Machine Learning –p3

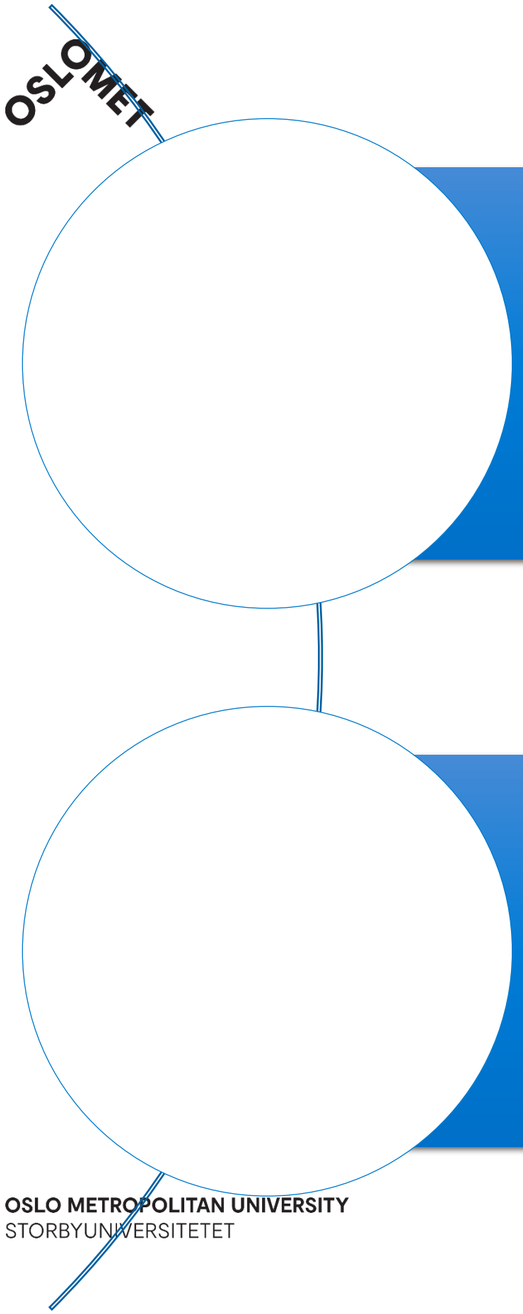
DAVE3625

Umair Mehmood Imam

DAVE3625- INTRO TO A.I. BY UMAIR M.I

OSLO METROPOLITAN UNIVERSITY
STORBYUNIVERSITETET

Types of supervised learning



Classification

- Support vector machines
- Decision Trees
- K-Nearest Neighbour
- Random Forest
- Logistic regression
- Naïve Bayes classifier
- …

Regression

- Linear Regression
- Polynomial Regression
- …

2. Naive Bayes classification algorithm (classification)

- A Naive Bayes classifier assumes that the presence of a particular feature in a class is unrelated to the presence of any other feature.
- Naive bayes does not take into account the relationship between features

For example:

- A dog will be considered a dog because it has 4 legs and a tail
  - But a Naive Bayes will consider everything a dog which has 4 legs or tail (e.g squirrel) not because its an animal. Being an animal is the relationship not legs / tail.
- A fruit may be considered to be an apple if it is red, round, and about 3 inches in diameter.

