

Ficha de Unidade Curricular – (Versão A3ES 2018-2023)

1. Caracterização da Unidade Curricular.

1.1. **Designação da unidade curricular (1.000 carateres).**
Aprendizagem Automática

1.2. **Sigla da área científica em que se insere (100 carateres).**
INF

1.3. **Duração¹ (100 carateres).**
Semestral

1.4. **Horas de trabalho² (100 carateres).**
162 horas

1.5. **Horas de contacto³ (100 carateres).**
20 horas T + 20 horas TP + 27.5 PL

1.6. **ECTS (100 carateres).**
6

1.7. **Observações⁴ (1.000 carateres).**

É recomendado a quem frequentar a disciplina, ter boas bases teóricas em probabilidade e estatística e em álgebra linear. Adicionalmente, a disciplina tem uma forte componente prática, recorrendo à linguagem de programação Python e a bibliotecas apropriadas (numpy, scipy, scikit-learn, entre outras), e é por isso aconselhável ter um bom domínio da linguagem Python.

1.7. **Remarks (1.000 carateres).**

It is recommended to those who attend the discipline, to have good theoretical bases in probability and statistics and in linear algebra. In addition, the discipline has a strong practical component using the Python programming language and appropriate libraries (numpy, scipy, scikit-learn, among others), and it is therefore advisable to have a good command of the Python language.

2. **Docente responsável e respetiva carga letiva na Unidade Curricular (preencher o nome completo) (1.000 carateres).**
Gonçalo Caetano Marques, 67.5 horas de contacto.

3. **Outros docentes e respetivas cargas letivas na unidade curricular (1.000 carateres).**

4. **Objetivos de aprendizagem (conhecimentos, aptidões e competências a desenvolver pelos estudantes). (1.000 carateres).**

Os estudantes que terminam com sucesso esta unidade curricular serão capazes de:

1. Conhecer o fundamentos teóricos de diversos métodos usados em aprendizagem automática, e implementar e aplicar estes métodos a problemas concretos
2. Compreender o funcionamento de algoritmos supervisionados de classificação e regressão e saber treinar e avaliar estes algoritmos.
3. Compreender o funcionamento e saber aplicar algoritmos não supervisionados de agrupamento e redução de dimensão.
4. Saber combinar ou modificar elementos chave de técnicas de aprendizagem automática para projetar novos algoritmos.
5. Perante novos problemas, escolher os métodos adequados e saber avaliar o desempenho dos mesmos.

4. **Intended learning outcomes (knowledge, skills and competences to be developed by the students). (1.000 characters).**

Students who successfully complete this course unit will be able to:

1. To know the theoretical foundations of several machine learning techniques, and to be able to implement and apply them to concrete problems.
2. To understand how supervised classification and regression algorithms work and to know how to train and evaluate these algorithms.
3. To understand how unsupervised clustering and dimension reduction algorithms work and to know how to apply them.
4. To know how to combine or modify key elements of machine learning techniques to design new algorithms.
5. To be able to choose the appropriate methods for new problems, and to evaluate their performances.

5. Conteúdos programáticos (1.000 carateres).

- I. Introdução à aprendizagem automática, áreas relacionadas e domínios de aplicação
- II. Conceitos fundamentais de classificação: regiões de decisão, funções discriminantes. Medidas de desempenho para a classificação multi-classe e binária.
- III. Conceitos fundamentais de aprendizagem supervisionada. Treino e avaliação de modelos, sobre aprendizagem e capacidade de generalização.
- IV. Pré-processamento de dados: normalizações e análise em componentes principais.
- V. Modelos lineares para regressão e classificação. Regressão linear e polinomial, discriminantes lineares, de Fisher e logísticos. Métodos de regularização.
- VI. Métodos de descida/subida de gradiente.
- VII. Teoria da decisão de Bayes.
- VIII. Técnicas de agrupamento: algoritmo k-médias, dendrogramas.
- IX. Trabalhar com dados de texto.

5. Syllabus (1.000 characters).

- I. Introduction to machine learning, related areas, and fields of application
- II. Key concepts of classification: decision regions, discriminant functions. Performance measures for multi-class and binary classification.
- III. Fundamental concepts of supervised learning. Model training and evaluation, overfitting and generalization capacity.
- IV. Data preprocessing: data normalization and principal component analysis.
- V. Linear models for regression and classification. Linear and polynomial regression, linear, Fisher and logistic discriminants. Regularization techniques.
- VI. Gradient descent methods.
- VII. Bayes decision theory.
- VIII. Clustering methods: k-means and dendrograms.
- IX. Working with text data.

