



**Data Science
Academy**

www.datascienceacademy.com.br

Introdução à Inteligência Artificial

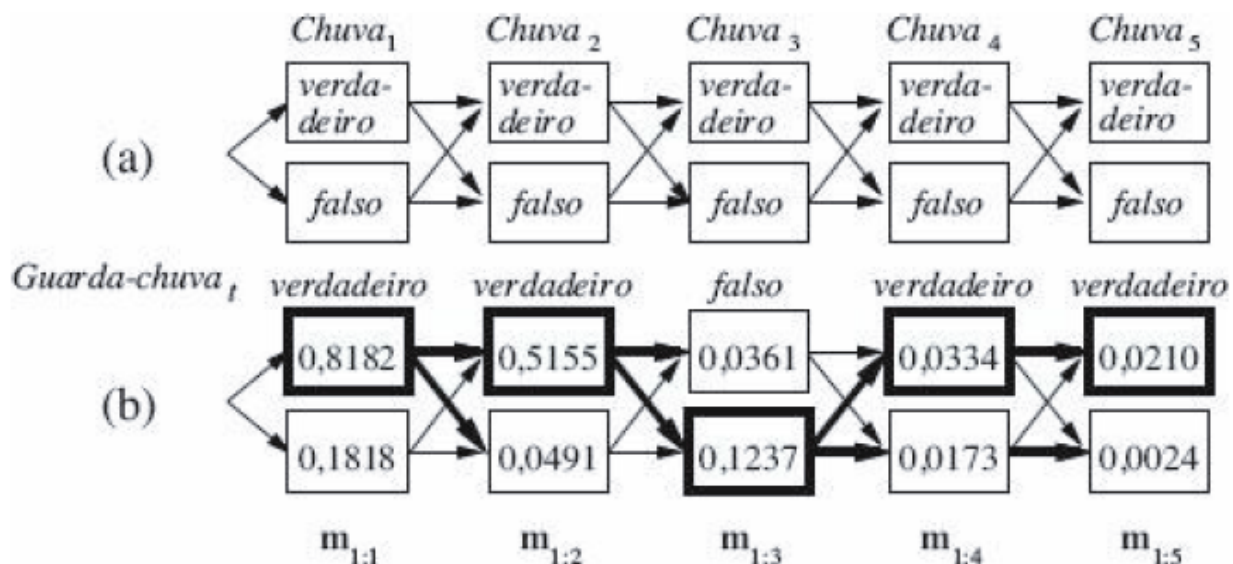
Como Descobrir a Sequência Mais Provável



Suponha que [verdadeiro, verdadeiro, falso, verdadeiro, verdadeiro] seja a sequência de guarda-chuva para os cinco primeiros dias de serviço do guarda de segurança. Qual é a sequência de condições do clima com maior probabilidade de explicar isso? A ausência do guarda-chuva no dia 3 significa que não estava chovendo ou será que o diretor se esqueceu de trazer o guarda-chuva? Se não tivesse chovido no dia 3, talvez (porque as condições do clima tendem a persistir) também não tivesse chovido no dia 4, mas o diretor pode ter trazido o guarda-chuva só por precaução. Ao todo, existem 25 possíveis sequências de condições do clima que poderíamos selecionar. Existe um meio de encontrar a mais provável que não seja enumerando todas elas?

Poderíamos tentar o procedimento de tempo linear a seguir: usar o algoritmo de suavização para descobrir a distribuição posterior para as condições do clima em cada período; em seguida, construir a sequência, utilizando em cada passo as condições do clima mais provavelmente concordantes com a distribuição posterior. Tal abordagem deve fazer soar um alarme na sua cabeça porque as distribuições posteriores calculadas pela suavização são distribuições sobre períodos de tempo isolados; por outro lado, para encontrar a sequência mais provável temos de considerar probabilidades conjuntas sobre todos os períodos de tempo. Os resultados podem de fato ser bastante diferentes.

