Hands-on Machine Learning Project

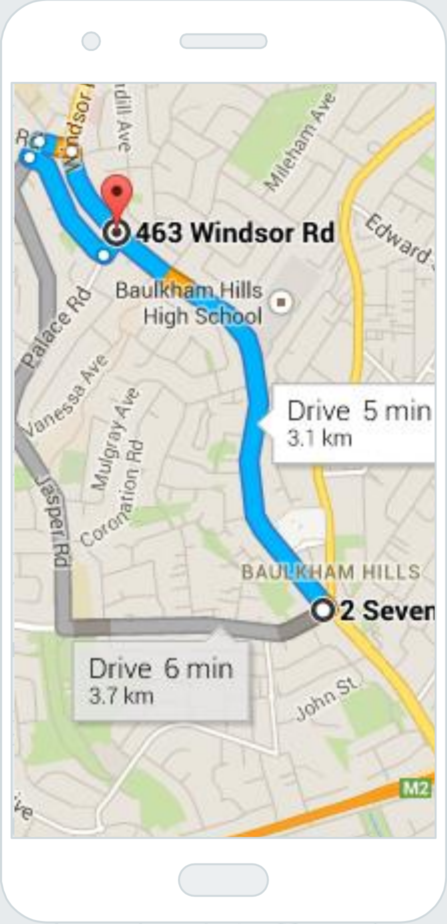
1.

Main Idea of ML

Let’s start with the explaining what is Machine learning.

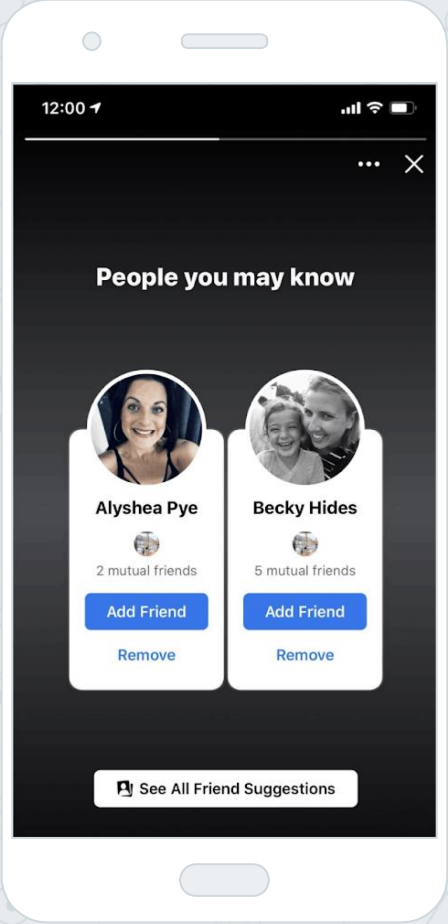
“Field of study that gives computers the ability to learn without being explicitly programmed“

- Arthur Samuel (1959)

