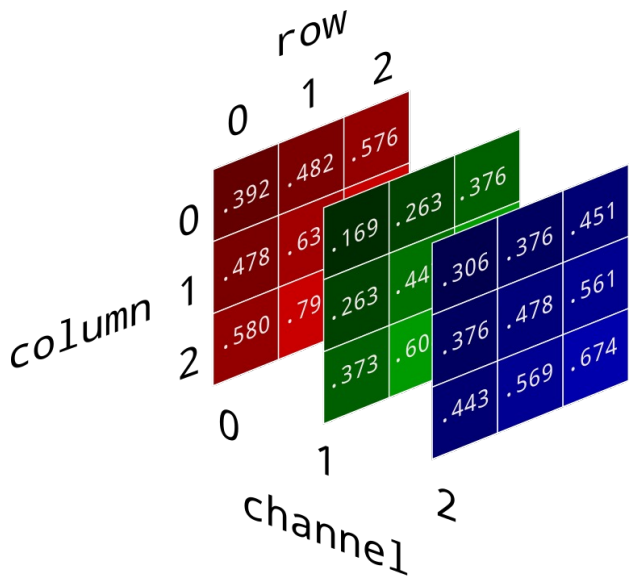
Output: Y

Complicated input!

Let's use a simpler example!

"Cooking?"

1 or 0



$X \in \mathbb{R}^{H \times W \times 3}$

$Y \in \{0,1\}$

(Categorical output)

(Real number: a value of a continuous quantity)

Image courtesy: https://www.kdnuggets.com/2019/12/convert-rgb-image-grayscale.html

But first some notations…

$\mathbb{X}$: $A\ set\ of\ input\ data$

$\mathbb{Y}$: $Associated\ set\ of\ target\ values\ (outputs)\ for\ supervised\ learning$

$x^{(k)}$: kth example (input) from a dataset

$y^{(k)}$: $Target\ (output) associated\ with\ x^{(k)} for\ supervised\ learning$

$\mathbb{R}$: A set of real numbers

Simpler example: How do we represent input/output?

Input: $\mathbb{X}$

"Temperature"

Regression

Target: $\mathbb{Y}$

"Profit made on selling lemonade"

$x^{(1)}$ 100.1 °F

$x^{(2)}$ 80.0 °F

$x^{(3)}$ 30.3 °F

$\mathbb{X} \in \mathbb{R}$

Function: f

f(X)→Y

$y^{(1)}$ $200.0

$y^{(2)}$ $180.5

$y^{(3)}$ $115.1

$\mathbb{Y} \in \mathbb{R}$

(Numerical output)

Do you see a trend here?

What is different about the output here?

Learning function f

Input: $\mathbb{X}$

"Temperature"

$x^{(1)} = 100.1$

$x^{(2)} = 80.0$

$x^{(3)} = 30.3$

$\mathbb{X} \in \mathbb{R}$

Regression

Function: f

f(X)→Y

Target: $\mathbb{Y}$

"Profit made on selling lemonade"

$y^{(1)} = 200.0$

$y^{(2)} = 180.5$

$y^{(3)} = 115.1$

$\mathbb{Y} \in \mathbb{R}$

(Numerical output)

(Image only for explaining concept, not drawn accurately)

Learning function f

Have you seen this equation before?

Regression

Input: $\mathbb{X}$

"Temperature"

$\mathbb{X} \in \mathbb{R}$

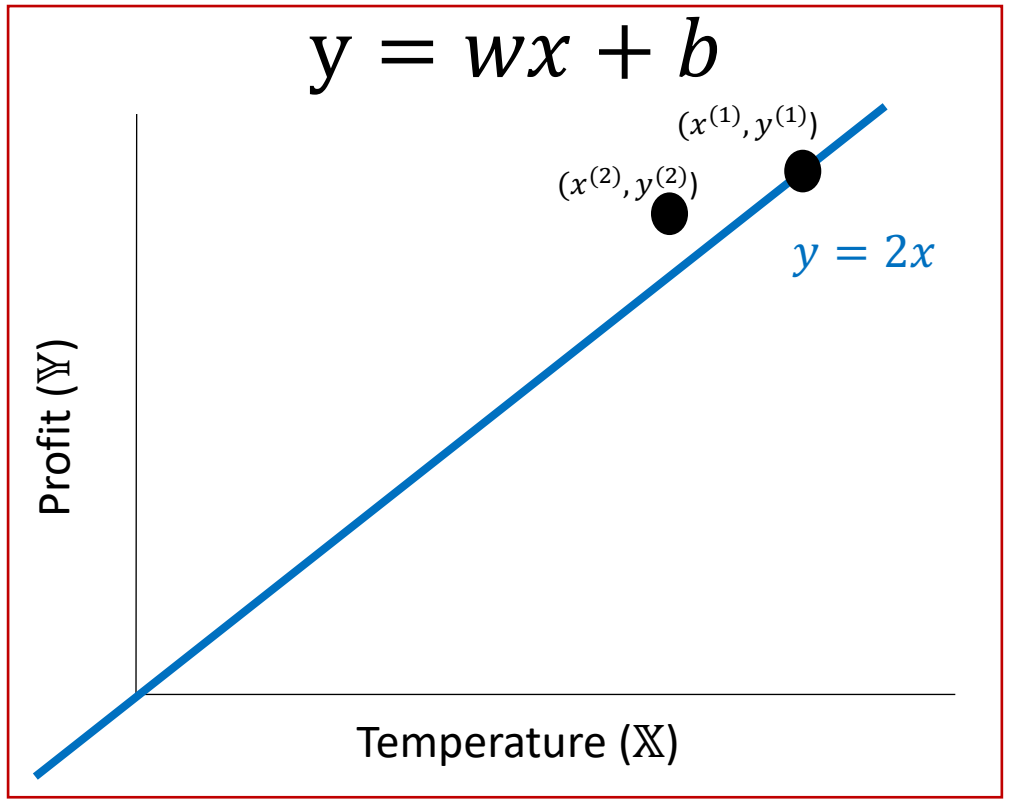
$x^{(1)} = 100.1$

$x^{(2)} = 80.0$

$x^{(3)} = 30.3$

Linear function

$$y = wx + b$$



Target: $\mathbb{Y}$

"Profit made on selling lemonade"

$y^{(1)} = 200.0$

$y^{(2)} = 180.5$

$y^{(3)} = 115.1$

$\mathbb{Y} \in \mathbb{R}$

(Numerical output)

(Image only for explaining concept, not drawn accurately)

Learning function f

Input: $\mathbb{X}$

"Temperature"

$\mathbb{X} \in \mathbb{R}$

Regression

Target: $\mathbb{Y}$

"Profit made on selling lemonade"

