



Universidade do Minho

Departamento de Informática

Mestrado [integrado] em Engenharia Informática

Perfil de Sistemas Inteligentes

Aprendizagem e Extração de Conhecimento

1º/4º Ano, 1º Semestre

Ano letivo 2017/2018

Trabalho Prático – 2ª Parte

Novembro, 2017

Tema

Extração de Conhecimento de dados estruturados.

Objetivos de aprendizagem

Com a realização deste enunciado prático pretende-se que os grupos de trabalho:

- Realizem um Projeto de Extração de Conhecimento, adotando uma metodologia para extração de conhecimento.

Enunciado

Para a realização da componente prática de avaliação, os grupos de trabalho deverão realizar as seguintes tarefas:

- Consultar, analisar e selecionar conjuntos de dados (“datasets”) de entre os que estão acessíveis a partir da página eletrónica da ferramenta WEKA (ou outras fontes);
- Selecionar e identificar os conjuntos de dados sobre os quais, cada grupo, desenvolverá o Projeto de Extração de Conhecimento e, ainda:
 - Para além dos conjuntos de dados selecionados por cada grupo:
 - Os grupos com número ímpar deverão analisar, também, o *dataset* em [Appliances energy prediction](#);
 - Os grupos com número par deverão analisar, também, o *dataset* em [Air Quality](#);
- Utilizar a ferramenta WEKA para proceder à análise dos dados (seleção de atributos, filtragem, visualização, preparação, classificação, segmentação, associação, etc.);
- Encontrar resultados, incluindo a identificação dos parâmetros mais adequados;
- Determinar quais os resultados da análise que se apresentam como os mais relevantes;
- Interpretar os resultados de modo a decidir sobre a forma como poderão ser úteis no contexto dos dados analisados.

Estas tarefas serão realizadas durante as aulas práticas e os resultados obtidos deverão ser objeto de um relatório que contenha uma enumeração de observações, como por exemplo:

- Objetivos: qual é o domínio e quais são os potenciais benefícios da extração de conhecimento sobre os dados analisados;
- Descrição do conjunto de dados: qual é o significado dos dados tratados e que tipo de pré-processamento foi aplicado sobre eles;
- Processo de extração de conhecimento: definição de parâmetros, escolha dos algoritmos de análise (classificação, segmentação, associação, etc.) e tempo dispendido;
- Resultados: sumário e seleção das regras que apresentam melhores resultados e maior utilidade;

- Recomendações: apresentação de sugestões sobre o modo de agir após a análise dos resultados do processo de extracção de conhecimento.

Todo o processo deverá ser acompanhado de exemplos e indicações que permitam reproduzir os passos realizados pela seleção e aplicação de uma metodologia de extração de conhecimento.

A data para a entrega do relatório final é o dia 10 de dezembro, 2017.

A data para a sessão de apresentação do trabalho é o dia 11 de dezembro, 2015, na sala DI-1.05, com início às 14h00m. Cada grupo disporá de 10 minutos para realizar a apresentação, utilizando os meios que considerar mais adequados.

Avaliação

A avaliação deste trabalho de grupo contará com os seguintes elementos:

- Pelo documento produzido (75%);
- Pela apresentação realizada do trabalho desenvolvido (25%).

Bibliografia

Indicam-se as referências de cada *dataset* indicado:

- S. De Vito, E. Massera, M. Piga, L. Martinotto, G. Di Francia, On field calibration of an electronic nose for benzene estimation in an urban pollution monitoring scenario, *Sensors and Actuators B: Chemical*, Volume 129, Issue 2, 22 February 2008, Pages 750-757, ISSN 0925-4005.
- Luis M. Candanedo, Veronique Feldheim, Dominique Deramaix, Data driven prediction models of energy use of appliances in a low-energy house, *Energy and Buildings*, Volume 140, 1 April 2017, Pages 81-97, ISSN 0378-7788.

Outras referências complementares deverão ser procuradas, estudadas e incluídas.