



**Data Science
Academy**

www.datascienceacademy.com.br

Introdução à Inteligência Artificial

Formulação Lógica da Aprendizagem

Em todas as abordagens para estudo da aprendizagem descritas nos capítulos anteriores, a ideia é construir uma função que tem o comportamento de entrada/saída observado nos dados. Em cada caso, os métodos de aprendizagem podem ser entendidos como a busca em um espaço de hipóteses para encontrar uma função apropriada, começando apenas por uma suposição muito básica sobre a forma da função, como “polinômio de segundo grau” ou “árvore de decisão” e, talvez, a preferência por uma hipótese mais simples. Isso significa dizer que, antes de poder aprender algo novo, primeiro você deve esquecer (quase) tudo o que sabe. Em muitos casos, métodos de aprendizagem podem tirar proveito do conhecimento a priori sobre o mundo. Na maioria dos casos aliás, o conhecimento a priori é representado como teorias lógicas gerais de primeira ordem; desse modo, pela primeira vez, juntamos o trabalho sobre representação do conhecimento e o de aprendizagem.

Vamos considerar o problema de aprendizagem do restaurante: aprender uma regra para decidir se devemos esperar por uma mesa. Os exemplos foram descritos por atributos como Alternativa, Bar, Sex/Sáb, e assim por diante. Em uma configuração lógica, um exemplo é um objeto descrito por uma sentença lógica; os atributos se tornam predicados unários. Vamos chamar genericamente o i -ésimo exemplo de X_i . Assim, podemos definir:

$$\text{Alternativa}(X_i) \wedge \neg \text{Bar}(X_i) \wedge \neg \text{Sex/Sáb}(X_i) \wedge \text{Faminto}(X_i) \wedge \dots$$

Usaremos a notação $D_i(X_i)$ para fazer referência à descrição de X_i , onde D_i pode ser qualquer expressão lógica que recebe um único argumento. A classificação do exemplo é dado por um literal utilizando o predicado objetivo, nesse caso

$$\text{VaiEsperar}(X_i) \quad \text{ou} \quad \neg \text{VaiEsperar}(X_i)$$

O conjunto de treinamento completo é então simplesmente a conjunção de todas as descrições de exemplos e literais objetivos. O objetivo da aprendizagem indutiva em geral é encontrar uma hipótese que classifique os exemplos bem e generalize bem para novos exemplos. Aqui estamos preocupados com hipóteses expressas em lógica; cada hipótese h_j terá a forma:

$$\forall x \text{ Objetivo}(x) \Leftrightarrow C_j(x)$$

onde $C_j(x)$ é uma definição candidata — alguma expressão envolvendo os predicados dos atributos. Por exemplo, uma árvore de decisão pode ser interpretada como uma expressão lógica dessa fórmula. Desse modo, a figura abaixo expressa a definição lógica a seguir (que chamaremos de h_j para referência futura):

$$\begin{aligned}\forall r \text{ VaiEsperar}(r) \Leftrightarrow & \text{Clientes}(r, \text{Alguns}) \\ & \vee \text{Clientes}(r, \text{Cheio}) \wedge \text{Faminto}(r) \wedge \text{Tipo}(r, \text{Francês}) \\ & \vee \text{Clientes}(r, \text{Cheio}) \wedge \text{Faminto}(r) \wedge \text{Tipo}(r, \text{Tailandês}) \\ & \quad \wedge \text{Sex/Sáb}(r) \\ & \vee \text{Clientes}(r, \text{Cheio}) \wedge \text{Faminto}(r) \wedge \text{Tipo}(r, \text{Hambúrguer})\end{aligned}$$

Cada hipótese prevê que certo conjunto de exemplos — ou seja, aqueles que satisfazem à sua definição candidata — será o conjunto de exemplos do predicado objetivo. Esse conjunto é chamado extensão do predicado.

Duas hipóteses com extensões diferentes são então logicamente inconsistentes uma em relação à outra porque elas discordam em suas previsões por pelo menos um exemplo. Se tiverem a mesma extensão, elas serão logicamente equivalentes.

O espaço de hipóteses é o conjunto de todas as hipóteses $\{h_1, \dots, h_n\}$ que o algoritmo de aprendizagem foi projetado para considerar. Por exemplo, o algoritmo APRENDIZAGEM-EM-ÁRVORE-DE-DECISÃO pode considerar qualquer hipótese de árvore de decisão definida em termos dos atributos fornecidos; portanto, seu espaço de hipóteses consiste em todas essas árvores de decisão. Presumivelmente, o algoritmo de aprendizagem acredita que uma das hipóteses seja correta; isto é, ele acredita na sentença

$$h_1 \vee h_2 \vee h_3 \vee \dots \vee h_n$$

À medida que os exemplos chegam, as hipóteses que não são consistentes com os exemplos podem ser eliminadas. Vamos examinar essa noção de consistência com mais atenção. É óbvio que, se a hipótese h_j é consistente com o conjunto de treinamento inteiro, ela tem de ser consistente com cada exemplo no conjunto de treinamento. O que significaria para ela o fato de ser inconsistente com um exemplo? Isso pode acontecer de duas maneiras:

1. Um exemplo pode ser falso negativo para a hipótese se a hipótese afirmar que ele deve ser negativo, mas de fato ele for positivo. Então, o novo exemplo X_{13} descrito por

$$\text{Clientes}(X_{13}, \text{Cheio}) \wedge \neg \text{Faminto}(X_{13}) \wedge \dots \wedge \text{VaiEsperar}(X_{13})$$

seria um falso negativo para a hipótese h_r dada anteriormente. A partir de h_r e da descrição do exemplo, podemos deduzir tanto $\text{VaiEsperar}(X_{13})$, que é a afirmação do exemplo, quanto $\neg \text{VaiEsperar}(X_{13})$, o que a hipótese prevê. A hipótese e o exemplo são então logicamente inconsistentes.



2. Um exemplo pode ser falso positivo para a hipótese se a hipótese afirmar que ele deve ser positivo, mas de fato ele for negativo

Se um exemplo é falso positivo ou falso negativo para uma hipótese, então o exemplo e a hipótese são logicamente inconsistentes um com o outro. Supondo-se que o exemplo seja uma observação correta do fato, a hipótese pode ser eliminada. Em termos lógicos, isso é exatamente análogo à regra de resolução de inferência, pela qual a disjunção de hipóteses corresponde a uma cláusula. Um sistema de inferência lógica comum poderia, em princípio, aprender a partir do exemplo, eliminando uma ou mais hipóteses. Vamos supor que o exemplo seja indicado pela sentença l_1 e que o espaço de hipóteses seja $h_1 \vee h_2 \vee h_3 \vee h_4$. Então, se l_1 é inconsistente com h_2 e h_3 , o sistema de inferência lógica pode deduzir o novo espaço de hipóteses $h_1 \vee h_4$.

Assim, podemos caracterizar a aprendizagem indutiva em uma configuração lógica como um processo de eliminação gradual de hipóteses que são inconsistentes com os exemplos, reduzindo as possibilidades. Como o espaço de hipóteses normalmente é vasto (ou até mesmo infinito, no caso da lógica de primeira ordem), não recomendamos tentar construir um sistema de aprendizagem usando a prova de teoremas baseada na resolução e uma enumeração completa do espaço de hipóteses. Em vez disso, descreveremos duas abordagens que encontram hipóteses logicamente consistentes com muito menos esforço.

Referências:

Livro: Inteligência Artificial

Autor: Peter Norvig