$x^{(1)} = 100.1$

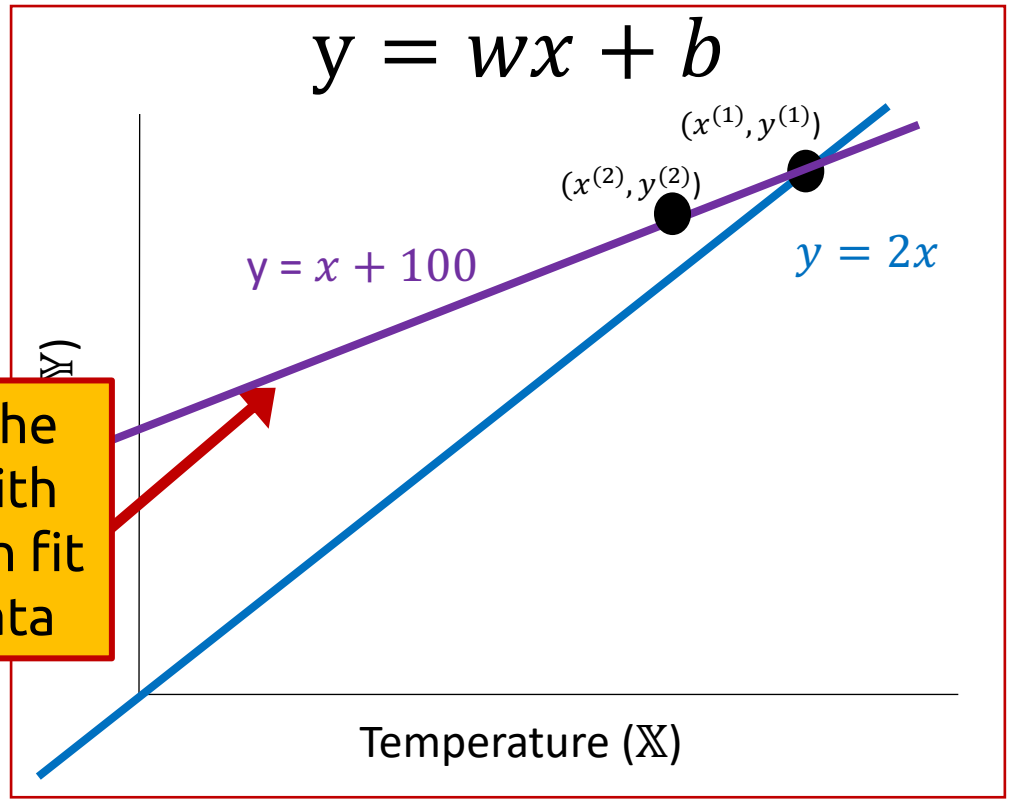
$x^{(2)} = 80.0$

$x^{(3)} =$

Only the line with bias can fit the data

Linear function

$$y = wx + b$$



$y^{(1)} = 200.0$

$y^{(2)} = 180.5$

$y^{(3)} = 115.1$

$\mathbb{Y} \in \mathbb{R}$

(Numerical output)

(Image only for explaining concept, not drawn accurately)

Learning function f

Input: $\mathbb{X}$

"Temperature"

$\mathbb{X} \in \mathbb{R}$

$x^{(1)} = 100.1$

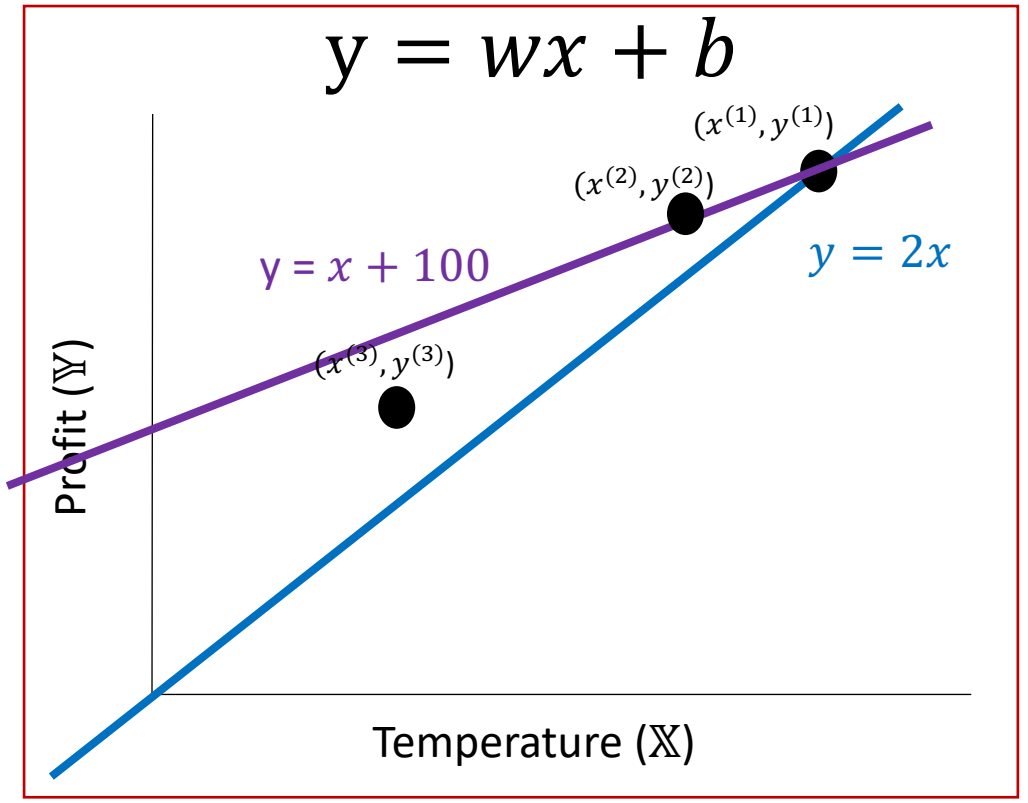
$x^{(2)} = 80.0$

$x^{(3)} = 30.3$

Regression

Linear function

$$y = wx + b$$



Target: $\mathbb{Y}$

"Profit made on selling lemonade"

$y^{(1)} = 200.0$

$y^{(2)} = 180.5$

$y^{(3)} = 115.1$

$\mathbb{Y} \in \mathbb{R}$

(Numerical output)

(Image only for explaining concept, not drawn accurately)

Learning function f

Input: $\mathbb{X}$

"Temperature"

Regression

Target: $\mathbb{Y}$

"Profit made on selling lemonade"

