

NCE/16/00007 — Apresentação do pedido - Novo ciclo de estudos

Apresentação do pedido

Perguntas A1 a A4

A1. Instituição de ensino superior / Entidade instituidora:

Universidade Nova De Lisboa

A1.a. Outras Instituições de ensino superior / Entidades instituidoras:

A2. Unidade(s) orgânica(s) (faculdade, escola, instituto, etc.):

Faculdade De Ciências E Tecnologia (UNL)

A3. Designação do ciclo de estudos:

Análise e Engenharia de Big Data

A3. Study programme name:

Big Data Analytics and Engineering

A4. Grau:

Mestre

Perguntas A5 a A10

A5. Área científica predominante do ciclo de estudos:

Informática

A5. Main scientific area of the study programme:

Informatics

A6.1. Classificação da área principal do ciclo de estudos (3 dígitos), de acordo com a Portaria n.º 256/2005, de 16 de Março (CNAEF):

480

A6.2. Classificação da área secundária do ciclo de estudos (3 dígitos), de acordo com a Portaria n.º 256/2005, de 16 de Março (CNAEF), se aplicável:

461

A6.3. Classificação de outra área secundária do ciclo de estudos (3 dígitos), de acordo com a Portaria n.º 256/2005, de 16 de Março (CNAEF), se aplicável:

N/A

A7. Número de créditos ECTS necessário à obtenção do grau:

120

A8. Duração do ciclo de estudos (art.º 3 DL-74/2006, de 26 de Março):

2 anos (4 semestres)

A8. Duration of the study programme (art.º 3 DL-74/2006, March 26th):

2 years (4 semesters)

A9. Número máximo de admissões:

25

A10. Condições específicas de ingresso:

O programa requer bases matemáticas e noções de computação e programação ao nível de um 1º ciclo em engenharia. Podem candidatar-se titulares do grau de licenciado, ou equivalente legal, nas áreas científicas da matemática, informática, economia, gestão, ciências naturais ou da engenharia, sendo as candidaturas sujeitas à apreciação curricular do candidato.

A10. Specific entry requirements:

The program requires the candidate to be knowledgeable of mathematics and computer programming at the level expected from a 1st cycle in general engineering.

Applicants should hold a B.Sc. degree, or equivalent, in the scientific areas of mathematics, computer science, economics, management, natural sciences or engineering, the applications being subject to the curriculum evaluation of the applicant.

Pergunta A11

Pergunta A11

A11. Percursos alternativos como ramos, variantes, áreas de especialização do mestrado ou especialidades do doutoramento em que o ciclo de estudos se estrutura (se aplicável):

Não

A11.1. Ramos, variantes, áreas de especialização do mestrado ou especialidades do doutoramento (se aplicável)

A11.1. Ramos, variantes, áreas de especialização do mestrado ou especialidades do doutoramento, em que o ciclo de estudos se estrutura (se aplicável) / Branches, options, specialization areas of the master or specialities of the PhD (if applicable)

Ramo, variante, área de especialização do mestrado ou especialidade do doutoramento:

Branch, option, specialization area of the master or speciality of the PhD:

<sem resposta>

A12. Estrutura curricular

Mapa I -

A12.1. Ciclo de Estudos:

Análise e Engenharia de Big Data

A12.1. Study Programme:

Big Data Analytics and Engineering

A12.2. Grau:

Mestre

A12.3. Ramo, variante, área de especialização do mestrado ou especialidade do doutoramento (se aplicável):

<sem resposta>

A12.3. Branch, option, specialization area of the master or speciality of the PhD (if applicable):

<no answer>

A12.4. Áreas científicas e créditos que devem ser reunidos para a obtenção do grau / Scientific areas and credits that must be obtained for the awarding of the degree

Área Científica / Scientific Area	Sigla / Acronym	ECTS Obrigatórios / Mandatory ECTS	ECTS Mínimos Optativos* / Minimum Optional ECTS*
Informática / Computer Science and Informatics	I	18	6
Matemática / Mathematics	M	12	6
Competências Complementares / Professional Skills	CC	3	0
Qualquer Área Científica / Any Scientific Area	QAC	0	6
Informática ou Matemática / Mathematics or Computer Science and Informatics	M/I	63	6

Perguntas A13 e A16

A13. Regime de funcionamento:

Diurno

A13.1. Se outro, especifique:

Não aplicável

A13.1. If other, specify:

Not applicable

A14. Local onde o ciclo de estudos será ministrado:

Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Monte da Caparica

A14. Premises where the study programme will be lectured:

Faculdade de Ciências e Tecnologia Universidade Nova de Lisboa Monte da Caparica

A15. Regulamento de creditação de formação e experiência profissional (PDF, máx. 500kB):

[A15._regcredcomp_dr114_16junho2016_FCT_UNL.pdf](#)

A16. Observações:

O Mestrado em Análise e Engenharia de Big Data (MAEBD) visa a educação de analistas, líderes de desenvolvimento do projeto e especialistas em inovação no campo emergente da Ciência de Dados e Engenharia, que tem vindo a criar atualmente novas oportunidades para a gestão de processos e otimização de produtos em todos os campos de atividade. Sendo uma colaboração equilibrada entre os departamentos de Informática e Matemática da FCT NOVA, o MAEBD reúne competências computacionais e matemáticas fundamentais para a resolução de problemas colocados por grandes volumes de dados, proporcionando uma compreensão dos princípios e fundamentos científicos e tecnológicos dos métodos atuais e emergentes.

