Bayesian Classification

Tushar B. Kute,
http://tusharkute.com

What is Probability?

- **Probability** is a measure of the likelihood of a random phenomenon or chance behavior.
- Probability describes the long-term proportion with which a certain **outcome** will occur in situations with short-term uncertainty.
- Example:
  - Simulate flipping a coin 100 times. Plot the proportion of heads against the number of flips. Repeat the simulation.

Probability

- Probability deals with experiments that yield random short-term results or outcomes, yet reveal long-term predictability.
- The long-term proportion with which a certain outcome is observed is the probability of that outcome.

Law of large numbers

- As the number of repetitions of a probability experiment increases, the proportion with which a certain outcome is observed gets closer to the probability of the outcome.

Probability and event

- In probability, an **experiment** is any process that can be repeated in which the results are uncertain.
- A **simple event** is any single outcome from a probability experiment. Each simple event is denoted e_i.
- The **sample space, *S*,** of a probability experiment is the collection of all possible simple events. In other words, the sample space is a list of all possible outcomes of a probability experiment.

The event

- An **event** is any collection of outcomes from a probability experiment.
- An event may consist of one or more simple events.
- Events are denoted using capital letters such as *E*.

Example:

- Consider the probability experiment of having two children.

- (a) Identify the simple events of the probability experiment.
- (b) Determine the sample space.
- (c) Define the event E = “have one boy”.

Denoting probability

- The **probability of an event**, denoted *P*(*E*), is the likelihood of that event occurring.

Properties of probabilities

- The probability of any event E, $P(E)$, must be between 0 and 1 inclusive. That is,

  $0 \le P(E) \le 1.$

- If an event is **impossible**, the probability of the event is 0.
- If an event is a **certainty,** the probability of the event is 1.
- If $S = \{e_1, e_2, \ldots, e_n\}$, then

  $P(e_1) + P(e_2) + \ldots + P(e_n) = 1.$

Unusual Event

- An **unusual event** is an event that has a low probability of occurring.

Method of probability

- Three methods for determining the probability of an event:

  (1) the classical method

  (2) the empirical method

  (3) the subjective method

Dependence and Independence

- Roughly speaking, we say that two events E and F are dependent if knowing something about whether E happens gives us information about whether F happens (and vice versa). Otherwise they are independent.
- For instance, if we flip a fair coin twice, knowing whether the first flip is Heads gives us no information about whether the second flip is Heads. These events are independent. On the other hand, knowing whether the first flip is Heads certainly gives us information about whether both flips are Tails. (If the first flip is Heads, then definitely it’s not the case that both flips are Tails.) These two events are dependent.

Dependence and Independence

- Mathematically, we say that two events E and F are independent if the probability that they both happen is the product of the probabilities that each one happens:

$$P(E, F) = P(E)P(F)$$

- In the example above, the probability of “first flip Heads” is 1/2, and the probability of “both flips Tails” is 1/4, but the probability of “first flip Heads and both flips Tails” is 0.

Conditional Probability

- When two events E and F are independent, then by definition we have:

$$P(E, F) = P(E)P(F)$$

- If they are not necessarily independent (and if the probability of F is not zero), then we define the probability of E "conditional on F" as:

$$P(E \mid F) = P(E, F) / P(F)$$

Conditional Probability

- You should think of this as the probability that E happens, given that we know that F happens.
- We often rewrite this as:

$$P(E, F) = P(E \mid F)P(F)$$

- When E and F are independent, you can check that this gives:

$$P(E \mid F) = P(E)$$

Example:

- One common tricky example involves a family with two (unknown) children.
- If we assume that:

  1. Each child is equally likely to be a boy or a girl
  2. The gender of the second child is independent of the gender of the first child

  then the event “no girls” has probability 1/4, the event “one girl, one boy” has probability 1/2, and the event “two girls” has probability 1/4.

Example:

- Now we can ask what is the probability of the event “both children are girls” (B) conditional on the event “the older child is a girl” (G)? Using the definition of conditional probability:

$$P(B \mid G) = P(B, G) / P(G) = P(B) / P(G) = 1/2$$

- since the event B and G (“both children are girls and the older child is a girl”) is just the event B. (Once you know that both children are girls, it’s necessarily true that the older child is a girl.)