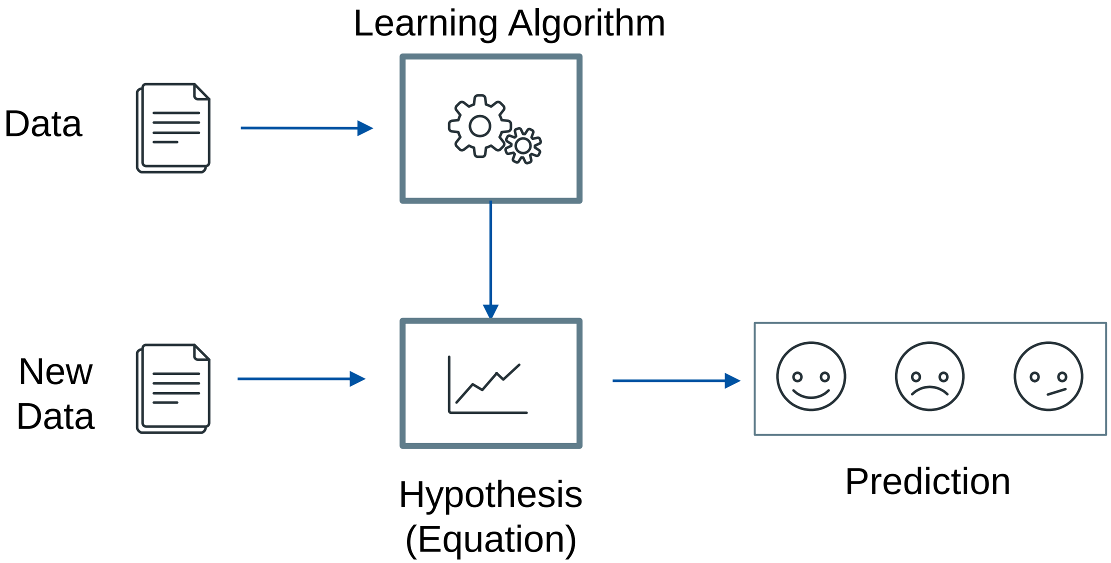

463 Windsor Rd
Baulkham Hills
High School
Drive 5 min
3.1 km
BAULKHAM HILLS
2 Seven
Drive 6 min
3.7 km
Mileham Ave
Edward
Palace Rd
Vanessa Ave
Mulgray Ave
Coronation Rd
Jasper Rd
John St
M2


Jackie
from danmollervfx


MTS 3:36 PM 62%
Messages (1) Stephen Details
offer the free ticket up.
And rock out for me.
Delivered
Will do man! Sounds goodbye
Good*
If you type a word the iPhone doesn't recognize like wowza
Send
"wowza"
wows
woes
space
return
123


12:00
People you may know
Alyshea Pye
2 mutual friends
Add Friend
Remove
Becky Hides
5 mutual friends
Add Friend
Remove
See All Friend Suggestions

Machine Learning – Main Idea



Data:
- Labeled Data
- Unlabeled Data

Prediction:
- Category
- Number
- Cluster

Machine Learning – Problems' Types

Supervised ML

Learn to predict target values from labelled data.

- Classification (target values are discrete classes)
- Regression (target values are continuous values)

Unsupervised ML

Find structure in unlabeled data.

- Clustering (find groups of similar instances in the data)
- Outlier Detection (finding unusual patterns)

Split the Data

In order to mimic the old/ new data we discussed, our number one step in any ML project is splitting the data.

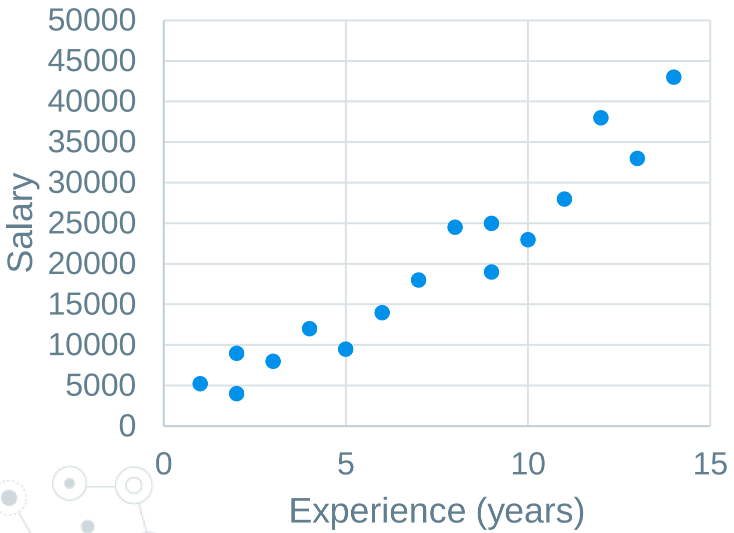
4.

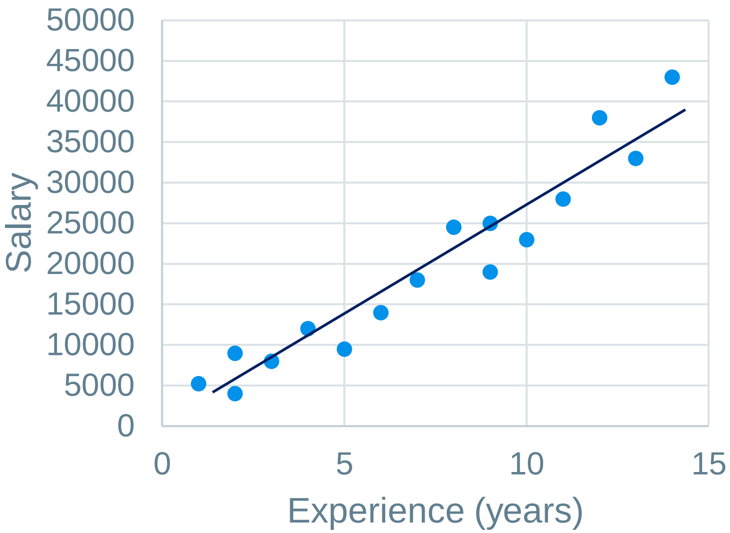
ML Algorithms

Now it’s time to learn some ML algorithms.

