Ciência de Dados

Conhecimento Incerto - Teorema de Bayes

Clayton Pereira

Universidade Estadual Paulista "Júlio de Mesquita Filho" (UNESP)
Faculdade de Ciências (FC) / Departamento de Computação (DCo)
Bauru, SP - Brasil

Sumário

1. Introdução ao Conhecimento Incerto
2. Probabilidade
   2.1 Conceitos Fundamentais
   2.2 Probabilidade da União de Eventos
   2.3 Probabilidade do Complemento
   2.4 Probabilidade Condicional
   2.5 Probabilidade da Intersecção de Dois Eventos
   2.6 Independência de Eventos
   2.7 Partição do Espaco Amostral
   2.8 Teorema da Probabilidade Total
3. Teorema de Bayes
4. Rede Bayesiana
5. Naïve Bayes

Introdução ao Conhecimento Incerto

Conhecimento Incerto

"Incerteza se origina de alguma deficiência de informação. A informação pode estar incompleta, pode ser vaga, pode ser imprecisa ou pode ser contraditória."

Klir, G.J. and Folger, T.A. "Fuzzy Sets, Uncertainty and Information", Prentice Hall, 1998.

Conhecimento Incerto

Algumas técnicas que podem ser utilizadas para tratar incertezas:

Conhecimento Incerto

Algumas técnicas que podem ser utilizadas para tratar incertezas:

- Ignorar as incertezas, ou seja, modelar o mundo negligenciando as incertezas.

Conhecimento Incerto

Algumas técnicas que podem ser utilizadas para tratar incertezas:

- Ignorar as incertezas, ou seja, modelar o mundo negligenciando as incertezas.
- Lógica Nebulosa: informacões possuem grau de verdade, forma precisa de se lidar com a imprecisão.

Conhecimento Incerto

Algumas técnicas que podem ser utilizadas para tratar incertezas:

- Ignorar as incertezas, ou seja, modelar o mundo negligenciando as incertezas.
- Lógica Nebulosa: informacões possuem grau de verdade, forma precisa de se lidar com a imprecisão.
- Raciocínio Probabilísco.

Conhecimento Incerto

- IA precisa lidar com incerteza em diversas situações.

Conhecimento Incerto

- IA precisa lidar com incerteza em diversas situações.
- **Exemplo:** Diagnóstico médico, previsões de tempo, reconhecimento de padrões.

Conhecimento Incerto

- IA precisa lidar com incerteza em diversas situações.
- **Exemplo:** Diagnóstico médico, previsões de tempo, reconhecimento de padrões.
- **Duas abordagens:** Probabilidade (Regra de Bayes) e Graus de Pertinência (Lógica Fuzzy).

Regra de Bayes vs. Lógica Fuzzy

- **Regra de Bayes:** Trabalha com probabilidades numéricas, ideal para situações onde as probabilidades a priori e as evidências podem ser conhecidas.

Regra de Bayes vs. Lógica Fuzzy

- **Regra de Bayes:** Trabalha com probabilidades numéricas, ideal para situações onde as probabilidades a priori e as evidências podem ser conhecidas.
- **Lógica Fuzzy:** Trabalha com graus de verdade e é útil em contextos onde a incerteza não pode ser quantificada facilmente, como em decisões humanas ou sistemas de controle.

Regra de Bayes vs. Lógica Fuzzy

- **Regra de Bayes:** Trabalha com probabilidades numéricas, ideal para situações onde as probabilidades a priori e as evidências podem ser conhecidas.
- **Lógica Fuzzy:** Trabalha com graus de verdade e é útil em contextos onde a incerteza não pode ser quantificada facilmente, como em decisões humanas ou sistemas de controle.

Exemplo comparativo: Previsão de chuva com base em dados históricos (Bayes) versus uma previsão qualitativa (nublado, parcialmente nublado, ensolarado – fuzzy).

Objetivos

- Como lidar com incerteza utilizando probabilidades?
- Como representar eficientemente a base de conhecimento de um sistema inteligente que utiliza raciocínio probabilístico?

Objetivos

- Como lidar com incerteza utilizando probabilidades?
- Como representar eficientemente a base de conhecimento de um sistema inteligente que utiliza raciocínio probabilístico?