Linear function

$$y = wx + b$$

$x^{(1)} = 100.1$ $y^{(1)} = 200.0$

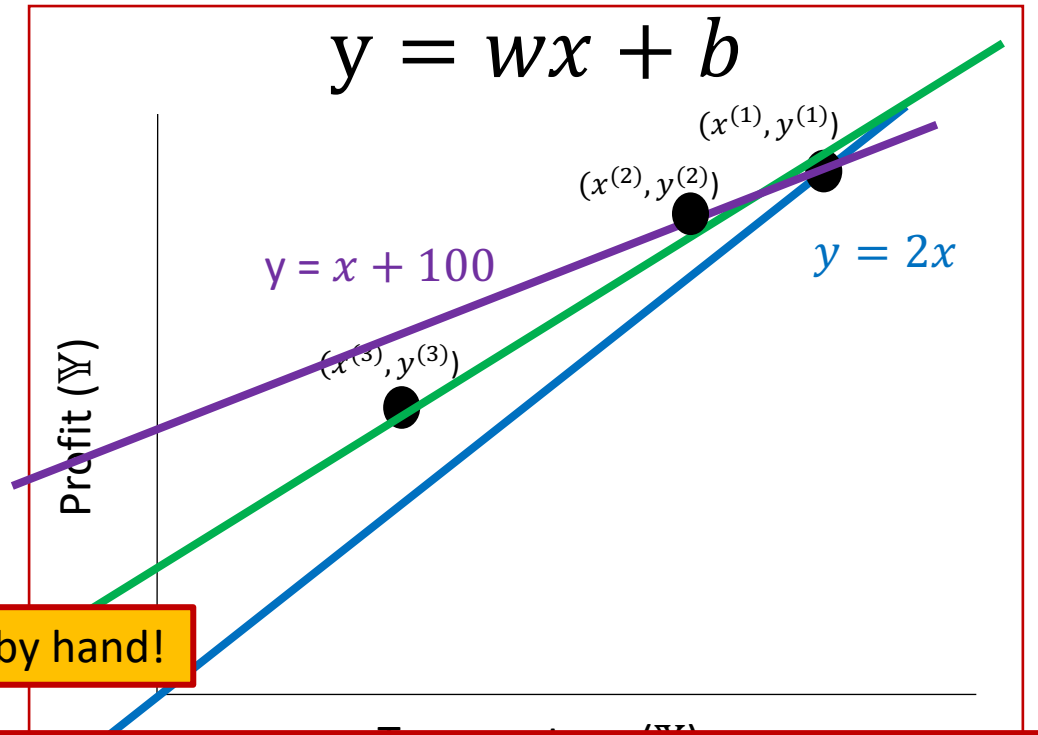
$x^{(2)} = 80.0$ $y^{(2)} = 180.5$

$x^{(3)} = 30.3$ $y^{(3)} = 115.1$

$\mathbb{X} \in \mathbb{R}$

$\mathbb{Y} \in \mathbb{R}$

(Numerical output)

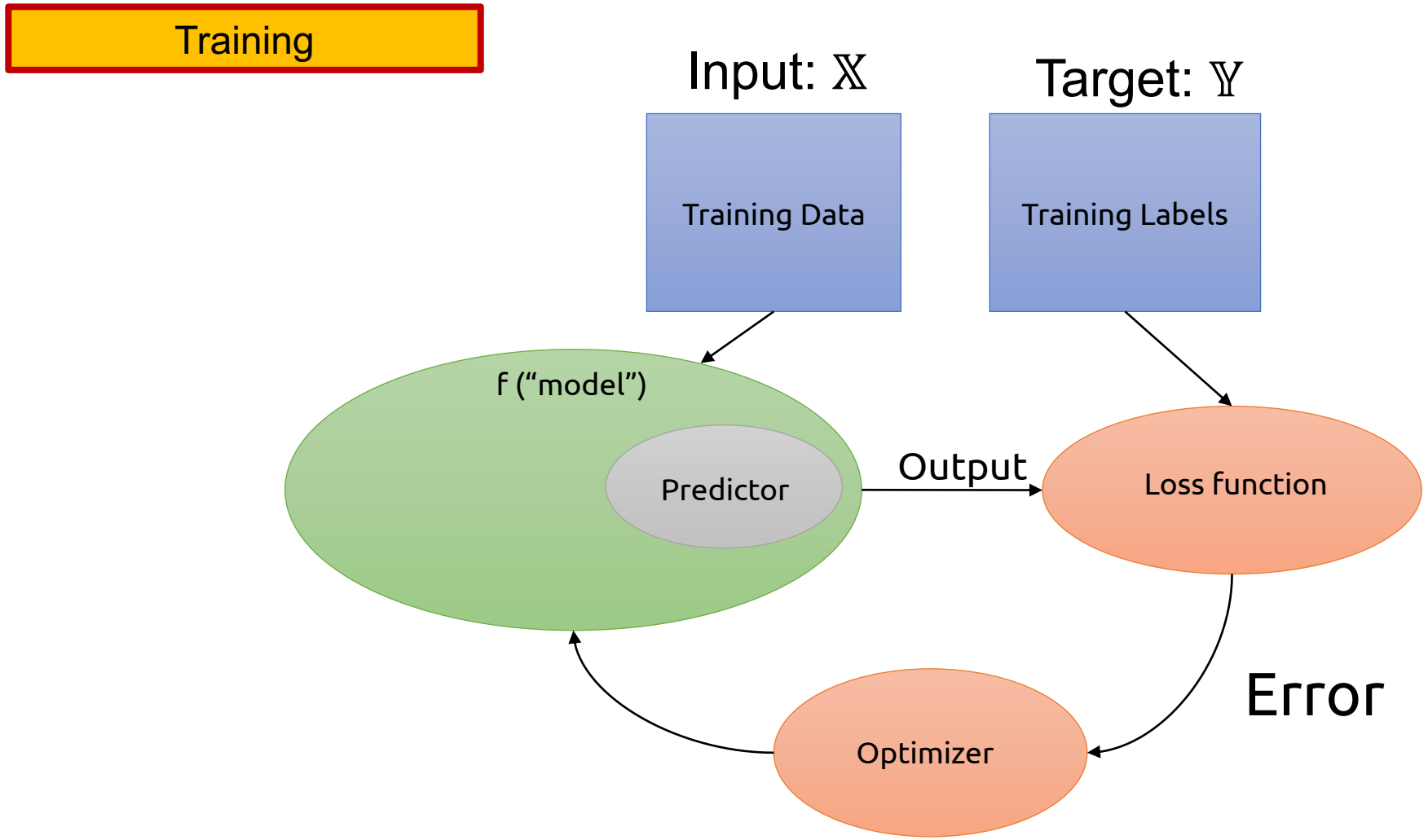


Very hard to learn these functions by hand!

Use machine learning to learn a good approximation of the function ***from data***

(Image only for explaining concept, not drawn accurately)

“Classic” Supervised Learning in Machine Learning



Any questions?