A área de Big Data, em rápida expansão, inclui temas de ciência de dados, tais como: avaliação e validade; previsão; análise de risco e decisão; análise de grafos e redes de grande dimensão; análise longitudinal; vários tópicos de aprendizagem automática, tais como a seleção de características, redes neuronais e redução de dimensionalidade; tópicos de “Data Warehousing” tais como pesquisa e recuperação da informação, mineração e aquisição de dados e conhecimentos; e ainda tópicos de Engenharia de Dados e Sistemas Computacionais, tais como análise de grandes volumes de dados, processamento de fluxo de dados (“streams”), e computação na nuvem (“cloud computing”). Esta área tem aplicações na gestão de atividades em geral, e em particular em áreas como a saúde, comércio eletrónico, finanças, energia, meio ambiente e urbanismo, telecomunicações, indústria farmacêutica e bioinformática, nos media e nas redes sociais, na previsão, no marketing, gestão de produtos e processos e inovação em todos os domínios da indústria e dos serviços.

O MAEBD acomoda vários perfis de pós-graduação entre Ciência de Dados e Engenharia de Dados, com um forte núcleo de fundamentos da ciência, engenharia e tecnologia. O 1º ano é preenchido com cursos formais, com um bloco obrigatório no 1º semestre de 5 UC, 3 na área de Informática e 2 na área de Matemática. No 2º semestre são exigidos 3 ECTS em Empreendedorismo, comum a todos os programas de 2º ciclo em FCT NOVA e 3 ECTS, num Seminário organizado em colaboração com as empresas parceiras (onde se apresentam cenários relevantes para a engenharia e análise de grandes dados em vários domínios da sociedade) e ainda 4 UC opcionais, de escolha livre a partir de uma oferta adequada, das quais pelo menos uma deve pertencer a cada uma das áreas científicas do programa (i.e. Informática e Matemática) e uma deve pertencer ao Bloco Livre FCT, que inclui UC de todas as áreas científicas da FCT NOVA. Este bloco de UC é definido anualmente pelo Conselho Científico da FCT NOVA. O 3º e 4º semestres são dedicados à dissertação, normalmente desenvolvida em colaboração com uma entidade externa, parceira do curso, que forneça um problema e cenário do mundo real, mas que pode igualmente ter o perfil de uma dissertação de investigação orientada.

A16. Observations:

The Master in Big Data Analytics and Engineering (MAEBD) targets the education of analysts, project development leaders and innovation specialists in the emergent field of Data Science and Engineering, which is nowadays creating novel opportunities for process management and product optimization in every field of activity. Building on an evenly balanced collaboration between the FCT NOVA Mathematics and Computer Science / Informatics Departments, the MAEBD brings together fundamental computational and mathematical skills for large scale datadriven problem solving while providing a principled understanding of scientific and technological underpinnings of current and emerging methods.

The rapidly expanding area of Big Data includes Data Science topics such as: evaluation and validity; prediction; risk analysis and decision; large graph and network analysis; longitudinal analysis; machine learning topics such as feature selection, neural networks, and dimensionality reduction; data warehousing topics such as information search and retrieval, data and knowledge mining and acquisition; and Data Engineering and Computing Infrastructures topics such as big data analytics, stream processing, and cloud computing. It finds applications in general business, health, ecommerce, finance, energy, environment and urbanism, telecom, pharmaceuticals, media, social networks, bioinformatics, forecasting, marketing, general process and product management and innovation in all fields of

industry and services.

The overall structure of the MAEBD curriculum accommodates a variety of graduate profiles between Data Science and Data Engineering, while ensuring a strong core on key science, engineering and technology principles. The first year is composed of formal courses. In the first semester the students are offered a compulsory block of 5 courses, 3 in the CS / Informatics area and 2 in the Mathematics area. In the second semester there is a compulsory 3 ECTS requirement in Entrepreneurship, common to all 2nd cycle programs at FCT NOVA, and a 3 ECTS seminar course, organized in collaboration with partner companies (where students learn from invited practitioner speakers about relevant scenarios of big data engineering and analytics in various societal domains) and also 4 elective courses from which students can choose freely from a suitable offer (to be defined in the full proposal). From these 4 courses at least one must belong to each of the two main scientific areas of the program (either Math or Informatics/CS), and one from the “Bloco Livre FCT”, that includes courses of all scientific areas of FCT NOVA, and is defined annually by the Scientific Council of FCT NOVA. The third and fourth semester are dedicated to the preparation and elaboration of the master project / thesis curricular unit, typically developed in collaboration with an external partner entity providing some real world scenario and problem, but which may also accommodate research oriented thesis.

Instrução do pedido

1. Formalização do pedido

1.1. Deliberações

Mapa II - Aprovação pelo Reitor da UNL, ouvido o Colégio de Diretores

1.1.1. Órgão ouvido:

Aprovação pelo Reitor da UNL, ouvido o Colégio de Diretores

1.1.2. Cópia de ata (ou extrato de ata) ou deliberação deste órgão assinada e datada (PDF, máx. 100kB):

[1.1.2._Despacho_Senhor_Reitor_Big_Data_11-10-2016.pdf](#)

Mapa II - Conselho Científico da FCT-UNL

1.1.1. Órgão ouvido:

Conselho Científico da FCT-UNL

1.1.2. Cópia de ata (ou extrato de ata) ou deliberação deste órgão assinada e datada (PDF, máx. 100kB):

[1.1.2._Declaração_CC_MAEED.pdf](#)

Mapa II - Conselho Pedagógico da FCT-UNL

1.1.1. Órgão ouvido:

Conselho Pedagógico da FCT-UNL

1.1.2. Cópia de ata (ou extrato de ata) ou deliberação deste órgão assinada e datada (PDF, máx. 100kB):

[1.1.2._Declaração_CP_MAEED.pdf](#)

1.2. Docente(s) responsável(eis) pela coordenação da implementação do ciclo de estudos

1.2. Docente(s) responsável(eis) pela coordenação da implementação do ciclo de estudos

A(s) respetiva(s) ficha(s) curricular(es) deve(m) ser apresentada(s) no Mapa V.