Probabilidade

Introdução à Probabilidade

- Probabilidade é usada para medir a incerteza.
- Em IA, a probabilidade é essencial para tomar decisões com base em informações incompletas.
- A Regra de Bayes nos permite atualizar nossas crenças à medida que novas informações chegam.

Experimento Aleatório

Um experimento aleatório, **E**, é um experimento em que o resultado não pode ser predito, mesmo que seja repetido várias vezes e nas mesmas condições.

Experimento Aleatório

Um experimento aleatório, **E**, é um experimento em que o resultado não pode ser predito, mesmo que seja repetido várias vezes e nas mesmas condições.

Experimentos:

Experimento Aleatório

Um experimento aleatório, **E**, é um experimento em que o resultado não pode ser predito, mesmo que seja repetido várias vezes e nas mesmas condições.

Experimentos:

- E1: Arremesso de moeda.

Experimento Aleatório

Um experimento aleatório, **E**, é um experimento em que o resultado não pode ser predito, mesmo que seja repetido várias vezes e nas mesmas condições.

Experimentos:

- E1: Arremesso de moeda.
- E2: Arremesso de dado.

Experimento Aleatório

Um experimento aleatório, **E**, é um experimento em que o resultado não pode ser predito, mesmo que seja repetido várias vezes e nas mesmas condições.

Experimentos:

- E1: Arremesso de moeda.
- E2: Arremesso de dado.
- E3: Retirar uma bola de uma urna contendo bolas numeradas de 1 a 100.

Experimento Aleatório

Um experimento aleatório, **E**, é um experimento em que o resultado não pode ser predito, mesmo que seja repetido várias vezes e nas mesmas condições.

Experimentos:

- E1: Arremesso de moeda.
- E2: Arremesso de dado.
- E3: Retirar uma bola de uma urna contendo bolas numeradas de 1 a 100.

Variáveis aleatórias:

- E1: Face da moeda para cima.

Experimento Aleatório

Um experimento aleatório, **E**, é um experimento em que o resultado não pode ser predito, mesmo que seja repetido várias vezes e nas mesmas condições.

Experimentos:

- E1: Arremesso de moeda.
- E2: Arremesso de dado.
- E3: Retirar uma bola de uma urna contendo bolas numeradas de 1 a 100.

Variáveis aleatórias:

- E1: Face da moeda para cima.
- E2: Face do dado para cima.

Experimento Aleatório

Um experimento aleatório, **E**, é um experimento em que o resultado não pode ser predito, mesmo que seja repetido várias vezes e nas mesmas condições.

Experimentos:

- E1: Arremesso de moeda.
- E2: Arremesso de dado.
- E3: Retirar uma bola de uma urna contendo bolas numeradas de 1 a 100.

Variáveis aleatórias:

- E1: Face da moeda para cima.
- E2: Face do dado para cima.
- E3: Número da bola.

Espaço Amostral Ω

É o conjunto de todos os valores que a variável aleatória pode assumir.

Exemplos de espaços amostrais:

Espaço Amostral Ω

É o conjunto de todos os valores que a variável aleatória pode assumir.

Exemplos de espaços amostrais:

- E1: $\{cara, coroa\}$.

Espaço Amostral Ω

É o conjunto de todos os valores que a variável aleatória pode assumir.

Exemplos de espaços amostrais:

- E1: $\{cara, coroa\}$.
- E2: $\{1, 2, 3, 4, 5, 6\}$.

Espaço Amostral Ω

É o conjunto de todos os valores que a variável aleatória pode assumir.

Exemplos de espaços amostrais:

- E1: $\{cara, coroa\}$.
- E2: $\{1, 2, 3, 4, 5, 6\}$.
- E3: $\{1, 2, 3, \ldots, 100\}$.

Eventos

Um evento é um subconjunto do espaço amostral.

Exemplo:

- E1: Jogar uma moeda três vezes e observar quantas vezes aparece cara e quantas vezes aparece coroa.

Eventos

Um evento é um subconjunto do espaço amostral.

Exemplo:

- E1: Jogar uma moeda três vezes e observar quantas vezes aparece cara e quantas vezes aparece coroa.
- Evento: Aparecer 2 caras (k) e uma coroa (c) em qualquer ordem.

Eventos

Um evento é um subconjunto do espaço amostral.

