

---

**Ano Letivo** 2016-17

---

**Unidade Curricular** INTRODUÇÃO À ANÁLISE DE DADOS

---

**Cursos** ENGENHARIA ELETRÓNICA E TELECOMUNICAÇÕES (Mestrado Integrado) (\*)  
ENGENHARIA INFORMÁTICA (2.º ciclo) (\*)

(\*) Curso onde a unidade curricular é opcional

---

**Unidade Orgânica** Faculdade de Ciências e Tecnologia

---

**Código da Unidade Curricular** 140064344

---

**Área Científica** CIÊNCIA DE COMPUTADORES, CIÊNCIAS DA COMPUTAÇÃO

---

**Sigla**

---

**Línguas de Aprendizagem** PT - Português

---

**Modalidade de ensino** Presencial

---

**Docente Responsável** António Eduardo de Barros Ruano

| DOCENTE                         | TIPO DE AULA | TURMAS  | TOTAL HORAS DE CONTACTO (*) |
|---------------------------------|--------------|---------|-----------------------------|
| António Eduardo de Barros Ruano | PL; T        | T1; PL1 | 30T; 30PL                   |

\* Para turmas lecionadas conjuntamente, apenas é contabilizada a carga horária de uma delas.

| ANO    | PERÍODO DE FUNCIONAMENTO* | HORAS DE CONTACTO | HORAS TOTAIS DE TRABALHO | ECTS |
|--------|---------------------------|-------------------|--------------------------|------|
| 4º, 1º | S2, S1                    | 30T; 30TP         | 168                      | 6    |

\* A-Anual; S-Semestral; Q-Quadrimestral; T-Trimestral

#### Precedências

Sem precedências

#### Conhecimentos Prévios recomendados

Conhecimentos de Algoritmos e Estruturas de Dados e de Probabilidades e Estatística

#### Objetivos de aprendizagem (conhecimentos, aptidões e competências)

Análise de dados descreve um conjunto de técnicas usadas para converter informação em conhecimento útil. Este conhecimento pode-nos ajudar a compreender melhor o nosso mundo, e em muitos contextos permite-nos tomar decisões melhores. Sendo este o objetivo global, nos últimos 20 anos tem-se vindo a assistir a uma diminuição de custos para recolher, armazenar e processar dados, criando assim uma motivação ainda mais forte para o uso de abordagens empíricas para resolver problemas. Esta unidade curricular visa apresentar uma gama de técnicas usadas em análise de dados e está estruturada em torno dos contornos gerais dos diferentes tipos de ferramentas, ou seja, descritivas, baseadas em análise inferencial, preditivas e prescritivas.

No final desta disciplina, o estudante saberá utilizar as ferramentas mais comuns em análise de dados, constituindo assim o background necessário a unidades curriculares mais avançadas neste tópico.

#### Conteúdos programáticos

1. Introdução
2. Estatística descritiva
3. Técnicas de regressão e classificação supervisionadas
4. Aprendizagem não supervisionada e desafios para Análise de grandes dados
5. Análise prescritiva

### **Metodologias de ensino (avaliação incluída)**

Nas aulas teóricas os conceitos serão expostos. Nas aulas teórico-práticas serão resolvidos problemas e executar-se-ão simulações com o objetivo de consolidar os conceitos introduzidos nas aulas teóricas. Cada grupo de alunos terá um problema prático atribuído, que deve ser resolvido ao longo da disciplina. Isto irá permitir que cada aluno (e a turma como um todo) possa entender a aplicabilidade dos conceitos introduzidos durante o curso. Em cada bloco do programa, haverá um mini-teste que deve ser preenchido on-line.

Há três componentes utilizadas na avaliação: um exame (na época normal), com um peso de 40% na avaliação final, um trabalho de grupo, com o peso de 40%, e mini-testes, no final de cada bloco, cuja média contribuirá com 20% para a nota final.

A soma ponderada dessas três componentes constituirá a nota final da disciplina. Se for 10 ou superior, o aluno passará. Caso contrário, o estudante é admitido ao exame de recurso, que só substituirá a nota do exame normal.