Pedro Manuel Corrêa Calvente Barahona (Informática) /Jorge Orestes Lasbarrères Cerdeira (Matemática)

2. Plano de estudos

Mapa III - - 1º Ano / 1º Semestre

2.1. Ciclo de Estudos:

Análise e Engenharia de Big Data

2.1. Study Programme:

Big Data Analytics and Engineering

2.2. Grau:

Mestre

2.3. Ramo, variante, área de especialização do mestrado ou especialidade do doutoramento (se aplicável):
<sem resposta>

2.3. Branch, option, specialization area of the master or speciality of the PhD (if applicable):
<no answer>

2.4. Ano/semestre/trimestre curricular:
1º Ano / 1º Semestre

2.4. Curricular year/semester/trimester:
1st Year / 1st Semester

2.5. Plano de Estudos / Study plan

Unidade Curricular / Curricular Unit	Área Científica / Scientific Area (1)	Duração / Duration (2)	Horas Trabalho / Working Hours (3)	Horas Contacto / Contact Hours (4)	ECTS	Observações / Observations (5)
Estatística Numérica Computacional / Computational Numerical Statistics	M	Semestral / Semester	168	TP:56	6	Obrigatória / Mandatory
Recuperação de Informação / Information Retrieval	I	Semestral / Semester	168	T:28;PL:28	6	Obrigatória / Mandatory
Estatística Multivariada / Multivariate Statistics	M	Semestral / Semester	168	TP:56	6	Obrigatória / Mandatory
Aprendizagem Automática / Machine Learning	I	Semestral / Semester	168	T:28;PL:28	6	Obrigatória / Mandatory
Sistemas para Processamento de Big Data / Systems for Big Data Processing	I	Semestral / Semester	168	T:28;PL:28	6	Obrigatória / Mandatory

(5 Items)

Mapa III - - 1º Ano / 2º Semestre

2.1. Ciclo de Estudos:
Análise e Engenharia de Big Data

2.1. Study Programme:
Big Data Analytics and Engineering

2.2. Grau:
Mestre

2.3. Ramo, variante, área de especialização do mestrado ou especialidade do doutoramento (se aplicável):
<sem resposta>

2.3. Branch, option, specialization area of the master or speciality of the PhD (if applicable):
<no answer>

2.4. Ano/semestre/trimestre curricular:
1º Ano / 2º Semestre

2.4. Curricular year/semester/trimester:
1st Year / 2nd Semester

2.5. Plano de Estudos / Study plan

Unidade Curricular / Curricular Unit	Área Científica / Scientific Area (1)	Duração / Duration (2)	Horas Trabalho / Working Hours (3)	Horas Contacto / Contact Hours (4)	ECTS	Observações / Observations (5)
Empreendedorismo / Entrepreneurship	CC	Semestral / Semester	84	TP:28	3	Obrigatória / Mandatory
Unidade Curricular do Bloco Livre A / Unrestricted Elective A	QAC	Semestral / Semester	168	depende da opção / selection dependent	6	Opção / Optional

Seminário / Seminar	I/M	Semestral / Semester	84	S:28	3	Obrigatória / Mandatory
Unidade de Especialização I / Specialization Unit I	I	Semestral / Semester	168	T:28 PL:28	6	Opção / Optional
Unidade de Especialização II / Specialization Unit II	M	Semestral / Semester	168	T:28 PL:28	6	Opção / Optional
Unidade de Especialização III / Specialization Unit III	I/M	Semestral / Semester	168	T:28 PL:28	6	Opção / Optional

(6 Items)

