



1. Descrição do Problema

A empresa Google Groups deseja obter um sistema inteligente que consiga classificar automaticamente mensagens segundo os seguintes assuntos específicos:

- comp.graphics
- comp.os.ms-windows.misc
- comp.sys.ibm.pc.hardware
- comp.sys.mac.hardware
- comp.windows.x
- rec.autos
- rec.motorcycles
- rec.sport.baseball
- rec.sport.hockey
- sci.crypt
- sci.electronics
- sci.med
- sci.space
- misc.forsale
- talk.politics.misc
- talk.politics.guns
- talk.politics.mideast
- talk.religion.misc
- alt.atheism
- soc.religion.christian

Para isso a empresa disponibilizou um base de dados com mensagens classificadas segundo tais grupos e pede que os interessados criem o modelo que a permita desenvolver tal sistema.

2. Dados

Os dados podem ser obtidos a partir do link a seguir:

<http://qwone.com/~jason/20Newsgroups/>