



UNIVERSIDADE DE COIMBRA
FACULDADE DE CIÊNCIAS E TECNOLOGIA
Departamento de Engenharia Informática

Enunciado do Projecto de Sistemas Inteligentes para a Gestão de Conhecimento

2013/2014 – 2º Semestre

Nota Importante: A fraude denota uma grave falta de ética e constitui um comportamento não admissível num estudante do ensino superior e futuro profissional licenciado. Qualquer tentativa de fraude pode levar à anulação da componente prática tanto do facilitador como do prevaricador.

1. Objectivos do trabalho

Este trabalho tem como principal objectivo a concepção e implementação de uma aplicação para gestão de conhecimento numa organização. O requisito essencial para este trabalho é:

- Que use pelo menos uma das técnicas apresentadas nas aulas, ou seja: “Information Retrieval”, “Case-Based Reasoning”, “Text Mining”, “Web Mining” ou “Social Mining”.

2. Funcionamento

O projecto vai ter três fases:

- **Escolha de Tema** – nesta fase vão ser apresentados alguns temas de projectos propostos pelo Professor, ao que os alunos devem organizar-se em grupos de 1 ou 2 elementos e escolher um tema. No final desta fase será elaborado um plano de trabalho para cada grupo e juntamente com o Professor será celebrado um Contrato Pedagógico. Os alunos podem também propor temas que serão discutidos com o Professor.
- **Execução do Projecto** – esta fase consiste na realização do projecto sendo que as linhas orientadoras são:
 - Este deve ser realizado também nas aulas específicas para a sua execução;
 - Serão realizadas reuniões semanais com o Professor de forma a dar acompanhamento ao projecto.

- **Defesa de Projecto** – esta fase compreende a apresentação e demonstração dos sistemas desenvolvidos. Cada grupo apresenta o seu projecto.

3. Contrato Pedagógico

O Contrato Pedagógico é fundamental para o entendimento entre os alunos e o Professor sobre os itens a serem avaliados. Desta forma o Contrato Pedagógico define quais os objectivos do projecto e como é que vão ser avaliados. Assim, como elementos de base para o Contrato Pedagógico temos:

- Objectivos.
- Metas intermédias para o projecto.
- Data de apresentação e entrega do projecto e respectivo relatório.

4. Temas para o Trabalho

Apresentam-se de seguida vários temas alternativos para o trabalho, sendo que existe também a hipótese de serem propostos temas pelos alunos (ver última entrada desta secção). Cada grupo deverá ter no máximo de dois elementos.

O trabalho tem 4 metas genéricas para todos os trabalhos, sendo que a data destas metas pode ser definida com o acordo do Professor. As metas são depois instanciadas para cada projeto concreto:

- M1 – Seleção e recolha da informação com que o sistema vai funcionar.
- M2 – Extração e indexação da informação.
- M3 – Pesquisa de informação.
- M4 – Sugestão e/ou análise de informação.

Para a execução do trabalho são sugeridas as seguintes ferramentas:

- Recolha de Informação (IR) – Lucene (<http://lucene.apache.org>)
- Text Mining/Web Mining (TM) – Mallet (<http://mallet.cs.umass.edu>) alternativamente o Mahout - <http://mahout.apache.org> ou o Weka - <http://www.cs.waikato.ac.nz/ml/weka/>)
- Raciocínio Baseado em Casos (CBR) – JColibri (<http://sourceforge.net/projects/jcolibri-cbr/>)

TEMA 1: Redireccionamento e Filtragem de Informação (<i>Information Routing</i>)
<p>DESCRIÇÃO:</p> <p>A ideia deste tema é a filtragem e redireccionamento de informação que está continuamente a chegar. Por exemplo, análise e redireccionamento de email, de notícias ou mensagens das redes sociais. O conceito base é análogo a ter um agente que recebe essa informação e que a analisa identificando quais os assuntos dessa informação, reencaminhando-a depois para as pessoas interessadas. Tendo o sistema conhecimento sobre os perfis de utilizadores para os quais pode</p>

reencaminhar a informação, esta tarefa é mais facilitada com a utilização de uma ontologia que permita a representação deste conhecimento.