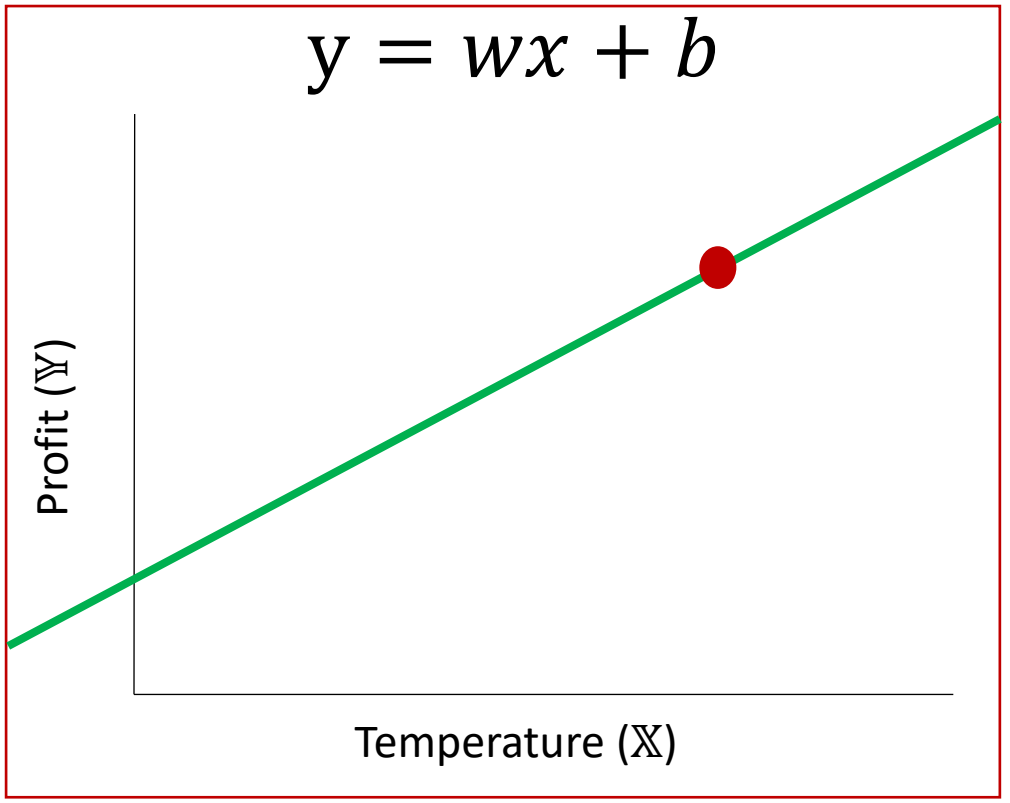
Testing our model

"Temperature"

$x' = 70$

Linear function

$$y = wx + b$$



"Profit made on selling lemonade"

Prediction

$y' = 175$

(Image only for explaining concept, not drawn accurately)

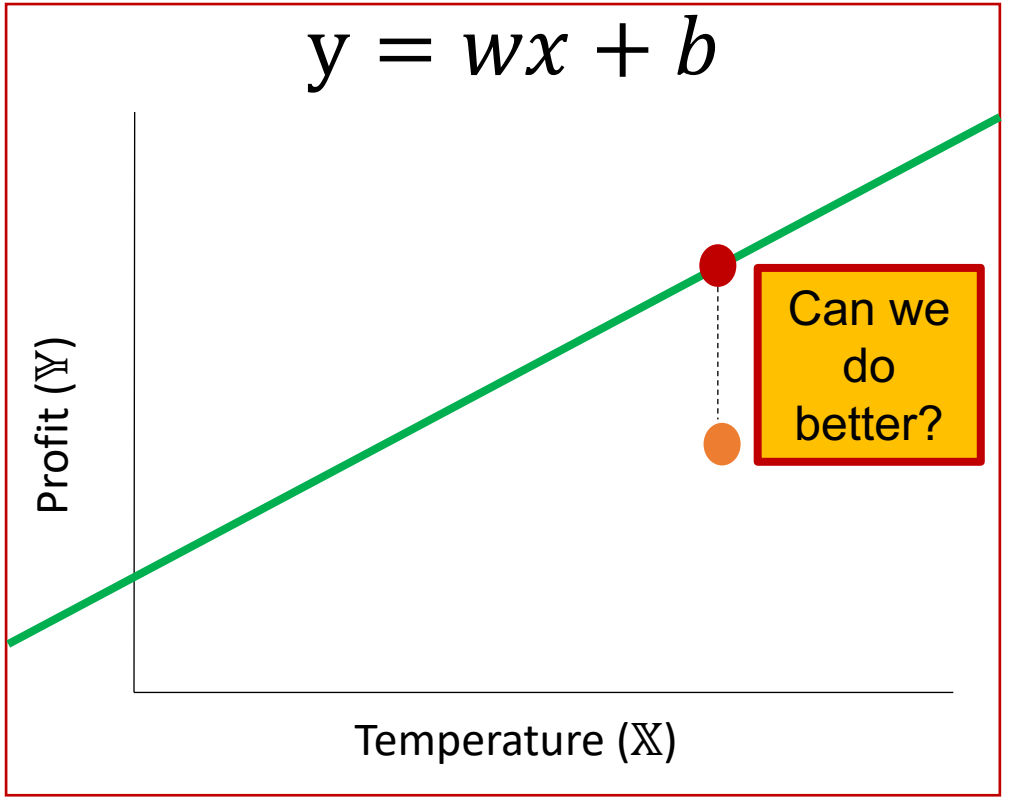
Testing our model

"Temperature"

"Profit made on selling lemonade"

Linear function

$$y = wx + b$$



$x' = 70$

$\hat{x} = 70$

Prediction

$y' = 175$

True observation

$\hat{y} = 140$

Real-world deployment

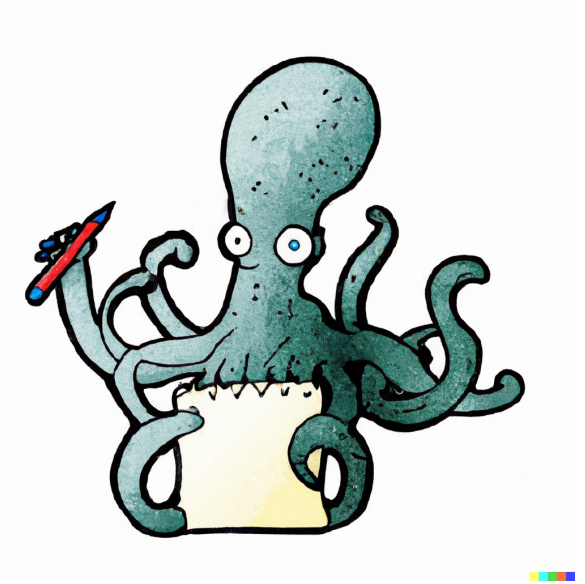
(Image only for explaining concept, not drawn accurately)

(An outlier is a data point that differs significantly from other observations)

Can we do better? – May be

How? Join at menti.com | use code 3688 8735

Option 1: Collect more data and retrain
Option 2: Try a different function
Option 3: Do both 1 and 2



DALL-E 2 prompt “a cartoon octopus writing notes”

Learning better models – Collect more data

Input: $\mathbb{X}$

"Temperature"

$x^{(1)} = 100.1$

$x^{(2)} = 80.0$

$x^{(3)} = 30.3$

.

.