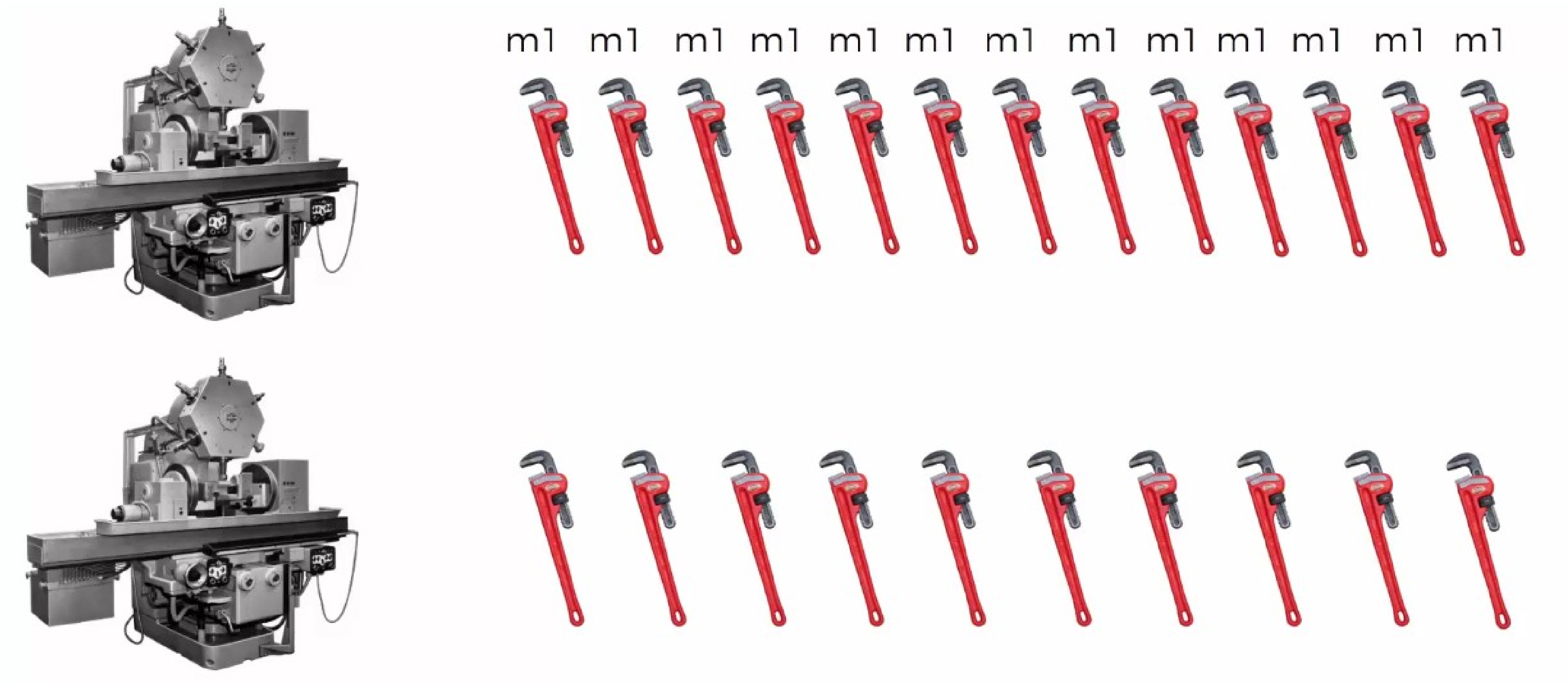
Example:

- We could also ask about the probability of the event "both children are girls" conditional on the event "at least one of the children is a girl" (L). Surprisingly, the answer is different from before!
- As before, the event B and L ("both children are girls and at least one of the children is a girl") is just the event B. This means we have:

$$P(B \mid L) = P(B, L) / P(L) = P(B) / P(L) = 1/3$$

- How can this be the case? Well, if all you know is that at least one of the children is a girl, then it is twice as likely that the family has one boy and one girl than that it has both girls.