neural net guesses memes
@ResNeXtGuesser

Image prediction: ping-pong ball
Confidence: 100.0%

7:00 PM · Nov 25, 2021

218.6K

boredpanda.com

neural net guesses memes
@ResNeXtGuesser

Image prediction: Chihuahua
Confidence: 36.22%

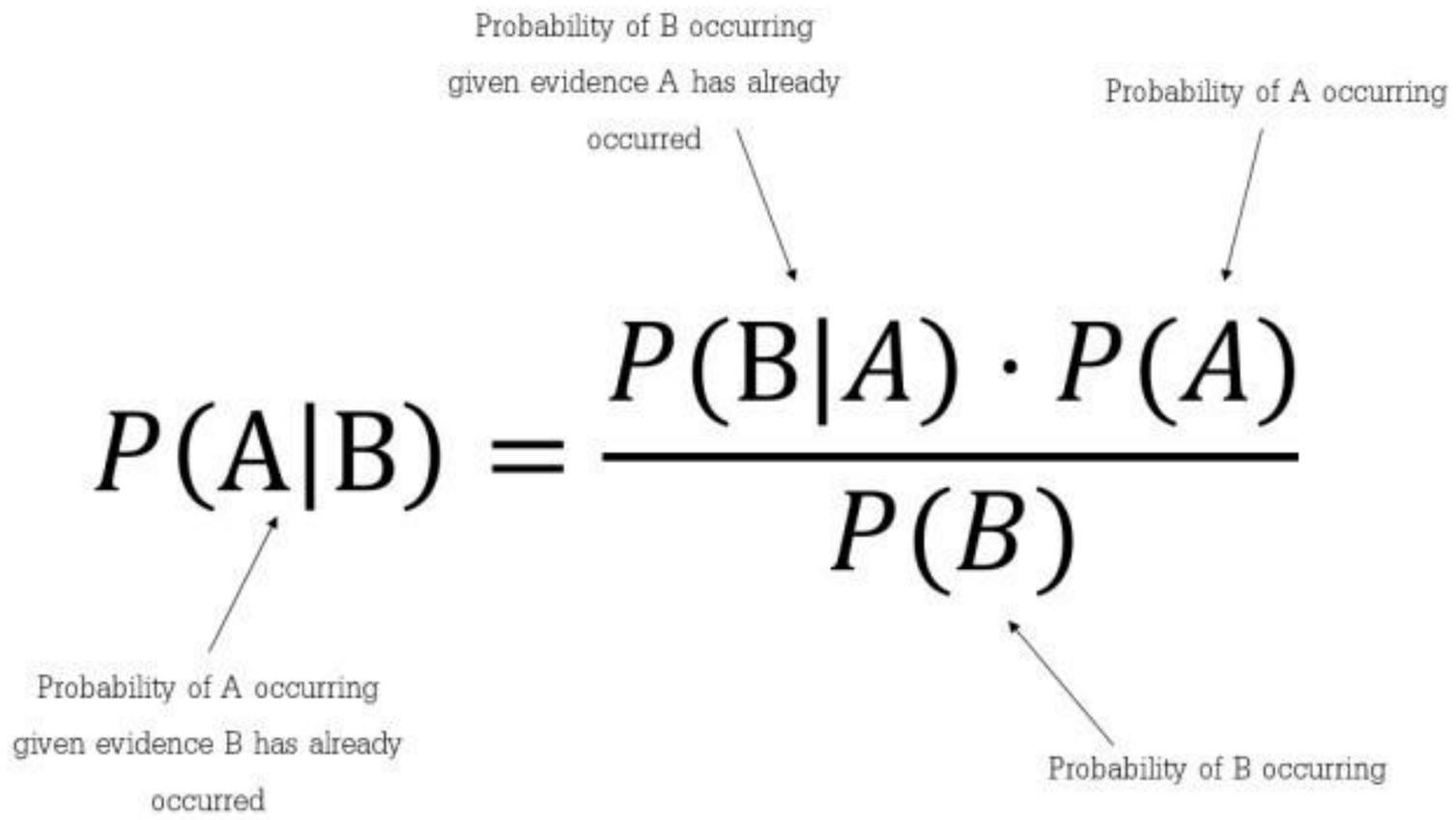
1:10 PM · Nov 21, 2021

94.7K

boredpanda.com

BAYES THEOREM

Bayes's theorem describes the probability of an event, based on prior knowledge of conditions that might be related to the event



Fruit	Long	Sweet	Yellow	Total
Banana	400	350	450	500
Orange	0	150	300	300
Other	100	150	50	200
Total	**500**	**650**	**800**	**1000**

We can see that

- 50% of the fruits are bananas
- 30% are oranges
- 20% are other fruits

We can also say

- From 500 bananas 400 (0.8) are Long, 350 (0.7) are Sweet and 450 (0.9) are Yellow
- Out of 300 oranges 0 are Long, 150 (0.5) are Sweet and 300 (1) are Yellow
- From the remaining 200 fruits, 100 (0.5) are Long, 150 (0.75) are Sweet and 50 (0.25) are Yellow

https://www.datasciencecentral.com/profiles/blogs/naive-bayes-for-dummies-a-simple-explanation

Banana:

$$P(Banana|Long, Sweet, Yellow)$$
$$= \frac{P(Long|Banana) \cdot P(Sweet|Banana) \cdot P(Yellow|Banana) \cdot P(Banana)}{P(Long) \cdot P(Sweet) \cdot P(Yellow)}$$
$$= \frac{0.8 \times 0.7 \times 0.9 \times 0.5}{P(evidence)}$$
$$= \frac{0.252}{P(evidence)}$$

Orange:

$$P(Orange|Long, Sweet, Yellow) = 0$$

Other Fruit:

$$P(Other|Long, Sweet, Yellow)$$
$$= \frac{P(Long|Other) \cdot P(Sweet|Other) \cdot P(Yellow|Other) \cdot P(Other)}{P(Long) \cdot P(Sweet) \cdot P(Yellow)}$$
$$= \frac{0.5 \times 0.75 \times 0.25 \times 0.2}{P(evidence)}$$
$$= \frac{0.01875}{P(evidence)}$$

Naive Bayes use in Natural Language Processing

OSLOMET

- Task:

Determine if the text “I feel terrible” is either mood category or not mood.

TEXT	CATEGORY
I am so angry.	mood
I feel like superstar.	mood
It is going to rain.	not mood
I want to cry.	mood
They will come together.	not mood

- Naive bayes will count 2 occurances for mood category and 0 for not mood category.