Existe um algoritmo de tempo linear para encontrar a sequência mais provável, mas ele exige um pouco mais de reflexão. Esse algoritmo se baseia na mesma propriedade de Markov que gerou algoritmos eficientes para filtragem e suavização. O modo mais fácil de pensar no problema é visualizar cada sequência como um caminho por um grafo cujos nós são os estados possíveis em cada período de tempo. Mostramos esse grafo para o mundo do guarda-chuva na Figura abaixo. Agora, considere a tarefa de descobrir o caminho mais provável por esse grafo, onde a probabilidade de qualquer caminho é o produto das probabilidades de transição ao longo do caminho pelas probabilidades das observações dadas em cada estado. Vamos nos concentrar em particular nos caminhos que alcançam o estado Chuva5 = verdadeiro. Devido à propriedade de Markov, verificamos que o caminho mais provável para o estado Chuva5 = verdadeiro consiste no caminho mais provável até algum estado no tempo 4, seguido por uma transição para Chuva5 = verdadeiro; e o estado no tempo 4 que se tornará parte do caminho para Chuva5 = verdadeiro será aquele que maximizar a probabilidade desse caminho. Em outras palavras, existe um relacionamento recursivo entre caminhos mais prováveis até cada estado x_{t+1} e caminhos mais prováveis até cada estado x_t . Podemos escrever esse relacionamento como uma equação que conecta as probabilidades dos caminhos.



(a) Sequências de estados possíveis para Chuva_t podem ser visualizadas como caminhos através de um grafo dos estados possíveis em cada período de tempo (os estados são mostrados como nós quadrados para evitar confusão com os nós de uma rede bayesiana).

(b) Operação do algoritmo de Viterbi para a sequência de observação de guarda-chuva [verdadeiro, verdadeiro, falso, verdadeiro, verdadeiro]. Para cada período de tempo t , mostramos os valores da mensagem $m_{1:t}$, que fornece a probabilidade de a melhor sequência alcançar cada estado no tempo t . Além disso, para cada estado, a seta grossa que leva até ele indica seu melhor predecessor. Seguindo as setas grossas de volta a partir do estado mais provável em $m_{1:5}$, temos a sequência mais provável.

$$\begin{aligned} \max_{\mathbf{x}_1 \dots \mathbf{x}_t} P(\mathbf{x}_1, \dots, \mathbf{x}_t, \mathbf{X}_{t+1} | \mathbf{e}_{1:t+1}) \\ = \alpha P(\mathbf{e}_{t+1} | \mathbf{X}_{t+1}) \max_{\mathbf{x}_t} \left(P(\mathbf{X}_{t+1} | \mathbf{x}_t) \max_{\mathbf{x}_1 \dots \mathbf{x}_{t-1}} P(\mathbf{x}_1, \dots, \mathbf{x}_{t-1}, \mathbf{x}_t | \mathbf{e}_{1:t}) \right) \end{aligned}$$

Desse modo, o algoritmo para calcular a sequência mais provável é semelhante à filtragem: ele percorre a sequência no sentido para a frente calculando a mensagem m em cada período de tempo utilizando a equação anterior. No final, ele terá a probabilidade para a sequência mais provável que alcança cada um dos estados finais. Pode-se, portanto, selecionar com facilidade a sequência mais provável de todas (o estado com contorno em negrito). Com a finalidade de identificar a sequência real, em vez de simplesmente calcular sua probabilidade, o algoritmo também precisará registrar, para cada estado, o melhor estado que conduz a isso, o que está indicado pelas setas em negrito na Figura (b). A sequência ótima é identificada seguindo-se os ponteiros de volta a partir do melhor estado final.



O algoritmo que acabamos de descrever denomina-se algoritmo de Viterbi, em homenagem a seu criador. Como o algoritmo de filtragem, sua complexidade é linear em t , a duração da sequência. Porém, diferentemente da filtragem, que utiliza espaço constante, seu requisito de espaço também é linear em t . Isso ocorre porque o algoritmo de Viterbi precisa manter os ponteiros que identificam a melhor sequência que leva a cada estado.

Referências:

Livro: Inteligência Artificial

Autor: Peter Norvig