Bayes Theorem



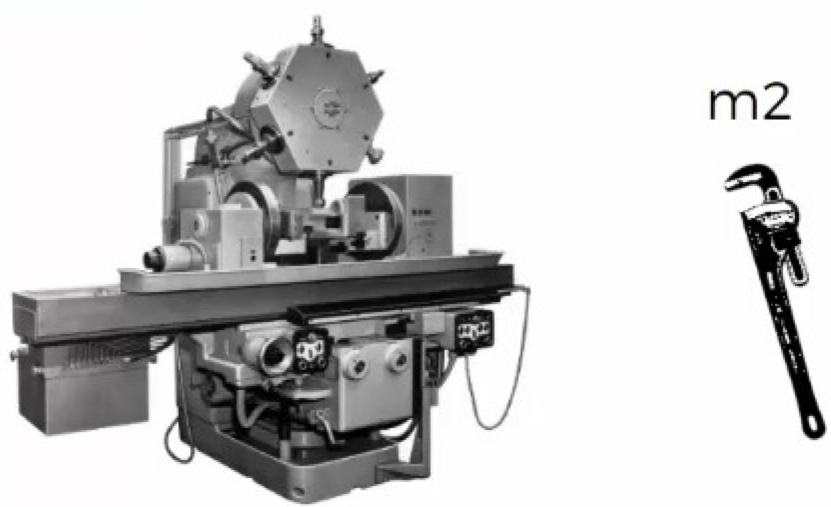
Example Reference: Super Data Science

Bayes Theorem

Defective Spanners

Bayes Theorem

What's the probability?



Bayes Theorem

$$P(A|B) = \frac{P(B|A) * P(A)}{P(B)}$$

Bayes Theorem

Mach1: 30 wrenches / hr -> P(Mach1) = 30/50 = 0.6
Mach2: 20 wrenches / hr -> P(Mach2) = 20/50 = 0.4

Out of all produced parts:
We can SEE that 1% are defective -> P(Defect) = 1%

Out of all defective parts:
We can SEE that 50% came from mach1 -> P(Mach1 | Defect) = 50%
And 50% came from mach2 -> P(Mach2 | Defect) = 50%

