

Session 9&10

Naive Bayes Classifier

Machine Learning | Zahra Amini

Telegram: @zahraamini_ai & Instagram:@zahraamini_ai & LinkedIn: @zahraamini-ai

<https://zil.ink/zahraamini>

: Naive Bayes Classifier

کردی از classifier های ساده‌تر که پایه‌ی احتمالات تضمین می‌کنند. برخلاف سادگی یک روش آماری تویا ات که با توجه به داده‌های وردی و الگوی رفتاری شان classification را نماید.

؟: تفاوت خم classifier با سایر Naive Bayes Classifier هاچیست؟

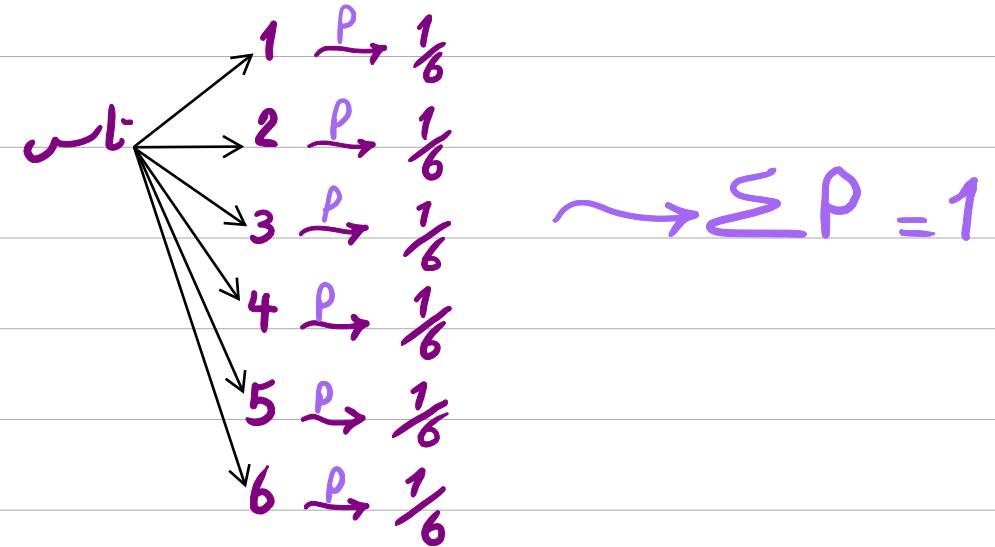
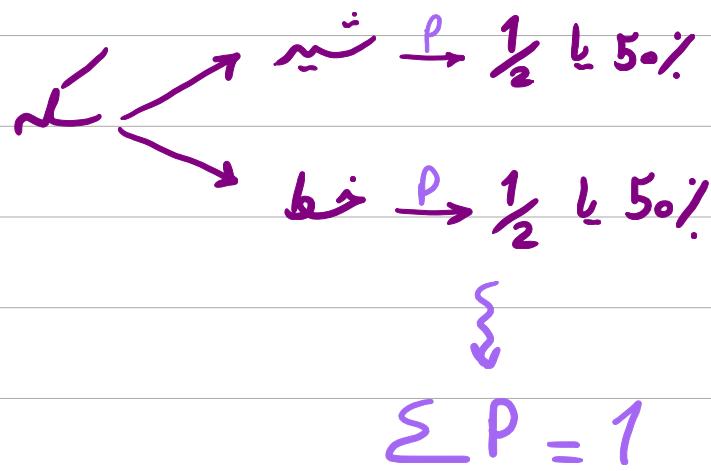
در دانش اولیه نیز مذکور است. یعنی چه؟ یعنی ساز تجربه قبلی شخص یا دانش اولیه فناوری مسأله خم در تضمین کردن استناده می‌کنم. \rightarrow به این دانش اولیه $Prior knowledge$ می‌کوییم.

: اطلاعاتی است که بعد از اندازه می‌گیری و بعد از توجه به داده‌ها بدست می‌آیند. $Prior knowledge$

اطلاعاتی که به صورت تجربی بدست می‌آیند.

