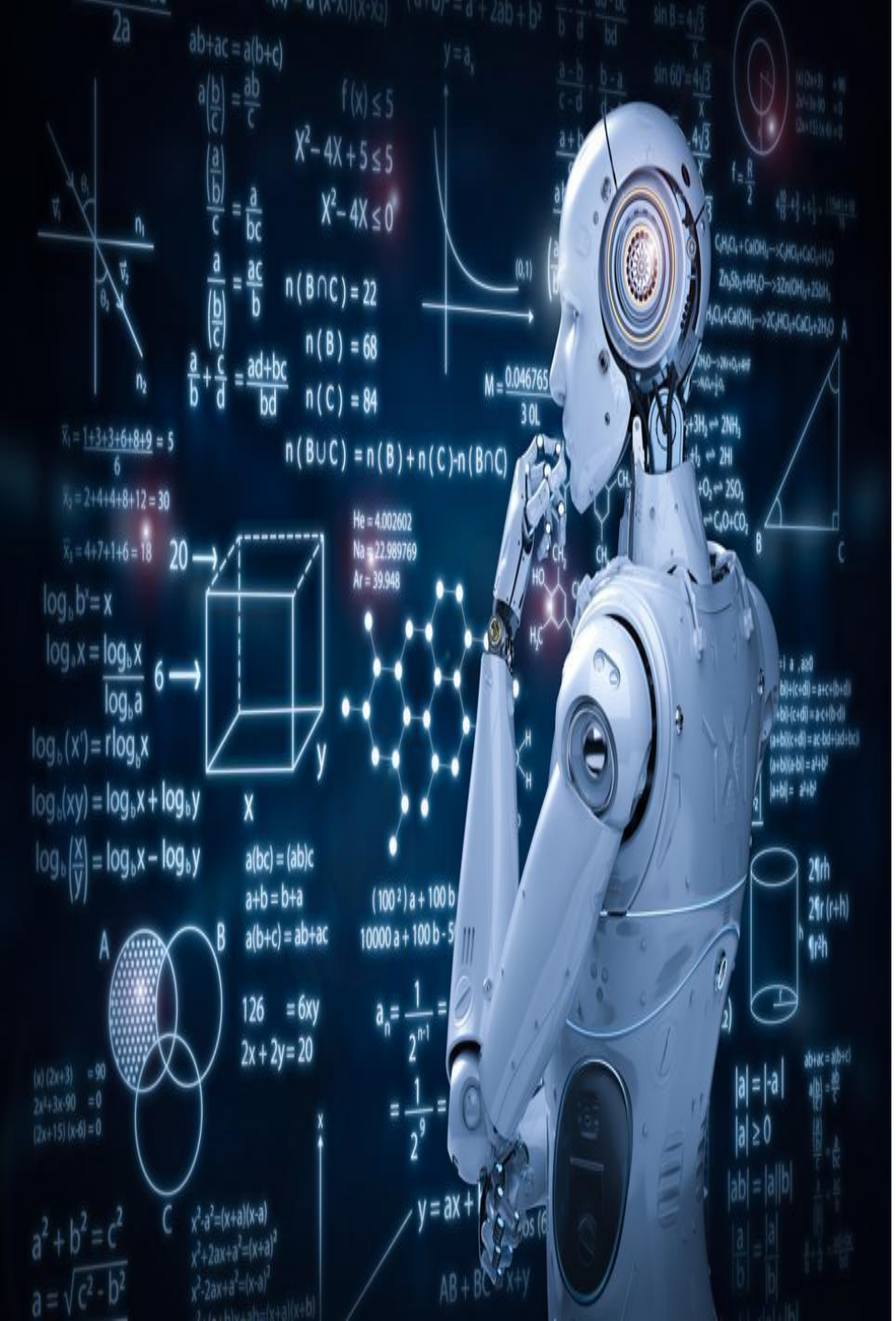


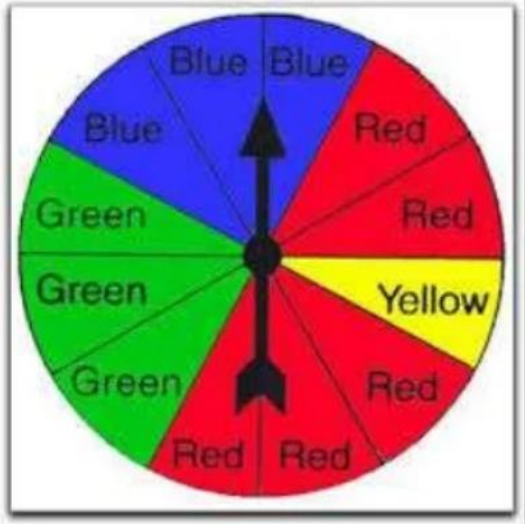
Practical Machine Learning

Day 10: Sep22 DBDA

Kiran Waghmare

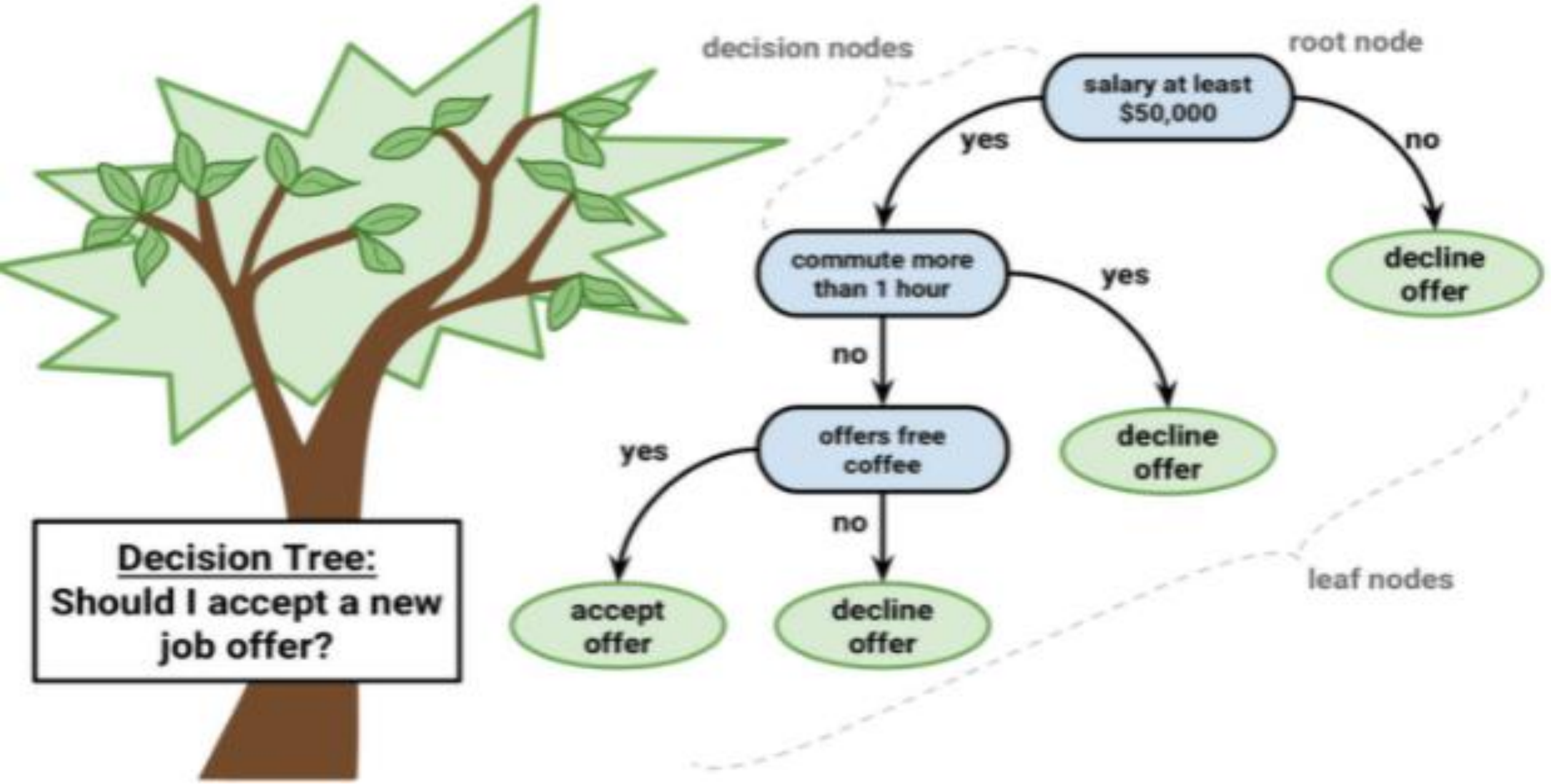
Agenda

