Naive Bayes Algorithm

↓

Bayes's Theorem

↓

Conditional Probability

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

Probability of A given B

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

$$P(B|A) * P(A) = P(A \cap B)$$

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

$P(Y | x_1, x_2, x_3)$

$P(N | x_1, x_2, x_3)$

target

$x_1, x_2, x_3 \longrightarrow$ Y / N

Outlook	Temp	Humidity	Windy	Play Golf
Rainy	Hot	High	FALSE	No
Rainy	Hot	High	TRUE	No
Overcast	Hot	High	FALSE	Yes
Sunny	Mild	High	FALSE	Yes
Sunny	Cool	Normal	FALSE	Yes
Sunny	Cool	Normal	TRUE	No
Overcast	Cool	Normal	TRUE	Yes
Rainy	Mild	High	FALSE	No
Rainy	Cool	Normal	FALSE	Yes
Sunny	Mild	Normal	FALSE	Yes
Rainy	Mild	Normal	TRUE	Yes
Overcast	Mild	High	TRUE	Yes
Overcast	Hot	Normal	FALSE	Yes
Sunny	Mild	High	TRUE	No

target

Sunny Hot Normal True → Yes/No

$$P(\underset{A}{Yes} \mid \underset{B}{Sunny, Hot, Normal, True}) = \frac{P(Sunny \mid yes) * P(Hot \mid yes) * P(Normal \mid yes) * P(True \mid yes) * P(yes)}{P(Sunny) * P(Hot) * P(Normal) * P(True)}$$

(N): $3/9 * 6/9 * 3/9 * 9/14$

(D): $2/9 * 2/9 * 2/9 * 9/14$

final prediction → Yes ⟹ 0.0105 ✓

$$P(\underset{A}{No} \mid \underset{B}{Sunny, Hot, Normal, True}) = \frac{P(Sunny \mid No) * P(Hot \mid No) * P(Normal \mid No) * P(True \mid No) * P(No)}{P(Sunny) * P(Hot) * P(Normal) * P(True)}$$

(N): $2/5 * 2/5 * 1/5 * 3/5 * 5/14$ ⟹ 0.0068 ✓

(D): $2/9 * 2/9 * 2/9 * 9/14$

$\frac{44}{2} > \frac{22}{2} \rightarrow D$

xx

Outlook

	Yes	No	P(Yes)	P(No)
Sunny	3	2	3/9	2/5
overcast	4	0	4/9	0
Rainy	2	3	2/9	3/5
	9	5		

Outlook

	Yes	No	P(yes)	P(no)
Sunny	3	2	2/9	3/5
Overcast	4	0	4/9	0/5
Rainy	3	2	3/9	2/5
Total	9	5	100%	100%

Temperature

	Yes	No	P(yes)	P(no)
Hot	2	2	2/9	2/5
Mild	4	2	4/9	2/5
Cool	3	1	3/9	1/5
Total	9	5	100%	100%

Humidity

	Yes	No	P(yes)	P(no)
High	3	4	3/9	4/5
Normal	6	1	6/9	1/5
Total	9	5	100%	100%

Wind

	Yes	No	P(yes)	P(no)
False	6	2	6/9	2/5
True	3	3	3/9	3/5
Total	9	5	100%	100%

Play		P(Yes)/P(No)
Yes	9	9/14
No	5	5/14
Total	14	100%

→ Training Phase

Prediction → faster → it has to just pick the values from the table

Is Naive Bayes also called as Lazy Learner algorithm ??

No

→ Training done upfront, fast prediction

xx

Continuous data

Gaussian Naive Bayes

→ Name to suggest the probability of both the classes.

https://www.geeksforgeeks.org/machine-learning/gaussian-naive-bayes/ → (Reference Link)

Classification
- Spam detection
- Sentiment analysis
- Medical diagnosis

xx

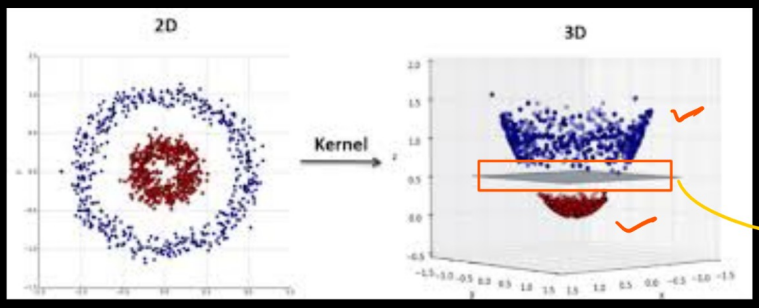
Advantages of Naive Bayes Classifier
Easy to implement and computationally efficient.
Effective in cases with a large number of features.
Performs well even with limited training data.
It performs well in the presence of categorical features.
For numerical features data is assumed to come from normal distributions

Disadvantages of Naive Bayes Classifier
Assumes that features are independent, which may not always hold in real-world data.
Can be influenced by irrelevant attributes.
May assign zero probability to unseen events, leading to poor generalization.

xx

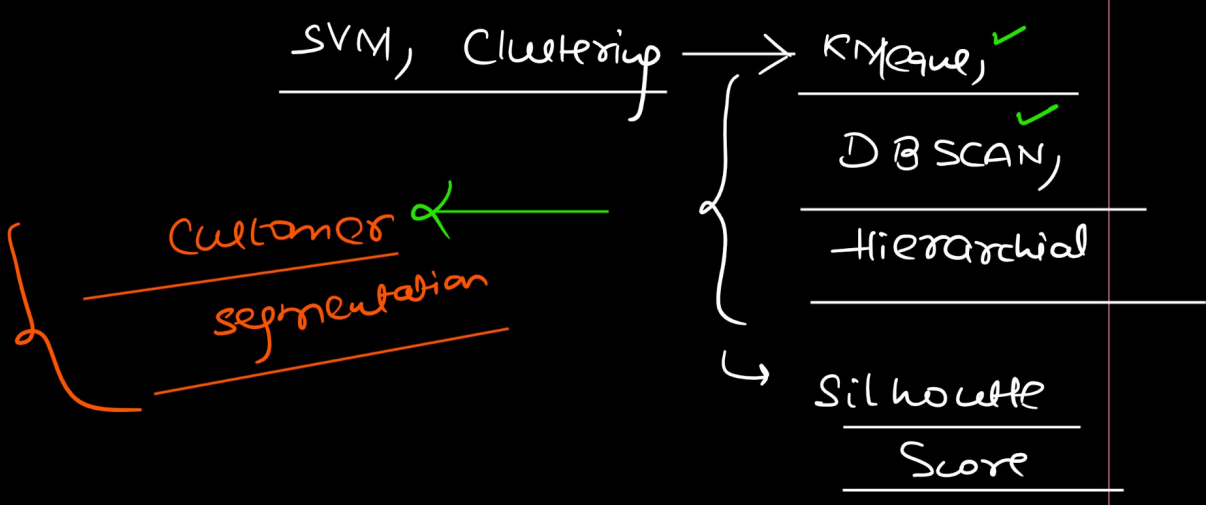
SVM (Support Vector Machine)
- SVC ← classification
- SVR ← Regression

Kernel Trick → RBF



Hyperplane

Non-linear data


SVM, Clustering
KMeans,
DBSCAN,
Hierarchial
Silhouette Score
Customer segmentation