Classifier with Prior knowledge:

فرض کنید سوالہ یا پرتاب یک کارہ است وہ خواہ یہ بانم ہر کارہ کو راپرتا بھی کنم جتنہ احتمال دارد شیر باشد یا خلہ.



$$P(C) \sim P(C_1), P(C_2), P(C_3), \dots, P(C_n)$$

فرض کنید یک پر شک متنصف بر اساس تجربہ اش یہ کوید (80٪ بیمارانی) که مراجعتی کرد سالم ہوتندو (20٪ ثانی مبتلا) به سرطان ہوتند.

$$\begin{cases} 1 & P(C_1) \geq P(C_2) \quad \checkmark \rightsquigarrow 80\% \geq 20\% \quad \checkmark \rightsquigarrow 80\% \text{ درست} \\ 2 & P(C_2) > P(C_1) \end{cases}$$

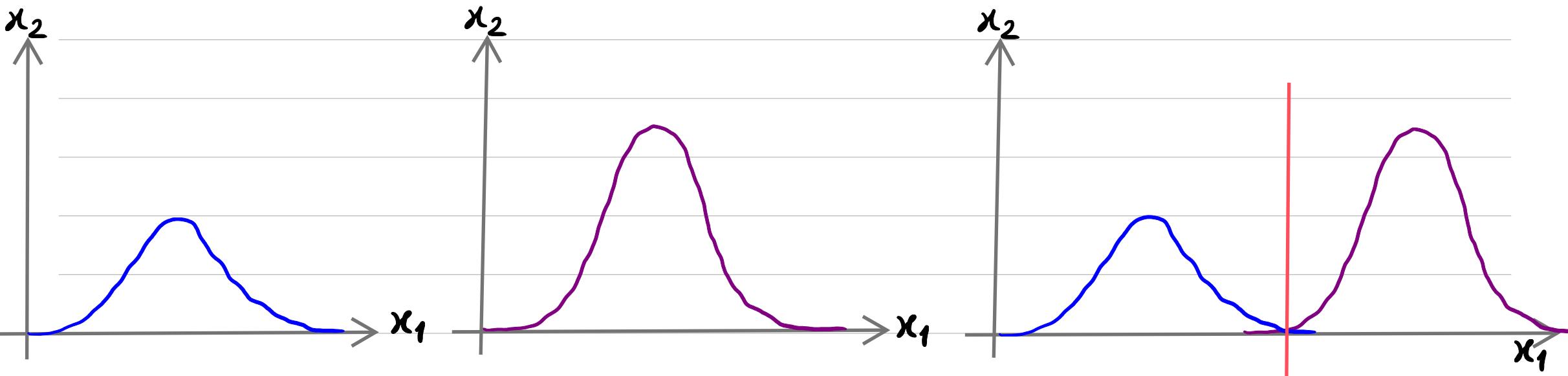
ج: این روش چه زمانی دقیق بالایی دارد؟

این روش زمانی باقت بالا عمل کند که اختلاف بین $P(c_1)$ و $P(c_2)$ زیاد باشد.

$$P(c_1) \gg P(c_2)$$

یا

$$P(c_2) \gg P(c_1)$$



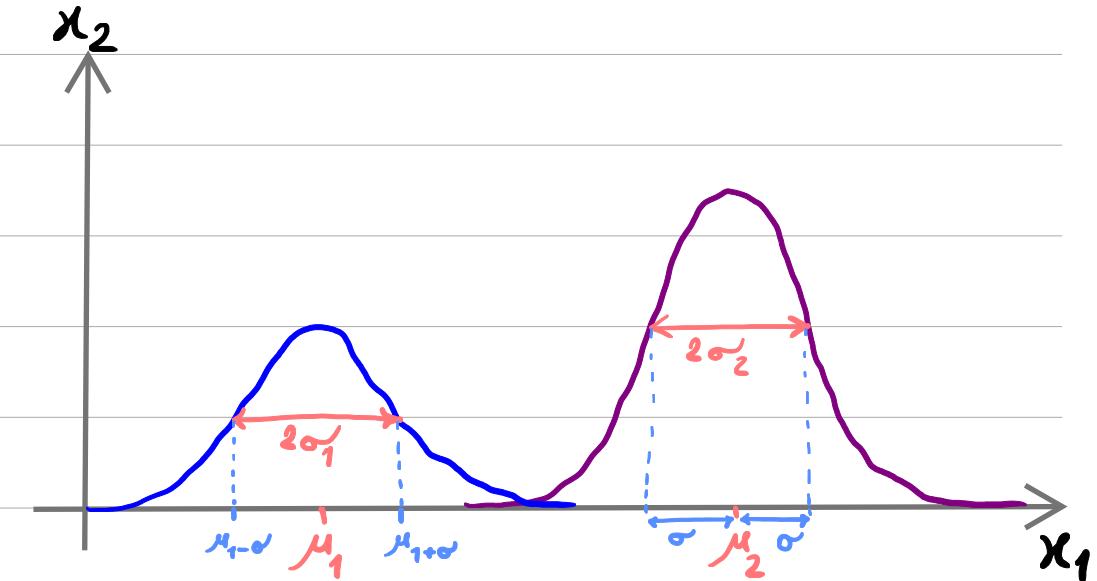
if $P(c_1) = P(c_2) \rightarrow$ random chance