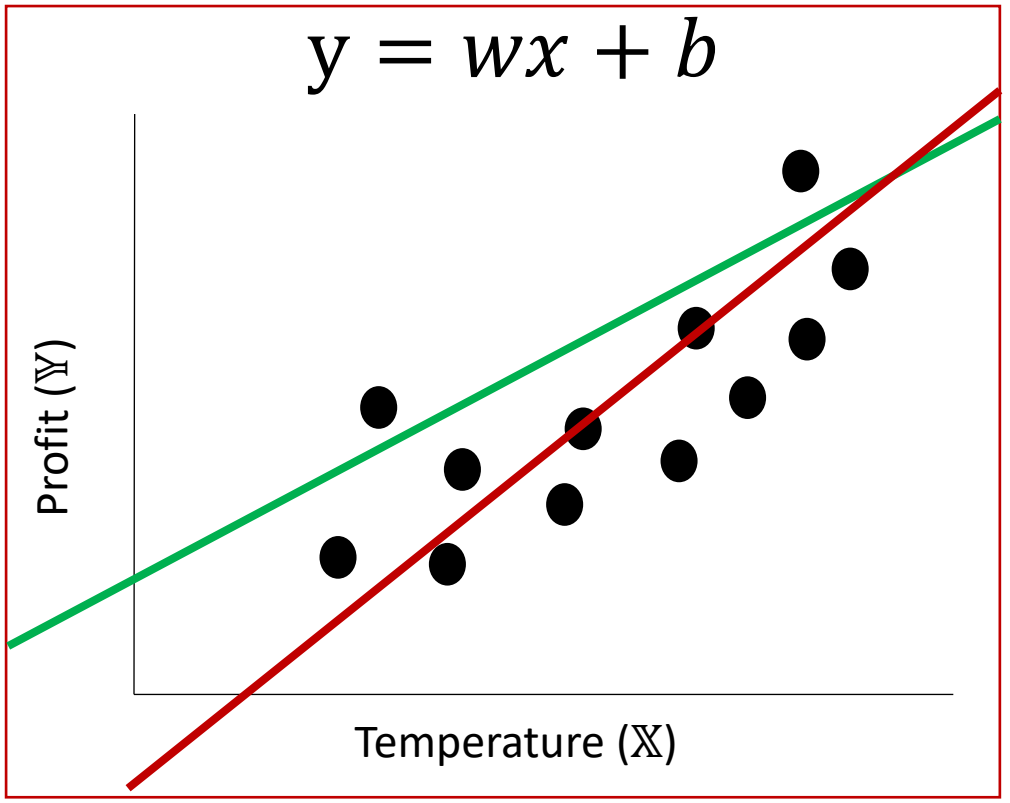
.

$x^{N} = \ldots$

$\mathbb{X} \in \mathbb{R}$

Linear function

$$y = wx + b$$



Target: $\mathbb{Y}$

"Profit made on selling lemonade"

$y^{(1)} = 200.0$

$y^{(2)} = 180.5$

$y^{(3)} = 115.1$

.

.

.

$y^{N} = \ldots$

$\mathbb{Y} \in \mathbb{R}$

(Numerical output)

(Image only for explaining concept, not drawn accurately)

Learning better models – Try different functions

Input: $\mathbb{X}$

"Temperature"

$\mathbb{X} \in \mathbb{R}$

$x^{(1)} = 100.1$

$x^{(2)} = 80.0$

$x^{(3)} = 30.3$

.

.

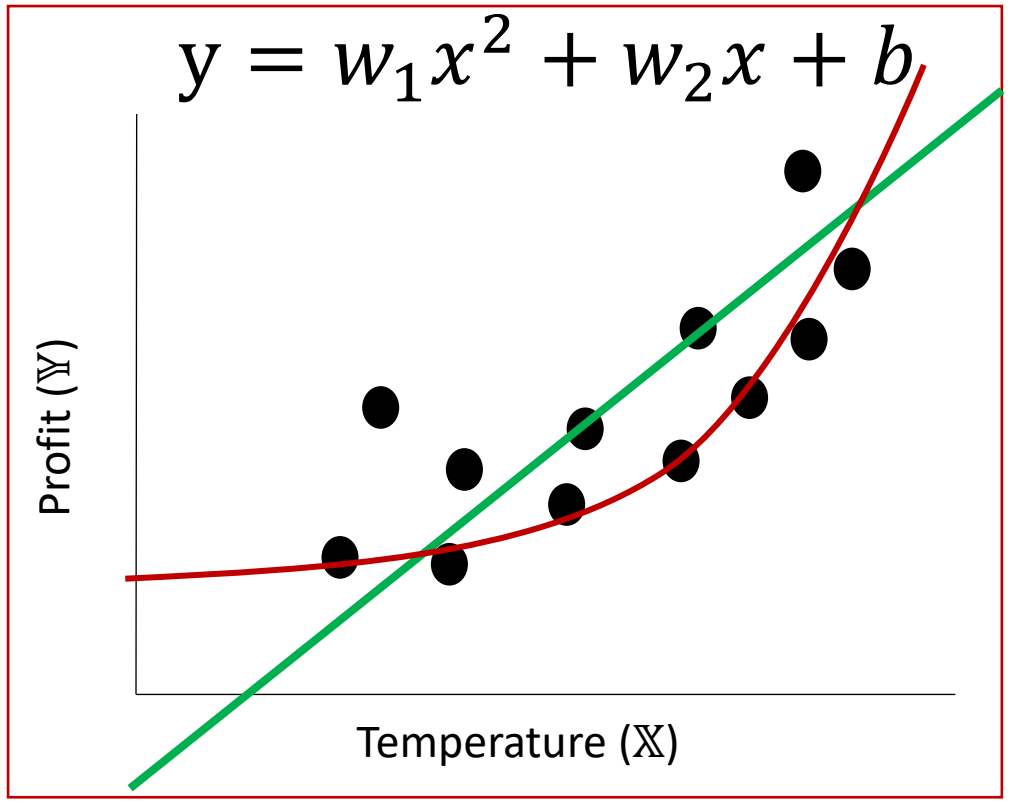
.

$x^{N} = \ldots$

Non-linear function

Polynomial function

$$y = w_1 x^2 + w_2 x + b$$



Target: $\mathbb{Y}$

"Profit made on selling lemonade"

$y^{(1)} = 200.0$

$y^{(2)} = 180.5$

$y^{(3)} = 115.1$

.

.

.

$y^{N} = \ldots$

$\mathbb{Y} \in \mathbb{R}$

(Numerical output)

(Image only for explaining concept, not drawn accurately)

How to know which function is the best?

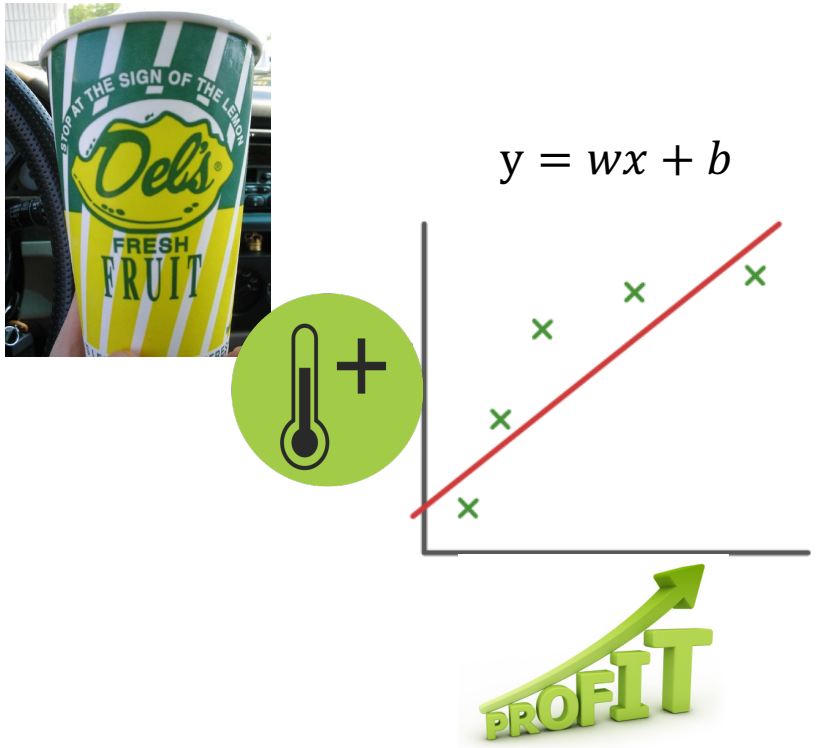
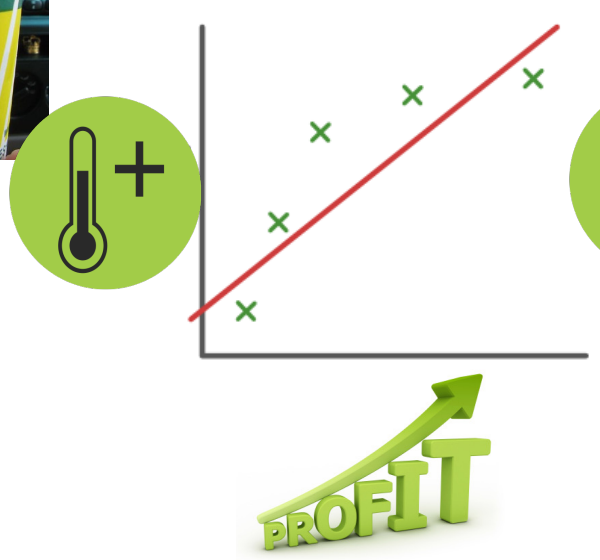