- Naïve Bayes
- Decision Tree



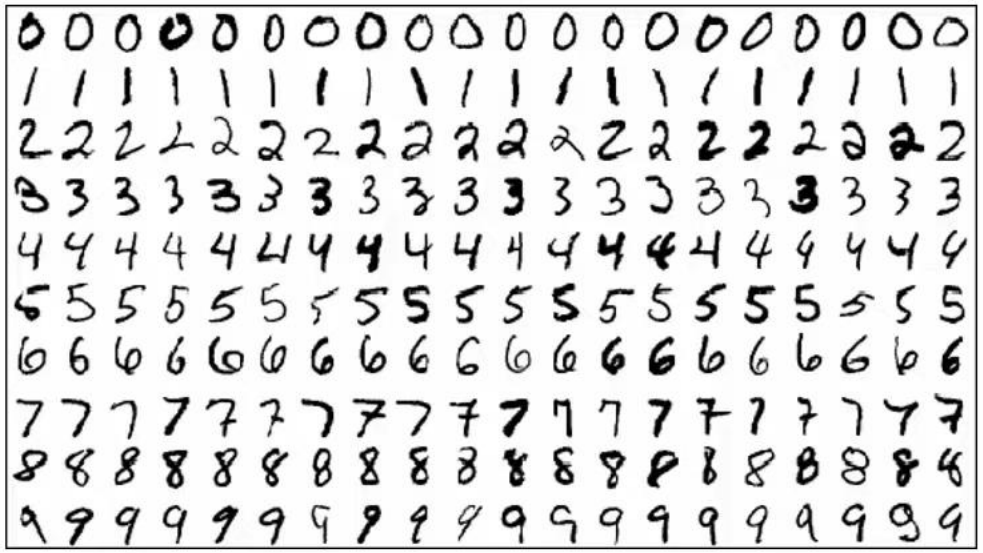
PROBABILITY

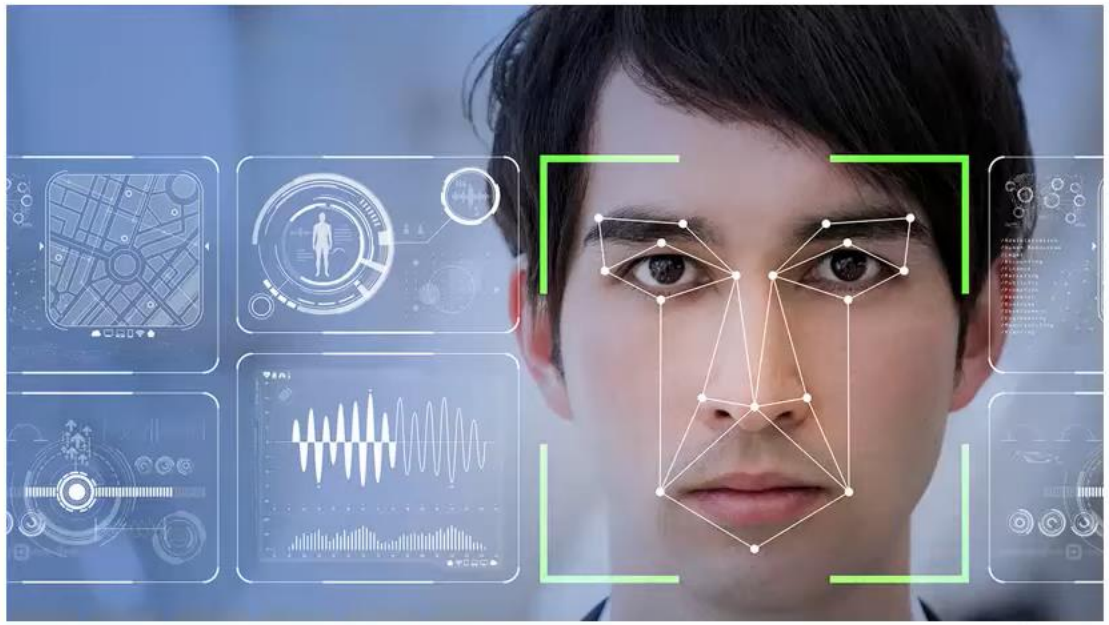
GETTING KNOWLEDGE READY


Decision Tree:
Should I accept a new job offer?
decision nodes
root node
salary at least $50,000
yes
no
commute more than 1 hour
yes
decline offer
no
offers free coffee
yes
decline offer
yes
no
accept offer
decline offer
leaf nodes