ج: در Likelihood il Prior knowledge علاوه بر Naive Bayes Classifier چه استنادهایی کنیم.

یا شباهت: می خواهیم شباهت مخونه ها را به هر کلاس کاریب کنیم.

$$f(x | C_j) = \frac{1}{\sqrt{2\pi\sigma_j^2}} e^{-\frac{(x_i - \mu_j)^2}{2\sigma_j^2}}$$

- 1 $P(C_1) \geq P(C_2)$
- 2 $P(C_2) > P(C_1)$



۲. حال اگر نو اطمین پارامترها را بجهتی $\hat{\mu}$ را به دست آوریم چه کنم؟

$$f(x|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_i - \mu)^2}{2\sigma^2}}$$

$$J = \sum_{i=0}^m f(x_i | \mu, \sigma^2)$$

برای به دست آوردن حدود مقدار بجهتی μ و σ کافیه از تابع هنزینه مشتق بگیریم.

$$\ln \sum_{i=0}^m f(x_i | \mu, \sigma^2)$$

سپس نسبت به پارامترها مشتق می‌گیریم و برابر ۰ قرار می‌دهیم.

$$\frac{\partial}{\partial \mu} J = \frac{\partial}{\partial \mu} \left(\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_i - \mu)^2}{2\sigma^2}} \right) = 0 \rightarrow \hat{\mu} = \frac{1}{m} \sum_{i=0}^m x_i$$

$$\frac{\partial}{\partial \sigma} J = \frac{\partial}{\partial \sigma} \left(\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_i - \mu)^2}{2\sigma^2}} \right) = 0 \rightarrow \hat{\sigma} = \sqrt{\frac{1}{m} \sum_{i=0}^m (x_i - \mu)^2}$$

به جموعه‌ی مراحلی که طی شردیم برای به دست آوردن پارامترها Maximum Likelihood Estimation (MLE) گویند.

① داده های هر کلاس چه رفتاری دارند؟ چه توزیع نرمالی دارند؟ مقدار هر کلاس را به روش MLE محاسبه می کنیم.
تحمیل می اذنم

$$\hat{\mu} = \frac{1}{m} \sum_{i=0}^m x_i$$

$$\hat{\sigma} = \sqrt{\frac{1}{m} \sum_{i=0}^m (x_i - \mu)^2}$$

لے برای تخمین پارامترها از
استناده x_{train} می کنیم.

② شباهت داده x را نسبت به هر کلاس محاسبه می کنیم \rightarrow محاسبه Likelihood \rightarrow محاسبه می خونه است x به ازای هر کلاس:

$$f(x | C_j) = \frac{1}{\sqrt{2\pi\sigma_j^2}} e^{-\frac{(x_i - \mu_j)^2}{2\sigma_j^2}}$$

③ نوبتی حم با شه نوبت : Bayes classifier

کلاس بندی بحینه است از هر دوی دانش اولی (Likelihood) و شباهت (Prior) و شباهت (Likelihood) در تعمیم کسری استفاده کند.

$$P(C_j|x) = \frac{P(x|C_j) * P(C_j)}{P(x)}$$

$$P(C_j|x) * P(x) = P(x|C_j) * P(C_j) \Rightarrow$$

شباهت
Likelihood

$$P(C_j|x) = \frac{P(x|C_j) * P(C_j)}{P(x)}$$

Prior

Normalize مترن

بایثی شود حاصل ($P(C_j|x)$) بین ۰ و ۱ باشد

اما اگر نباشد هم در تعمیم متأثیری ندارد.