Image courtesy: https://www.geeksforgeeks.org/underfitting-and-overfitting-in-machine-learning/

How to know which function is the best?

$y = wx + b$



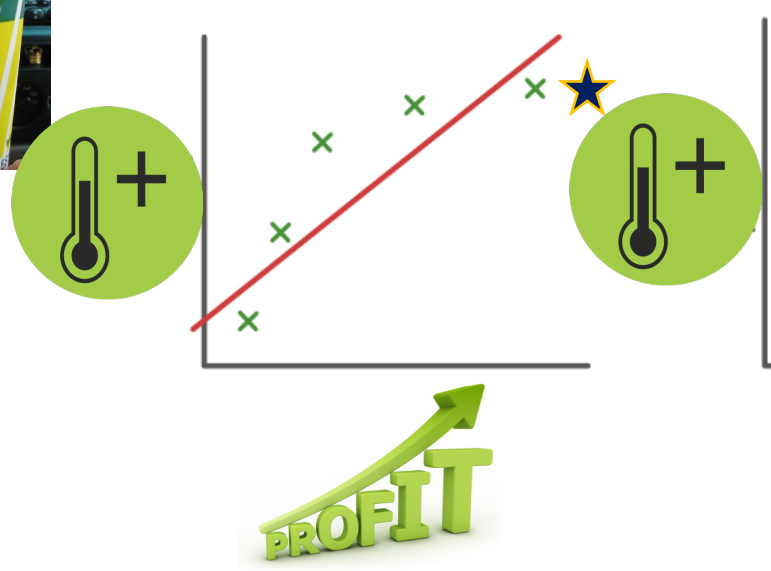
$y = w_1 x^2 + w_2 x + b$

Image courtesy: https://www.geeksforgeeks.org/underfitting-and-overfitting-in-machine-learning/

How to know which function is the best?

Underfit

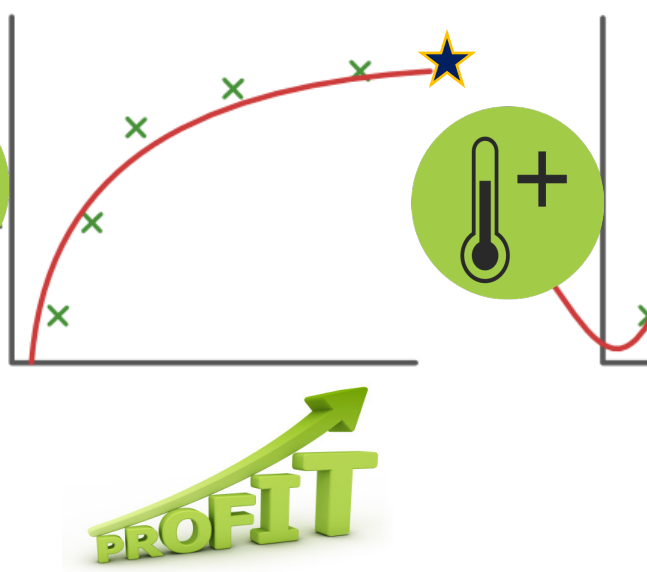
$y = wx + b$



"My model is not doing that well on the given data and new data" ☹

Good fit

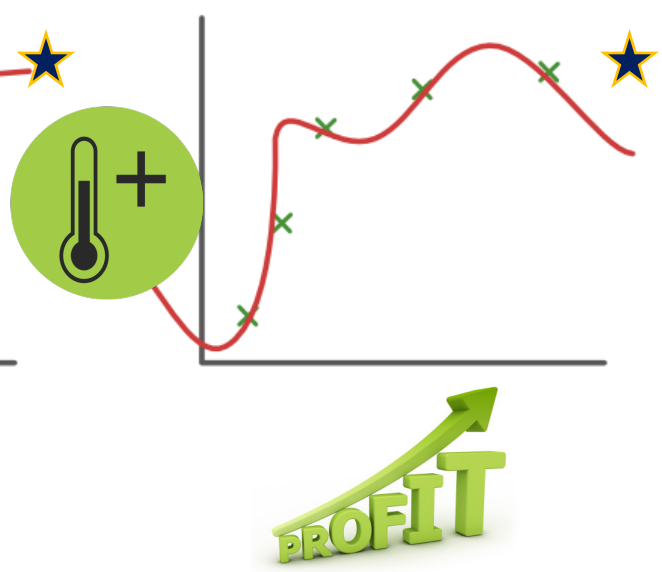
$y = w_1x^2 + w_2x + b$



"My model is doing well on the given data AND the new data point!! ☺

Overfit

$y = w_1x^4 + w_2x^3 + w_3x^2 + w_4x + b$



"My model is doing really well on the given data!! ☺

"The performance is bad on new data point" ☹

Image courtesy: https://www.geeksforgeeks.org/underfitting-and-overfitting-in-machine-learning/