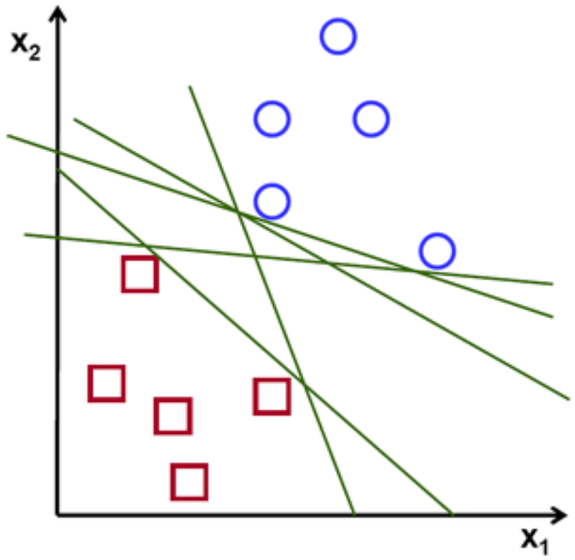
- Hence classifying the text as mood

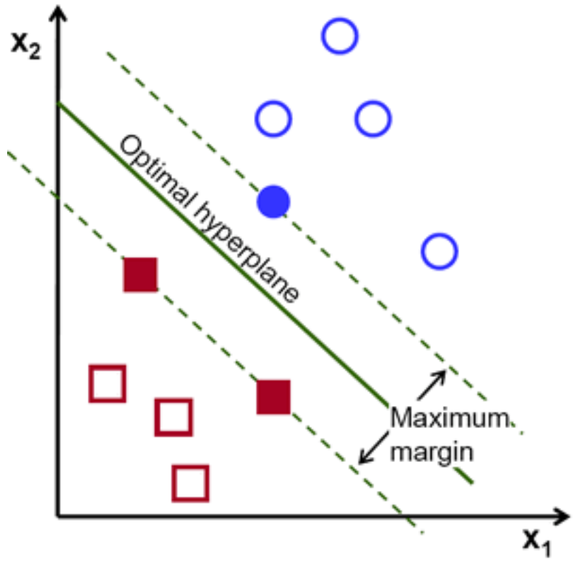
I	mood
feel	mood
so	
terrible	

https://medium.com/@montjoile/easy-and-quick-explanation-naive-bayes-algorithm-99cb5f3f4e9c

3. Support vector machines (classification)

- Supprt vector machines classifies the data by finding a clear seperation between the data points
  - It looks for a hyperplane
- Can be used both in classification and regression



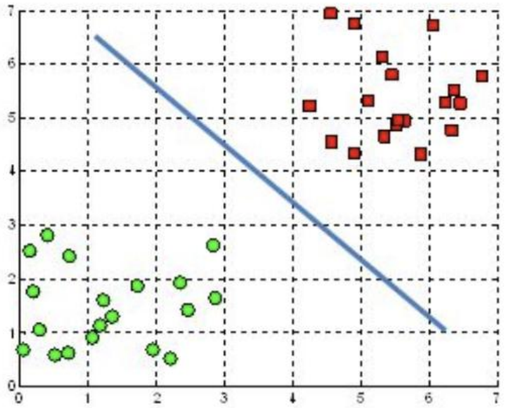


https://towardsdatascience.com/support-vector-machine-introduction-to-machine-learning-algorithms-934a444fca47

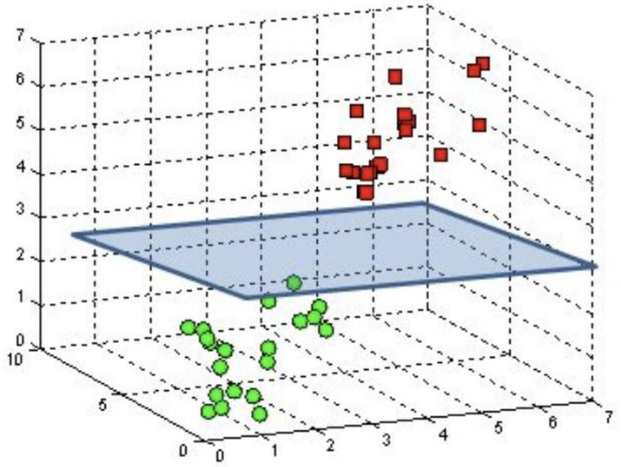
Hyperplanes

- Hyperplanes are boundries that help classify the data points
- We can use a 2 dimensional hyperplane (if number of features is 2)
- or a 3 dimensional hyperplane (if number of features is 3)
- We can also have n dimensional hyperplane (for more features)

A hyperplane in $\mathbb{R}^2$ is a line



A hyperplane in $\mathbb{R}^3$ is a plane



https://towardsdatascience.com/support-vector-machine-introduction-to-machine-learning-algorithms-934a444fca47

Uses of support vector machines

- Used to detect cancerous cells
- Used to predict driving routes
- Face detection
- Image classification
- Handwriting detection