$$P(x) = \sum_{i=1}^c P(x|C_i) * P(C_i)$$

Naive Bayes Classifier

Prior $\rightarrow P(C_1), P(C_2), P(C_3), \dots, P(C_n)$

① **کاربی پارامترها:**

$$\hat{\mu} = \frac{1}{m} \sum_{i=0}^m x_i \quad \longrightarrow \quad \mu_1, \mu_2, \mu_3, \dots, \mu_c$$

$$\hat{\sigma} = \sqrt{\frac{1}{m} \sum_{i=0}^m (x_i - \mu)^2} \quad \longrightarrow \quad \sigma_1, \sigma_2, \sigma_3, \dots, \sigma_c$$

② **کاربی مونه تست x بازای هر کلاس:** Likelihood

$P(x|C_1), P(x|C_2), P(x|C_3), \dots, P(x|C_j)$

$$f(x|C_j) = \frac{1}{\sqrt{2\pi\sigma_j^2}} e^{-\frac{(x_i - \mu_j)^2}{2\sigma_j^2}}$$

③ محاسبه احتمال پسین Posterior مخونه ت است به ازای هر کلاس:

$$P(C_j|x) = \frac{P(x|C_j) * P(C_j)}{P(x)}$$

⑤ تفسیم لیری: x به کلاسی تعلق دارد که بیشترین احتمال Posterior را بست به آن کلاس داشته باشد.

| Gender | height(feet) | Weight(lbs) | Foot size (inch) |
|--------|--------------|-------------|------------------|
| M | 6 | 180 | 12 |
| M | 5.9 | 190 | 11 |
| M | 5.58 | 170 | 12 |
| M | 5.92 | 165 | 10 |
| F | 5 | 100 | 6 |
| F | 5.5 | 150 | 8 |
| F | 5.42 | 130 | 7 |
| F | 5.75 | 150 | 9 |

$$\mu = \frac{1}{m} \sum_{i=1}^m x_i$$

$$\mu_{\text{height male}} = \frac{1}{4} \times (6 + 5.9 + 5.58 + 5.92) = 5.855$$

$$\mu_{\text{height female}} = \frac{1}{4} \times (5 + 5.5 + 5.42 + 5.75) = 5.4175$$

$$\text{Var}(x) = \text{Var}(x - \mu) = \sigma^2 = \frac{\sum_{i=1}^m (x_i - \mu)^2 - (\sum_{i=1}^m (x_i - \mu))^2 / m}{m-1}$$

$$\sigma^2 = \frac{\cancel{[(6-5.855)^2 + (5.92-5.855)^2 + (5.58-5.855)^2 + (5.92-5.855)^2]} - \cancel{[(6-5.855) + (5.92-5.855) + (5.58-5.855) + (5.92-5.855)]^2 / 4}}{4-1}$$

$$= \frac{0.1051}{3} = 0.035033$$

| | height(feet) | | Weight(lbs) | | Foot size (inch) | |
|--------|--------------|-------------------------|-------------|----------------------|------------------|-------------------------|
| Gender | mean | Variance | mean | Variance | mean | Variance |
| M | 5.855 | 3.5033×10^{-2} | 176.25 | 1.2292×10^2 | 11.25 | 9.1667×10^{-1} |
| F | 5.4175 | 9.7225×10^{-2} | 132.5 | 5.5833×10^2 | 7.5 | 1.6667 |

$$X = [6, 130, 8]$$

$$\text{Posterior (male)} = \frac{P(\text{male}) P(\text{height}|\text{male}) * P(\text{weight}|\text{male}) * P(\text{foot}|\text{male})}{P(x)}$$

$$\text{Posterior (female)} = \frac{P(\text{female}) P(\text{height}|\text{female}) * P(\text{weight}|\text{female}) * P(\text{foot}|\text{female})}{P(x)}$$

