Redes Bayesianas

Índice

[Fórmula de Bayes 2](#_Toc140070982)

[Independência Condicional 4](#_Toc140070983)

[Redes Bayesianas 4](#_Toc140070984)

[Modelo 5](#_Toc140070985)

[Cálculos com Redes Bayesianas 7](#_Toc140070986)

[Exemplo 1: Cálculo de Probabilidades Conjuntas 9](#_Toc140070987)

[Exemplo 2: Inferência Causal 9](#_Toc140070988)

[Exemplo 3: Inferência Causal 11](#_Toc140070989)

[Exemplo 4: Inferência Causal 12](#_Toc140070990)

[Exemplo 5: Inferência de Diagnóstico 12](#_Toc140070991)

Probabilidade – Revisão

Ideia base de um classificador Bayesiano:

* Estimar a probabilidade de uma dada classificação ou **hipótese** H, conhecidos os valores dos atributos de um novo exemplo
* Formalmente, calcular:
  + p(Hi| **X**ij), com Hi = Hipótese de ordem i, **X**j = vector de atributos (xi1, xi2…xin) e p(H|X) = probabilidade *a posteriori* ou probabilidade condicionada (p de H suposto **X**)
* Para isso bastaria:
  + Dispor de um dataset de treino composto pelas diversas combinações possíveis de **todos** os atributos com **todas** as classes, i.e.:
    - Para cada combinação possível dos atributos, contar as ocorrências de cada classe, Hi e as respetivas ocorrências da combinação **X**ij

## Fórmula de Bayes



**Exemplo**



Uma imagem com texto, Tipo de letra, file, captura de ecrã

Descrição gerada automaticamente

**Exemplo igual ao anterior, mas com mais tuplos**

**Uma imagem com texto, captura de ecrã, Tipo de letra, algebra

Descrição gerada automaticamente**

O exemplo anterior mostra que:

* Para calcular diretamente P(H|X) das tabelas, é necessário conhecer um número enorme de exemplos, dado que têm de se considerar:
  + Todas as combinações possíveis dos N atributos
  + Multiplicar essas combinações por todas as classes
  + Ter um número “elevado” de instâncias de cada combinação de atributos/classe para se poder aproximar a probabilidade através das frequências relativas calculadas

**Exemplo**

5 atributos, cada um com 3 valores, 4 classes:

* 35\*4=972
* Se se pretendesse uma média de 100 exemplos por cada combinação de atributos / classe, ter-se-ia necessidade de cerca de 972 \* 100 = 97200 exemplos !

## Independência Condicional

Formalmente, sendo X, Y e Z variáveis aleatórias, X é independente de Y se:



# Redes Bayesianas

Este classificador não exige a independência condicional de todos os atributos dentro de uma mesma classe.

Pelo contrário, este modelo permite traduzir dependências condicionais, mas apenas as “importantes”.

* Situa-se assim entre as exigências da Fórmula de Bayes original (dependência condicional de todos os atributos, da classe) e as do classificador Naive Bayes (independência condicional de todos os atributos, da classe)

Aspeto geral de uma rede Bayesiana é o de um grafo acíclico e dirigido

* As dependências são traduzidas por setas, entre atributos
* Os atributos constituem nós e contêm tabelas de probabilidade condicionada que os associam aos nós pai
* O grafo não pode conter quaisquer ciclos fechados

## Modelo

Sejam 3 variáveis aleatórias A, B e C

* Se A for independente de B, e A e B influenciarem, ambas, C, a rede correspondente terá o seguinte aspeto:

Uma imagem com file, Tipo de letra, branco

Descrição gerada automaticamente

Uma imagem com círculo, esboço, branco, preto e branco

Descrição gerada automaticamente

Na rede seguinte:

* D é um antepassado de B
* B é um descendente de D

A cada nó são assinaladas as seguintes tabelas de probabilidade:

* Se um nó não tem pais, a tabela é do tipo p(X)
* Se um nó tem apenas 1 pai, a tabela é do tipo P(X|Y)
* Se um nó tem vários pais, a tabela é do tipo p(X|Y1, Y2…Yn)

De acordo com este modelo, a independência condicional entre atributos feita pelo Bayes Classifier pode ser assim visualizada:

Uma imagem com texto, file, Tipo de letra

Descrição gerada automaticamente

Uma imagem com texto, diagrama, file, Esquema

Descrição gerada automaticamente

Desenhar um Rede Bayesiana envolve 2 fases:

* Desenhar a rede
* Estimar as probabilidades condicionadas envolvidas (através de históricos ou opinião de peritos)

Para desenhar a rede há duas opções:

* Captar as relações de causa-efeito e traçar as setas das causas para os efeitos
* Proceder sistematicamente (ver rede anterior):
  + Ordenar as variáveis. P.e.: E,D,HD,Hb,CP,BP
  + P(D|E) pode simplificar-se para p(D)
  + P(HD|E,D) não pode simplificar-se
  + P(Hb|HD,E,D) pode simplificar-se para p(Hb|D)
  + P(CP|Hb,HD,E,D) pode simplificar-se para p(CP|Hb,HD)
  + P(BP|CP,Hb,HD,E,D) pode simplificar-se para p(BP|HD)
* Esta abordagem pode gerar modelos diferentes consoante a ordenação original das variáveis, mas garante ausência de ciclos
* Dificuldade: para *d* atributos há *d!* grafos distintos