Question:
What is the probability that a part produced by mach2 is defective = ? -> P(Defect | Mach2) = ?

Bayes Theorem

$$P(Defect \mid Mach2) = \frac{P(Mach2 \mid Defect) * P(Defect)}{P(Mach2)}$$

$$P(Defect \mid Mach2) = \frac{0.5 * 0.01}{0.4} = 0.0125$$

That's intuitive

$$P(Defect \mid Mach2) = \frac{P(Mach2 \mid Defect) * P(Defect)}{P(Mach2)} = 1.25\%$$

Let's look at an example:

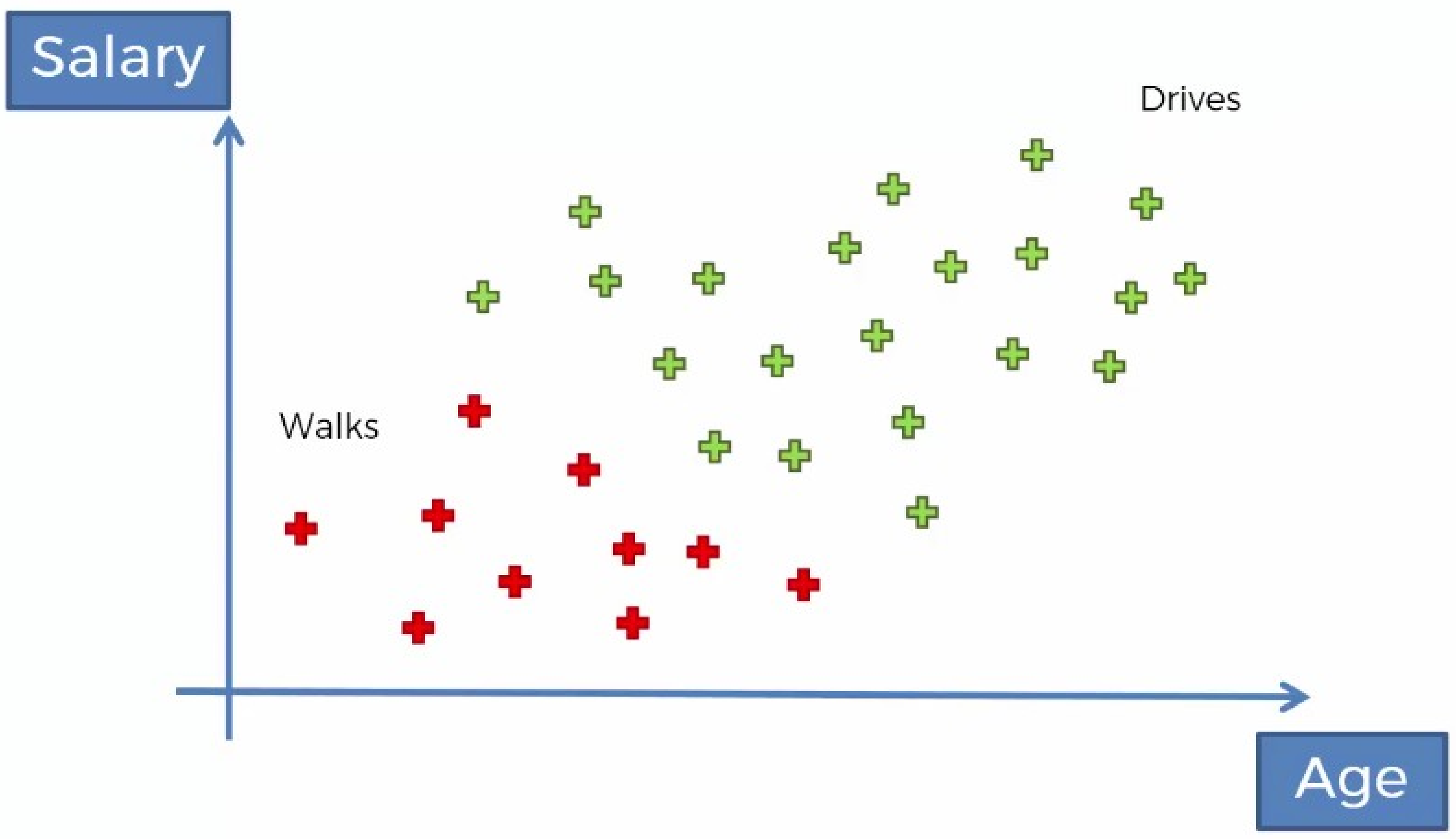
- 1000 wrenches
- 400 came from Mach2
- 1% have a defect = 10
- of them 50% came from Mach2 = 5
- % defective parts from Mach2 = 5/400 = 1.25%