---

### **Bibliografia principal**

1. Aparentamentos On-line
2. Hastie, Trevor, et al. *The elements of statistical learning*. Vol. 2. No. 1. New York: Springer, 2009.
3. 2. Montgomery, Douglas C., and George C. Runger. *Applied statistics and probability for engineers*. John Wiley & Sons, 2010
4. Simon Haykin, *Neural Networks: a Comprehensive Foundation*, Prentice-Hall, 1994 (1ª ed.), 1999 (2ª ed.)
5. Daniel Keim, Jörn Kohlhammer, Geoffrey Ellis and Florian Mansmann (Eds.); *Mastering the Information Age: Solving Problems with Visual Analytics*, 2010.
6. Matthew North; *Data mining for the masses*, 2012.

**Academic Year** 2016-17

**Course unit** INTRODUÇÃO À ANÁLISE DE DADOS

**Courses** ELECTRONIC ENGINEERING AND TELECOMMUNICATIONS (Integrated Master's) (\*)  
INFORMATICS ENGINEERING (\*)

(\*) Optional course unit for this course

**Faculty / School** Faculdade de Ciências e Tecnologia

**Main Scientific Area** CIÊNCIA DE COMPUTADORES, CIÊNCIAS DA COMPUTAÇÃO

**Acronym**

**Language of instruction** PT - Portuguese

**Teaching/Learning modality** Presential

**Coordinating teacher** António Eduardo de Barros Ruano

| Teaching staff                  | Type  | Classes | Hours (*) |
|---------------------------------|-------|---------|-----------|
| António Eduardo de Barros Ruano | PL; T | T1; PL1 | 30T; 30PL |

\* For classes taught jointly, it is only accounted the workload of one.

**Contact hours**

| T  | TP | PL | TC | S | E | OT | O | Total |
|----|----|----|----|---|---|----|---|-------|
| 30 | 30 | 0  | 0  | 0 | 0 | 0  | 0 | 168   |

T - Theoretical; TP - Theoretical and practical ; PL - Practical and laboratorial; TC - Field Work; S - Seminar; E - Training; OT - Tutorial; O - Other

**Pre-requisites**

no pre-requisites

**Prior knowledge and skills**

Knowledge of Algorithms and Data Structures, and a basic course of Probability and Statistics

**The students intended learning outcomes (knowledge, skills and competences)**

**Syllabus**

1. Introduction
2. Descriptive Statistics
3. Inferential Statistics
4. Supervised Learning with Regression and Classification techniques
5. Unsupervised Learning and Challenges for Big Data Analytics
6. Prescriptive analytics

**Teaching methodologies (including evaluation)**

In the T classes the concepts will be exposed. In the TP classes problems will be solved and simulations will be executed with the aim of consolidating the concepts introduced in the T classes. Each group of students will have a practical problem assigned which must be solved throughout the course. This will enable each student (and the class as a whole) to understand the applicability of the concepts introduced during the course. At each block of the programme there will be an assignment which must be completed on-line.

There are two components used in the evaluation: one exam (in the normal epoch), with a weight of 30% in the final evaluation, and one individual work, with a weight of 70% for the final grade.

The weighted sum of these two components will constitute the final mark of the discipline. If it is 10 or higher, the student will pass. Otherwise, the student is admitted to the "recurso" examination, which will only replace the grade of the normal exam.

### Main Bibliography

1. On-line lecture notes
2. Hastie, Trevor, et al. *The elements of statistical learning*. Vol. 2. No. 1. New York: Springer, 2009.
3. Montgomery, Douglas C., and George C. Runger. *Applied statistics and probability for engineers*. John Wiley & Sons, 2010
4. Simon Haykin, *Neural Networks: a Comprehensive Foundation*, Prentice-Hall, 1994 (1ª ed.), 1999 (2ª ed.)
5. Daniel Keim, Jörn Kohlhammer, Geoffrey Ellis and Florian Mansmann (Eds.); *Mastering the Information Age: Solving Problems with Visual Analytics*, 2010.
6. Matthew North; *Data mining for the masses*, 2012.