Mapa III - - 1º Ano / 2º Semestre – Unidades de Especialização I, II e III

2.1. Ciclo de Estudos:

Análise e Engenharia de Big Data

2.1. Study Programme:

Big Data Analytics and Engineering

2.2. Grau:

Mestre

2.3. Ramo, variante, área de especialização do mestrado ou especialidade do doutoramento (se aplicável):

<sem resposta>

2.3. Branch, option, specialization area of the master or speciality of the PhD (if applicable):

<no answer>

2.4. Ano/semestre/trimestre curricular:

1º Ano / 2º Semestre – Unidades de Especialização I, II e III

2.4. Curricular year/semester/trimester:

1st Year / 2nd Semester - Specialization Units I, II and III

2.5. Plano de Estudos / Study plan

Unidade Curricular / Curricular Unit	Área Científica / Scientific Area (1)	Duração / Duration (2)	Horas Trabalho / Working Hours (3)	Horas Contacto / Contact Hours (4)	ECTS	Observações / Observations (5)
Métodos Bayesianos / Bayesian Methods	M	Semestral / Semester	168	TP:56	6	Opção / Optional
Decisão e Risco / Decision and Risk	M	Semestral / Semester	168	TP:56	6	Opção / Optional
Visualização Interativa de Dados / Interactive Data Visualization	I	Semestral / Semester	168	T:28 PL:28	6	Opção / Optional
Análise de Grandes Grafos / Large Graph Analytics	M	Semestral / Semester	168	TP:56	6	Opção / Optional
Processamento de Streams / Stream Processing	I	Semestral / Semester	168	T:28 PL:28	6	Opção / Optional
Algoritmos para Redes Complexas / Algorithms for Complex Networks	I	Semestral / Semester	168	T:28 PL:28	6	Opção / Optional
Tópicos de Otimização / Topics in Optimization	M	Semestral / Semester	168	TP:56	6	Opção / Optional
Aprendizagem com Dados Não Estruturados / Learning from Unstructured Data	I	Semestral / Semester	168	T:28 PL:28	6	Opção / Optional
Prospecção de Dados / Data Analytics and Text Mining	I	Semestral / Semester	168	T:28 PL:28	6	Opção / Optional

(9 Items)

Mapa III - - 2º Ano

2.1. Ciclo de Estudos:

Análise e Engenharia de Big Data

2.1. Study Programme:*Big Data Analytics and Engineering***2.2. Grau:***Mestre***2.3. Ramo, variante, área de especialização do mestrado ou especialidade do doutoramento (se aplicável):***<sem resposta>***2.3. Branch, option, specialization area of the master or speciality of the PhD (if applicable):***<no answer>***2.4. Ano/semestre/trimestre curricular:***2º Ano***2.4. Curricular year/semester/trimester:***2nd year***2.5. Plano de Estudos / Study plan**

Unidade Curricular / Curricular Unit	Área Científica / Scientific Area (1)	Duração / Duration (2)	Horas Trabalho / Working Hours (3)	Horas Contacto / Contact Hours (4)	ECTS	Observações / Observations (5)
Dissertação / Dissertation (1 Item)	I/M	Anual / Annual	1680	OT:40	60	Obrigatória / Mandatory

3. Descrição e fundamentação dos objetivos, sua adequação ao projeto educativo, científico e cultural da instituição, e unidades curriculares**3.1. Dos objetivos do ciclo de estudos****3.1.1. Objetivos gerais definidos para o ciclo de estudos:**

O objetivo do Mestrado em Análise e Engenharia de Big Data da FCT NOVA é formar especialistas em análise e engenharia de dados em grande escala, habilitados a liderar e desenvolver atividades de análise, projeto, inovação e investigação, em contexto empresarial e/ou académico, e com bases para aceder a um 3º Ciclo. Como disposto no DL74/2006, os formados pelo curso deverão:

- Conhecer princípios e técnicas, e desenvolver capacidades analíticas de conceção e desenvolvimento, assim como inovação tecnológica e científica.*
- Saber aplicar os conhecimentos na resolução de novos problemas, incluindo em contextos multidisciplinares.*
- Saber pesquisar e integrar conhecimentos, avaliar e comparar soluções criticamente.*
- Refletir de forma útil sobre implicações éticas e sociais da atividade, numa perspetiva de Ciência, Tecnologia e Sociedade, assim como comunicar resultados de forma clara e rigorosa;*
- Demonstrar competências e atitudes colaborativas, que suportem a aprendizagem ao longo da vida.*

3.1.1. Generic objectives defined for the study programme:

The objective of the FCT NOVA Big Data Analytics and Engineering Master programme is to educate specialists prepared to lead, and develop analysis, project, and innovation activities, in a corporate or academic context, and with basis to pursue a 3rd cycle. As stipulated in DL74/2006, program graduates should:

- Be knowledgeable of principles, techniques and analytical skills supporting design and conceptual engineering activities, as well as technological and scientific innovation.*
- Be able to apply knowledge in the solution of new problems, including in multidisciplinary contexts.*
- Be able to research and integrate knowledge, and critically evaluate and compare solutions.*
- Demonstrate ability to reflect on the ethical and social consequences of their activity, with Science, Technology and Society mindset, and to communicate results with clarity and technical rigour.*
- Develop collaborative attitude and teamwork competencies, able to support life long learning skills.*

3.1.2. Objetivos de aprendizagem (conhecimentos, aptidões e competências) a desenvolver pelos estudantes:

- Compreender profundamente os fundamentos matemáticos e computacionais, as metodologias, as tecnologias e o estado-da-arte para a ciência e engenharia de Big Data;*
- Integrar de forma coesa os aspetos matemáticos, computacionais e tecnológicos necessários à ciência e engenharia de Big Data, e saber utilizá-los na formulação de questões, e na procura de respostas, modelos, e soluções computacionais eficazes para problemas de análise complexos em várias áreas de atividade social e interdisciplinar*
- Combinar capacidades de desenvolvimento técnico com o contexto em que este se insere, num espírito de auto-*

atualização contínua;

- Enfrentar a complexidade de forma proativa e criativa;*
- Trabalhar em equipa, saber documentar objetivos, soluções e resultados a especialistas e a não especialistas;*
- Endereçar aspetos ético-profissionais, incluindo responsabilidades sociais e ambientais;*
- Relatar, sintetizar, e argumentar temas técnicos, demonstrar visão crítica perante soluções propostas.*

