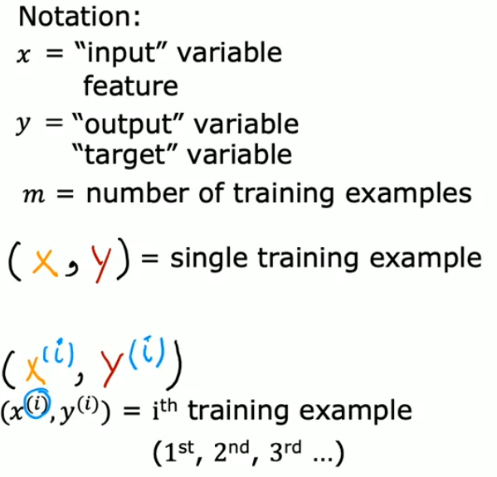
**Machine Learning**

Linear Regression model:

It is a model that is usually used in prediction.

A table with numbers and numbers

Description automatically generatedYou can make these predictions using one parameter or multiple parameters e.g.: - you want to predict the price of a house with respect to only the Area of it[single variable linear regression], or with respect to Area, number of floors, number of rooms, etc…. [multiple linear regression].





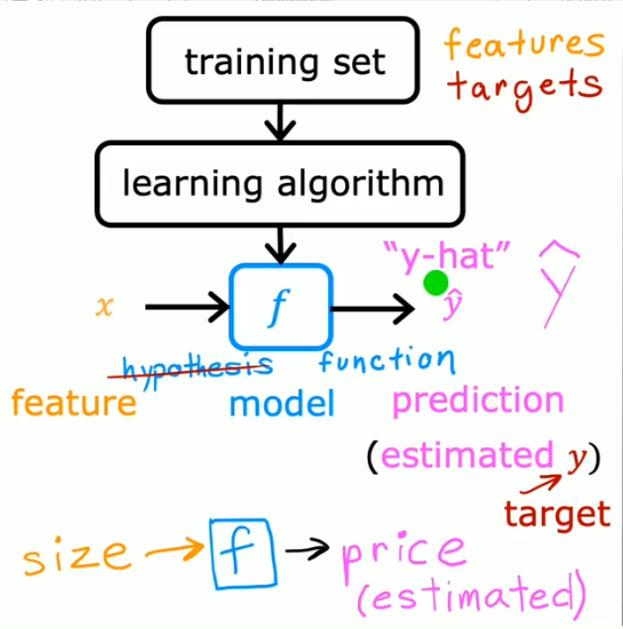
Output or Target

Input or Features



\_

ML process

 Y hat is the prediction

Y is the actual value of the target

X is the input

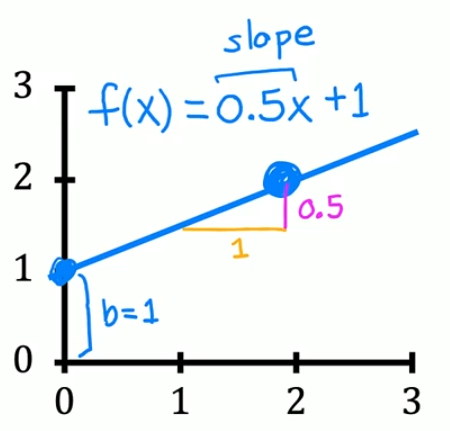
H or F is the function or the hypothesis that predict the output from the input

for the model all you have to do is to get the function that predicts the output with the most accurate result.

So in our example we use the straight line which equation is: f(x) = wx+b

we need to determine w and b

before anything what is w and b

w is the slope of the line and b is the bias (y-intercept)

**\_**

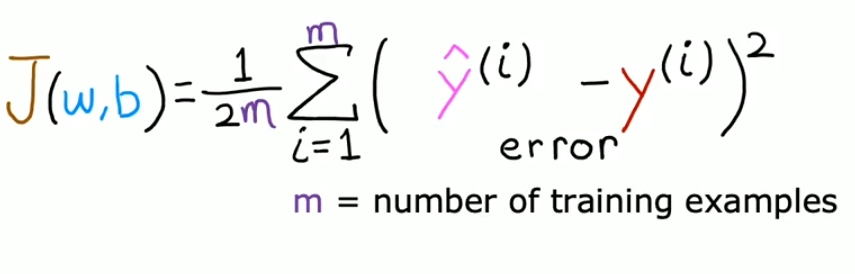
**Cost function (J)**

It is the function that is used to makes the predictions close to the true values of target

By getting the best values of weights of the line.

