

Modelação e inferência em Redes Bayesianas

O tópico destas aulas é modelação de redes Bayesianas e sua utilização para obter inferências. Será utilizada a ferramenta Netica (<http://www.norsys.com>), devendo os resultados ser confirmados analiticamente.

Exercício Tutorial

Para se familiarizar com o Netica resolva o seguinte Exercício Tutorial. É sabido que as causas principais de cancro do pulmão são o tabaco e a exposição ao gás radão. Estima-se que 1% da população esteja exposta ao gás radão e que 30% da população tenha pais fumadores. A probabilidade de se ser fumador dado que os pais o são é de 40%, sendo 20% a probabilidade de se ser fumador quando os pais não o são. Quando os pais são fumadores, tem-se 80% de probabilidade de se ser fumador passivo, enquanto a probabilidade de se ser fumador passivo noutras circunstâncias é de 30%. Nos casos em que não há exposição ao gás radão, a probabilidade de se ter cancro no pulmão dado que se é fumador e fumador passivo é 6%; quando se é só fumador a probabilidade é 4% enquanto para os fumadores passivos é 1%; nos restantes casos a probabilidade de se ter cancro no pulmão é de 0,05%. A exposição ao gás radão decuplica (aumenta 10x) a probabilidade de se ter cancro de pulmão para os fumadores, passivos ou não. Nas restantes situações, com exposição ao gás radão, a probabilidade de se ter cancro do pulmão aumenta para o dobro.

1. Modele a situação anterior com uma rede de Bayes, indicando as variáveis aleatórias, seus domínios, topologia da rede e tabelas de probabilidade condicionada.
2. Determine a probabilidade do acontecimento conjunto em que: não se é fumador passivo, não se tem pais fumadores nem se foi exposto ao gás radão.
3. Determine a probabilidade de se ter cancro no pulmão dado que não se tem pais fumadores mas que se esteve exposto ao gás radão.

Exercícios Práticos

Modele cada uma das seguintes situações com uma Rede de Bayes, indicando as variáveis aleatórias, seus domínios, topologia da rede e tabelas de probabilidade condicionada, e use-a, recorrendo ao Nética, para calcular a resposta às perguntas indicadas.

I - A federação de futebol de um país europeu decidiu investir num sistema que lhe permita raciocinar sobre a probabilidade dos seus clubes de topo virem a vencer a Liga dos Campeões. No caso deste país, a probabilidade de um clube vencer a Liga dos Campeões é de 20% se tiver vencido o Campeonato Nacional, e de 1% caso contrário. Vencer o Campeonato Nacional pode ser o resultado de ter o melhor plantel (conjunto de jogadores) a nível nacional, ou de comprar os árbitros. Se tiver o melhor plantel, a probabilidade de vencer o Campeonato Nacional é de 70% independentemente de ter ou não comprado os árbitros. Se não tiver o melhor plantel, tem uma probabilidade de 20% de vencer o Campeonato Nacional no caso de ter comprado os árbitros, descendo para 5% caso contrário. Se um clube comprar os árbitros, existe uma probabilidade de 30% de algum dos seus dirigentes ser preso. Mesmo não comprando os árbitros, existe uma probabilidade de 2% de algum dos seus dirigentes ser preso. Para os clubes considerados, estima-se que a probabilidade de ter o melhor plantel a nível nacional é de 20%, e a probabilidade de comprar os árbitros de 30%.

1. Qual a probabilidade de uma destas equipas vencer o Campeonato Nacional?
2. Sabendo que uma destas equipas venceu a Liga dos Campeões, e que nenhum dos seus dirigentes foi preso, qual a probabilidade desse clube ter o melhor plantel?

II - A Direção Geral de Contribuições e Impostos (DGCI) classifica os rendimentos dos contribuintes em baixos, médios ou elevados. Estatisticamente, sabe-se que 30% dos rendimentos são baixos e que 64% são médios. Os contribuintes de rendimentos baixos não entregam a declaração de impostos; os contribuintes de rendimentos médios e elevados entregam a sua declaração de impostos em 90% e 60% dos casos, respectivamente. Dos contribuintes que entregam a declaração, 3% são fiscalizados, subindo esta percentagem de fiscalização para 10% para aqueles que não entregam qualquer declaração. Dado o bom funcionamento do sistema de fiscalização, 99% dos contribuintes sujeitos a fiscalização pagam o montante correto de impostos independentemente do seu nível rendimentos. Para os contribuintes não fiscalizados, verifica-se que 10% dos contribuintes de rendimentos baixos não paga o valor correto de impostos, aumentando esta percentagem para 20% e 25% no caso de contribuintes de rendimentos médios e elevados, respectivamente.