Exemplo:

- E1: Jogar uma moeda três vezes e observar quantas vezes aparece cara e quantas vezes aparece coroa.
- Evento: Aparecer 2 caras (k) e uma coroa (c) em qualquer ordem.
- Subconjunto: $\{(kkc), (kck), (ckk)\}$.

Eventos

Um evento é um subconjunto do espaço amostral.

Exemplo:

- E1: Jogar uma moeda três vezes e observar quantas vezes aparece cara e quantas vezes aparece coroa.
- Evento: Aparecer 2 caras (k) e uma coroa (c) em qualquer ordem.
- Subconjunto: $\{(kkc), (kck), (ckk)\}$.

- E2: Jogar um dado e observar o número na face de cima.

Eventos

Um evento é um subconjunto do espaço amostral.

Exemplo:

- E1: Jogar uma moeda três vezes e observar quantas vezes aparece cara e quantas vezes aparece coroa.
- Evento: Aparecer 2 caras (k) e uma coroa (c) em qualquer ordem.
- Subconjunto: $\{(kkc), (kck), (ckk)\}$.

- E2: Jogar um dado e observar o número na face de cima.
- Evento: Número é par.

Eventos

Um evento é um subconjunto do espaço amostral.

Exemplo:

- E1: Jogar uma moeda três vezes e observar quantas vezes aparece cara e quantas vezes aparece coroa.
- Evento: Aparecer 2 caras (k) e uma coroa (c) em qualquer ordem.
- Subconjunto: $\{(kkc), (kck), (ckk)\}$.

- E2: Jogar um dado e observar o número na face de cima.
- Evento: Número é par.
- Subconjunto: $\{2, 4, 6\}$.

Eventos

Um evento é um subconjunto do espaço amostral.

Exemplo:

- E1: Jogar uma moeda três vezes e observar quantas vezes aparece cara e quantas vezes aparece coroa.
- Evento: Aparecer 2 caras (k) e uma coroa (c) em qualquer ordem.
- Subconjunto: $\{(kkc), (kck), (ckk)\}$.

- E2: Jogar um dado e observar o número na face de cima.
- Evento: Número é par.
- Subconjunto: $\{2, 4, 6\}$.

- E3: Uma urna contém bolas numeradas de 1 a 100.

Eventos

Um evento é um subconjunto do espaço amostral.

Exemplo:

- E1: Jogar uma moeda três vezes e observar quantas vezes aparece cara e quantas vezes aparece coroa.
- Evento: Aparecer 2 caras (k) e uma coroa (c) em qualquer ordem.
- Subconjunto: $\{(kkc), (kck), (ckk)\}$.

- E2: Jogar um dado e observar o número na face de cima.
- Evento: Número é par.
- Subconjunto: $\{2, 4, 6\}$.

- E3: Uma urna contém bolas numeradas de 1 a 100.
- Evento: Bola com número maior que 80.

Eventos

Um evento é um subconjunto do espaço amostral.

Exemplo:

- E1: Jogar uma moeda três vezes e observar quantas vezes aparece cara e quantas vezes aparece coroa.
- Evento: Aparecer 2 caras (k) e uma coroa (c) em qualquer ordem.
- Subconjunto: $\{(kkc), (kck), (ckk)\}$.

- E2: Jogar um dado e observar o número na face de cima.
- Evento: Número é par.
- Subconjunto: $\{2, 4, 6\}$.

- E3: Uma urna contém bolas numeradas de 1 a 100.
- Evento: Bola com número maior que 80.
- Subconjunto: $\{81, 82, 83, \ldots, 100\}$.

Variável Aleatória

- É aquela que assuma valores em um espaço amostral Ω e para a qual está determinada a probabilidade de ocorrência de cada um dos elementos do espaço amostral.

Variável Aleatória

- É aquela que assuma valores em um espaço amostral Ω e para a qual está determinada a probabilidade de ocorrência de cada um dos elementos do espaço amostral.
- Assim, se o espaço amostral for contínuo, é necessário conhecer a função distribuição de probabilidade (fdp) para caracterizar a variável aleatória.

Variável Aleatória

- É aquela que assuma valores em um espaço amostral Ω e para a qual está determinada a probabilidade de ocorrência de cada um dos elementos do espaço amostral.
- Assim, se o espaço amostral for contínuo, é necessário conhecer a função distribuição de probabilidade (fdp) para caracterizar a variável aleatória.
- Se o espaço amostral for discreto, é necessário conhecer a função massa de probabilidade (fmp) para caracterizar a variável aleatória.