3.1.2. Intended learning outcomes (knowledge, skills and competences) to be developed by the students:

- Deeply understand the mathematical and computational foundations, the methodologies, technologies and the state-of-the art for Big Data Science and Engineering;*
- Integrate the mathematical, computational and technological instruments necessary to Big Data science and engineering, and leverage them to formulate questions, and find answers, models, and efficient computational solutions for complex analysis problems in various areas of social and interdisciplinary activity ;*
- Combine solid technical development skills with the context in which they are embedded, promoting a lifelong learning attitude;*
- Deal with the complexity with proactivity and creativity;*
- Work in a team, know how to document objectives, solutions and results, both to specialists and non-specialists;*
- Address ethical-professional aspects, including social / environmental responsibilities;*
- Relate, synthesise, and argue technical themes, develop a critical view before proposed solutions;*

3.1.3. Inserção do ciclo de estudos na estratégia institucional de oferta formativa face à missão da instituição:

A FCT NOVA é uma instituição de ensino superior de Ciência e de Engenharia que tem como missão desenvolver:

- a) Investigação científica competitiva no plano internacional, privilegiando áreas interdisciplinares, incluindo a investigação orientada para a resolução de problemas que afetam a sociedade;*
 - b) Um ensino de excelência, com ênfase crescente na investigação realizada, veiculado por programas académicos competitivos a nível nacional e internacional;*
 - c) Uma base alargada de participação interinstitucional orientada para a integração das diferentes culturas científicas, com vista à criação de sinergias inovadoras para o ensino e para a investigação;*
 - d) Uma forte ligação à sociedade, transferência de conhecimentos, tecnologias e serviços, quer no plano interno, quer internacional, capaz de contribuir para o desenvolvimento social e para a qualificação dos recursos humanos.*
- O curso de Mestrado em Análise e Engenharia de Big Data (MAEBD) da FCT NOVA contribui para a formação de especialistas com competências de projeto e investigação, desenvolvimento e inovação numa área de enorme relevância e atualidade, através de colaboração entre os Departamentos de Informática e Matemática. O aumento da disponibilidade de Big Data cria oportunidades para otimizar processos de gestão e produtos em termos de custos, produtividade e competitividade. Os fundamentos e tecnologias de exploração de Big Data contribuem para a criação de produtos e serviços inovadores, competitivos, geradores de vantagens económicas em todas as áreas. Focando-se em técnicas avançadas de base matemática, estatística, e computacional, como a análise multivariada, a aprendizagem automática, a ciência de redes complexas, o processamento paralelo e distribuído, a modelação, recuperação e visualização de informação, o MAEBD acomoda vários tipos de perfis entre a ciência de dados e a engenharia de dados, competências chave num sector em expansão para as quais existe grande procura de recursos humanos, mas lacuna de oferta educativa a nível nacional.*
- É uma aposta estratégica da FCT NOVA continuar a contribuir para o desenvolvimento da educação avançada e investigação em áreas emergentes e interdisciplinares onde detém especiais competências científicas e experiência pedagógica, tendo em particular sido pioneira na área da Engenharia Informática em Portugal. É de referir o alinhamento do MAEBD com o Perfil Curricular FCT, que favorece o desenvolvimento de competências transversais, na área da comunicação, ciência tecnologia e sociedade, e empreendedorismo, potenciando a ligação da escola à sociedade em geral através da oferta de atividades colaborativas com o exterior, por exemplo através da oferta de estágios e dissertações em regime académico-empresarial.*
- Finalmente, o curso contribui ainda para formar estudantes preparados para aceder a um 3º ciclo (Doutoramento), potenciando a investigação realizado no Departamento de Informática e/ou Matemática e nos seus Centros de Investigação.*

3.1.3. Insertion of the study programme in the institutional training offer strategy against the mission of the institution:

FCT NOVA is a higher education institution in the Science and Engineering area. Its mission comprises:

- a) Internationally competitive scientific research, with an interdisciplinary emphasis, including research targeted to address societal challenges.*
 - b) Excellence in teaching, supported on research activities; implemented by national and internationally competitive academic programs;*
 - c) A broad basis for inter-institutional participation, towards the integration of different scientific cultures, fostering the creation of synergies for research and education.*
 - d) A strong social commitment, promoting technology, service and knowledge transfer, able to contribute for the social development and for the high quality qualification of human resources.*
- The Master in Big Data Analysis and Engineering (MAEBD) now proposed contributes for the education of specialists competent for project development and innovation activities in the key emergent field of Data Science and Engineering, combining competences of the Math and Computer Science Departments. The increasing availability of Big Data is creating novel opportunities for process management and product optimisation, in terms of cost, productivity and competitiveness. The principles and technologies for exploiting Big Data strongly contribute to the improvement and creation of innovative products and services in every area. Focusing on advanced mathematical, statistical and computing techniques, such as multivariate analysis, machine learning, network science, parallel and distributed processing, and information modeling, retrieval and visualisation, the MAEBD program accommodates a range of profiles between data science and data engineering. These are key roles in an expanding sector worldwide, for which there is an impressive increase in demand, but currently a lack of educational offer at the national level. Therefore, FCT NOVA is strategically committed to continue contributing for the development of advanced education*

and research in its fields of pedagogical and scientific competence, in particular as the pioneering school in the Informatics and Computer Science in Portugal. The MAEBD program is perfectly aligned with the "FCT Curricular Profile", which favours the emergence of various transferable skills in areas such as communication, science, technology and society, and entrepreneurship, leveraging the connections between the school and the society at large, in particular, with firms and companies, for instance by offering to students internships and master dissertations in academic-industrial collaboration.

Finally, the program also contributes for the education of strong candidates to a 3rd cycle (PhD program) in Computer Science or Mathematics, leveraging on the research pursued in the involved Departments and their associated research units.

