

Informação Geral:

Unidade Orgânica	Escola Superior de Tecnologia e Gestão	Ano Letivo	2015/2016
Curso	Licenciatura em Engenharia Informática	Grau	Licenciatura
UC/Módulo	Engenharia do Conhecimento	ECTS	6
Área Científica	Engenharia Informática - Sistemas de Informação	Carácter	Obrigatório
Horas Totais	162	T 30	TP 0
		PL 45	TC 0
		S 0	E 0
		OT 0	

T - Ensino Teórico; TP - Teórico Prático; PL - Prático e Laboratorial; TC - Trabalho de Campo; S - Seminário; E - Estágio; OT - Orientação Tutorial

Docente responsável:

José Vitor Martins Ramos

Docentes que lecionam a unidade curricular:

José Vitor Martins Ramos (75,00 horas semanais de contacto: T: 30,00; PL: 45,00;)

Maria Beatriz Guerra Piedade (45,00 horas semanais de contacto: PL: 45,00;)

Pré-requisitos:

Não tem

Idioma:

Português

Enquadramento:

Na sociedade de informação atual a capacidade para extrair conhecimento a partir de dados representa uma vantagem competitiva para as organizações, pois cada vez mais decisões são tomadas com base na análise de dados.

No âmbito desta unidade curricular pretende-se transmitir aos estudantes conceitos acerca do processo de data mining, abordar diferentes metodologias e algoritmos, bem como utilizar ferramentas computacionais para a resolução de problemas práticos de data mining.

Objetivos de aprendizagem:

- C1. Compreensão da importância do data mining nas organizações
- C2. Compreensão do processo de data mining
- C3. Compreensão das metodologias para o processo de data mining
- C4. Conhecimento dos problemas típicos de data mining
- C5. Conhecimento das técnicas de data mining
- C6. Aplicar os conhecimentos adquiridos na criação e avaliação de modelos
- C7. Utilizar ferramentas de software de data mining
- C8. Utilizar técnicas e ferramentas de data mining para resolver problemas reais das organizações
- C9. Capacidade de estudar e aprender autonomamente
- C10. Desenvolvimento de raciocínio abstrato, análise crítica e avaliação de soluções
- C11. Capacidade de pesquisar informação em diferentes meios e formatos e ser capaz de a utilizar de forma eficaz
- C12. Capacidade de estruturar e elaborar documentação
- C13. Capacidade de trabalhar em equipa

Programa:**Conteúdos Programáticos:**

1. Introdução ao Data Mining
2. Metodologias para o Processo de Data Mining
3. Exploração de Dados
4. Preparação de Dados
5. Redução de Dados
6. Algoritmos de Data Mining
7. Avaliação e Seleção de Modelos
8. Tópicos Avançados de Data Mining
9. Privacidade, Segurança e Desafios

Fundamentação da coerência dos conteúdos programáticos com os objetivos/competências da unidade curricular:

1. Introdução ao Data Mining - (C1;C2;C4;C9;C11)
2. Metodologias para o Processo de Data Mining - (C3;C9;C11)
3. Exploração de Dados - (C4;C5;C6;C7;C8;C9;C11;C12;C13;C14)
4. Preparação de Dados - (C4;C5;C6;C7;C8;C9;C11;C12;C13;C14)
5. Redução de Dados - (C4;C5;C6;C7;C8;C9;C11;C12;C13;C14)
6. Algoritmos de Data Mining - (C4;C5;C6;C7;C8;C9;C11;C12;C13;C14)
7. Avaliação e Seleção de Modelos - (C4;C5;C6;C7;C8;C9;C11;C12;C13;C14)
8. Tópicos Avançados de Data Mining - (C4;C5;C6;C9;C11;C12;C13;C14)
9. Privacidade, Segurança e Desafios - (C1;C9)

Metodologia de Ensino / Aprendizagem:**Presencial:**

Teórico (T): exposição e compreensão dos conteúdos programáticos.
Prático e Laboratorial (PL): Resolução de problemas e aplicação de conhecimentos.