6. Demonstração da coerência dos conteúdos programáticos com os objetivos de aprendizagem da unidade curricular (1.000 carateres).

A metodologia expositiva utilizada para explicar a matéria teórica é reforçada pela exemplificação com problemas concretos, e permite aos alunos perceber como aplicar a matéria dada em situações reais. Assim o aluno adquire os conhecimentos para formalizar um problema concreto, escolher os métodos adequados a aplicar e proceder à sua correta avaliação.

A capacidade aplicar, implementar, e avaliar os métodos estudados é desenvolvida com estudo de casos reais, fichas individuais e de trabalhos práticos.

6. Evidence of the syllabus coherence with the curricular unit's intended learning outcomes (1.000 characters).

The methodology used to explain the theoretical material is reinforced by the exemplification with concrete problems, and allows students to understand how to apply the given material in real situations. Thus the student acquires the knowledge to formalize a concrete problem, to choose the

appropriate methods to apply and to carry out its correct evaluation.
The ability to apply, implement, and evaluate the methods studied is developed with case studies, individual files and practical assignments.

7. Metodologias de ensino (avaliação incluída) (1.000 caracteres).

Ensino teórico-prático, estando previstas 30 aulas a que correspondem 67,5 horas de contacto (15 aulas de 3 horas e 15 de 1,5 horas). O tempo total de trabalho do estudante é de 162 horas. As aulas de carácter teórico destinam-se à exposição e discussão dos principais conteúdos programáticos, incentivando a interatividade e colocação de questões. Os tópicos principais são ainda explorados através da realização de trabalhos práticos de aprendizagem automática. Os resultados da aprendizagem são avaliados com base em fichas individuais (FI) de exercícios e trabalhos práticos (TP) realizados ao longo do semestre, e uma discussão oral sobre ambos. A classificação final (CF) é obtida segundo a fórmula: $CF=FI*0.4+TP*0.6$.

Para ambas as componentes (FI e TP), o aluno deverá obter uma classificação mínima de 9.5 valores.

7. Teaching methodologies (including assessment) (1.000 characters).

Theoretical-practical teaching, with 30 classes corresponding to 67.5 contact hours (15 lessons of 3 hours and 15 of 1,5 hours). The total work time of the student is 162 hours. Theoretical classes are designed to expose and discuss the main programmatic contents, encouraging interactivity and questioning. The main topics are still explored through the realization of computer-based, machine-learning projects. The learning outcomes are evaluated based on individual series (IS) of exercises and projects (P) done during the semester, and an oral discussion on both components. The final grade (FG) is obtained according to the following formula: $FG=IS*0.4+P*0.6$.

In both components (IS and P), the student must have a minimum grade of 9.5.

8. Demonstração da coerência das metodologias de ensino com os objetivos de aprendizagem da unidade curricular (3.000 caracteres).

Os objetivos da unidade curricular são obtidos através de aulas teóricas e práticas, de material de apoio e respetiva bibliografia, da realização de exercícios práticos, e de casos de estudo selecionados pelo docente.

A realização dos trabalhos práticos é acompanhada pelo docente durante as horas de contacto para assegurar o correto desenvolvimento dos conhecimentos e das competências dos estudantes.

8. Evidence of the teaching methodologies coherence with the curricular unit's intended learning outcomes (3.000 characters).

The objectives of the curricular unit are obtained through theoretical and practical classes, bibliographic and other resources, practical exercises and case studies selected by the teacher.

The practical work is carried out by the teacher during the contact hours to ensure the correct development of the knowledge and skills acquired by the students.

9. Bibliografia de consulta/existência obrigatória (1.000 caracteres).

Peter Duda, Richard Hart, David Stork, "Pattern Classification", Wiley, 2007.

Christopher Bishop, "Pattern Recognition and Machine Learning (Information Science and Statistics), Springer-Verlag, 2006.

Andreas Müller, Sarah Guido, "Introduction to Machine Learning with Python", O'Reilly, 2017

¹ Anual, semestral, trimestral, ...

² Número total de horas de trabalho.

³ Discriminadas por tipo de metodologia adotado (T - Ensino teórico; TP - Ensino teórico-prático; PL - Ensino prático e laboratorial; TC - Trabalho de campo; S - Seminário; E - Estágio; OT - Orientação tutorial; O - Outro).

⁴ Assinalar sempre que a unidade curricular seja optativa.