Classification: Naive Bayes

Posterior

$$\boxed{\mathrm{P}(y|X)} = \frac{\boxed{\mathrm{P}(y)}\boxed{\mathrm{P}(X|y)}}{\mathrm{P}(X)}$$

Likelihood

Prior

$$\mathrm{P}(y|X) \propto \mathrm{P}(y, X_1, ..., X_m) \propto \mathrm{P}(y) \prod_{i=1}^{m} \mathrm{P}(X_i|y)$$

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

Unnormalized Posterior

$$\hat{y} = arg \max_{y} \boxed{\mathrm{P}(y) \prod_{i=1}^{m} \mathrm{P}(X_i|y)}$$

Gaussian Naive Bayes

Assume $\mathrm{P}(\mathrm{X}|\mathrm{Y})$ is *Normal*

$$\hat{y} = arg \max_{y} \ \mathrm{P}(y) \prod_{i=1}^{m} \mathrm{P}(X_i|y)$$

Gaussian Naive Bayes

Assume P(X|Y) is *Normal*

Then, training is:

1. Estimate P(Y = k); $\boldsymbol{\pi}_k$ = count(Y = k) / Count(Y = *)
2. MLE to find parameters ($\boldsymbol{\mu}$, $\boldsymbol{\sigma}$) for each class of Y.
   (the "class conditional distribution")

$$\hat{y} = arg\max_{y} \; \mathrm{P}(y) \prod_{i=1}^{m} \mathrm{P}(X_i|y)$$

Gaussian Naive Bayes

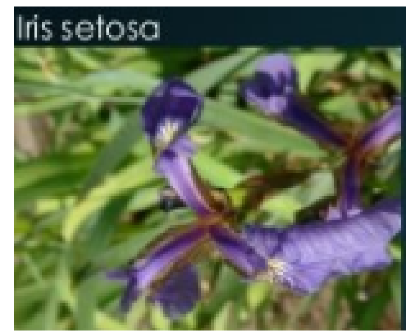
Assume P(X|Y) is *Normal*

Then, training is:

1. Estimate P(Y = k); $\boldsymbol{\pi}_{\mathrm{k}}$ = count(Y = k) / Count(Y = *)
2. MLE to find parameters ($\boldsymbol{\mu}$, $\boldsymbol{\sigma}$) for each class of Y.
   (the "class conditional distribution")

$$\hat{y} = arg \max_{y} \ \mathrm{P}(y) \prod_{i=1}^{m} \mathrm{P}(X_i|y)$$

Gaussian Naive Bayes



Assume P(X|Y) is *Normal*

Then, training is:

1. Estimate P(Y = k); $\boldsymbol{\pi}_k$ = count(Y = k) / Count(Y = *)
2. MLE to find parameters ($\boldsymbol{\mu}$, $\boldsymbol{\sigma}$) for each class of Y. (the "class conditional distribution")

$$\hat{y} = arg\max_{y} \mathrm{P}(y)\prod_{i=1}^{m}\mathrm{P}(X_i|y)$$

Example Project

https://docs.google.com/presentation/d/1jD-FQhOTaMh82JRc-p81TY1QCUbtpKZGwe5U4A3gml8/