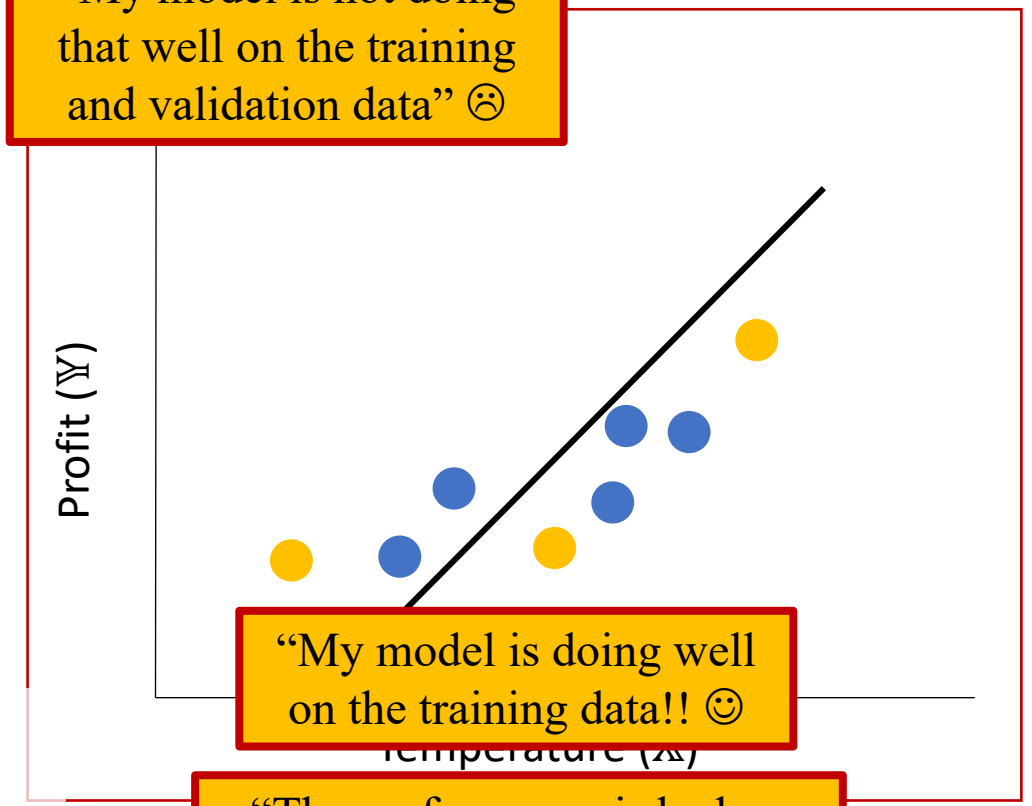
How to train your ~~dragon~~ model

How can we use the data we have to test for the fit?

	Input: $\mathbb{X}$ "Temperature"	Target: $\mathbb{Y}$ "Profit made on selling lemonade"
Training set	$x^{(1)} = 100.1$	$y^{(1)} = 200.0$
	$x^{(2)} = 80.0$	$y^{(2)} = 180.5$
	$x^{(3)} = 30.3$	$y^{(3)} = 115.1$
	.	.
Validation set	.	.
	.	.
	.	.
	.	.
Test set	.	.
	.	.
	$x^N = \ldots$	$y^N = \ldots$

Training phase

"My model is not doing that well on the training and validation data" ☹



"My model is doing well on the training data!! ☺

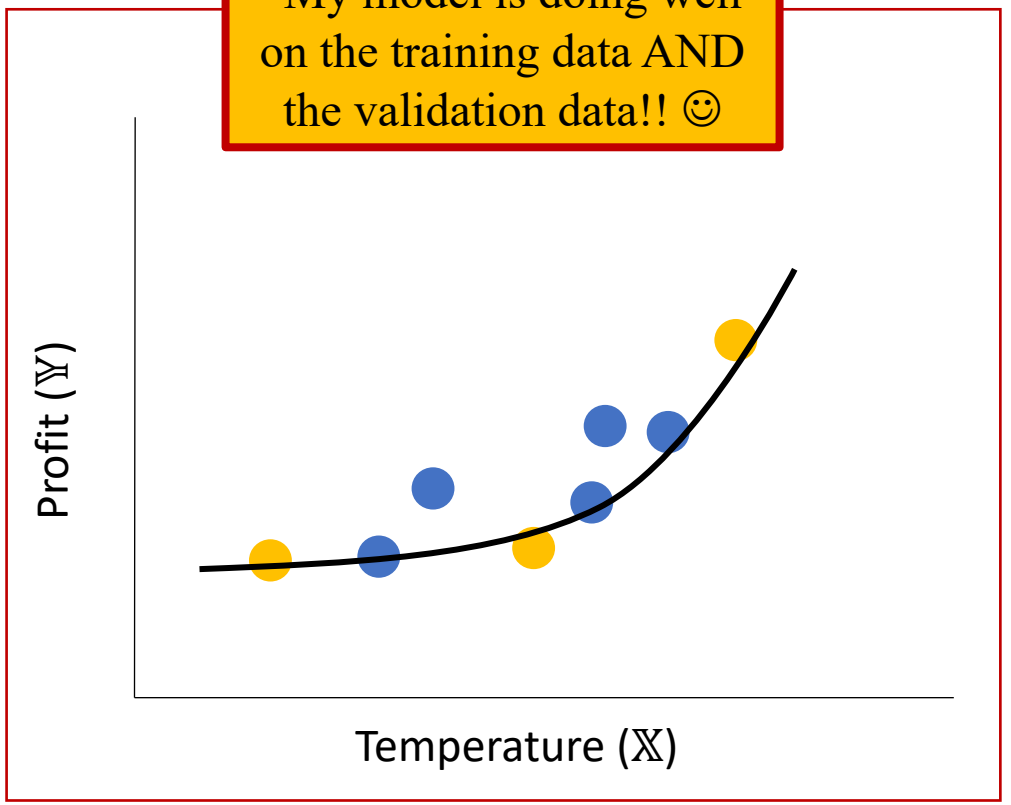
"The performance is bad on validation data" ☹

How to train your ~~dragon~~ model

	Input: $\mathbb{X}$	Target: $\mathbb{Y}$
	"Temperature"	"Profit made on selling lemonade"
Training set	$x^{(1)} = 100.1$	$y^{(1)} = 200.0$
	$x^{(2)} = 60.0$	$y^{(2)} = 160.5$
	$x^{(3)} = 30.3$	$y^{(3)} = 115.1$
	.	.
Validation set	.	.
	.	.
	.	.
	.	.
Test set	.	.
	.	.
	$x^N = \ldots$	$y^N = \ldots$