3.2. Adequação ao projeto educativo, científico e cultural da Instituição

3.2.1. Projeto educativo, científico e cultural da Instituição:

De acordo com os seus Estatutos, a Faculdade de Ciências e Tecnologia tem identidade e missão idênticas às da UNL, dirigidas às áreas de Ciências e de Engenharia. Na sua missão, enquanto instituição universitária que se pretende de referência, inclui-se o desenvolvimento de investigação competitiva no plano internacional, privilegiando áreas interdisciplinares e a investigação orientada para a resolução dos problemas que afetam a sociedade, bem como a oferta de ensino de excelência, com ênfase crescente em segundos e terceiros ciclos, mas fundado em primeiros ciclos sólidos, veiculado por programas académicos competitivos a nível nacional e internacional, erigindo o mérito como medida essencial da avaliação. Fundamentalmente, a política de ensino e investigação tem por objetivo promover a qualidade e reconhecimento destas atividades, devendo a investigação ser progressivamente incorporada nas estruturas curriculares dos ciclos de estudos, proporcionando uma oferta educativa atualizada e substancialmente diferenciadora. Por outro lado, a Faculdade dispõe de uma política de qualidade que visa assegurar a melhoria contínua das suas atividades, por forma a aumentar, de modo sustentado, a sua eficiência e corresponder às expectativas decorrentes do seu objeto social. Neste âmbito, o projeto educativo tem contemplado não só a criação de novas áreas de estudo, decorrentes da evolução da economia associada às mudanças sociais, como também a introdução de métodos de ensino e de avaliação conducentes a uma aprendizagem mais eficiente e a reestruturação da oferta formativa existente. Neste último caso, salienta-se a recente introdução, em todos os cursos de Licenciatura, de Mestrado e de Mestrado Integrado, de competências complementares, designadamente soft skills, contacto com empresas ou investigação e empreendedorismo, configurando o designado "Perfil Curricular FCT" como fator diferenciador dos diplomados da instituição e elemento facilitador da sua inserção na vida ativa (<http://www.fct.unl.pt/perfil-curricular-fct>). Com o intuito de progredir para uma escola "research oriented", a Faculdade tem vindo a adotar uma política de incentivos para o desenvolvimento de atividades de investigação, potenciando o mérito dos seus docentes como referencial e, ainda, uma política promotora de transferência da tecnologia e do conhecimento gerados para a Sociedade através de parcerias com empresas, licenciamento de propriedade industrial e apoio à criação de empresas spin-off. A Faculdade atribui grande importância às atividades culturais que disponibiliza aos seus estudantes, considerando que valorizam a qualidade dos serviços educativos que oferece e que constituem elemento diferenciador para a notoriedade da Escola. Assim, para cada ano letivo é programado um extenso conjunto de atividades culturais de alto nível (palestras, conferências, debates, exposições de arte) com a intervenção de personalidades detentoras de elevado prestígio nacional e internacional.

3.2.1. Institution's educational, scientific and cultural project:

According to its Statutes, the Faculty of Sciences and Technology identity and mission in the Sciences and Engineering areas are similar to those of the UNL – Universidade NOVA de Lisboa. As a higher education institution striving to be a reference, it includes the development of competitive research at international level that privileges interdisciplinary areas and research aimed at solving social problems, as well as an educational excellence offer increasingly focused on second and third cycles, but founded on solid first cycles with competitive academic programs at both national and international levels, adopting merit as the essential measure of assessment. Basically the policy for teaching and research aims at promoting quality and recognition of those activities, increasingly incorporating research in the curricular structures of the study cycles, enabling an updated educational offer expected to be positively discriminated. On the other hand, the Faculty is enforcing a quality policy for the continuous improvement of its activities in order to increment its efficiency in a sustainable process leading to a better achievement of its social responsibilities. Therefore, its educational project includes not only the creation of new study areas that can follow economical evolution associated to social changes but also the introduction of teaching and assessment methods aimed at improving the learning efficiency, and the restructuring of the existent educational offer. About this last issue, it is worth mentioning the recent introduction, in all first and second study cycles and Integrated Master programs, of common competences, namely soft skills, undergraduate practice or research opportunities and entrepreneurship, leading to the so-called "Perfil Curricular FCT" (FCT Curricular Profile) as a differentiating feature of the institution graduates and a facilitator of their insertion in the active life (<http://www.fct.unl.pt/perfil-curricular-fct>). As the Faculty aims to become a research oriented school, a policy of incentives to research development is being adopted fostering the merit of its academic staff and, also, a policy aims at promoting the technology and knowledge transfer to the Society through partnerships with companies, licensing of industrial property and support to the creation of spin-off companies. Cultural activities are looked as an important aspect of the Faculty's educational offer that contributes to a positive discrimination of the School. For each academic year a set of high-level cultural activities is scheduled, such as seminars, conferences, debates and art exhibitions, with the cooperation of prominent individualities holding high national and international prestige.

