

O resultado foi bom ou ruim

Transcrição

Voltando na aba "*Preprocess*" do Weka, se olharmos o número de pessoas que fizeram depósitos e o número de pessoas que não fizeram depósitos, veremos que esse número é similar.

Talvez um número um pouco maior de pessoas não fizeram um depósito após serem submetidas à campanha de telemarketing. Como saberemos se valerá a pena construir um classificador, treinando um algoritmo, se os acertos só ocorreram para 85% das vezes, por exemplo? Será que não valeria mais a pena só chutar os valores, apostando na classe mais popular?

Se tivéssemos uma base de dados com gatos, cachorros e wekas (o passarinho símbolo do Weka) e um número maior de exemplos de gatos, poderíamos chutar sempre nos gatos, e teríamos sempre gato como resposta. Claro, se construíssemos um classificador que fizesse isso, ele não seria muito inteligente.

Mas existe um classificador no Weka que faz justamente isso de escolher a classe mais popular e usar sempre essa classe como resposta, **ZeroR**. Vamos testá-lo usando a base de dados de treinamento, para compará-la consigo mesma. Voltando para a aba "*Classify*" temos podemos clicar em "*Choose*" e em "*rules*" e escolheremos o **ZeroR**. Veremos que ele acertará em 52% das vezes, pois ele chutou exatamente o número referente às respostas "*No*" (não) para o depósito.

Então, temos esse tipo de classificador como uma base que nos mostrará o mínimo que queremos acertar. Nossa intenção, portanto, é construir uma classificação que acerte um valor **maior** do que esse.