Pros and cons

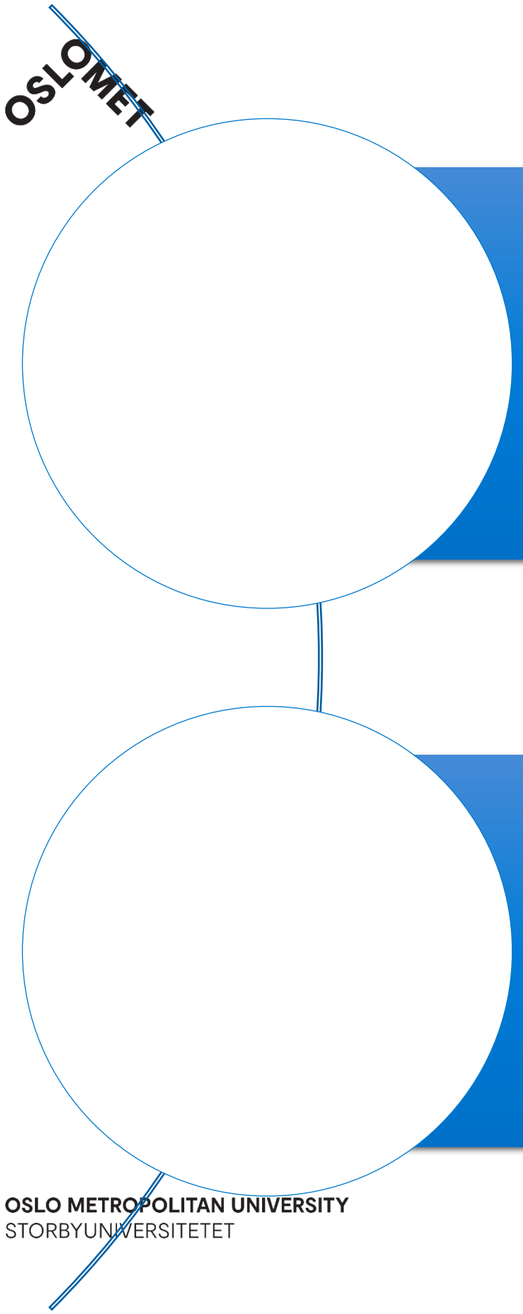
- Pros
  - Effective on data sets with multiple features
  - Effective in cases where number of features is greater than the data points
  - Memory efficient

- Cons
  - They donot provide probability estimates. Those are calculated using an expensive five fold cross vsalidation
  - Works best on small sample sets because of its high training time

Other classification algorithms

- Decision trees
- K nearest neigbour
- Random Forest
- ..

Types of supervised learning



Classification

- Support vector machines
- Decision Trees
- K-Nearest Neighbour
- Random Forest
- Logistic regression
- Naïve Bayes classifier
- …

Regression

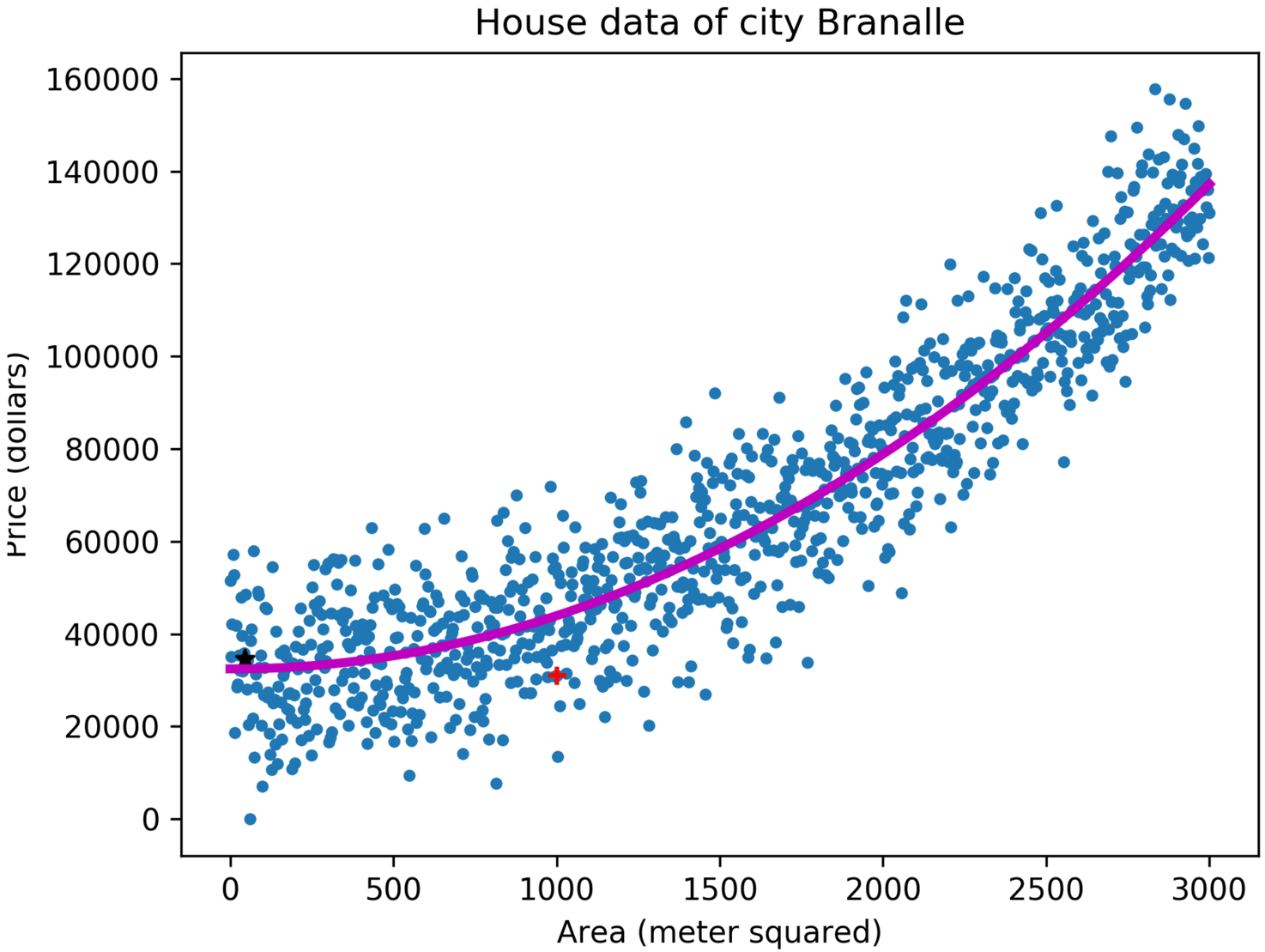
- Linear Regression
- Polynomial Regression
- …

