Gaussian Naive Bayes

Gaussian → is used continous data

Naive → Assumption

Bayes → Bayes Theorem

Bayes

$$P(y=t \mid x) = \frac{P(y=t) \cdot P(x \mid y=t)}{\cancel{P(x)}}$$

$$p(x \mid y=t) = \prod_{j=1}^{\text{num features}} P(x_j \mid y=k)$$

Classification

data — yes — dog — P

no — cat — P

P + P = 1

proportional

$$P(y=t \mid x) \propto \boxed{P(y=t) \cdot \prod_{j=1}^{\text{num features}} P(x_j \mid y=k)}$$

Gaussian distrib.

$$\frac{n_k}{n}$$

$$N(x_j;\ \mu_{kj},\ \sigma_{kj}^2)$$

mean of feature j within all samples of class k

variance

$$\frac{1}{\sqrt{2\pi\sigma_{kj}^2}} \cdot e^{-\frac{(x_j - \mu_{kj})^2}{2\sigma_{kj}^2}}$$

for computational purposes

$$\log(P(y=t \mid x)) \propto \log(P(y=t)) + \sum_{j=1} \underbrace{\log(P(x_j \mid y=t))}$$

Our Goal (Classification)

$$\boxed{\hat{y} = \operatorname{argmax}\left[\text{logprobs}\right]}$$

$$\log(\mathcal{N}(x_j, \mu_{t_j}, \sigma^2_{t_j}))$$

$$\log\left(\frac{1}{\sqrt{2\pi\sigma^2_{t_j}}} + e^{-\frac{(x_j-\mu_{t_j})^2}{2\sigma^2_{t_j}}}\right)$$

$$\log\left(\left(2\pi\sigma^2_{t_j}\right)^{-\frac{1}{2}}\right) - \frac{(x_j-\mu_{t_j})^2}{2\sigma^2_{t_j}}$$

$$-\frac{1}{2}\log(2\pi\sigma^2_{t_j}) - \frac{(x_j-\mu_{t_j})^2}{2\sigma^2_{t_j} + \underbrace{\text{small num}}} + \text{prior}$$

we add so that zero division does not occur