Exercise

Quick exercise:

P(Defect | Mach1) = ?

Example:

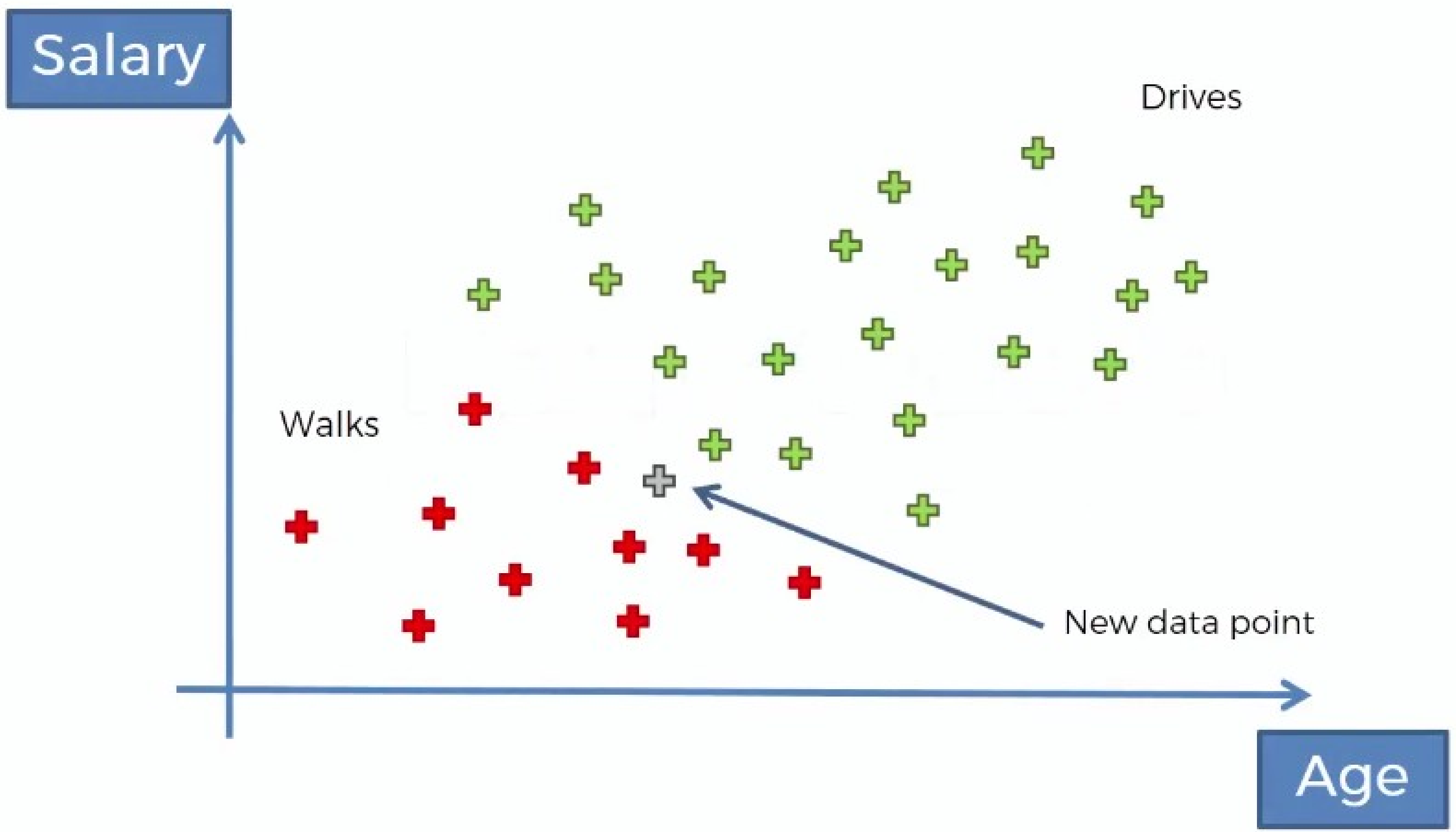


Step-1

$$P(A|B) = \frac{P(B|A) * P(A)}{P(B)}$$

$$P(Walks|X) = \frac{P(X|Walks) * P(Walks)}{P(X)}$$

Step-1



Step-1

#4 Posterior Probability

#3 Likelihood

#1 Prior Probability

$$P(Walks|X) = \frac{P(X|Walks) * P(Walks)}{P(X)}$$

#2 Marginal Likelihood

Step-2

#4 Posterior Probability

#3 Likelihood

#1 Prior Probability

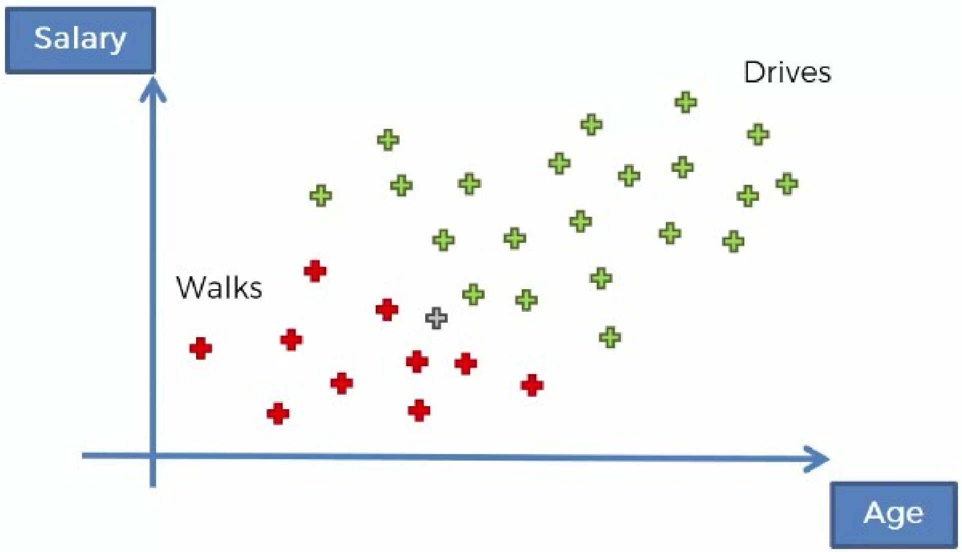
#2 Marginal Likelihood

$$P(Drives|X) = \frac{P(X|Drives) * P(Drives)}{P(X)}$$