Regression

- Regression models are used to predict a continuous value
- Examples
  - Predicting prices of a house given a set of features (size, price, location etc)
  - Predicting sales revenue of a company based on previous sales figures

- Question: Are the following use cases regression or classification ?
  - Customer Churn: when a customer leaves the company
  - Stock market price prediction
  - Spam email
  - Prediction of price of an oil
  - Salary prediction
  - Age prediction
  - Gender prediction

- Churn: Classification
- Stock market price prediction: Classification
- Spam email: classification
- Prediction of price of an oil: Regression
- Salary prediction: Regression
- Age prediction: Regression
- Gender prediction: Classification



https://www.quora.com/What-is-regression-in-machine-learning

- Regression is used for prediction, forecasting, time series modeling, and determining the causal effect between variables

Regression Algorithm	Classification Algorithm
In Regression, the output variable must be of continuous nature or real value.	In Classification, the output variable must be a discrete value.
The task of the regression algorithm is to map the input value (x) with the continuous output variable(y).	The task of the classification algorithm is to map the input value(x) with the discrete output variable(y).
Regression Algorithms are used with continuous data.	Classification Algorithms are used with discrete data.
In Regression, we try to find the best fit line, which can predict the output more accurately.	In Classification, we try to find the decision boundary, which can divide the dataset into different classes.
Regression algorithms can be used to solve the regression problems such as Weather Prediction, House price prediction, etc.	Classification Algorithms can be used to solve classification problems such as Identification of spam emails, Speech Recognition, Identification of cancer cells, etc.
The regression Algorithm can be further divided into Linear and Non-linear Regression.	The Classification algorithms can be divided into Binary Classifier and Multi-class Classifier.