3.2.2. Demonstração de que os objetivos definidos para o ciclo de estudos são compatíveis com o projeto educativo, científico e cultural da Instituição:

A Faculdade de Ciências e Tecnologia da NOVA (FCT NOVA) inclui na sua missão o desenvolvimento de investigação competitiva no plano internacional, bem como a oferta de ensino de excelência, com ênfase crescente em segundos e

terceiros ciclos, fundado em primeiros ciclos sólidos, veiculado por programas académicos competitivos a nível nacional e internacional. Os objetivos do Mestrado em Análise e Engenharia de Big Data (MAEBD) da FCT NOVA aqui proposto, integra de forma coesa e flexível os objetivos de um 2º ciclo acessível a detentores de um 1º ciclo nas áreas de ciências, engenharias, economia ou afins, e visa a formação de especialistas habilitados a desenvolver atividades profissionais de projeto, liderança, e inovação, em muitos casos em contexto de investigação, com bases para aceder a um 3º Ciclo, de acordo o Artº 15 do DL74/2006. Esta orientação é implementada numa lógica de continuidade, oferecendo a diplomados de 1º ciclo uma formação sólida, coerente e relevante em termos de mobilidade, numa área de grande procura e empregabilidade a nível mundial. Os objetivos do MAEBD não se sobrepõem com o de outra oferta educativa existente, sendo que o plano curricular combina de forma equilibrada conteúdos das áreas científicas da Matemática e da Informática, organizados em unidades curriculares especialmente adaptadas e / ou especificamente desenhados tendo em conta os objetivos do curso, e a possível heterogeneidade na formação de base dos alunos recrutados.

Todos estes objetivos estão claramente alinhados com o projeto educativo, científico e cultural da FCT NOVA. O curso adequa-se ao "Perfil Curricular da FCT", conjunto de características comuns a todos os cursos da escola que favorecem o desenvolvimento de competências transversais, potenciam a ligação à sociedade, e desenvolvem uma cultura de inovação, empreendedorismo, e desenvolvimento científico. Em integração com a investigação realizada no Departamento de Informática, de Matemática, e seus Centros, o curso contribui ainda para formar estudantes bem preparados para aceder a cursos de 3º ciclo em Informática ou Matemática, em particular nos seus programas doutorais. São inúmeras as áreas utilizadoras dos mestres formados pelo MAEBD, podendo referir-se o mundo dos negócios em geral, saúde, comércio eletrónico, finanças, energia, ambiente, telecomunicações, media, redes sociais, bioinformática, previsão, marketing, inovação, e gestão de processos e produtos em todas as áreas da indústria e serviços.

É uma aposta estratégica da FCT NOVA continuar a contribuir decisivamente para o desenvolvimento da educação avançada e investigação científica e inovação em áreas emergentes, sendo escola pioneira na área de Informática no País.

3.2.2. Demonstration that the study programme's objectives are compatible with the Institution's educational, scientific and cultural project:

The Faculdade de Ciências e Tecnologia da NOVA (FCT NOVA) includes in its mission the development of internationally competitive research, excellence in teaching with an emphasis on 2nd and 3rd cycles, while being founded in solid 1st cycles, implemented in nationally and internationally competitive academic programs.

The Master in Big Data Analysis and Engineering (MAEBD) integrates, in a tightly coupled but flexible way, the goals of a 2nd cycle suitable for 1st cycle graduates in a science, engineering, economics or close areas, towards the education of specialists able to develop professional activities involving project, leadership, innovation, often in a R&D context, and potential candidates for a PhD, according to the Artº15 of DL74/2006. These guidelines are implemented in a progressive scheme, so that incoming 1st cycle graduates may receive a solid, coherent and relevant education, leveraging mobility mechanism, in a field of very high demand and employability worldwide. The objectives of the MAEBD do not overlap any other currently existing educational offer at FCT, its curricular plan combining, in a very well balanced way, content from Mathematics and Informatics. Such content has been organised in specifically adapted or designed curricular units taking into account the program objectives and expected differences in the basic education of the incoming students.

These objectives are clearly aligned with the educational, scientific and cultural project of FCT NOVA.

The program conforms to the "FCT Curricular Profile", which favours the emergence of various transferable skills in areas such as communication, science, technology and society, and entrepreneurship, leveraging the connections between the school and the society at large, and promotes a culture of innovation, entrepreneurship, and scientific development. In harmony with the education and research activities of the participating Departments, the master program contributes to the education of potentially strong candidates to local PhD programs. There are many areas in which MAEBD graduates may develop professional activities: general business, health, e-commerce, finance, energy, environment and urbanism, telecom, pharmaceutical, media, social networks, bioinformatics, forecasting, marketing, general process and product management and innovation in all fields of industry and services.

It is a strong strategic commitment of FCT NOVA to continue contributing for the development of advanced education and research in its field of pedagogical and scientific expertise, in particular as the pioneering higher education Informatics program at the national level.

3.3. Unidades Curriculares

Mapa IV - Estatística Numérica Computacional / Computational Numerical Statistics

3.3.1. Unidade curricular:

Estatística Numérica Computacional / Computational Numerical Statistics

3.3.2. Docente responsável (preencher o nome completo) e respetivas horas de contacto na unidade curricular:

Marta Cristina Vieira Faias Mateus – TP: 20h

3.3.3. Outros docentes e respetivas horas de contacto na unidade curricular:

Isabel Cristina Maciel Natário – TP: 20h

Pedro José dos Santos Palhinhas Mota – TP: 16h

3.3.4. Objetivos de aprendizagem (conhecimentos, aptidões e competências a desenvolver pelos estudantes):

A programação é fundamental no desenvolvimento de metodologias estatísticas, frequentemente assentes em métodos numéricos. O objetivo desta UC é aprender estatística computacional, com ênfase em métodos estatísticos e software mais relevantes para problemas de inferência de grande dimensionalidade; compreender e ser capaz de aplicar os seguintes métodos: Monte Carlo, técnicas de reamostragem (Bootstrap e Jackknife), modelação e inferência computacional, algoritmos do tipo Newton-Raphson, algoritmo EM, técnicas de amostragem-reamostragem e simulação iterativa (Método de Monte Carlo via Cadeias de Markov, MCMC). Ser capaz de utilizar o software estatístico R-project na aplicação dos métodos anteriores, em casos de estudo, recorrendo a bibliotecas estatísticas desenvolvidas especificamente para os métodos ou adaptando-as de modo a resolver os problemas de modo eficiente. Ser ainda capaz de escrever relatórios cientificamente bem documentados sobre as análises efetuadas.