Autónoma:

Estudo: Leitura de apontamentos e bibliografia.
E-Learning: Materiais disponibilizados no Moodle, participação nos fóruns.
Projeto: Realização de um projeto em equipa por forma a promover a organização do trabalho e o desenvolvimento de capacidades de autonomia, iniciativa e análise crítica.

Recursos Específicos:

Laboratório de Sistemas de Informação
Plataforma de gestão e distribuição de conteúdos (Moodle)
Software específico (Weka; RapidMiner; IBM Watson Analytics)
Plataforma de competições de data mining (Kaggle)

Avaliação:**Descrição:**

Avaliação Periódica
Prova Escrita Teórica (PET)
Projeto: 2 Entregas
- Entrega 1: 25%
- Entrega 2: 75%

Classificação Final: CF=50%PET+50%Projeto

Avaliação Final
Prova Escrita Teórica (PET)
Prova Escrita Prática (PEP)
Classificação Final: CF=50%PET+50%PEP

Observações

- Projeto em grupo: 2 estudantes, de preferência
- Defesas individuais do projeto por amostragem
- Projeto, PET, PEP: mínimo de 9,5 valores em [0;20]
- A classificação obtida numa das componentes de avaliação transita para as restantes épocas de avaliação do presente ano letivo
- A melhoria de classificação, na época de recurso, pode ser feita a uma ou a ambas as componentes de avaliação
- A classificação obtida numa das componentes de avaliação no ano letivo anterior não é guardada

Número de elementos de avaliação final:

2

Demonstração da coerência das metodologias de ensino com os objetivos de aprendizagem da unidade curricular:

Ensino Presencial
Teórico - (C1;C2;C3;C4;C5;C6)
Prático e Laboratorial - (C2;C3;C4;C5;C6;C7;C8)
Orientação Tutorial - (C4;C5;C6)

Aprendizagem Autônoma
Estudo - (C1;C2;C3;C4;C5;C6)
e-Learning - (C9;C11)
Projeto - (C10;C11;C12;C13;C14)

Bibliografia:**Recomendada:**

Data Mining: Concepts, Models, Methods and Algorithms; Mehmed Kantardzic; John Wiley & Sons, 2011
Discovering Knowledge in Data: An Introduction to Data Mining; Daniel T. Larose; John Wiley & Sons, 2014
Data Mining: Concepts and Techniques; J. Han, M. Kamber e J. Pei; Morgan Kaufmann Publishers, 2012
Introduction to Data Mining; Pang-Ning Tan, Michael Steinbach e Vipin Kumar; Addison-Wesley, 2005
Extração de Conhecimento de Dados – Data Mining; J. Gama et al., Edições Sílabo, 2012
Data Mining: Practical Machine Learning Tools and Techniques; Ian Witten e Eibe Frank; Morgan Kaufmann, 2011
Mining of Massive Datasets, J. Leskovec et al., Cambridge University Press, 2014
CRISP-DM 1.0: Step-by-Step Data Mining Guide, P. Chapman et al., SPSS, 2000

Complementar:

Weka Manual for Version 3.7.13, R. Bouckaert et al., University of Waikato, 2015
RapidMiner Studio Manual, RapidMiner, 2016
RapidMiner 7.0 - How to Extend RapidMiner, S. Kirstein et al., Rapidminer, 2016
RapidMiner Advanced Charts Manual, RapidMiner, 2014
Exploring Data with RapidMiner, A. Chisholm, Packt Publishing, 2013
IBM Watson Analytics - Workshop Guide, R. Hayden et al, IBM, 2015
Learning IBM Watson Analytics, J. Miller, Packt Publishing, 2016
PMML in Action, A. Guazzelli et al., CreateSpace Independent Publishing Platform, 2012