$$0.5 \quad 1.5789 \quad 5.9887 \times 10^{-6} \quad 1.3112 \times 10^{-3}$$

$$P(x) = P(\text{male}) * P(\text{height}|\text{male}) * P(\text{weight}|\text{male}) * P(\text{foot}|\text{male})$$

$$* P(\text{female}) * P(\text{height}|\text{female}) * P(\text{weight}|\text{female}) * P(\text{foot}|\text{female}) = 3.32 \times 10^{-12}$$
$$0.5 \quad 2.23 \times 10^{-1} \quad 1.6789 \times 10^{-2} \quad 2.8669 \times 10^{-1}$$

$$P(\text{male}) = 0.5$$

$$P(\text{height} | \text{male}) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(6-\mu)^2}{2\sigma^2}\right) \approx 1.5789$$

$$\mu = 5.855$$

$$\text{variance} = \sigma^2 = 3.5033 \times 10^{-2}$$

$$P(\text{Weight} | \text{male}) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(130-\mu)^2}{2\sigma^2}\right) = 5.9881 \times 10^{-6}$$

$$P(\text{foot} | \text{male}) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(8-\mu)^2}{2\sigma^2}\right) = 1.3112 \times 10^{-3}$$

$$\sum_x P(x) \rightarrow 6.1984 \times 10^{-9}$$

$$P(\text{female}) = 0.5$$

$$P(\text{height} | \text{female}) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(6-\mu)^2}{2\sigma^2}\right) \approx 2.23 \times 10^{-1}$$

$$\mu = 5.4175 \quad \text{variance} = \sigma^2 = 9.7225 \times 10^{-2}$$

$$P(\text{Weight} | \text{female}) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(130-\mu)^2}{2\sigma^2}\right) = 1.6789 \times 10^{-2}$$

$$P(\text{foot} | \text{female}) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(8-\mu)^2}{2\sigma^2}\right) = 2.8669 \times 10^{-1}$$

$$\underbrace{\pi}_{x} \rightarrow 5.3778 \times 10^{-4}$$

Posterior(male) < Posterior(female) $\rightsquigarrow x \in \text{Female}$

Bernoulli Naive Bayes

$$P(x|C_j) = \begin{cases} 1-P & x=0 \\ P & x=1 \end{cases}$$

برای داده های با کمتر از استفاده کا شود discrete

$$P(x|C_j) = P(x_i|C_j)x_i + (1 - P(x_i|C_j))(1 - x_i)$$

1. Prior

2. Likelihood

3. Posterior

4. Make a decision

ابزاریا: ex

| Weather | temperature | humidity | Windy | Play? |
|---------|-------------|----------|-------|-------|
| Rainy | Hot | yes | No | No |
| Rainy | Hot | yes | yes | No |
| Cloudy | Hot | yes | No | Yes |
| Sunny | Moderate | yes | No | Yes |
| Sunny | Cold | No | No | Yes |
| Sunny | Cold | No | yes | No |
| Cloudy | Cold | No | yes | Yes |
| Rainy | Moderate | yes | No | No |
| Rainy | Cold | No | No | Yes |
| Sunny | Moderate | No | No | Yes |
| Rainy | Moderate | No | yes | Yes |
| Cloudy | Moderate | yes | yes | Yes |
| Cloudy | Hot | No | No | Yes |
| Sunny | Moderate | yes | yes | No |

Weather tem humidity windy
 $X = [Rainy, Cold, yes, yes]$

①:

$$P(\text{yes}) = \frac{9}{14}$$

$$P(\text{No}) = \frac{5}{14}$$

↳ class

$$②. P(\text{Weather}=\text{Rainy} \mid \text{class}=\text{yes}) = \frac{2}{9} \quad P(\text{Weather}=\text{Rainy} \mid \text{class}=\text{No}) = \frac{3}{5}$$

$$P(\text{tem.} = \text{cold} \mid \text{class} = \text{yes}) = \frac{3}{9} \quad P(\text{tem.} = \text{cold} \mid \text{class} = \text{No}) = \frac{1}{5}$$

$$P(\text{hum.} = \text{yes} \mid \text{class} = \text{yes}) = \frac{3}{9} \quad P(\text{hum.} = \text{yes} \mid \text{class} = \text{No}) = \frac{4}{5}$$