Thomas Bayes

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





Ads in 2

Examples of Classification in Data Analytics

- Life Science: Predicting tumor cells as benign or malignant
- Security: Classifying credit card transactions as legitimate or fraudulent
- Prediction: Weather, voting, political dynamics, etc.
- Entertainment: Categorizing news stories as finance, weather, entertainment, sports, etc.
- Social media: Identifying the current trend and future growth

Classification Techniques

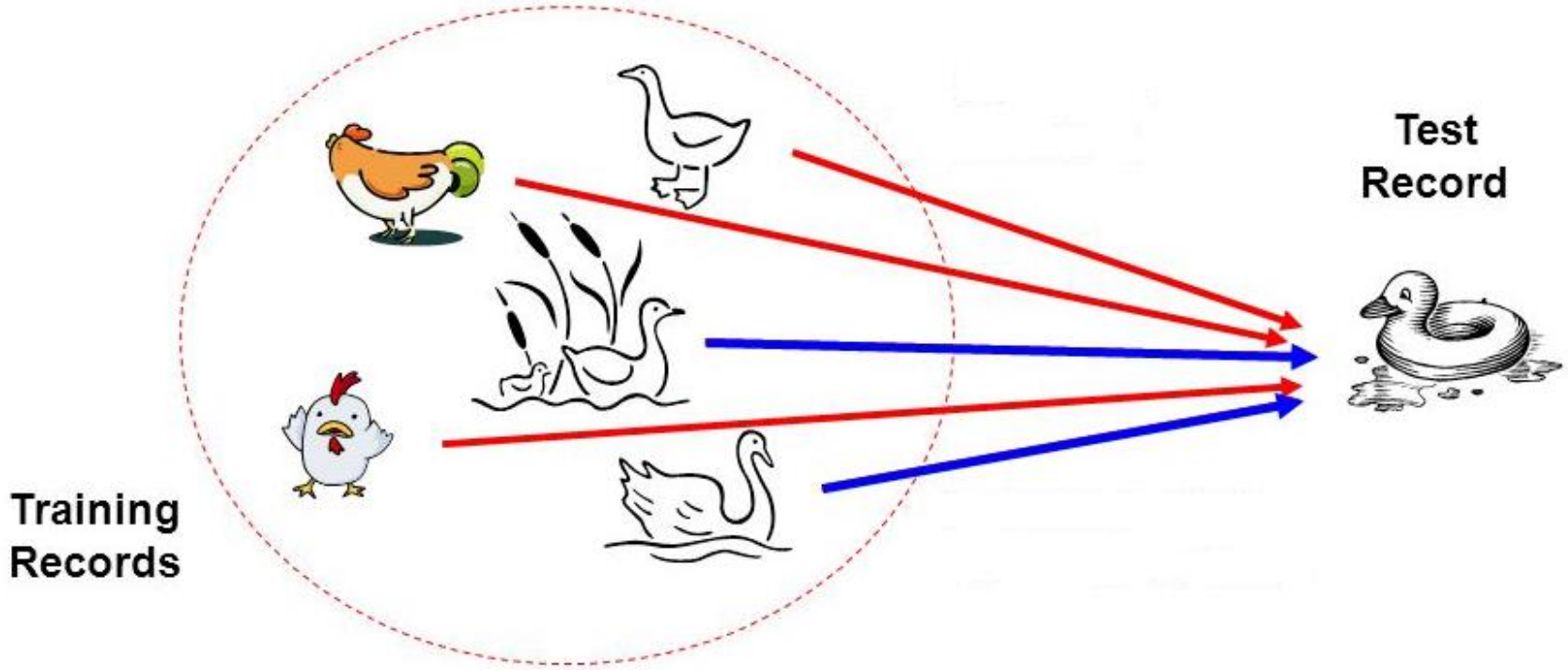
- A number of classification techniques are known, which can be broadly classified into the following categories:

1. Statistical-Based Methods
   - Regression
   - Bayesian Classifier
2. Distance-Based Classification
   - K-Nearest Neighbours
3. Decision Tree-Based Classification
   - ID3, C 4.5, CART
5. Classification using Machine Learning (SVM)
6. Classification using Neural Network (ANN)

Bayesian Classifier

Bayesian Classifier

- Principle
  - If it walks like a duck, quacks like a duck, then it is probably a duck



Bayesian Classifier

- **A statistical classifier**
  - Performs *probabilistic prediction*, *i.e.,* predicts class membership probabilities
- **Foundation**
  - Based on Bayes' Theorem.
- **Assumptions**
  1. The classes are mutually exclusive and exhaustive.
  2. The attributes are independent given the class.
- **Called "Naïve" classifier because of these assumptions.**
  - Empirically proven to be useful.
  - Scales very well.