Espaço Amostral Contínuo e Função de Distribuição de Probabilidade (fdp)

- **Espaço Contínuo:** Se o espaço amostral é contínuo (infinitos possíveis valores), como as alturas de pessoas, a probabilidade de um valor específico é 0, e, em vez disso, usamos a função de distribuição de probabilidade (fdp) para descrever a distribuição dos valores.

Espaço Amostral Contínuo e Função de Distribuição de Probabilidade (fdp)

- **Espaço Contínuo:** Se o espaço amostral é contínuo (infinitos possíveis valores), como as alturas de pessoas, a probabilidade de um valor específico é 0, e, em vez disso, usamos a função de distribuição de probabilidade (fdp) para descrever a distribuição dos valores.
- **fdp:** Indica a probabilidade da variável aleatória assumir um valor em um intervalo específico.

Espaço Amostral Contínuo e Função de Distribuição de Probabilidade (fdp)

- **Espaço Contínuo:** Se o espaço amostral é contínuo (infinitos possíveis valores), como as alturas de pessoas, a probabilidade de um valor específico é 0, e, em vez disso, usamos a função de distribuição de probabilidade (fdp) para descrever a distribuição dos valores.
- **fdp:** Indica a probabilidade da variável aleatória assumir um valor em um intervalo específico.

Exemplo: Para uma variável aleatória que representa a altura de pessoas, a fdp nos dá a probabilidade de uma pessoa ter uma altura entre 1,70m e 1,80m.

Espaço Amostral Discreto e Função Massa de Probabilidade (fmp)

- **Espaço Discreto:** Se o espaço amostral é discreto (com um número finito ou contável de resultados possíveis), como o lançamento de um dado ou a contagem de carros em uma rua, usamos a função massa de probabilidade (fmp).

Espaço Amostral Discreto e Função Massa de Probabilidade (fmp)

- **Espaço Discreto:** Se o espaço amostral é discreto (com um número finito ou contável de resultados possíveis), como o lançamento de um dado ou a contagem de carros em uma rua, usamos a função massa de probabilidade (fmp).
- **fmp:** indica a probabilidade de cada valor possível ocorrer.

Espaço Amostral Discreto e Função Massa de Probabilidade (fmp)

- **Espaço Discreto:** Se o espaço amostral é discreto (com um número finito ou contável de resultados possíveis), como o lançamento de um dado ou a contagem de carros em uma rua, usamos a função massa de probabilidade (fmp).
- **fmp:** indica a probabilidade de cada valor possível ocorrer.

Exemplo: Para o lançamento de um dado, a fmp atribui a probabilidade de $1/6$ a cada resultado (1, 2, 3, 4, 5, 6).

Probabilidade

"Quantidade do que é provável, do que tem a possibilidade de acontecer."

- Para cada evento, é atribuído um número real no intervalo [0, 1], dizendo qual a probabilidade de ocorrer tal evento.
- Se o espaço amostral consiste de N elementos igualmente prováveis e o evento A corresponde a um subconjunto de r elementos do espaço amostral, então a probabilidade de ocorrer A é dada por: $P(A) = r/N$.

Assim, temos:

- $0 \leq P(A) \leq 1$.
- $P(\Omega) = 1$.
- $P(\emptyset) = 0$.

Exemplo: Uma urna contém 10 bolas numeradas de 0 a 9. Em um experimento, é preciso selecionar uma bola da urna e anotar seu número. Deseja-se encontrar a probabilidade dos eventos:

- número da bola é 5.
- número da bola é ímpar.
- número da bola é múltiplo de 3.

O espaço amostral é $S = 0, 1, 2, \ldots, 9$ e os resultados correspondentes aos eventos acima são:

$A = \{5\}$, $B = \{1, 3, 5, 7, 9\}$, e $C = \{3, 6, 9\}$.

Se for suposto que os resultados são equiprováveis, então:

- $P(A) = P(5) = 1/10$.
- $P(B) = P(1) + P(3) + P(5) + P(7) + P(9) = 1/10 + 1/10 + 1/10 + 1/10 + 1/10 = 5/10$.
- $P(C) = P(3) + P(6) + P(9) = 1/10 + 1/10 + 1/10 = 3/10$.