## Cálculos com Redes Bayesianas

Uma rede Bayesiana permite 3 tipos de cálculos base:

1. Probabilidade conjunta de qualquer acontecimento



1. Inferência
   1. Causal (das causas para os efeitos)
   2. Diagnóstico (dos efeitos para as causas)

Uma imagem com Tipo de letra, escrita à mão, texto, branco

Descrição gerada automaticamente



Existem ainda outros tipos de cálculos baseados neste, nomeadamente **Intercausal** (estudo da influência de novas causas sobre um efeito comum) e **Misto** (mistura de várias destas formas num só cálculo)

Note-se por exemplo:



Isto significa que, conhecidas as probabilidades conjuntas relativas a uma rede, pode aplicar-se a equação acima para realizar uma inferência de diagnóstico

Por outro lado, como vimos, para calcular estas probabilidades conjuntas basta aplicar:



Que permite calculá-las a partir das probabilidades condicionadas que figuram nos nós da rede.

Uma rede Bayesiana permite:

* Reduzir muito o número de probabilidades que é necessária conhecer inicialmente para implementar a rede
* Que essas probabilidades sejam do tipo condicionada e não conjuntas (importante no caso de a rede ser modelizada por recurso a opinião de peritos e não por análise de dados)

**Exemplo**

**Uma imagem com texto, captura de ecrã, file, Tipo de letra

Descrição gerada automaticamente**

# Exemplo 1: Cálculo de Probabilidades Conjuntas

Uma casa possui um alarme que toca quando há assaltos, mas por vezes, também quando há um tremor de terra. Quando toca, os vizinhos João e Maria telefonam ao dono, segundo as probabilidades assinaladas na rede Bayesiana da figura.

Calcular a probabilidade de John e Mary telefonarem ambos, o alarme tocar e não ocorrer nenhum roubo nem tremor de terra:

Uma imagem com texto, Tipo de letra, branco, algebra

Descrição gerada automaticamenteUma imagem com diagrama, texto, esboço, file

Descrição gerada automaticamente

# Exemplo 2: Inferência Causal

* Um indivíduo X pratica exercício.
* De acordo com a rede Bayesiana da figura, deverá ser classificado como candidato ou não candidato a uma doença de coração?

Uma imagem com texto, diagrama, file, Esquema

Descrição gerada automaticamente

1. A rede especifica que **p(Exercise=Yes) = 0.7**
2. Por se tratar de um nó sem pai, esta probabilidade significa que, retirada uma amostra da população, 70% dos indivíduos praticam exercício
3. Se nada for dito em contrário, é com esta probabilidade que se lida no calculo de outras probabilidades da rede
4. Porém, neste caso o enunciado especifica o valor do atributo Exercise ao afirmar que X pratica exercício: ou seja, p(Exercise=Yes) = 1 (logo, p(Exercise=No) = 0)
5. Contudo, para dieta nada se sabe (o enunciado é omisso) e, portanto, X pode ou não ter uma dieta saudável
6. O problema consiste em calcular p(HD=Yes), sabendo que para X p(Exercise=Yes) = 1 e nada sabendo de p(Diet=Healhty), pelo que assumem as probabilidades dadas nas tabelas de Diet. Ou seja:

Uma imagem com texto, Tipo de letra, algebra

Descrição gerada automaticamenteUma imagem com texto, Tipo de letra, número, recibo

Descrição gerada automaticamente

Falta agora determinar a probabilidade de **NÃO ter doença do coração**:

Uma imagem com texto, Tipo de letra

Descrição gerada automaticamente

* Como 0.6 > 0.4 o indivíduo X é classificado como **não candidato** a doença do coração
* Notar que o segundo cálculo é o complementar do primeiro, pelo que poderia ter-se feito assim:

Uma imagem com texto, Tipo de letra, captura de ecrã, file

Descrição gerada automaticamente

# Exemplo 3: Inferência Causal

Sem qualquer informação adicional, e no cenário descrito por esta rede, inferir qual a probabilidade de uma pessoa ter uma doença do coração (HD).

Uma imagem com texto, captura de ecrã, Tipo de letra, algebra

Descrição gerada automaticamente

Como 0.51 > 0.49, na ausência de informação adicional há uma ligeira probabilidade a favor de não se contrair doença do coração (HD).

# Exemplo 4: Inferência Causal

* Qual a probabilidade de uma pessoa ter tensão alta (BP=H) sabendo que ela pratica exercício e tem dieta saudável?

Uma imagem com texto, captura de ecrã, Tipo de letra

Descrição gerada automaticamenteUma imagem com texto, diagrama, file, Esquema

Descrição gerada automaticamente

# Exemplo 5: Inferência de Diagnóstico

* Qual a probabilidade de doença do coração para um doente que tem tensão arterial elevada, não faz exercício nem tem uma dieta saudável?

Uma imagem com texto, diagrama, file, Esquema

Descrição gerada automaticamenteUma imagem com texto, Tipo de letra, escrita à mão, captura de ecrã

Descrição gerada automaticamente