Linear Regression

- It's a regression Algorithm – used in supervised ML problems.
- the objective is to find the **regression line** that best fits the data.
Experience





Linear Regression - Equation

- The regression line has an equation of the form

$$\boxed{\text{“}Y = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \cdots\text{”}}$$

$\boldsymbol{X}$ is the features

$\boldsymbol{Y}$ is the predicted value

θ_0 is the bias term ($\boldsymbol{y}$ when $\boldsymbol{x} = 0$)

Training Phase

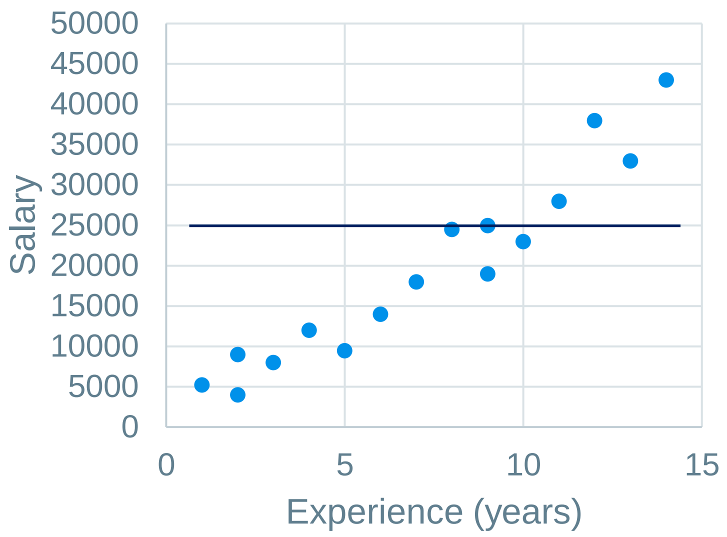
$$\text{“}\boldsymbol{Y} = \theta_0 + \theta_1 \boldsymbol{x_1} + \theta_2 \boldsymbol{x_2} + \cdots\text{”}$$

Testing Phase

$$\text{“}Y = \boldsymbol{\theta_0} + \boldsymbol{\theta_1 x_1} + \boldsymbol{\theta_2 x_2} + \cdots\text{”}$$

Linear Regression - Initialize theta

- We start by initializing the theta values to any random value include zeros



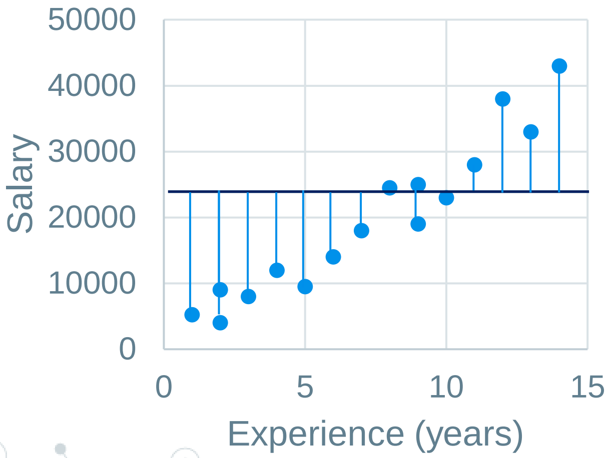
- We find that this value doesn’t fit our data, so we need to find the best fit values.

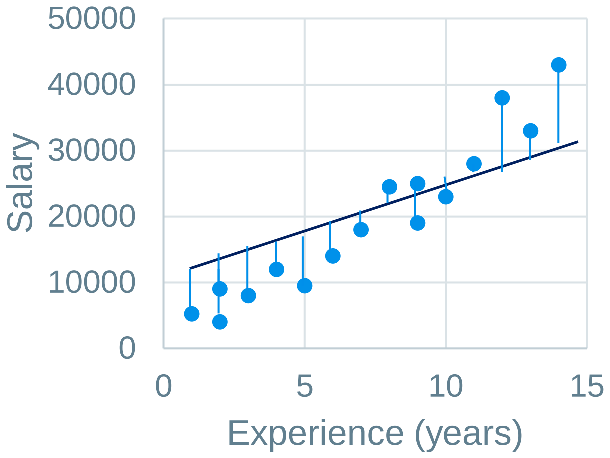
Linear Regression - Cost function

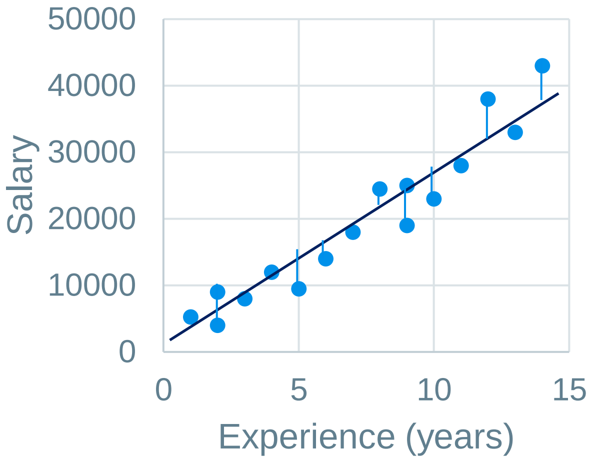
we use the MSE as a cost function.