Probabilidade da União de Eventos

$$P(A \cup B) = P(A) + P(B) - P(A \cap B). \tag{1}$$

- $P(A \cup B)$: Representa a probabilidade de ocorrer **A** ou **B** (ou ambos). É a união dos dois eventos.

Probabilidade da União de Eventos

$$P(A \cup B) = P(A) + P(B) - P(A \cap B). \tag{1}$$

- $P(A \cup B)$: Representa a probabilidade de ocorrer **A** ou **B** (ou ambos). É a união dos dois eventos.
- $P(A)$: Representa a probabilidade de ocorrer o evento **A**.

Probabilidade da União de Eventos

$$P(A \cup B) = P(A) + P(B) - P(A \cap B). \tag{1}$$

- $P(A \cup B)$: Representa a probabilidade de ocorrer **A** ou **B** (ou ambos). É a união dos dois eventos.
- $P(A)$: Representa a probabilidade de ocorrer o evento **A**.
- $P(B)$: Representa a probabilidade de ocorrer o evento **B**.

Probabilidade da União de Eventos

$$P(A \cup B) = P(A) + P(B) - P(A \cap B). \tag{1}$$

- $P(A \cup B)$: Representa a probabilidade de ocorrer **A** ou **B** (ou ambos). É a união dos dois eventos.
- $P(A)$: Representa a probabilidade de ocorrer o evento **A**.
- $P(B)$: Representa a probabilidade de ocorrer o evento **B**.
- $P(A \cap B)$: Representa a probabilidade de ocorrer tanto A quanto B simultaneamente (a interseção dos dois eventos).

Probabilidade da União de Eventos

$$P(A \cup B) = P(A) + P(B) - P(A \cap B). \tag{2}$$

Exemplo: Considere o experimento, o lançamento de um dado e os seguintes eventos:

- A: sair o número 2, 3, 4.
- B: sair número par.
- C: sair número ímpar.

Determinar $P(A \cup B)$ e $P(A \cup C)$.

Probabilidade da União de Eventos Disjuntos

Se A e B são disjuntos, então:

$$P(A \cup B) = P(A) + P(B). \tag{3}$$

Dois eventos A e B são disjuntos (ou mutuamente exclusivos) se não podem ocorrer simultaneamente. Isso significa que a interseção deles é vazia.

$$A \cup B = \emptyset \tag{4}$$

(não há resultados comuns entre A e B)

Probabilidade da União de Eventos Disjuntos

Exemplo: No lançamento de duas moedas temos:

- A: uma cara.
- B: duas coroas.

Qual a probabilidade de sair duas coroas ou uma cara?

Exemplo: As probabilidades de uma pessoa que deseja adquirir um carro novo escoler um Chevrolet, um Ford ou um Honda são 0.17, 0.22, e 0.08, respectivamente. Supondo que ela compre apenas um carro, qual é a probabilidade de ser uma das três marcas?

Probabilidade do Complemento

$$P(A^c) = 1 - P(A). \tag{5}$$

Exemplo: Um dado é lançado 10 vezes, qual a probabilidade de sair:

- A: pelo menos um 6.

Probabilidade Condicional

A probabilidade de um evento particular acontecer depende do resultado de algum outro evento.

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, P(B) > 0. \tag{6}$$

Exemplo: Um número é sorteado ao acaso entre os inteiros $1, 2, 3, \ldots, 15$. Se o número sorteado for ímpar, qual a probabilidade de que seja o número 9?

Probabilidade da Intersecção de Dois Eventos

A probabilidade condicional permite-nos calcular diretamente a probabilidade da intersecção de dois eventos. Assim,

$$P(A \cap B) = P(B)P(A|B). \tag{7}$$

Exemplo: Considere os seguintes eventos:

- A: retirar uma carta de copas do baralho.
- B: retirar um às do baralho.

Determine a probabilidade desses eventos ocorrerem simultâneamente.

