**Machine Learning Algorithm**

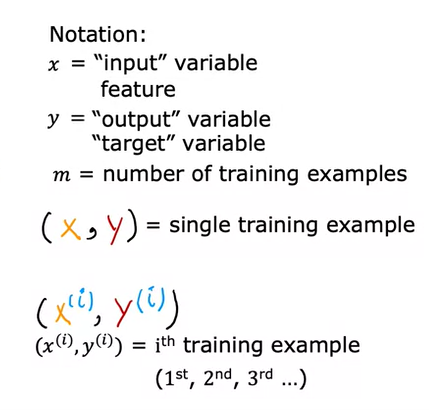
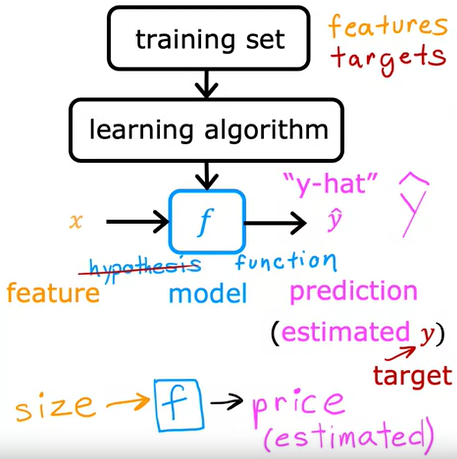
1. Supervised learning

More commonly .

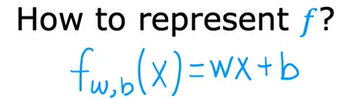
Input to output **Labled.**

„learn from being given „right answers” “.

Data has “right” answers.

1. Regression (predict a number)
   1. Predicts a number  
       

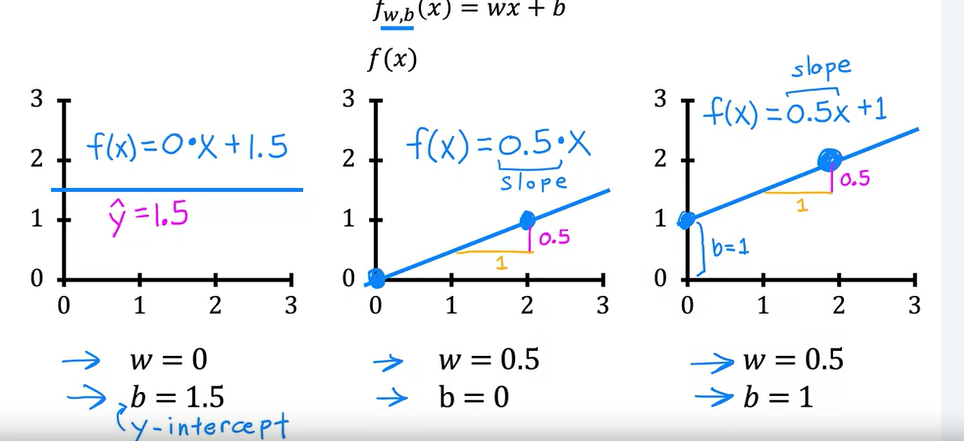
**Linear Regression with one feature**



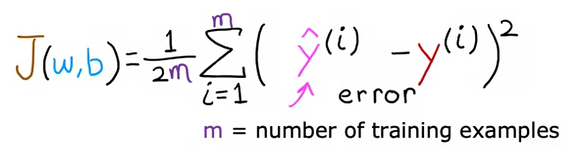
f is a linear function – means only has one variable (**univariate linear regression**)

**w = slope**

**b = interecept**

****

**Cost function**

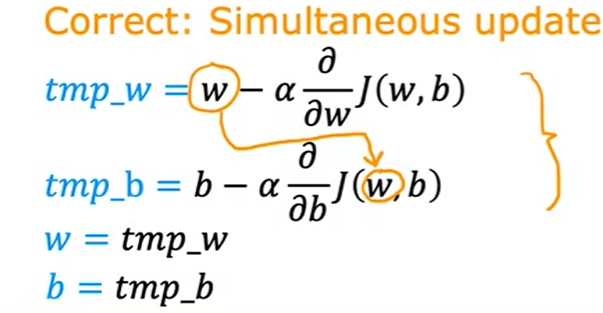
****

**Goal:** minimize J(w,b) to find the best fitting f(x)

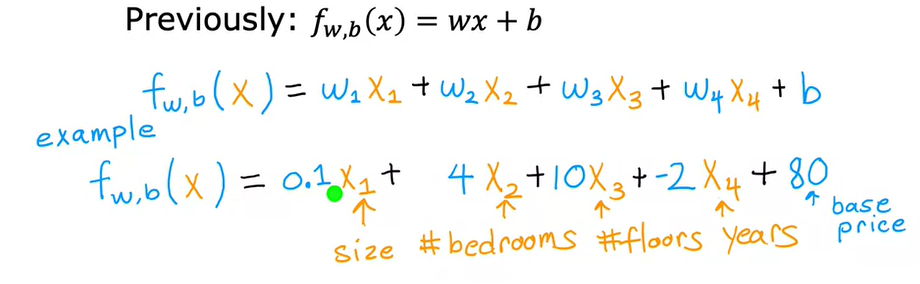
**Gradient Descent**

Find the local minima for the cost function J to minimize it.