Background

- There are three methods to establish a classifier

a) Model a classification rule directly

Examples: k-NN, decision trees, perceptron, SVM

b) Model the probability of class memberships given input data

Example: multi-layered perceptron with the cross-entropy cost

c) Make a probabilistic model of data within each class

Examples: naive Bayes, model based classifiers

- *a*) and *b*) are examples of discriminative classification
- *c*) is an example of generative classification
- *b*) and *c*) are both examples of probabilistic classification

Probability Basics

- Prior, conditional and joint probability
  - Prior probability: $P(X)$
  - Conditional probability: $P(X_1 \mid X_2), P(X_2 \mid X_1)$
  - Joint probability: $\mathbf{X} = (X_1, X_2), P(\mathbf{X}) = P(X_1, X_2)$
  - Relationship: $P(X_1, X_2) = P(X_2 \mid X_1)P(X_1) = P(X_1 \mid X_2)P(X_2)$
  - Independence: $P(X_2 \mid X_1) = P(X_2), P(X_1 \mid X_2) = P(X_1), P(X_1, X_2) = P(X_1)P(X_2)$
- Bayesian Rule

$$P(C \mid \mathbf{X}) = \frac{P(\mathbf{X} \mid C)P(C)}{P(\mathbf{X})} \qquad Posterior = \frac{Likelihood \times Prior}{Evidence}$$

BAYES THEOREM

- Bayes theorem is the cornerstone of Bayesian learning methods because it provides a way to calculate the posterior probability **P(h|D), from**
- **the prior** probability **P(h),**
- **Probability over the data set P(D) and**
- **Current probability P(D(h)**

$$P(h|D) = \frac{P(D|h)p(h)}{P(D)}$$

Maximum A Posteriori (MAP) Hypothesis

- The learner considers some set of candidate hypotheses H and is interested in finding the most probable hypothesis **h ϵ H** given the observed data **D (or at least one of the maximally probable if there** are several).
- Any such maximally probable hypothesis is called a **maximum a posteriori (MAP) hypothesis.**
- **We can determine the MAP hypotheses by using** Bayes theorem to calculate the posterior probability of each candidate hypothesis.

Maximum A Posteriori (MAP) Hypothesis

- More precisely, we will say that h_{MAP} **is a MAP hypothesis provided**

$$h_{MAP} \equiv \underset{h \in H}{\operatorname{argmax}} P(h|D)$$

$$= \underset{h \in H}{\operatorname{argmax}} \frac{P(D|h)P(h)}{P(D)}$$

$$= \underset{h \in H}{\operatorname{argmax}} P(D|h)P(h)$$

Bayesian Classifier

- In many applications, the relationship between the attributes set and the class variable is non-deterministic.
  - In other words, a test cannot be classified to a class label with certainty.
  - In such a situation, the classification can be achieved probabilistically.

- The Bayesian classifier is an approach for modelling probabilistic relationships between the attribute set and the class variable.
- More precisely, Bayesian classifier use Bayes' Theorem of Probability for classification.
- Before going to discuss the Bayesian classifier, we should have a quick look at the Theory of Probability and then Bayes' Theorem.

NAIVE BAYES CLASSIFIER – Example -1

Day	*Outlook*	*Temperature*	*Humidity*	*Wind*	*PlayTennis*
D1	Sunny	Hot	High	Weak	No
D2	Sunny	Hot	High	Strong	No
D3	Overcast	Hot	High	Weak	Yes
D4	Rain	Mild	High	Weak	Yes
D5	Rain	Cool	Normal	Weak	Yes
D6	Rain	Cool	Normal	Strong	No
D7	Overcast	Cool	Normal	Strong	Yes
D8	Sunny	Mild	High	Weak	No
D9	Sunny	Cool	Normal	Weak	Yes
D10	Rain	Mild	Normal	Weak	Yes
D11	Sunny	Mild	Normal	Strong	Yes
D12	Overcast	Mild	High	Strong	Yes
D13	Overcast	Hot	Normal	Weak	Yes
D14	Rain	Mild	High	Strong	No

$$\langle Outlook = sunny, Temperature = cool, Humidity = high, Wind = strong \rangle$$

