



# Aprendizagem por regra

## Conteúdo

- Teoria sobre indução de regras - algoritmo OneR
- Árvores de decisão x regras
- Regras (CN2) com Orange - base risco de crédito
- Regras (CN2) com Orange - base do crédito
- Regras (CN2) com Orange/interface gráfica - base censo
- Classificador base - (Majority Learner)

## Base Original ( Risco de Crédito )

Histórico	Divida	Garantias	Renda	Risco
ruim	alta	nenhuma	0_15	alto
desconhecida	alta	nenhuma	15_35	alto
desconhecida	baixa	nenhuma	15_35	moderado
desconhecida	baixa	nenhuma	acima_35	alto
desconhecida	baixa	nenhuma	acima_35	baixo
desconhecida	baixa	adequada	acima_35	baixo
ruim	baixa	nenhuma	0_15	alto
ruim	baixa	adequada	acima_35	moderado
Boa	baixa	nenhuma	acima_35	baixo
Boa	alta	adequada	acima_35	baixo
Boa	alta	nenhuma	0_15	alto
Boa	alta	nenhuma	15_35	moderado

Boa	alta	nenhuma	acima_35	baixo
ruim	alta	nenhuma	15_35	alto

## Regras - OneR

Regras	Resultado
Se renda $\Rightarrow$ 35.000 e História de credito = Boa	Risco = Baixo
Se a renda $\Rightarrow$ 35.000 e a História de credito = Desconhecida	Risco = Baixo
Default (Padrão)	Risco = Alto

Pessoa 1	Pessoa 2	Pessoa 3
História = boa	História = ruim	História = ruim
Divida = alta	Divida = Alta	Divida = Alta
Garantias = nenhuma	Garantias = Adequada	Garantias = Nenhuma
Renda = >35	Renda = < 15	Renda = < 15
Resultado = Baixo	Resultado = Alto	Resultado = Alto

## Indução de regras - Algoritmo OneR

- Testar coisas simples primeiro
- Um atributo faz todo trabalho

Atributo	Quantidade	Regras	Erro	Total erro
História de Credito	Boa = 5	SE Boa ENTÃO Baixo	2/5	6/14
	Desconhecida = 5	SE Desconhecida ENTÃO Alto	3/5	
	Ruim = 4	SE Ruim ENTÃO Alto	1/4	
Divida	Alta = 7	SE Alta ENTÃO Alto	3/7	7/14
	Baixa = 7	SE Baixa ENTÃO Baixo	4/7	
Garantia	Nenhuma = 11	SE Nenhuma ENTÃO Alto	5/11	6/14
	Adequada = 3	SE Adequada ENTÃO Baixo	1/3	
Renda	< 15 = 3	SE <15 ENTÃO Alto	0	4/14
	>=15<35 = 4	SE >= 15 < 35 ENTÃO Alto	2/4	
	>35 = 7	SE >35 ENTÃO Baixo	2/7	

Atributo	Quantidade	Regras	Erro	Total erro
História de Crédito	Boa = 5	SE Boa ENTÃO Baixo	2/5	6/14
	Desconhecida = 5	SE Desconhecida ENTÃO Alto	3/5	
	Ruim = 4	SE Ruim ENTÃO Alto	1/4	
Dívida	Alta = 7	SE Alta ENTÃO Alto	3/7	7/14
	Baixa = 7	SE Baixa ENTÃO Baixo	4/7	
Garantia	Nenhuma = 11	SE Nenhuma ENTÃO Alto	5/11	6/14
	Adequada = 3	SE Adequada ENTÃO Baixo	1/3	
Renda	< 15 = 3	SE <15 ENTÃO Alto	0	4/14
	>=15<35 = 4	SE >= 15 < 35 ENTÃO Alto	2/4	
	>35 = 7	SE >35 ENTÃO Baixo	2/7	

Para achar a regra, vamos usar como exemplo a Renda

- Se <15 então Alto = 3
- Se <15 então Moderado = 0
- Se <15 então Baixo = 0

Pegamos a expressão com o maior valor nesse caso **Se <15 então Alto = 3** e adicionamos o erro na tabela de **Erro**

- Se  $\geq 15 < 35$  Então Alto = 2
- Se  $\geq 15 < 35$  Então Moderado = 2
- Se  $\geq 15 < 35$  Então Baixo = 0

Pegamos a expressão com o maior valor nesse caso **Se  $\geq 15 < 35$  Então Alto = 2 (podia ser a outra também)** e adicionamos o erro na tabela de **Erro**

- Se  $> 35$  Então Alto = 1
- Se  $> 35$  Então Moderado = 1
- Se  $> 35$  Então Baixo = 5

Pegamos a expressão com o maior valor nesse caso **Se  $> 35$  Então Baixo = 5** e adicionamos o erro na tabela de **Erro**, fazemos a soma total dos erros