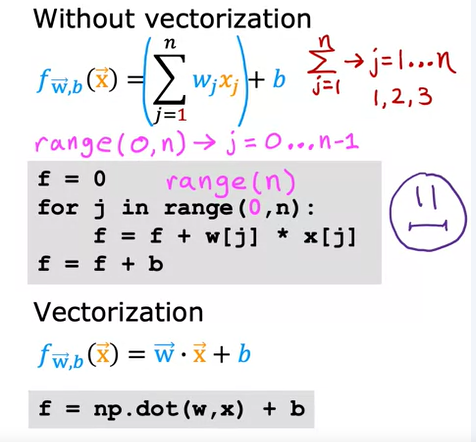
Like walking a hill to the valley



**Multiple features**

Instead of using just one x as a feature like the size of a house we are now going to use multiple features

(0.1, 4, 10, -2, 80 are randomly chosen)



**Features and Parameters**

The range of feature numbers and the parameters have impact on the accuracy of the prediction:

It has to be like:  
**Large range: small parameters**

**Small range: large parameters**

Otherwise the prediction is not accurate!!

Ein Bild, das Text, Screenshot, Diagramm, Schrift enthält.

Automatisch generierte Beschreibung

We need feature scaling!

**Feature scaling**

Mean normalization

Ein Bild, das Text, Schrift, Reihe, weiß enthält.

Automatisch generierte Beschreibung

Notice mü\_1 is the average in this case 600

Z-score normalization

Ein Bild, das Schrift, Reihe, Diagramm, Design enthält.

Automatisch generierte Beschreibung

Notice mü\_1 is the average and sigma\_1 is the standard deviation

Ein Bild, das Text, Schrift, Screenshot enthält.

Automatisch generierte Beschreibung

Choosing the learning rate

* Alpha can be to big -> use a smaller one
* Use a minus sign at the update of w
* With a small enough alpha the cost function J should decrease on every iteration
* Values of alpha to try: 0.0001, 0.0003 , 0.001, 0.003 , 0.1, 0.03, 1, 10, 1000

**Feature Engineering**

**Ein Bild, das Text, Screenshot, Design enthält.

Automatisch generierte Beschreibung**

Here we can combine the frontage and the depth to one new feature “area” which will optimize the model

Polynomial regression:

Ein Bild, das Text, Reihe, Diagramm, Schrift enthält.

Automatisch generierte Beschreibung

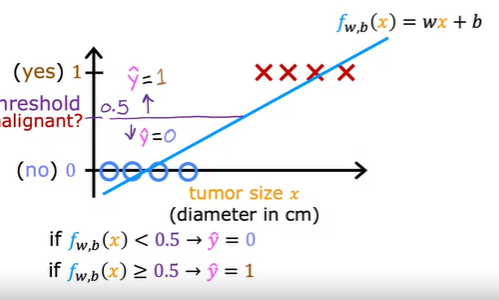
Selecting Features: 

* less weight value implies less important/correct feature, and in extreme, when the weight becomes zero or very close to zero, the associated feature is not useful in fitting the model to the data.
* above, after fitting, the weight associated with the 𝑥2x2 feature is much larger than the weights for 𝑥x or 𝑥3x3 as it is the most useful in fitting the data.

1. Classification
   1. predicts categories
      1. is email spam? yes or no (2 options)
      2. tumor recognition? malignant or non malignant

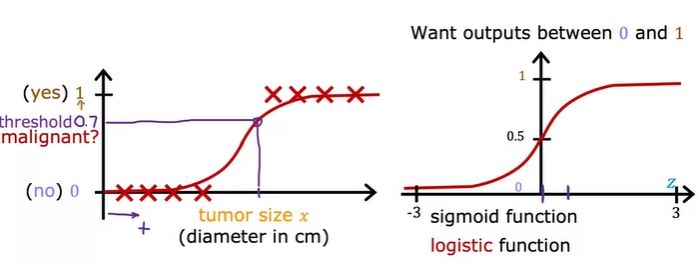
**binary classification**: y can only be one of two values (negative class / positive class)

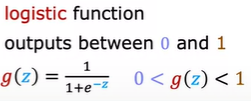
Threshold



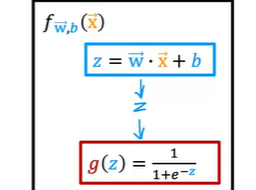
**Logistric Regression**

Sigmoid: output is 1 or 0





But what is z in logistic regression?





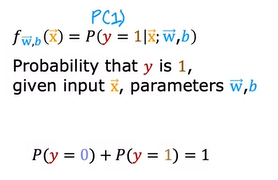
“probability” that class is 1

Example:

x is tumor size  
y is 0 or 1

f\_w,b(x) = 0.7 means that the chance is 70% the tumor is malignant (1)

AND that means:



p(y = 0) = 1-P(y=1) => 1-0.7 = 0.3

1. Unsupervised learning

**No labels**

Find pattern / interesting informations in unlabeled data