Independência de Eventos

Dois eventos são considerados independentes quando a ocorrência de um deles não depende da ocorrência do outro, isto é,

$$P(A|B) = P(A) \text{e} P(B|A) = P(B). \quad (8)$$

Logo, o teorema do produto para dois eventos independentes é dado por:

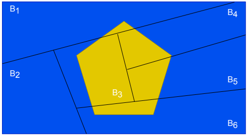
$$P(A \cap B) = P(A)P(B). \quad (9)$$

Exemplo: Suponha que a probabilidade de uma pessoa ser do tipo sanguíneo O é de 40%, ser A é de 30%, e ser B é de 20%. Suponha ainda que a probabilidade de RH+ é de 90% e que o fator independe do tipo sanguíneo. Nestas condições, qual a probabilidade de uma pessoa tomada ao acaso da população ser:

- A: O e RH+?
- B: AB ou RH-?

Partição do Espaço Amostral

Definição: $B_1, B_2, B_3, \ldots, B_n$ formam uma partição do espaço amostral se:



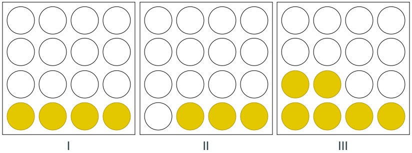
$B_i \cap B_j = \emptyset.$
$\bigcup_{i=1}^{n} B_i = \Omega.$
$P(B_i) \geq 0, i = 1, \ldots, n.$

Seja A um evento no espaço amostral Ω e seja $B_1, B_2, B_3, \ldots, B_n$ partições do espaço amostral Ω. Podemos escrever A considerando tais partições:

$$A = \bigcup_{i=1}^{n} A \cap B_i. \tag{10}$$

$$P(A) = \sum_i P(A|B_i)P(B_i), i = 1, 2, \ldots \quad (11)$$

Exemplo: Moedas de ouro e prata são colocadas em 3 urnas, conforme figura:



Cada urna tem uma probabilidade de escolha: $P(I) = \frac{1}{2}$, $P(II) = \frac{1}{4}$, e $P(III) = \frac{1}{4}$. Qual é a probabilidade de se escolher uma moeda de ouro?

Teorema da Probabilidade Total

Exemplo: Em uma indústria, bolas são feitas em duas máquinas. A máquina 1 é responsável por 75% da produção. Das bolas produzidas na máquina 1, 20% é azul e o restante é vermelha. A máquina 2 é responsável por 25% da produção. Das bolas produzidas na máquina 2, 30% é azul e o resto é vermelha.

- Qual a probabilidade de ser produzida uma bola azul?
- Sabendo-se que uma bola azul foi produzida, qual a probabilidade de ela ter sido feita na máquina 1?

Teorema de Bayes

Teorema de Bayes

O **Teorema de Bayes** descreve a probabilidade de uma hipótese com base em evidências observadas. Ele permite atualizar probabilidades a partir de novas informações e é amplamente utilizado em áreas como medicina, aprendizado de máquina e estatística.

Teorema de Bayes

O **Teorema de Bayes** descreve a probabilidade de uma hipótese com base em evidências observadas. Ele permite atualizar probabilidades a partir de novas informações e é amplamente utilizado em áreas como medicina, aprendizado de máquina e estatística.

“Probabilidade é uma opinião metódica; e a inferência de dados nada mais é do que a revisão de tal opinião com novas informações relevantes”. (Thomas Bayes, 1701 - 1761)

Rede Bayesiana

Representação do Conhecimento

Como representar o conhecimento em sistemas inteligentes que utilizam o raciocínio probabilístico?

Uma alternativa é usar **Redes Bayesianas**.

Redes Bayesianas

Rede Bayesiana é uma ferramenta gráfica para raciocínio e representação de conhecimento frente a incertezas.

Ela é uma representação compacta da distribuição de probabilidades conjuntas do universo do problema.

Redes Bayesianas

Formalmente, é um grafo acíclico direcionado.

Os nós (também chamados de variáveis) são os eventos que queremos modelar.

Os arcos ligando os nós indicam a presença de uma dependência condicional entre eles.

Cada nó possui uma tabela de probabilidades, dizendo as chances de ocorrer o evento representado por ele (inclusive probabilidade condicional).

Vantagens

- A maior vantagem do raciocínio probabilístico, em relação ao raciocínio fundamentado em lógica dedutiva, é permitir chegar a decisões racionais mesmo quando não há informação suficiente para provar qualquer das hipóteses.
- Raciocínio probabilístico é capaz de tratar grau de incerteza de domínios bem variados, além de propor representações parcimoniosas.
- Redes Bayesianas podem ser diretamente geradas a partir de dados coletados de processos do mundo real.