HTTPS://WWW.JAVATPOINT.COM/REGRESSION-VS-CLASSIFICATION-IN-MACHINE-LEARNING

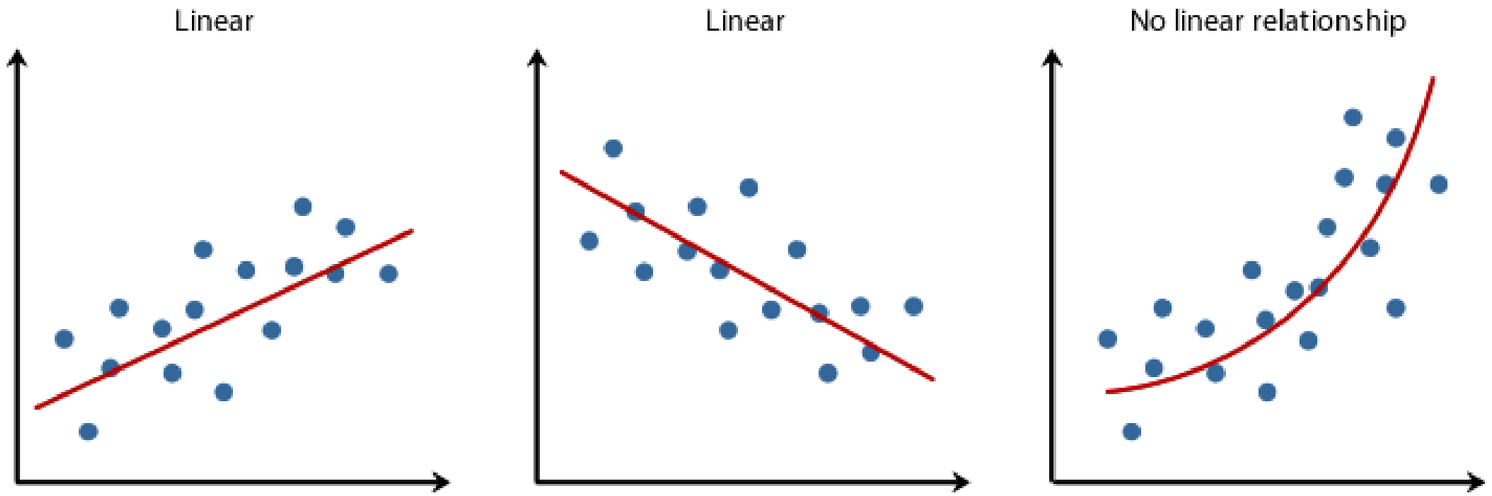
Types of regression

- Linear regression
- Logistic regression
- Polynomial regression
- Support vector regression
- Decision tree regression
- Random forest regression
- Ridge regression
- Lasso regression

1. Linear Regression

- It’s a linear model
  - an equation that describes a relationship between two quantities that show a constant rate of change.
  - e.g the older I get, the wiser I will be (hopefully)
  - Attending all lectures in Intro to A.I course will result in passing the exam

- There is always an input variable (x) and an output variable (y)
  - Y can be calculated from a linear combination of input variables (x)



