Classification: Naive Bayes

\[ P(y|X) = \frac{P(y)P(X|y)}{P(X)} \]

**Prior**

\[ P(y|X) \propto P(y, X_1, \ldots, X_m) \propto P(y) \prod_{i=1}^{m} P(X_i|y) \]

**Likelihood**

**Maximum a Posteriori (MAP):** Pick the class with the maximum posterior probability.

\[ \hat{y} = \arg \max_y P(y) \prod_{i=1}^{m} P(X_i|y) \]
Gaussian Naive Bayes

Assume $P(X|Y)$ is Normal

\[ \hat{y} = \arg \max_y P(y) \prod_{i=1}^{m} P(X_i|y) \]
Gaussian Naive Bayes

Assume $P(X|Y)$ is *Normal*

Then, training is:

1. Estimate $P(Y = k); \quad \pi_k = \text{count}(Y = k) / \text{Count}(Y = *)$
2. MLE to find parameters $(\mu, \sigma)$ for each class of $Y$.
   (the “class conditional distribution”)

$$
\hat{y} = \arg \max_y P(y) \prod_{i=1}^{m} P(X_i|y)
$$
Gaussian Naive Bayes

Assume $P(X|Y)$ is Normal

Then, training is:

1. Estimate $P(Y = k); \quad \pi_k = \text{count}(Y = k) / \text{Count}(Y = *)$
2. MLE to find parameters $(\mu, \sigma)$ for each class of $Y$.
   (the “class conditional distribution”)

$$\hat{y} = \text{arg max}_y P(y) \prod_{i=1}^{m} P(X_i|y)$$
Gaussian Naive Bayes

Assume $P(X|Y)$ is Normal

Then, training is:

1. Estimate $P(Y = k); \quad \pi_k = \text{count}(Y = k) / \text{Count}(Y = \ast)$
2. MLE to find parameters $(\mu, \sigma)$ for each class of $Y$.
   (the “class conditional distribution”)

\[
\hat{y} = \arg \max_y \ P(y) \prod_{i=1}^{m} P(X_i|y)
\]
Example Project

https://docs.google.com/presentation/d/1jD-FQhOTaMh82JRC-p81TY1QCUbtpKZGwe5U4A3gml8/