$$P(\text{windy} = \text{yes} \mid \text{class} = \text{yes}) = \frac{3}{9} \quad P(\text{windy} = \text{yes} \mid \text{class} = \text{No}) = \frac{3}{5}$$

$$③. \text{Posterior} \quad P(C_j | x) = \frac{P(x | C_j) * P(C_j)}{P(x)}$$
$$P(\text{class} = \text{yes} | x) = \frac{P(x | \text{class} = \text{yes}) * P(\text{yes})}{P(x)} = \frac{\frac{2}{9} * \frac{3}{9} * \frac{3}{5} * \frac{3}{9} * \frac{9}{14}}{0.021} = 0.238$$

$$P(x) = P(\text{Weather} = \text{Rainy}) * P(\text{tem} = \text{cold}) * P(\text{hum.} = \text{yes}) * P(\text{windy} = \text{yes}) =$$
$$\frac{5}{14} * \frac{4}{14} * \frac{7}{14} * \frac{6}{14} = 0.021$$

$$P(\text{class}=\text{No} | x) = \frac{P(x | \text{class}=\text{No}) * P(\text{No})}{P(x)} = \frac{\frac{3}{5} * \frac{1}{5} * \frac{4}{5} * \frac{3}{5} * \frac{5}{14}}{0.021} = 0.952$$

$$P(\text{class}=\text{yes} | x) < P(\text{class}=\text{No} | x) \Rightarrow x \in \text{No}$$

| Confident | Studied | Sick | Result |
|-----------|---------|------|--------|
| Yes | No | No | Fail |
| Yes | No | Yes | Pass |
| No | Yes | Yes | Fail |
| No | Yes | No | Pass |
| Yes | Yes | Yes | Pass |

مثال زیر جزو کارهای ایجاد مجموعه داده است: ex

confident studied sick

$$x = [\text{Yes}, \text{Yes}, \text{No}]$$

مثال کاری زیر جزو کدام طاس است؟: ex

| Confident | Studied | Sick | Result |
|-----------|---------|------|--------|
| Yes | No | No | Fail |
| yes | No | yes | Pass |
| No | Yes | yes | Fail |
| No | yes | No | Pass |
| yes | yes | yes | Pass |

confident studied sick

$$x = [\text{yes}, \text{yes}, \text{No}]$$

$$\textcircled{1} \quad P(\text{Pass}) = ? \frac{3}{5}$$

$$P(\text{Fail}) = ? \frac{2}{5}$$

$$\textcircled{2} \quad P(\text{Confident=yes} \mid \text{Result=Pass}) = ? \frac{2}{3}$$

$$P(\text{Confident=yes} \mid \text{Result=Fail}) = ? \frac{1}{2}$$

$$P(\text{Studied=yes} \mid \text{Result=Pass}) = ? \frac{2}{3}$$

$$P(\text{Studied=yes} \mid \text{Result=Fail}) = ? \frac{1}{2}$$

$$P(\text{Sick=No} \mid \text{Result=Pass}) = ? \frac{1}{3}$$

$$P(\text{Sick=No} \mid \text{Result=Fail}) = ? \frac{1}{2}$$

③ Posterior:

$$P(C_j|x) = \frac{P(x|C_j) * P(C_j)}{P(x)}$$

$$P(\text{Result}=\text{Pass} | x) = \frac{P(x | \text{Result}=\text{Pass}) * P(\text{Result}=\text{Pass})}{P(x)} = \frac{\frac{2}{3} * \frac{2}{3} * \frac{1}{3} * \frac{3}{5}}{\frac{18}{125}}$$

$$P(x) = P(\text{confident}=\text{yes}) * P(\text{studied}=\text{yes}) * P(\text{Sick}=\text{No}) = 0.611$$

$$\hookrightarrow P(x) = \frac{3}{5} * \frac{3}{5} * \frac{2}{5} = \frac{18}{125}$$

$$P(\text{Result}=\text{Fail} | x) = \frac{P(x | \text{Result}=\text{Fail}) * P(\text{Result}=\text{Fail})}{P(x)} = \frac{\frac{1}{2} * \frac{1}{2} * \frac{1}{2} * \frac{2}{5}}{\frac{18}{125}} = 0.34$$

$$\Rightarrow 0.611 > 0.34 \rightarrow x \in \text{Pass}$$