Funcionalidades Requeridas:

- Classificação de Notícias/Emails/Posts (TM)
- Pesquisa de Notícias/Emails/Posts (IR)
- Sugestão de Notícias/Emails/Posts com base no Perfil do Utilizador (CBR)
- Sugestão de Notícias/Emails/Posts Relacionadas (CBR)

TEMA 2: Gestão Documental numa Organização

DESCRIÇÃO:

Este tema consiste no desenvolvimento de um sistema que faz a gestão de documentos de uma forma mais inteligente. Tendo um conjunto de documentos para gerir, o sistema deve ser capaz de classificar e indexar documentos de uma determinada organização. Ou seja, capaz de lidar com situações de homonímia (duas palavras iguais com sentidos diferentes), fazer sugestões ao utilizador no âmbito de pesquisas que não são óbvias, por exemplo, se o utilizador estiver a pesquisar documentos sobre Java e o sistema não tiver nenhum, pode sugerir documentos semelhantes, tais como um documento em C#.

Funcionalidades Requeridas:

- Clustering de Documentos (TM)
- Classificação de Documentos (TM)
- Pesquisa de Documentos (IR)
- Sugestão de Documentos Semelhantes (IR)

TEMA 3: Helpdesk Inteligente

DESCRIÇÃO:

A ideia deste tema é utilizar uma base de casos descrevendo problemas de software (bugs) e como resolvê-los. O utilizador (que em princípio será um programador) poderá fazer perguntas ao sistema expondo o problema que enfrenta. O sistema poderá fazer-lhe um conjunto de perguntas para poder especificar melhor a situação com que o utilizador se está a deparar. Sendo que depois irá sugerir-lhe qual a causa para o problema, bem como qual a solução. O sistema pode ainda não ter uma resposta, no entanto deve ser capaz de aprender a resolver novos problemas.

Funcionalidades Requeridas:

- Clustering de Bugs (TM)
- Classificação de Bugs (TM)
- Pesquisa de Bugs (IR)
- Sugestão de Bugs Semelhantes (CBR)

TEMA 4: Análise de Comentários Online

DESCRIÇÃO:

Este tema tem como objectivo analisar comentários online sobre produtos ou serviços, analisá-los e apresentar os resultados dessa análise. Essa análise pode ser: o nome do produto ao qual o comentário se refere (por exemplo considere o comentário: “O iPad 5 é fabuloso, mas o melhor que tem é a nova câmara”, neste caso seria ‘iPad 5’), as características que são comentadas (‘câmara’), se o comentário é positivo, negativo ou neutro (neste caso seria positivo), etc. Para além desta informação o sistema deve ser capaz de ter as seguintes funcionalidades:

- Clustering de Comentários (TM)
- Classificação de Comentários (TM)
- Pesquisa de Comentários (IR)
- Análise de Comentários (TM)

Poderão ser aceites outros temas para o trabalho que sejam propostos, desde que tenham complexidade idêntica aos temas anteriores. Para aprovação de um tema que não os propostos anteriormente, os alunos em causa deverão contactar o docente da cadeira com a antecedência necessária para a sua aprovação.

5. Critérios de avaliação

Os trabalhos serão avaliados considerando os seguintes critérios:

- Acompanhamento do projecto e respectivo cumprimento das metas intercalares [4 valores].
- Relatório final [5 valores].
- Apresentação final e defesa do projecto [5 valores].

6. Entrega do trabalho

Escreva um **relatório** descrevendo os aspectos mais importantes do que realizou. A estrutura para o relatório deverá ser a seguinte:

1. Introdução.
2. Descrição da arquitectura e design do sistema.
3. Descrição das várias metas do projeto.
4. Descrição da base de conhecimento utilizada.
5. Descrição dos vários processos usados.
6. Descrição detalhada dos módulos mais importantes do sistema.
7. Estrutura e organização do código (não é a listagem do código).

8. Conclusão.

9. Referências.

O relatório deverá ter o formato de artigo. No material de apoio da cadeira está lá um ficheiro de exemplo do formato a ser utilizado. Este ficheiro pode ser utilizado como base para a escrita do relatório.

Submeta o relatório em formato PDF por email para o endereço do docente (pgomes@dei.uc.pt). **Data de Entrega: 11 de Junho**

Serão realizadas defesas dos trabalhos na semana da entrega.

Guarde cópia de tudo, por razões de segurança. Caso tenha dúvidas pode enviá-las por email para o docente da cadeira.