1. Qual a probabilidade dos rendimentos serem elevados e de se ter pago o valor correto dos impostos?
2. Sabendo que um contribuinte que pagou o montante correto dos impostos foi fiscalizado, qual a probabilidade dos seus rendimentos serem elevados?

III - Considere o seguinte cenário. A energia para o funcionamento de um Rover Marciano é originada em baterias de longa duração e painéis solares de alto rendimento. A produção de energia dos painéis solares depende de vários factores, nomeadamente se é dia ou noite, se os painéis estão abertos e se existe poeira no ar (as tempestades em Marte são das mais violentas no Sistema Solar). Obviamente, durante a noite Marciana não existe produção de energia; nesta altura do ano Marciano o dia tem uma duração aproximada de 20 horas no local onde o Rover se encontra. Os painéis solares do Rover encontram-se quase sempre abertos, mesmo durante a noite. A maior parte do tempo existe alguma poeira no ar e, ocasionalmente, há muita poeira na atmosfera; a probabilidade de haver muita poeira no ar é cerca de metade da probabilidade de não haver nenhuma poeira. A probabilidade de se conseguir produzir energia eléctrica com a atmosfera com muita poeira é baixa; contudo, se a atmosfera estiver limpa, verifica-se exactamente o oposto.

A visibilidade é afectada essencialmente pelo facto de ser dia ou noite e pela existência de poeira na atmosfera. A visibilidade é nula durante a noite. Nos restantes casos quanto mais poeira existir na atmosfera mais reduzida é a visibilidade. Quando há muita poeira na atmosfera então só em 20% dos casos é que a visibilidade é normal, sendo pouca na maioria das situações. Quando a visibilidade é nula, o Rover adormece. Nas outras situações, a probabilidade do Rover se encontrar adormecido é quase nula (de vez em quando acontecem uns reboots...).

Os sistemas de comunicações com a Terra são alimentados apenas pela bateria. Quando a bateria se encontra carregada e não existe muita poeira na atmosfera, é quase sempre possível comunicar com a Terra; a existência de alguma poeira na atmosfera afecta ligeiramente as comunicações. Quando a bateria se encontra descarregada só é possível a comunicação quando não há nenhuma poeira na atmosfera, mesmo assim num número diminuto de vezes.

A tensão eléctrica no motor do Rover depende directamente do estado da bateria e da produção de energia pelos painéis solares. A tensão do motor é garantida essencialmente pelos painéis solares, servindo a bateria de backup para casos de emergência não podendo ser utilizada durante muito tempo. Quando os painéis solares não produzem energia, a tensão no motor nunca é elevada. Se a bateria estiver descarregada e os painéis não produzem energia, então quase sempre a tensão no motor é fraca. Quando a bateria está carregada ou os painéis produzem energia eléctrica, então em mais de 80% dos casos a tensão é normal.

Quando a tensão no motor é fraca e a visibilidade normal, o Rover desloca-se à velocidade normal apenas em 5% dos casos, encontrando-se quase sempre parado. Quando as condições de visibilidade são piores, a probabilidade de estar parado ou em velocidade lenta aumentam. A probabilidade de o Rover se deslocar a velocidade lenta, com as condições de visibilidade nulas e tensão fraca, é de 1%. Por questões de segurança, o Rover encontra-se muitas vezes parado quando as condições de visibilidade são abaixo do normal. Normalmente, o deslocamento do Rover a velocidade rápida só se verifica em condições de tensão normal ou elevada e visibilidade não nula. Quando a tensão é elevada há tendência para o Rover se deslocar mais rapidamente.

1. Qual a probabilidade de ser dia dado que a tensão no motor é normal e que os painéis estão abertos?
2. Qual a probabilidade da bateria estar descarregada dado que a tensão no motor é normal, existe produção de energia solar pelos painéis, não existe poeira na atmosfera e as comunicações estão a funcionar?
3. Qual a probabilidade do Rover estar a dormir?

Exercícios Analíticos

Para cada um dos exercícios anteriores, calcule, de forma analítica i.e. sem recorrer ao Nética, as probabilidades indicadas.