HTTPS://TOWARDSDATASCIENCE.COM/LINEAR-REGRESSION-AND-ITS-MATHEMATICAL-IMPLEMENTATION-29D520A75EDE

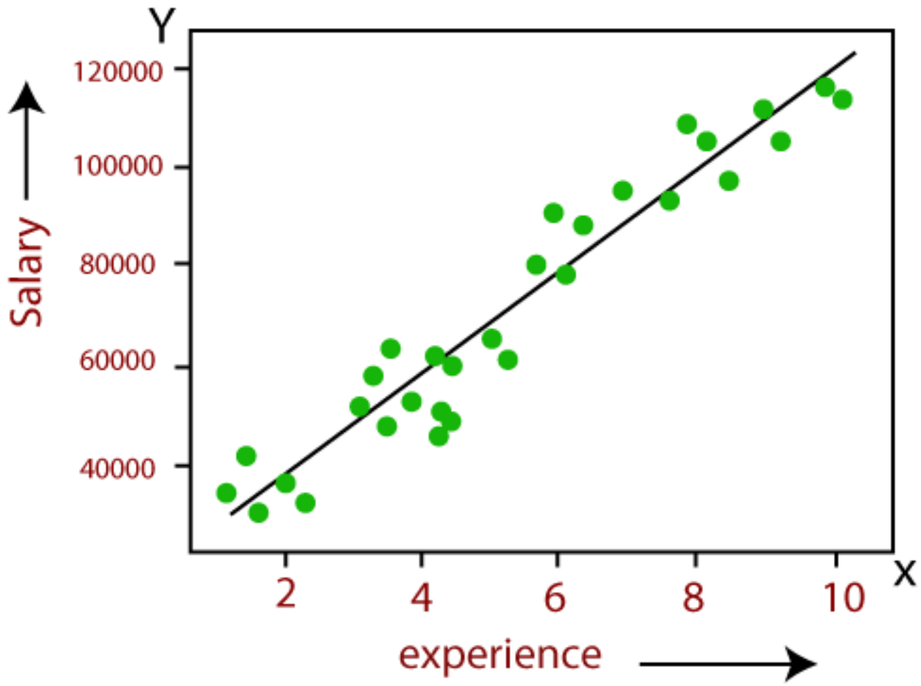
$$Y = a+bX$$

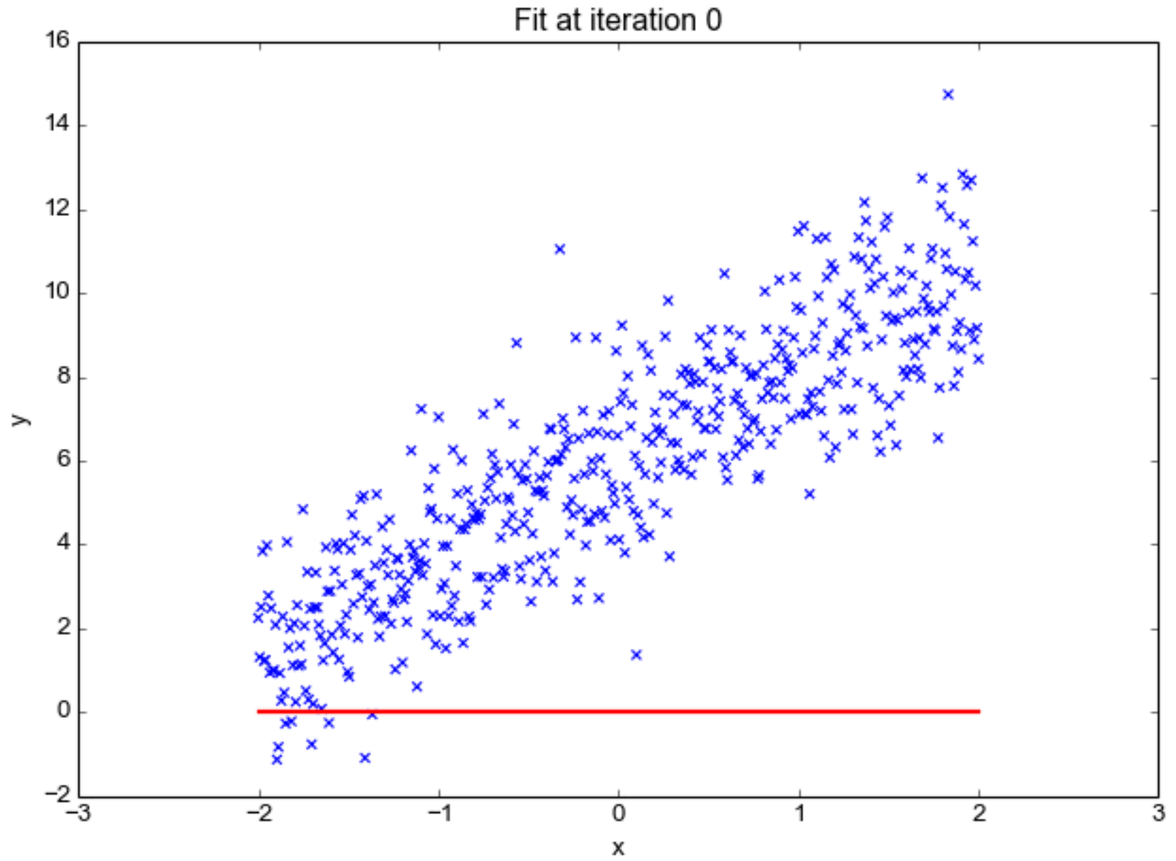
Y = Salary

X = Employees age

A,B: Coefficients of the equation

Coefficients are estimates of the unknown. These are calculated for a regression model and help us predict value of Y for each value of X



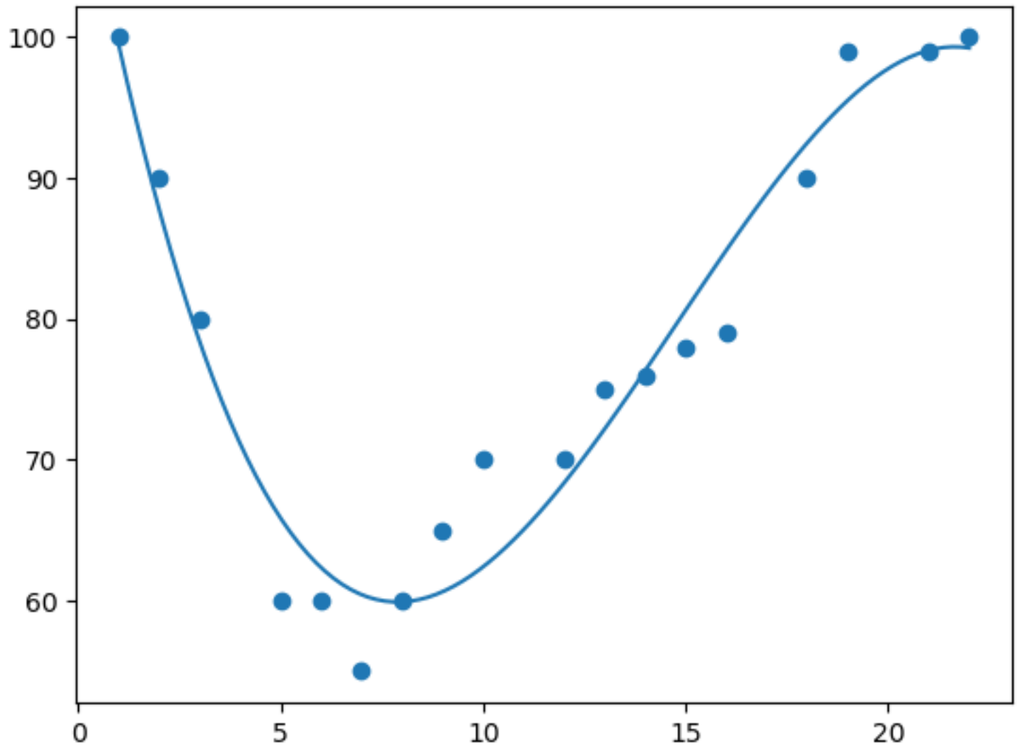

Fit at iteration 0
16
14
12
10
8
6
4
2
0
−2
y
−3
−2
−1
0
1
2
3
x

- Simple Linear regression
  - When there is a single input variable (x)
- Multiple linear regression
  - Multiple input variables

