Machine Learning

Naïve Bayes Classifier

Mathematical Classification

Class Objective

By the end of this class, students will be able to:

- Apply Bayes Theorem for classification
- Use the Naïve (conditional independence) assumption
- Compute posteriors **step-by-step** for a new instance
- Classify an **unseen** data point mathematically

Key terms

- **Prior**: a starting guess for how likely each class is, before seeing the new example.
- **Likelihood**: how likely the observed features are under a given class (from the training data).
- **Posterior**: the updated probability of a class after seeing the example (what we want).
- **Normalize**: scaling scores so they add to 1 and become proper probabilities.

Bayes Theorem (Core Idea)

For class v and evidence x:

$$P(v \mid x) = \frac{P(x \mid v)\, P(v)}{P(x)}$$

Where:

- $P(v)$ = prior probability of class
- $P(x \mid v)$ = likelihood
- $P(v \mid x)$ = posterior probability
- $P(x)$ = evidence (normalization)

Naïve Bayes Assumption

If $x = (a_1, a_2, \ldots, a_n)$, then:

$$P(x \mid v) = \prod_{i=1}^{n} P(a_i \mid v)$$

So the decision rule becomes:

$$v_{NB} = \arg \max_{v \in \{Yes, No\}} P(v) \prod_i P(a_i \mid v)$$

Naïve Bayes assumes each feature (like Outlook, Temperature) gives independent evidence about the class. We compute a score for each class by multiplying the prior and the feature probabilities, then pick the class with the higher score.

Dataset

Outlook	Temperature	Humidity	Windy	PlayTennis
Sunny	Hot	High	Weak	No
Sunny	Hot	High	Strong	No
Overcast	Hot	High	Weak	Yes
Rainy	Mild	High	Weak	Yes
Rainy	Cool	Normal	Weak	Yes
Rainy	Cool	Normal	Strong	No
Overcast	Cool	Normal	Strong	Yes
Sunny	Mild	High	Weak	No
Sunny	Cool	Normal	Weak	Yes
Rainy	Mild	Normal	Weak	Yes

Outlook	Temperature	Humidity	Windy	PlayTennis
Sunny	Mild	Normal	Strong	Yes
Overcast	Mild	High	Strong	Yes
Overcast	Hot	Normal	Weak	Yes
Rainy	Mild	High	Strong	No

Dataset: PlayTennis (Summary)

Class counts (14 total):

- $PlayTennis = Yes$: 9
- $PlayTennis = No$: 5

So:

$$P(Yes) = \frac{9}{14}, \quad P(No) = \frac{5}{14}$$

Conditional Probabilities (From Training Data)

Outlook

Outlook	$P(\cdot \mid Yes)$	$P(\cdot \mid No)$
sunny	$2/9$	$3/5$
overcast	$4/9$	$0/5$
rain	$3/9$	$2/5$

Temperature

Temp	$P(\cdot \mid Yes)$	$P(\cdot \mid No)$
hot	$2/9$	$2/5$
mild	$4/9$	$2/5$
cool	$3/9$	$1/5$

Conditional Probabilities (Continued)

Humidity

Humidity	$P(\cdot \mid Yes)$	$P(\cdot \mid No)$
high	$3/9$	$4/5$
normal	$6/9$	$1/5$

Wind

Wind	$P(\cdot \mid Yes)$	$P(\cdot \mid No)$
strong	$3/9$	$3/5$
weak	$6/9$	$2/5$

Problem Instance

Classify:

$$x = (Outlook = sunny,\ Temperature = cool,\ Humidity = high,\ Wind = strong)$$

We compute **unnormalized** scores:

$$Score(v) = P(v) \prod_i P(a_i \mid v)$$

These are not yet probabilities — they are proportional to the posterior. We divide by the sum of scores to get real probabilities that sum to 1.

Score for Yes (Step-by-step)

$$Score(Yes) = P(Yes) \cdot P(sunny \mid Yes) \cdot P(cool \mid Yes) \cdot P(high \mid Yes) \cdot P(strong \mid Yes)$$

Substitute values:

$$= \frac{9}{14} \cdot \frac{2}{9} \cdot \frac{3}{9} \cdot \frac{3}{9} \cdot \frac{3}{9} = \frac{1}{189} \approx 0.00529$$

Score for No (Step-by-step)

$$Score(No) = P(No) \cdot P(sunny \mid No) \cdot P(cool \mid No) \cdot P(high \mid No) \cdot P(strong \mid No)$$

Substitute values:

$$= \frac{5}{14} \cdot \frac{3}{5} \cdot \frac{1}{5} \cdot \frac{4}{5} \cdot \frac{3}{5} = \frac{18}{875} \approx 0.02057$$

Normalize to Get Posterior Probabilities

$$P(Yes \mid x) = \frac{Score(Yes)}{Score(Yes) + Score(No)}$$

$$P(No \mid x) = \frac{Score(No)}{Score(Yes) + Score(No)}$$

Using computed scores:

- $P(Yes \mid x) = \frac{125}{611} \approx 0.2046$
- $P(No \mid x) = \frac{486}{611} \approx 0.7954$

✅ **Prediction: PlayTennis = No**

New Unseen Instance

Now classify a new data point:

$$x_{new} = (Outlook = rain,\ Temperature = mild,\ Humidity = high,\ Wind = weak)$$

Compute scores again.

New Instance: Score(Yes)

$$Score(Yes) = \frac{9}{14} \cdot \frac{3}{9} \cdot \frac{4}{9} \cdot \frac{3}{9} \cdot \frac{6}{9}$$

$$= \frac{4}{189} \approx 0.02116$$

New Instance: Score(No)

$$Score(No) = \frac{5}{14} \cdot \frac{2}{5} \cdot \frac{2}{5} \cdot \frac{4}{5} \cdot \frac{2}{5}$$

$$= \frac{16}{875} \approx 0.01829$$

New Instance: Normalize + Decision

$$P(Yes \mid x_{new}) = \frac{0.02116}{0.02116 + 0.01829} \approx 0.5365$$

$$P(No \mid x_{new}) \approx 0.4635$$

✅ **Prediction for unseen data: PlayTennis = Yes**

Summary

- Naïve Bayes uses:
  - Priors $P(v)$
  - Likelihoods $P(a_i \mid v)$
  - Product rule (naïve independence)
- For PPT instance $(sunny, cool, high, strong)$: **No**
- For unseen instance $(rain, mild, high, weak)$: **Yes**