



(http://www2.servico=9)

(http://www.correiobrasiliense.com.br/)

ACOMPANHE E COMPARTILHE NOSSAS CAUSAS

JUNTOS PODEMOS FAZER UM MUNDO MELHOR

Software desenvolvido em São Paulo filtra dados digitais

O programa separa os dados a partir de associações próprias

postado em 24/08/2015 06:00

Para muitos, organizar as pastas e os arquivos virtuais do computador é uma tarefa continuamente adiada. Assim também acontece com a obrigação de ler e separar as mensagens acumuladas na caixa de entrada do e-mail, uma ideia que se torna mais insuportável a cada novo recado que chega. Imagine, então, como difícil seria examinar todo o conteúdo publicado em sites da internet, como portais de notícias, blogs e redes sociais. Um desafio impossível para humanos. Para máquinas, porém, um trabalho que pode ser cumprido sem dificuldades. Um software em desenvolvimento no Instituto de Ciências Matemáticas e de Computação (ICMC) da Universidade de São Paulo (USP) em São Carlos consegue classificar automaticamente grande quantidade de textos digitais.

Trata-se de um algoritmo que identifica os termos usados em cada tipo de texto e analisa a relação entre as palavras para classificar um novo documento. Tudo é feito de acordo com os exemplos dados por humanos. Se uma biblioteca virtual tiver vários tipos de arquivos científicos, por exemplo, bastaria cadastrar no programa alguns trabalhos relacionados a cada assunto. A partir de alguns exemplos de cada categoria, o programa conclui a organização por conta própria.

A maioria dos programas de classificação automática de textos considera a frequência com que certas palavras-chave aparecem nos documentos. No entanto, os algoritmos desenvolvidos pelo aluno de doutorado do ICMC Rafael Rossi também são capazes de interpretar as redes formadas por associações entre termos, o que permite ao computador identificar padrões não assimilados em outros tipos de representações, tornando o software mais eficiente. Por meio do aprendizado de máquina, o sistema pode se aperfeiçoar na sua tarefa, imitando o discernimento de um humano sem que ele tenha de ser especialmente programado.

“O que propomos é considerar a similaridade entre termos em uma coleção de documentos. Se tenho a palavra ‘banco’ e ‘dados’ no mesmo documento, elas serão similares. Se, por outro lado, eu tiver os termos ‘banco’ e ‘redes’, que são áreas distintas, elas não serão similares. Assim, definimos o que chamamos de valor de relevância. Seria o peso ou a força que um termo tem para determinado documento. O objetivo é usar a relação de similaridade para definir essa relevância”, explica Rafael Rossi.

A matéria completa está disponível aqui

(http://impresso.correioweb.com.br/app/noticia/cadernos/tecnologia/2015/08/24/interna_tecnologia,180131/classificacao-requintada.shtml), para assinantes. Para assinar, clique aqui (<https://www2.correiobrasiliense.com.br/seguro/digital/assine.php>).

PUBLICIDADE

Dólar a R\$ 4,00?

A Tendência de Alta do Dólar Vai Durar Até Quando? Saiba Agora, Aqui



Comentários

Os comentários não representam a opinião do jornal;
a responsabilidade é do autor da mensagem.

Digite seu comentário

Faça seu login para comentar

e-mail senha

Login | Esqueci minha senha | Quero me cadastrar (<http://www.correiobraziliense.com.br/cadastro/>)

- Notícias + lidas + comentadas
- 18:10 - 09/09/2015
Novo iPhone e iPad Pro empolgam fãs, mas não contêm queda de ações da Apple
(http://www.correiobraziliense.com.br/app/noticia/tecnologia/2015/09/09/interna_tecnologia,497964/lancamentos-da-apple-nao-sao-suficientes-para-reanimar-mercado.shtml)
- 16:37 - 09/09/2015
WhatsApp corrige falha que colocava em risco 200 milhões de usuários
(http://www.correiobraziliense.com.br/app/noticia/tecnologia/2015/09/09/interna_tecnologia,497949/whatsapp-corrige-falha-que-colocava-em-risco-200-milhoes-de-usuarios.shtml)

PUBLICIDADE

Eu quero testar meu...

Francês	Inglês
Italiano	Holandês
Alemão	Espanhol

Babel.com

- Fotos Vídeos



9/9 - Acompanhe as transformações das novas gerações do iPhone

10/9 - Homo Naledi: Nova espécie do gênero humano é descoberta na África do Sul
(http://www.correiobraziliense.com.br/app/galeria-de-fotos/2015/09/10/interna_galeriafotos,5705/10-9-homo-naledi-nova-especie-do-genero-humano-e-descoberta-na-africa-do-sul.shtml)



(http://www.correiobraziliense.com.br/app/galeria-de-fotos/2015/09/09/interna_galeriafotos,5704/9-9-acompanhe-as-transformacoes-das-novas-geracoes-do-iphone.shtml)

PUBLICIDADE

Eu quero testar meu...  Babbel.com	Francês ▶	Inglês ▶	Italiano ▶ ⓘ
	Holandês ▶	Alemão ▶	Espanhol ▶