$$\text{Error} = \sum_{i=1}^{n}(actual\_output - predicted\_output) ** 2$$

Where the error is the distance between the actual point to the regression line.



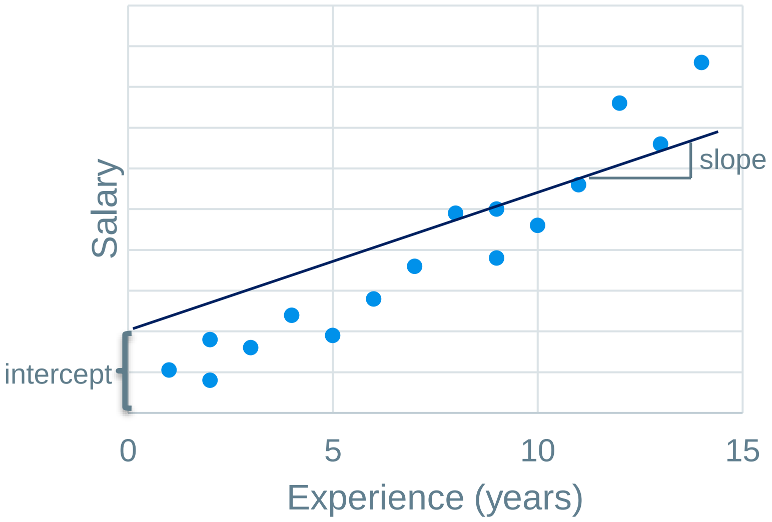




To find the best line we need to minimize the cost function, so we use the gradient descent algorithm.

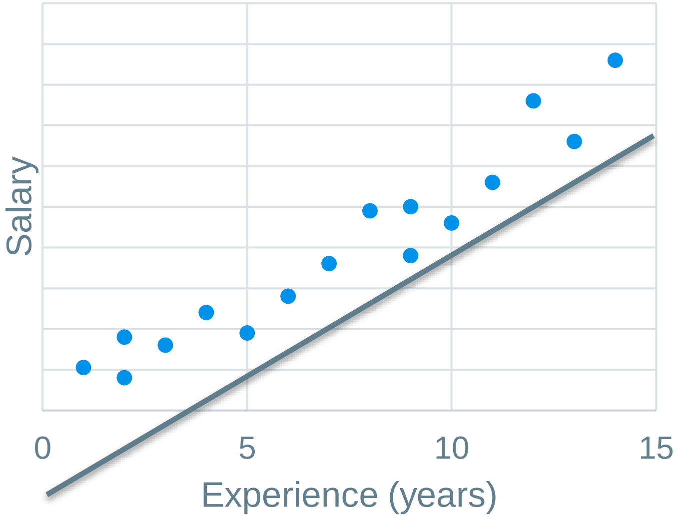
Gradient Descent

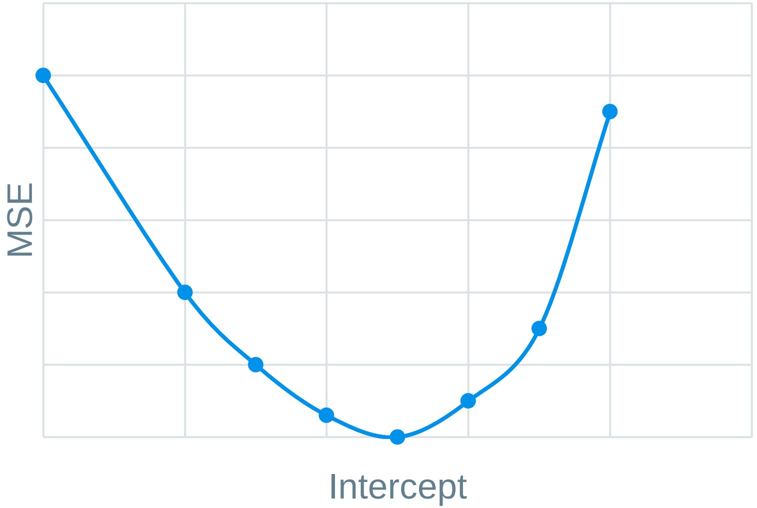
- Each line represented using two information: intercept, slop.