A Practice Example

Example 8.4

Class:
C1:buys_computer = 'yes'
C2:buys_computer = 'no'

Data instance
X = (age <=30,
Income = medium,
Student = yes
Credit_rating = fair)

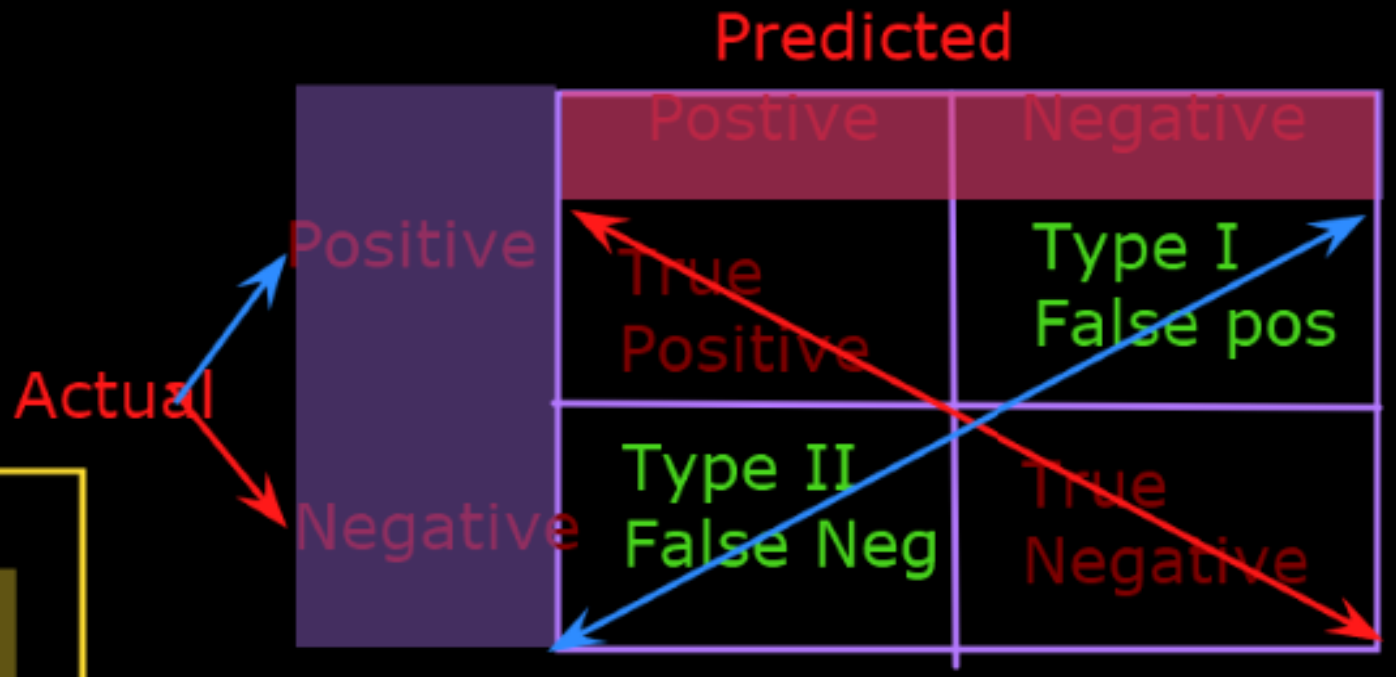
age	income	student	credit_rating	_comp
<=30	high	no	fair	no
<=30	high	no	excellent	no
31…40	high	no	fair	yes
>40	medium	no	fair	yes
>40	low	yes	fair	yes
>40	low	yes	excellent	no
31…40	low	yes	excellent	yes
<=30	medium	no	fair	no
<=30	low	yes	fair	yes
>40	medium	yes	fair	yes
<=30	medium	yes	excellent	yes
31…40	medium	no	excellent	yes
31…40	high	yes	fair	yes
>40	medium	no	excellent	no

==========
Day 10: Classification Algorithm
==========
Date: 12/01/2023
Topics:

-Classification Algorithm
-Naive Bayes
-Decision Tree

Accuracy
Sensitivity
Specificity

Precision
Recall=Sensitivity



```
==================================================
Day 10: Classification Algorithm
==================================================
Date: 12/01/2023
Topics:
------------------
    -Classification Algorithm
    -Naive Bayes
    -Decision Tree
```

	-ve	+ve
-ve	TN	FN
+ve	FP	TP

Accuracy
Sensitivity
Specificity

Precision
Recall=Sensitivity