There are multiple cost functions in machine learning:

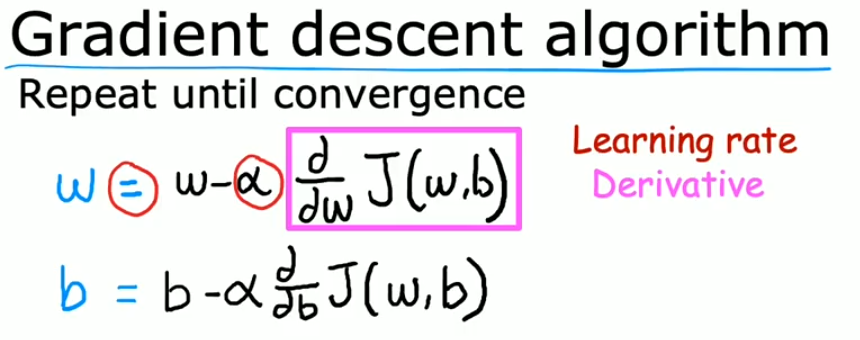
1. **Mean Squared Error (MSE)**
2. **Distance-Based Error**
3. **Root Mean Squared Error**
4. **Cross-Entropy Function**
5. **Mean Absolute Error**
6. **Kullback-Liebler (KL) Divergence**

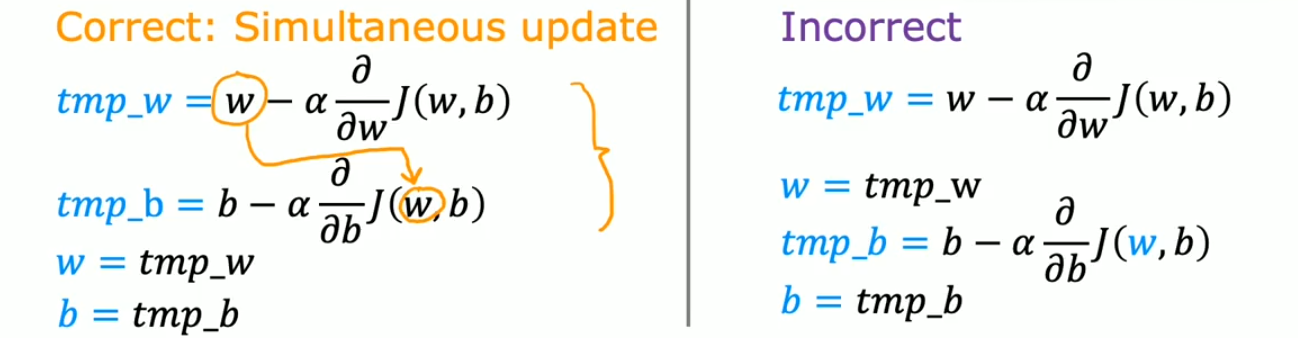


\_

Ok we have a function that gets the amount of error of the model compared to the actual target values

