

NAÏVE BAYES

Cristiane Neri Nobre

Introdução

Veja o link:

<https://www.youtube.com/watch?v=npl1uz8CMMY>

Introdução

História do crédito	Dívida	Garantias	Renda Anual	Risco
Ruim	Alta	Nenhuma	<15000	Alto
Desconhecida	Alta	Nenhuma	>=15000 a <=35000	Alto
Desconhecida	Baixa	Nenhuma	>=15000 a <=35000	Moderado
Desconhecida	Baixa	Nenhuma	>35000	Alto
Desconhecida	Baixa	Nenhuma	>35000	Baixo
Desconhecida	Baixa	Adequada	>35000	Baixo
Ruim	Baixa	Nenhuma	<15000	Alto
Ruim	Baixa	Adequada	>35000	Moderado
Boa	Baixa	Nenhuma	>35000	Baixo
Boa	Alta	Adequada	>35000	Baixo
Boa	Alta	Nenhuma	<15000	Alto
Boa	Alta	Nenhuma	>=15000 a <=35000	Moderado
Boa	Alta	Nenhuma	>35000	Baixo
Ruim	Alta	Nenhuma	>=15000 a <=35000	Alto

Gerando a tabela de probabilidades

	História de crédito			Dívida		Garantia		Renda anual		
Risco de crédito	Boa 5	Desconhe cida 5	Ruim 4	Alta 7	Baixa 7	Nenhuma 11	Adequada 3	<15000 3	>=15000 a <=35000 4	>35000 7
Alto 6/14	1/6	2/6	3/6	4/6	2/6	6/6	0/6	3/6	2/6	1/6
Moderado 3/14	1/3	1/3	1/3	1/3	2/3	2/3	1/3	0/3	2/3	1/3
Baixo 5/14	3/5	2/5	0/5	2/5	3/5	3/5	2/5	0/5	0/5	5/5

Como avaliar um novo registro, dada esta tabela?

Registro de teste:

História: Boa
Dívida: Alta
Garantias = Nenhuma
Renda = >35000

Para esta instância, qual é a probabilidade do risco ser:

Alto?
Moderado?
Baixo

	História de crédito			Dívida		Garantia		Renda anual		
	Boa	Desconhecida	Ruim	Alta	Baixa	Nenhuma	Adequada	<15000	>=15000 a <=35000	>35000
Risco de crédito	5	5	4	7	7	11	3	3	4	7
Alto 6/14	1/6	2/6	3/6	4/6	2/6	6/6	0/6	3/6	2/6	1/6
Moderado 3/14	1/3	1/3	1/3	1/3	2/3	2/3	1/3	0/3	2/3	1/3
Baixo 5/14	3/5	2/5	0/5	2/5	3/5	3/5	2/5	0/5	0/5	5/5

Como avaliar um novo registro, dada esta tabela?

Registro de teste:

História: Boa

Dívida: Alta

Garantias = Nenhuma

Renda = >35000

$$P(\text{alto}) = 6/14 * 1/6 * 4/6 * 6/6 * 1/6$$

$$P(\text{alto}) = 0,0079$$

$$P(\text{alto}) = 0,0079/0,0645 * 100 = 12,24\%$$

Soma:

$$0,0079 + 0,0052 + 0,0514 = 0,0645$$

$$P(\text{moderado}) = 3/14 * 1/3 * 1/3 * 2/3 * 1/3$$

$$P(\text{moderado}) = 0,0052$$

$$P(\text{moderado}) = 0,0052/0,0645 * 100 = 8,06\%$$

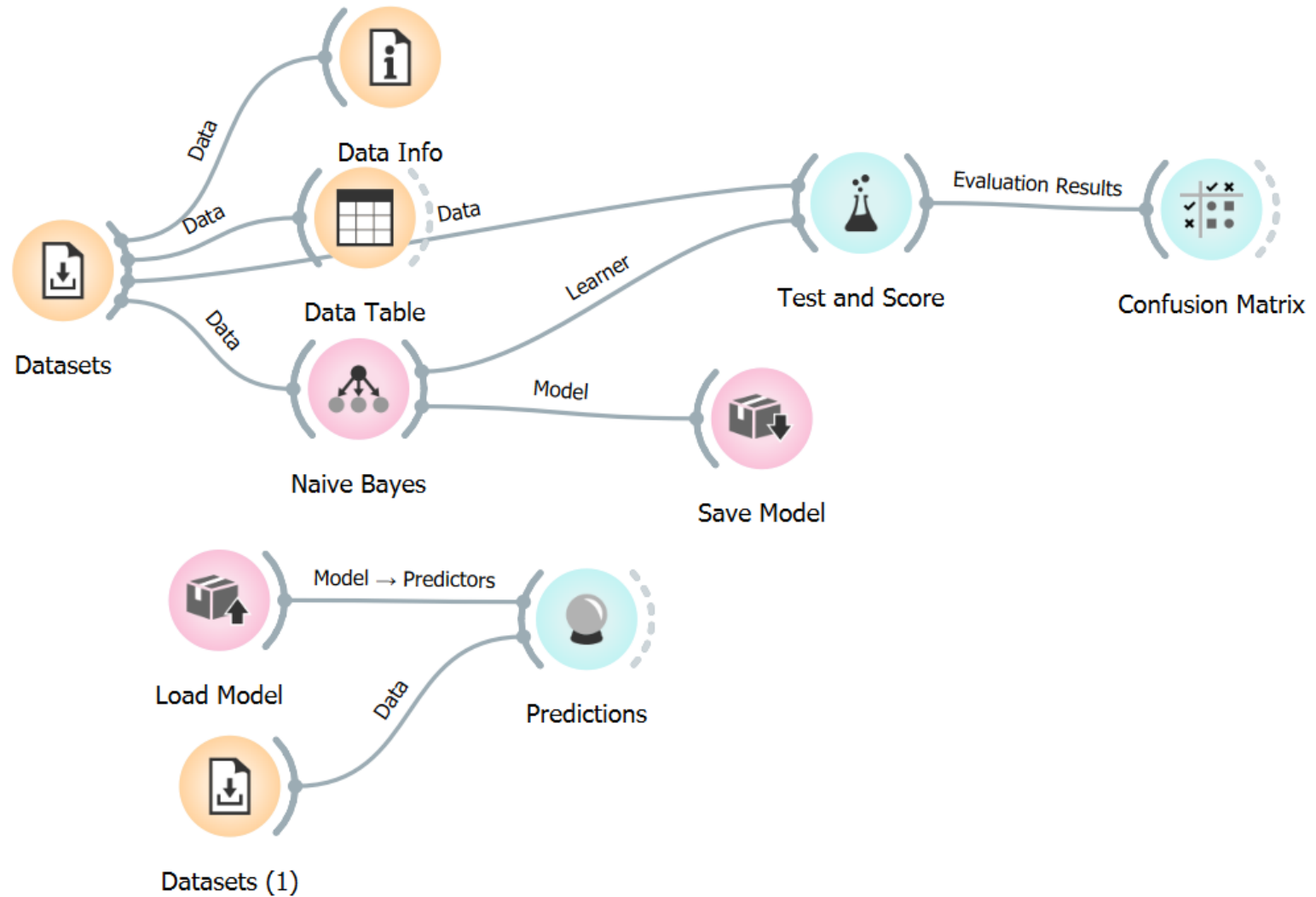
$$P(\text{baixo}) = 5/14 * 3/5 * 2/5 * 3/5 * 5/5$$

$$P(\text{baixo}) = 0,0514$$

$$P(\text{baixo}) = 0,0514/0,0645 * 100 = 79,68\%$$

	História de crédito			Dívida		Garantia		Renda anual		
	Boa	Desconhecida	Ruim	Alta	Baixa	Nenhuma	Adequada	<15000	>=15000 a <=35000	>35000
Risco de crédito	5	5	4	7	7	11	3	3	4	7
Alto 6/14	1/6	2/6	3/6	4/6	2/6	6/6	0/6	3/6	2/6	1/6
Moderado 3/14	1/3	1/3	1/3	1/3	2/3	2/3	1/3	0/3	2/3	1/3
Baixo 5/14	3/5	2/5	0/5	2/5	3/5	3/5	2/5	0/5	0/5	5/5

Naïve Bayes no Orange



Naïve Bayes no Orange

Vamos fazer um exercício para praticar?

Considere a tabela:

Dia	Aparência	temperatura	Umidade	Ventando	Jogar
d1	Sol	Quente	Alta	Não	Não
d2	Sol	Quente	Alta	Sim	Não
d3	Nublado	Quente	Alta	Não	Sim
d4	Chuva	Agradável	Alta	Não	Sim
d5	Chuva	Fria	Normal	Não	Sim
d6	Chuva	Fria	Normal	Sim	Não
d7	Nublado	Fria	Normal	Sim	Sim
d8	Sol	Agradável	Alta	Não	Não
d9	Sol	Fria	Normal	Não	Sim
d10	Chuva	Agradável	Normal	Não	Sim
d11	Sol	Agradável	Normal	Sim	Sim
d12	Nublado	Agradável	Alta	Sim	Sim
d13	Nublado	Quente	Normal	Não	Sim
d14	Chuva	Agradável	Alta	Sim	Não

Naïve Bayes no Orange

Qual a probabilidade de Jogar ou não Jogar para o seguinte registro:

Aparência = Sol

Temperatura = Fria

Umidade = Normal

Ventando = Não



Dia	Aparência	temperatura	Umidade	Ventando	Jogar
d1	Sol	Quente	Alta	Não	Não
d2	Sol	Quente	Alta	Sim	Não
d3	Nublado	Quente	Alta	Não	Sim
d4	Chuva	Agradável	Alta	Não	Sim
d5	Chuva	Fria	Normal	Não	Sim
d6	Chuva	Fria	Normal	Sim	Não
d7	Nublado	Fria	Normal	Sim	Sim
d8	Sol	Agradável	Alta	Não	Não
d9	Sol	Fria	Normal	Não	Sim
d10	Chuva	Agradável	Normal	Não	Sim
d11	Sol	Agradável	Normal	Sim	Sim
d12	Nublado	Agradável	Alta	Sim	Sim
d13	Nublado	Quente	Normal	Não	Sim
d14	Chuva	Agradável	Alta	Sim	Não

Entenda o teorema de Bayes

<https://www.youtube.com/watch?v=I643PqSrETM>

<https://www.youtube.com/watch?v=Wp4zRV5tvzY>