FORMAÇÃO CIENTISTA DE DADOS

CLASSIFICAÇÃO



CUSTO

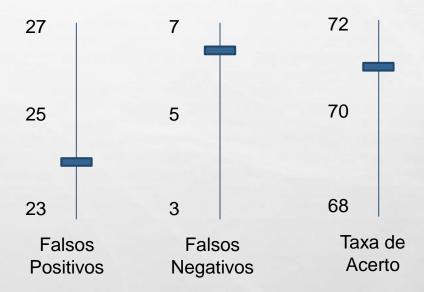
		Classi	Classificação	
ais		Sim	Não	
riginais	Sim	450	50	
Ori	Não	250	255	

Taxa de acerto: 70%.

Falso Positivos: 250 (25%) Maus pagadores, mas que eu vou ceder empréstimo. Prejuízo!

Falsos Negativos: 50 (5%) Bons pagadores, mas que eu não vou ceder empréstimo. Perdi Negócios!

CUSTO



CUSTO

		Custo	Perdas
Falsos Positivos	25%	R\$ 5.000,00	1.250.000,00
Falsos Negativos	5 %	R\$ 1.500,00	75.000,00
Taxa de Acerto	70%		-
Perdas Totais			1.325.000,00

		Custo	Perdas
Falsos Positivos	23%	R\$ 5.000,00	1.150.000,00
Falsos Negativos	8%	R\$ 1.500,00	120.000,00
Taxa de Acerto	69%		-
Perdas Totais			1.270.000,00
Diferença			55.000,00

COMO "MELHORAR" UM MODELO?

- TESTANDO DIFERENTES ALGORITMOS
- PARAMETRIZANDO ALGORITMOS
- SELECIONANDO E TRATANDO DADOS
- SELEÇÃO DE ATRIBUTOS

MÉTRICAS

		Treino		
ais		Aprovado	Reprovado	
riginais	Aprovado		1004	208
Ori	Reprovado		202	294

Métrica	Fórmula	Descrição	Cálculo
Acertos	(VP+VN)/Total	Total de Acertos	75,99
Erros	(FP+FN)/Total	Total de Erros	24,00
Precisão	VP/(VP+FP)	Quantos registros de fato são positivos	83,25
Lembrança ou Positivos Verdadeiros	VP/(VP+FN)	Positivos corretamente previstos	82,83
Negativos Verdadeiros	VN/(VN+FP)	Total de Negativos Verdadeiros	59,27
Positivos Falsos	FP/(VN+FP)	Total de Positivos Falsos	40,72
Negativos Falsos	FN/(VP+FN)	Total de Negativos Falsos	17,16