Ref: https://sphweb.bumc.bu.edu/otlt/MPH-Modules/BS/BS704-EP713_MultivariableMethods/

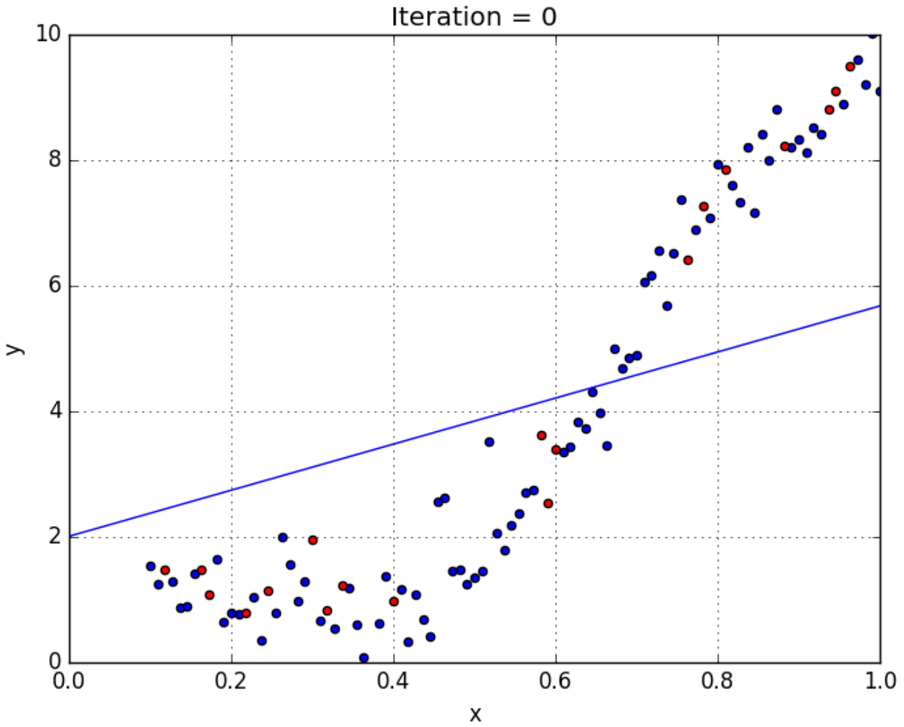
1. Polynomial Regression

- It is also a linear model
- But it is never a striaght line

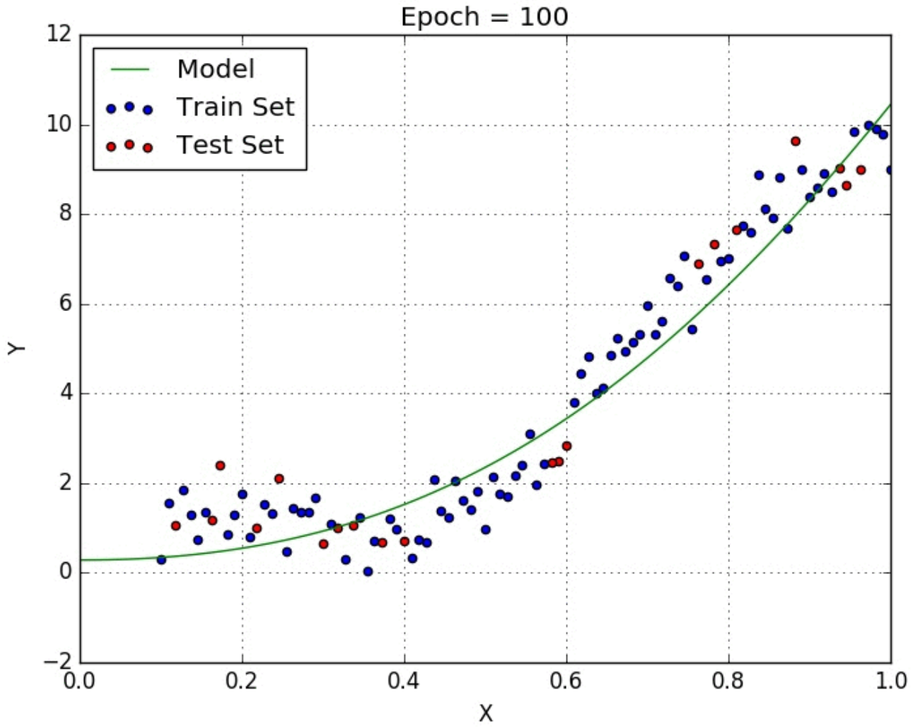


https://www.w3schools.com/python/python_ml_polynomial_regression.asp

Linear regression



Polynomial regression



https://towardsdatascience.com/introduction-to-linear-regression-and-polynomial-regression-f8adc96f31cb