



# 7. Naïve Bayes

# Table of Contents

- Naïve Bayes with sklearn.
  - Gaussian Naïve Bayes
  - Multinomial Naïve Bayes
  - Custom implementation





# Naïve Bayes

## ● Problem setting

- Data:  $D = \{(X, Y)^n\}_{n=1}^N$
- Input features:  $X = (x_1, \dots, x_k)$
- Output:  $Y \in \{1, \dots, L\}$
- Hypothesis:  $\operatorname{argmax}_Y P(Y|X) = \hat{Y}$

## ● Naïve Bayes Model

- $\hat{Y} = \operatorname{argmax}_Y P(Y|X)$
- $P(Y|X) = \frac{P(X|Y)P(Y)}{P(X)} = \frac{P(X|Y)P(Y)}{\sum P(X|Y)P(Y)}$
- $P(X|Y) = P(x_1, \dots, x_k|Y) = \prod_i P(x_i|Y)$ 
  - Assuming that all features are independent.

# Sklearn.naïve\_bayes.GaussianNB

- Assuming that each feature is independent and can be modeled as Gaussian distribution.
- Let's do with Irish dataset.

$$P(x_i | y) = \frac{1}{\sqrt{2\pi\sigma_y^2}} \exp \left( -\frac{(x_i - \mu_y)^2}{2\sigma_y^2} \right)$$

# Sklearn.naïve\_bayes.MultinomialNB

- Assuming that features are under multinomial distribution (ex, word frequency).
- Let's do with Text dataset.



# Custom implementation

- If features are categorical, previous two dataset is not available.
- Can you implement your own version?