Step-3

$$P(Walks|X)\ v.s.\ P(Drives|X)$$

Naive Bayes – Step-1

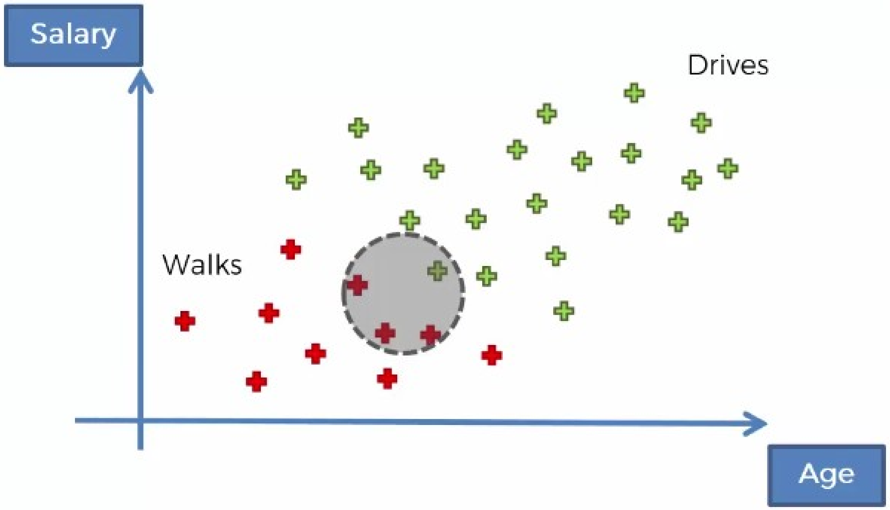


#1. P(Walks)

$$P(Walks) = \frac{Number\ of\ Walkers}{Total\ Observations}$$

$$P(Walks) = \frac{10}{30}$$

Naive Bayes – Step-2

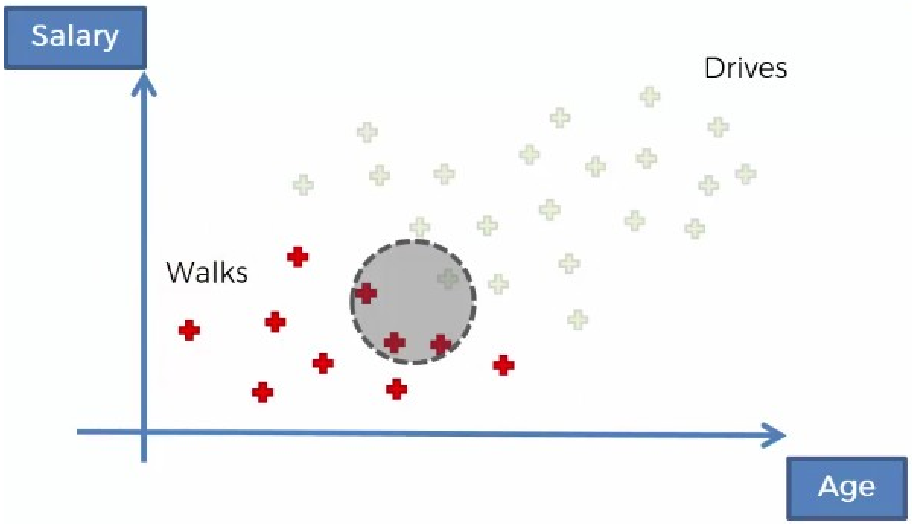


#2. P(X)

$$P(X) = \frac{Number\ of\ Similar\ Observations}{Total\ Observations}$$

$$P(X) = \frac{4}{30}$$

Naive Bayes – Step-3



#3. P(X|Walks)

$$P(X|Walks) = \frac{\text{Number of Similar Observations Among those who Walk}}{\text{Total number of Walkers}}$$