Training phase

"My model is doing well on the training data AND the validation data!! ☺

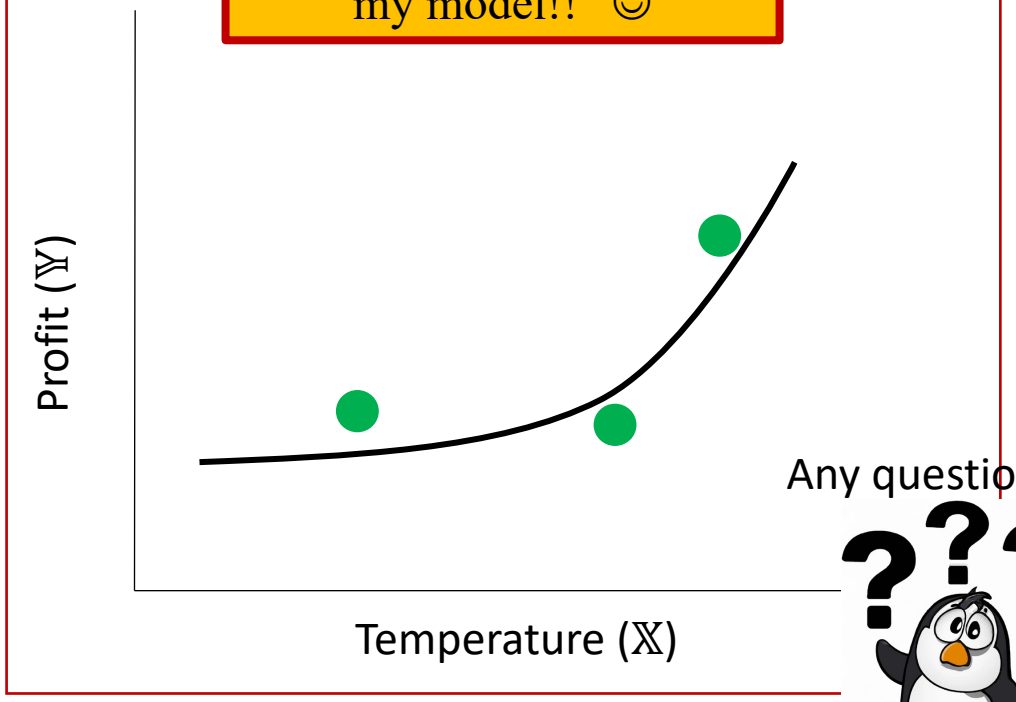


How to train your ~~dragon~~ model

	Input: $\mathbb{X}$	Target: $\mathbb{Y}$
	"Temperature"	"Profit made on selling lemonade"
Training set	$x^{(1)} = 100.1$	$y^{(1)} = 200.0$
	$x^{(2)} = 80.0$	$y^{(2)} = 180.5$
	$x^{(3)} = 30.3$	$y^{(3)} = 115.1$
	.	.
Validation set	.	.
	.	.
	.	.
Test set	.	.
	.	.
	$x^N = \ldots$	$y^N = \ldots$

Testing phase

"I have more confidence in the generalizability of my model!!" ☺



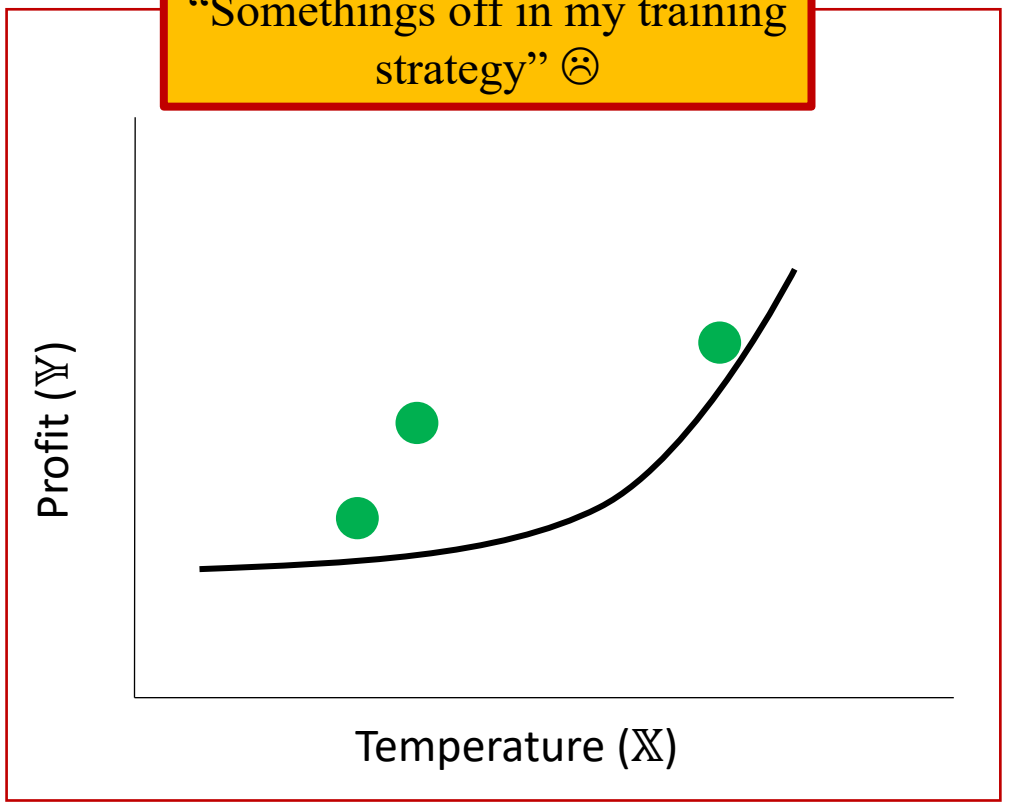
Any questions?

How to train your ~~dragon~~ model

	Input: $\mathbb{X}$	Target: $\mathbb{Y}$
	"Temperature"	"Profit made on selling lemonade"
Training set	$x^{(1)} = 100.1$	$y^{(1)} = 200.0$
	$x^{(2)} = 80.0$	$y^{(2)} = 180.5$
	$x^{(3)} = 30.3$	$y^{(3)} = 115.1$
	.	.
Validation set	.	.
	.	.
	.	.
Test set	.	.
	$x^N = \ldots$	$y^N = \ldots$

Testing phase

"Somethings off in my training strategy" ☹



Real world data tends to be complicated!

Input: $\mathbb{X}$

"Temperature" "Stand Hours" "Sunny?"

$$x_1^{(1)} = 100.1 \quad x_2^{(1)} = 8 \quad x_3^{(1)} = 1$$
$$x_1^{(2)} = 80.0 \quad x_2^{(2)} = 4 \quad x_3^{(2)} = 1$$
$$x_1^{(3)} = 30.3 \quad x_2^{(3)} = 8 \quad x_3^{(3)} = 0$$
$$\vdots \qquad \vdots \qquad \vdots$$
$$x_i^{(k)} = \ldots$$

$\mathbb{X} \in \mathbb{R}^3$

Now our function needs to capture the relationships of the combined feature space of the input and the output!

Target: $\mathbb{Y}$

"Profit made on selling lemonade"

$$y^{(1)} = 200.0$$
$$y^{(2)} = 180.5$$
$$y^{(3)} = 115.1$$
$$\vdots$$
$$y^{(k)} = \ldots$$

$\mathbb{Y} \in \mathbb{R}$

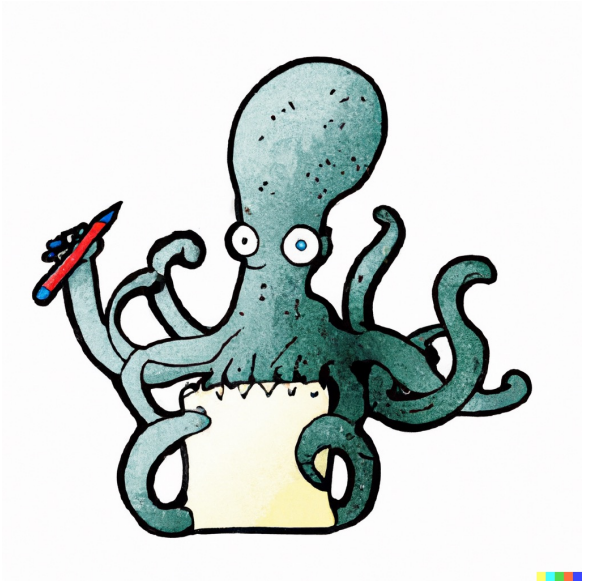
(Numerical output)

(Image only for explaining concept, not drawn accurately)

Recap

How to represent inputs and outputs

- Represent input and output as numbers
- Classification – predicting categorical outputs
- Regression – predicting numerical outputs



Supervised Learning

- Learn a function that approximates the data well
- Get more data!
- Try different models
- Pick a good model