

### Informação Geral:

<b>Unidade Orgânica</b>	Escola Superior de Tecnologia e Gestão	<b>Ano Letivo</b>	2017/2018
<b>Curso</b>	Licenciatura em Engenharia Informática (D) [9119]	<b>Grau</b>	Licenciatura
<b>Ano Curricular</b>	3	<b>Período</b>	S2
<b>UC/Módulo</b>	Engenharia do Conhecimento	<b>ECTS</b>	6
<b>Área Científica</b>	Engenharia Informática - Sistemas de Informação	<b>Carácter</b>	Obrigatório
<b>Horas Totais</b>	162	<b>T</b> 30	<b>TP</b> 0
		<b>PL</b> 45	<b>TC</b> 0
		<b>S</b> 0	<b>E</b> 0
		<b>OT</b> 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 Vítor Martins Ramos

### Docentes que lecionam a unidade curricular:

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

Maria Beatriz Guerra Piedade ( 90,00 horas semanais de contacto: PL: 45,00; PL: 45,00; )

Rui Miguel de Carvalho Leal de Oliveira ( 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  
C14. Capacidade de realizar projetos

## 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 (RapidMiner; IBM Watson Analytics; Weka)  
Plataforma de competições de data mining (Kaggle)

## Avaliação:

### Descrição:

Avaliação Periódica  
Prova Escrita (PE)  
Projeto: 2 Entregas  
- 1ª Entrega: 35%  
- 2ª Entrega: 65%  
Classificação Final: CF=50%PE+50%Projeto

Avaliação Final  
Prova Escrita (PE)  
Classificação Final: CF=100%PE

### Observações

- Projeto em grupo: 2 estudantes, de preferência
- Defesas individuais do projeto por amostragem
- Não há mínimos nos elementos de avaliação
- A classificação obtida nos elementos de avaliação da época de avaliação periódica não transita para as restantes épocas de avaliação
- A melhoria de classificação, na época de recurso, é feita através da realização de uma Prova Escrita

<b>Número de elementos de avaliação final:</b>	1
<b>Número de elementos de avaliação contínua/periódica:</b>	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)

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 3.9.2, R. Bouckaert et al., University of Waikato, 2017  
 RapidMiner Studio Manual, RapidMiner, 2018  
 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