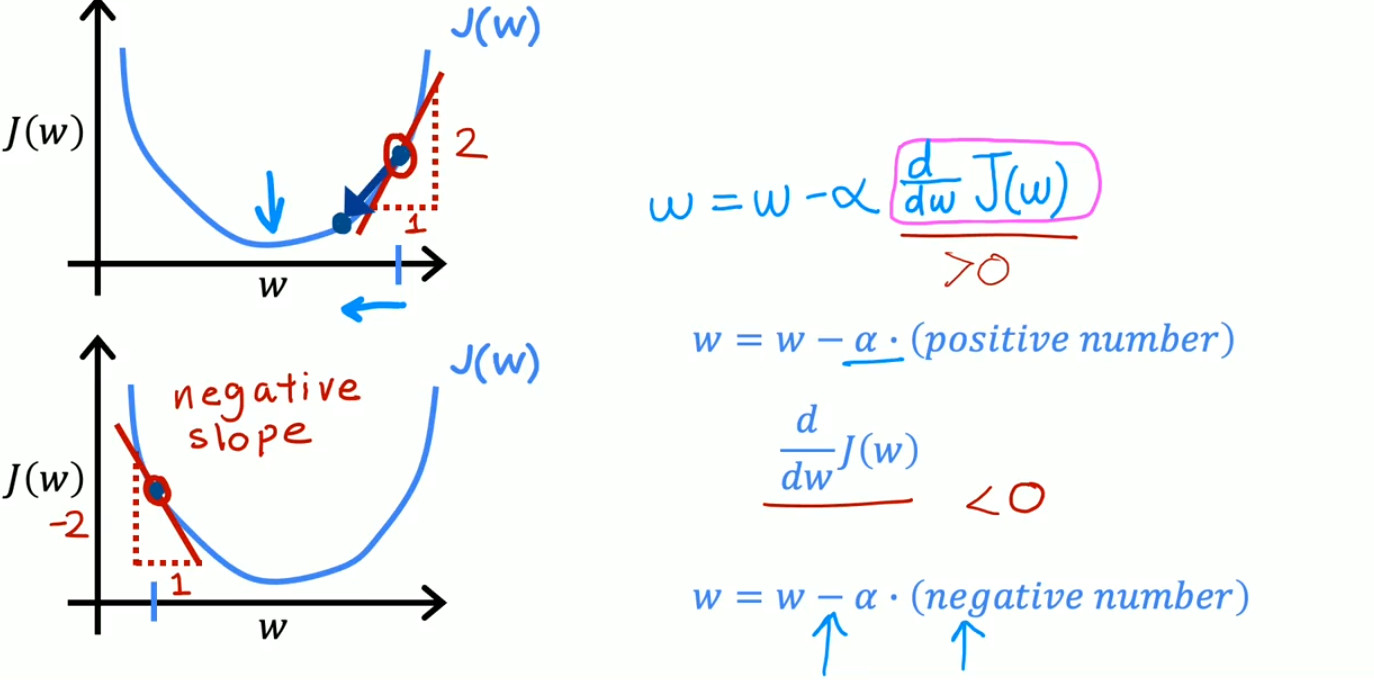
We need to get the minimum value of it

And here the gradient descent shines

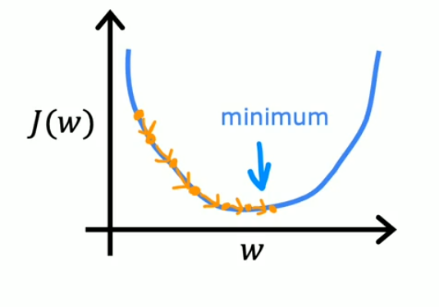


Tips about gradient descent

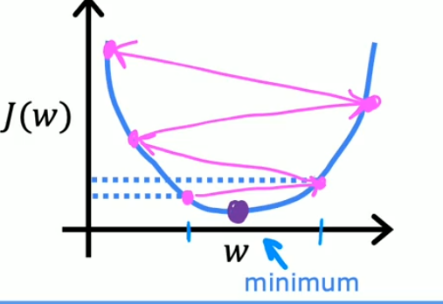
When the tangent line is positive then the new value of w will decrease and vice versa, if the line is negative then the new value increases.



\_

Learning rate (alpha symbol):

When you chose a small learning rate number the model will change values of weight with small steps so the model will be so slow but it will be accurate



In the other hand if it is too large, the model will take a bigger step in the change of values and may causes that the values diverges from the minimum cost

The final form of the gradient descent algorithm

