

Ficha 8

Scripting no Processamento de Linguagem Natural

17 de Maio de 2015

Escreva um módulo em Perl que lhe permita reconhecer spam. Esse módulo deve conter um método que receba uma mensagem e que devolva a classe correspondente. Se possível, deverá também devolver o grau de confiança associado a essa classificação. Utilize o dataset disponibilizado em http://www.di.uminho.pt/~rcm/aulas/spln1415/spam_corpus.tar.bz2 para validar a capacidade do seu módulo.

Avaliação

Quem Grupos de duas pessoas

Quando Até ao fim do dia **30 de Maio de 2015**

O quê Deverá ser entregue a distribuição do módulo (i.e., usando o **make dist**) contendo o seguinte material:

- Código do módulo
- Código da aplicação que utiliza o módulo para o reconhecimento de entidades em Português
- Testes (e.g., usando **Test::More**)
- Documentação (usando **PerlPod**)
- Relatório (usando **PerlPod**)

Como A avaliação do módulo irá incidir sobre a estratégia de implementação, treino e validação do sistema bem como sobre a capacidade de classificação das mensagens de spam. Para além disto, a avaliação irá também olhar para a qualidade do código escrito, dos testes, da documentação e do relatório produzido.

Recursos

Consulte em <http://www3.di.uminho.pt/~rcm/aulas/spln1415>.