Limitações

- Requer conhecimento a priori.
- Número de variáveis do problema é alto.
- A relação entre as variáveis pode ser muito complexa.
- Redes Bayesianas não modelam comportamentos dinâmicos.

Naïve Bayes

Teoria da Decisão Bayesiana

- $P(C_1|\mathbf{x})$: probabilidade da observação com atributos $\mathbf{x}$ pertencer à classe C_1.
- $P(C_2|\mathbf{x})$: probabilidade da observação com atributos $\mathbf{x}$ pertencer à classe C_2.

Teoria da Decisão Bayesiana

- Se $P(C_1|\mathbf{x}) > P(C_2|\mathbf{x})$ então classifique $\mathbf{x} \in C_1$.
- Se $P(C_2|\mathbf{x}) > P(C_1|\mathbf{x})$ então classifique $\mathbf{x} \in C_2$..

Teoria da Decisão Bayesiana

Na decisão Bayesiana, a principal limitação é que na maioria das vezes, não conhecemos a distribuição de probabilidade:

$$P(C_i|\mathbf{x}) = \frac{p(\mathbf{x}|C_i)P(C_i)}{p(x)}. \tag{12}$$

A sua estimação é bastante complicada, principalmente se a dimensão dos dados é muito alta.

Naïve Bayes

Naïve Bayes: assume independência entre os atributos:

$$P(C_i|\mathbf{x}) = \frac{p(\mathbf{x}|C_i)P(C_i)}{p(x)}, \tag{13}$$

$$p(\mathbf{x}|C_i) = \prod_{j=1}^{d} p(x_j|C_i), i = 1, 2, \ldots, k. \tag{14}$$

Classificação usa a regra de Bayes:

$$C_m = \operatorname*{arg\,max}_{i \in \{1,\ldots,M\}} p(C_i) \prod_{j=1}^{d} p(x_j|C_i). \tag{15}$$

Naïve Bayes

Exemplo: Como diagnosticar um paciente?

$\mathbf{x}$ = (Calafrios=Sim, Coriza=Não, Cefaleia=média, Febre=Não).

x_1 = Calafrios	x_2 = Coriza	x_3 = Cefaleia	x_4 = Febre	Gripe
Sim	Não	Média	Sim	Não
Sim	Sim	Não	Não	Sim
Sim	Não	Forte	Sim	Sim
Não	Sim	Média	Sim	Sim
Não	Não	Não	Não	Não
Não	Sim	Forte	Sim	Sim
Não	Sim	Forte	Não	Não
Sim	Sim	Média	Sim	Sim

Naïve Bayes

Exemplo: Como diagnosticar um paciente?

$\mathbf{x}$ = (Calafrios=Sim, Coriza=Não, Cefaleia=média, Febre=Não).

Precisamos calcular:

P(Gripe=Sim—$\mathbf{x}$) e P(Gripe=Não—$\mathbf{x}$).

x_1 = Calafrios	x_2 = Coriza	x_3 = Cefaleia	x_4 = Febre	Gripe
Sim	Não	Média	Sim	Não
Sim	Sim	Não	Não	Sim
Sim	Não	Forte	Sim	Sim
Não	Sim	Média	Sim	Sim
Não	Não	Não	Não	Não
Não	Sim	Forte	Sim	Sim
Não	Sim	Forte	Não	Não
Sim	Sim	Média	Sim	Sim

Naïve Bayes: atributos contínuos

Para cada atributo x_j, a sua distribuição de probabilidade é assumida como normal:

$$p(x_j|C_i) = \frac{1}{\sqrt{(2\pi\sigma_{C_i})}} \exp\left[-\frac{1}{2}\left(\frac{x_j - \mu_{C_i}}{\sigma_{C_i}}\right)^2\right]. \tag{16}$$

Assuma independência e calcule a distribuição conjunta para cada nova observação q:

$$p(x_q|C_i) = \prod_{j=1}^{d} p(x_{qj}|C_i), i = 1, 2, \ldots, k. \tag{17}$$

Classifique de acordo com a classe mais provável.

$$\hat{y} = \underset{i=1,\ldots,k}{\arg\max}\, P(C_i) \prod_{j=1}^{d} p(x_{qj}|C_i). \tag{18}$$

Perguntas?

Obrigado pela atenção!

RECOGNA
LABORATORY