

IF769 - Mineração na Web

Lucas Antonio Sales dos Santos

7 de maio de 2018

1 Introdução

A disciplina de Mineração na Web atua na área de Big Data, tem como foco a recuperação de informação, a disciplina possui 3 módulos, onde no primeiro módulo tem o conteúdo focado em Recuperação de informação(Clássica), onde são introduzidos a Recuperação de informação pelo método de indexação e busca de documentos e ainda como conteúdo, a avaliação de sistemas de recuperação de informação. No segundo módulo os conteúdos são Recuperação inteligente de informação, mineração de texto, extração de informação e etc. E no terceiro módulo são feitos projetos de mineração na web pelos alunos.

2 Relevância

A disciplina é de extrema importância, pois a Web é hoje a maior fonte de informação eletrônica que temos, mas por não termos controle sobre ela, achar informações relevantes pode se tornar algo que em muitas vezes é frustrante, e muitos esforços de pesquisa têm sido feitos para sanar esse problema. E um deles é a utilização de técnicas de mineração de dados para a descoberta de informações na web. Mas, utilizar e compreender os dados disponíveis na Web não é tarefa fácil, pois esses dados são muito mais sofisticados e dinâmicos do que os sistemas de armazenamento de banco de dados tradicionais. Enquanto estes utilizam de estruturas de armazenamento bem definidas e estruturadas, a Web não possui qualquer controle sobre a estrutura ou o tipo dos documentos que virtualmente armazena, por isso é muito importante que esta disciplina seja ofertada.

3 Relação com outras disciplinas

Disciplina	Relação
IF695 -BANCO DE DADOS AVANÇADOS+ ELETIVO	Relação direta pois o banco de dados armazena os dados e esses dados serão minerados para utilização.
IF694 -BANCO DE DADOS DISTRIB. E MOVEIS+ ELETIVO	Relação direta pois a mineração de dados vai acessar os bancos de dados.
IF693 -SIST.GERENC. DE BANCO DE DADOS+ ELETIVO	Relação direta pois a mineração de dados vai utilizar dados armazenados por esses bancos de dados.
IF697 -TOP.AVANC.GERENC.DADOS INFORMACAO+ ELETIVO	Banco de dados e mineração estão diretamente relacionados.
IF696 -INTEG. DADOS WEB E WAREHOUSE+ ELETIVO	Banco de dados e mineração estão diretamente relacionados.
IF692 -PROJETO DE BANCO DE DADOS+ ELETIVO	Banco de dados e mineração estão diretamente relacionados.
IF698 -SEMIN. GERENC. DADOS E INFORMACAO+ ELETIVO	Banco de dados e mineração estão diretamente relacionados.
IF685 -GERENCIAMENTO DADOS E INFORMACAO OBRIGATÓRIO	Banco de dados e mineração estão diretamente relacionados.

4 Referências

<http://trec.nist.gov/> - The TREC (Text REtrieval Conference) series is co-sponsored by the NIST
<http://infolab.stanford.edu/backrub/google.html>
<https://www.nist.gov/itl>
<http://www.cs.utexas.edu/users/mooney/ir-course/>