Gradient Descent

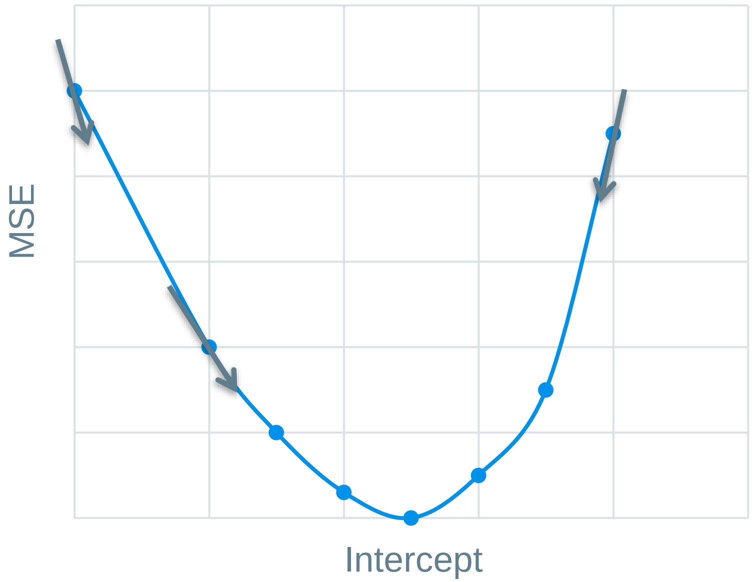
- To simplify the algorithm let’s keep the same slop and only change the intercept.





Gradient descent

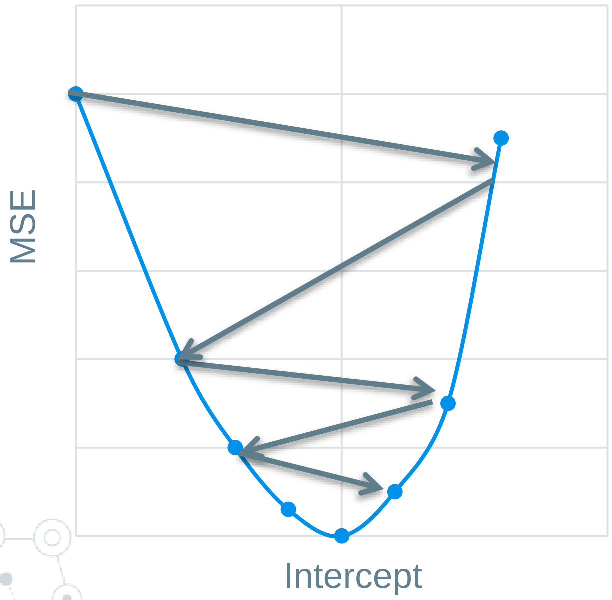
At each step, we calculate the slope of the cost function, we get the direction we need to move towards, in order to reach the local minimum.

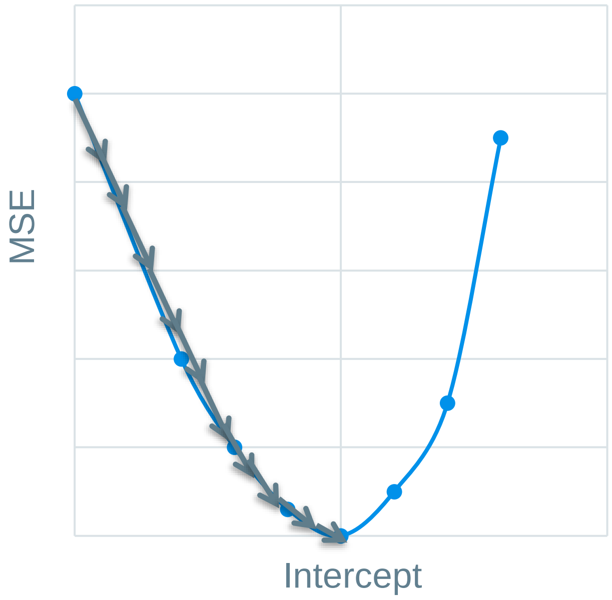


Gradient descent

To choose the step size we use the hyperparameter α, which is the learning rate.

If this step size, alpha, is too large, we will overshoot the minimum, If alpha is too small, we will take too many iterations to get to the minimum. So, alpha needs to be *just right*.





Simple Linear Regression

Experience	Salary

Multi Linear Regression

Experience	Title	Major	…..	Salary