



Machine Learning

Classificação Bayesiana

Prof. Hugo de Paula

Classificação Bayesiana (*Naïve Bayes*)

Este método é baseado em classificador estatístico.

Trabalha com probabilidades de ocorrência de cada classe para cada valor de atributo.

Supõe que variáveis são condicionalmente independentes.

Classificação Bayesiana (*Naïve Bayes*)

Independência condicional

X é condicionalmente independente de Y dado Z, se a distribuição de probabilidade de X é independente de Y, dado o valor de Z.

$$P(X|Y, Z) = P(X|Z)$$

Classificação Bayesiana (*Naïve Bayes*)

$P(C|X)$ probabilidade do registro **X** ser da classe **C**

$$P(C|X) = P(C) \prod_{i=1}^n P(X_i|C)$$

Seleciona $P(C|X)$ máximo.

Correção de Laplace evita alta influência de valores com probabilidade 0.

Classificação Bayesiana: exemplo

Aparência	Temperatura	Umidade	Vento	Jogar
Ensolarado	Quente	Alta	Fraco	Não
Ensolarado	Quente	Alta	Forte	Não
Nublado	Quente	Alta	Fraco	Sim
Chuvoso	Moderado	Alta	Fraco	Sim
Chuvoso	Frio	Normal	Fraco	Sim
Chuvoso	Frio	Normal	Forte	Não
Nublado	Frio	Normal	Forte	Sim
Ensolarado	Moderado	Alta	Fraco	Não
Ensolarado	Frio	Normal	Fraco	Sim
Chuvoso	Moderado	Normal	Fraco	Sim
Ensolarado	Moderado	Normal	Forte	Sim
Nublado	Moderado	Alta	Forte	Sim
Nublado	Quente	Normal	Fraco	Sim
Chuvoso	Moderado	Alta	Forte	Não

O objetivo é identificar quais as condições ideais para se jogar um determinado jogo.

Classificação Bayesiana: exemplo

Aparência	
$P(\text{sol} \text{sim}) = 2/9$	$P(\text{sol} \text{não}) = 3/5$
$P(\text{nublado} \text{sim}) = 4/9$	$P(\text{nublado} \text{não}) = 0$
$P(\text{chuvoso} \text{sim}) = 3/9$	$P(\text{chuvoso} \text{não}) = 2/5$
Temperatura	
$P(\text{quente} \text{sim}) = 2/9$	$P(\text{quente} \text{não}) = 2/5$
$P(\text{moderado} \text{sim}) = 4/9$	$P(\text{moderado} \text{não}) = 2/5$
$P(\text{frio} \text{sim}) = 3/9$	$P(\text{frio} \text{não}) = 1/5$
Humidade	
$P(\text{alta} \text{sim}) = 3/9$	$P(\text{alta} \text{não}) = 4/5$
$P(\text{normal} \text{sim}) = 6/9$	$P(\text{normal} \text{não}) = 2/5$
Vento	
$P(\text{forte} \text{sim}) = 3/9$	$P(\text{forte} \text{não}) = 3/5$
$P(\text{fraco} \text{sim}) = 6/9$	$P(\text{fraco} \text{não}) = 2/5$

Jogar

$$P(\text{sim}) = 9/14$$

$$P(\text{não}) = 5/14$$

Classificação Bayesiana: exemplo

Dado $\mathbf{X} = \langle \text{chuvoso}, \text{quente}, \text{alta}, N \rangle$

$$\begin{aligned} P(\mathbf{X}|\mathbf{S}) \cdot P(\mathbf{S}) &= P(\text{chuvoso}|\mathbf{S}) \cdot P(\text{quente}|\mathbf{S}) \cdot P(\text{alta}|\mathbf{S}) \cdot P(\text{fraco}|\mathbf{S}) \cdot P(\mathbf{S}) \\ &= 3/9 \cdot 2/9 \cdot 3/9 \cdot 6/9 \cdot 9/14 = 0.010582 \end{aligned}$$

$$\begin{aligned} P(\mathbf{X}|\mathbf{N}) \cdot P(\mathbf{N}) &= P(\text{chuvoso}|\mathbf{N}) \cdot P(\text{quente}|\mathbf{N}) \cdot P(\text{alta}|\mathbf{N}) \cdot P(\text{fraco}|\mathbf{N}) \cdot P(\mathbf{N}) \\ &= 2/5 \cdot 2/5 \cdot 4/5 \cdot 2/5 \cdot 5/14 = 0.018286 \end{aligned}$$

Amostra classificada como **não jogar**

Aviso legal

O material presente nesta apresentação foi produzido a partir de informações próprias e coletadas de documentos obtidos publicamente a partir da Internet. Este material contém ilustrações adquiridas de bancos de imagens de origem privada ou pública, não possuindo a intenção de violar qualquer direito pertencente à terceiros e sendo voltado para fins acadêmicos ou meramente ilustrativos. Portanto, os textos, fotografias, imagens, logomarcas e sons presentes nesta apresentação se encontram protegidos por direitos autorais ou outros direitos de propriedade intelectual.

Ao usar este material, o usuário deverá respeitar todos os direitos de propriedade intelectual e industrial, os decorrentes da proteção de marcas registradas da mesma, bem como todos os direitos referentes a terceiros que por ventura estejam, ou estiveram, de alguma forma disponíveis nos slides. O simples acesso a este conteúdo não confere ao usuário qualquer direito de uso dos nomes, títulos, palavras, frases, marcas, dentre outras, que nele estejam, ou estiveram, disponíveis.

É vedada sua utilização para finalidades comerciais, publicitárias ou qualquer outra que contrarie a realidade para o qual foi concebido. Sendo que é proibida sua reprodução, distribuição, transmissão, exibição, publicação ou divulgação, total ou parcial, dos textos, figuras, gráficos e demais conteúdos descritos anteriormente, que compõem o presente material, sem prévia e expressa autorização de seu titular, sendo permitida somente a impressão de cópias para uso acadêmico e arquivo pessoal, sem que sejam separadas as partes, permitindo dar o fiel e real entendimento de seu conteúdo e objetivo. Em hipótese alguma o usuário adquirirá quaisquer direitos sobre os mesmos.

O usuário assume toda e qualquer responsabilidade, de caráter civil e/ou criminal, pela utilização indevida das informações, textos, gráficos, marcas, enfim, todo e qualquer direito de propriedade intelectual ou industrial deste material.



PUC Minas
Virtual

© PUC Minas • Todos os direitos reservados, de acordo com o art. 184 do Código Penal e com a lei 9.610 de 19 de fevereiro de 1998.
Proibidas a reprodução, a distribuição, a difusão, a execução pública, a locação e quaisquer outras modalidades de utilização sem a devida autorização da Pontifícia Universidade Católica de Minas Gerais.