$$P(X|Walks) = \frac{3}{10}$$

Combining altogether

#4 Posterior Probability

#3 Likelihood

#1 Prior Probability

$$P(Walks|X) = \frac{\frac{3}{10} * \frac{10}{30}}{\frac{4}{30}} = 0.75$$

#2 Marginal Likelihood

Naive Bayes – Step-4

#4 Posterior Probability

#3 Likelihood

#1 Prior Probability

$$P(Drives|X) = \frac{P(X|Drives) * P(Drives)}{P(X)}$$

#2 Marginal Likelihood

Naive Bayes – Step-5

#4 Posterior Probability

#3 Likelihood

#1 Prior Probability

$$P(Drives|X) = \frac{\frac{1}{20} * \frac{20}{30}}{\frac{4}{30}} = 0.25$$

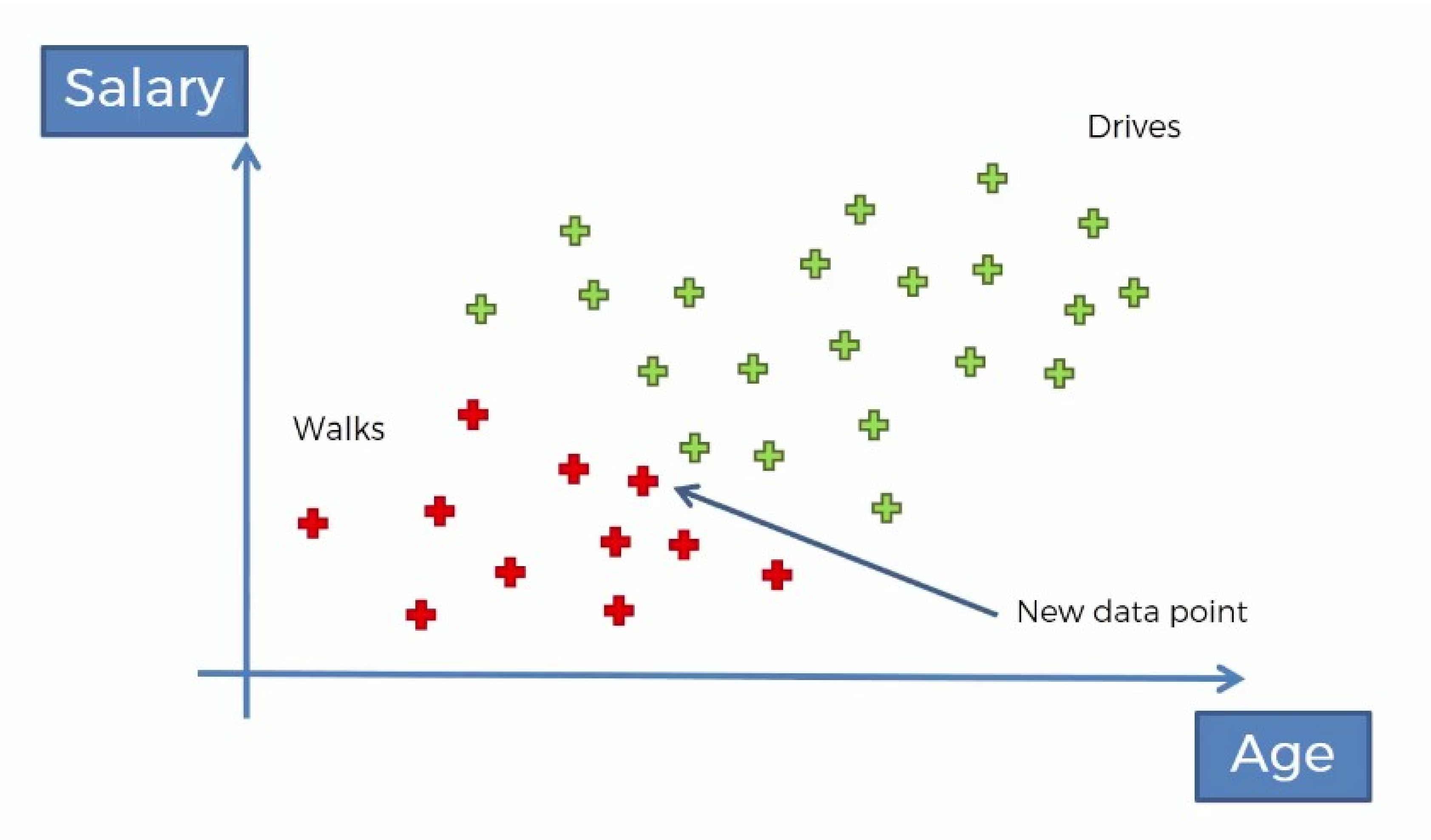
#2 Marginal Likelihood

Types of model

$$P(Walks|X) \; v.s. \; P(Drives|X)$$

$$0.75 \; v.s. \; 0.25$$

Final Classification



Thank you

@mitu_skillologies

/mITuSkillologies

@mitu_group

/company/mitu-skillologies

MITUSkillologies

Web Resources
https://mitu.co.in
http://tusharkute.com

contact@mitu.co.in

tushar@tusharkute.com