3.3.4. Intended learning outcomes (knowledge, skills and competences to be developed by the students):

Programming plays a fundamental role in the development of statistical methodologies, which frequently are based on numerical methods. The objective of this CU is to learn computational statistics, focusing on statistical methods and software more relevant for high dimensionality inference problems; to be able to understand and apply the following methods: Monte Carlo, resampling techniques (Bootstrap and Jackknife), computational modelling and inference, algorithms of Newton-Raphson type, EM algorithm, sampling-resampling techniques and iterative simulation (Monte Carlo via Markov Chain, MCMC method). To be able to use the statistical software R-project in applying the previous methods, in case study problems, by using the adequate available libraries or by properly modifying them, if needed. To be able to write scientifically well documented reports about the analyses performed.

3.3.5. Conteúdos programáticos:

1. R-project
2. Geração de números aleatórios, geração de variáveis aleatórias discretas e contínuas
3. Métodos de Monte Carlo
4. Técnicas de redução de variância
5. Técnicas de reamostragem: Bootstrap e Jackknife
6. Método de Newton-Raphson
7. Métodos de amostragem de importância
8. Métodos de Monte Carlo via Cadeias de Markov (MCMC)
9. Algoritmo EM
10. Método dos Scores de Fisher (modelos lineares generalizados)
11. Aplicações dos métodos em vários contextos (regressões logística, Poisson, Gaussiana, Gama, séries temporais, modelos hierárquicos, etc.)
12. Uso das técnicas aprendidas e adaptação das bibliotecas do R a casos de estudo de índole prática
13. Elaboração de relatórios onde se documenta e suporta cientificamente, recorrendo à estatística, a análise realizada e as conclusões relativas a cada caso de estudo

3.3.5. Syllabus:

1. R-project
2. Pseudo-random number generation (discrete and continuous)
3. Monte Carlo methods
4. Variance reduction techniques
5. Resampling techniques: Bootstrap and Jackknife
6. Newton-Raphson method
7. Importance sampling methods
8. Monte Carlo via Markov Chain (MCMC) methods
9. EM algorithm
10. Fisher scoring method (generalized linear models).
11. Application of the methods in different contexts (logistic, Poisson, Gaussian, Gamma regression, time series, hierarchical models, etc.)
12. Utilization of the above techniques and adaptation of the R libraries to practical case studies
13. Writing of well documented reports, where properly justified statistical analyses of a study case can be used to take and support conclusions

3.3.6. Demonstração da coerência dos conteúdos programáticos com os objetivos de aprendizagem da unidade curricular:

Principia-se pela introdução do software R-project, linguagem na qual serão programados os métodos numéricos em estatística que serão ensinados (ponto 1 do conteúdo programático); nos pontos 2 e 3 estudam-se a geração de números aleatórios e os métodos Monte Carlo, o ponto 4 introduz as técnicas de redução de variância de modo a tornar a estimação robusta, os pontos 5, 7 e 8 são dedicados às técnicas de amostragem e reamostragem. Os pontos 6, 9 e 10 expõem o método de Newton-Raphson, o algoritmo EM e o método dos scores de Fisher, essenciais à modelação estatística. Nestes tópicos estudam-se os métodos e técnicas e também a teoria que os suporta pelo que se cobrem em simultâneo os dois primeiros objetivos enunciados. No fim de cada capítulo resolvem-se exemplos de aplicação das técnicas estudadas com recurso ao R pelo que se cobre também parte do terceiro objetivo enunciado. Nos pontos 11 a 13 aplicam-se os diferentes métodos a alguns modelos específicos (regressões logística, Poisson, Gaussiana, Gama, séries temporais, modelos hierárquicos, etc.) e elaboram-se os respetivos relatórios cobrindo assim os terceiro e quarto objetivos.

3.3.6. Evidence of the syllabus coherence with the curricular unit's intended learning outcomes:

We begin with an introduction to software R-project, the language in which the numerical methods in statistics that are taught here are to be programmed (point 1 of the syllabus); points 2 and 3 are devoted to pseudo-random number

generation and Monte-Carlo methods, point 4 introduces variance reduction techniques for addressing robustness of the estimates, points 5, 7 and 8 are dedicated to sampling and resampling techniques. Points 6, 9 and 10 detail the Newton-Raphson method, the EM algorithm and the Fisher scoring method, important parts of statistical modelling. For all these topics we study the methods, the techniques and also the theory that supports them and, therefore, we cover the first and second objectives.

At the end of each chapter, practical examples are solved using the R and, thus, part of the third objective is also covered.

In points 11 to 13, the different methods are applied to specific models (logistic, Poisson, Gaussian, Gamma regression, time series, hierarchical models, etc.) and the respective reports are written, covering the third and fourth objectives.

3.3.7. Metodologias de ensino (avaliação incluída):

Aulas teórico-práticas onde a par com a exposição dos conceitos e resultados fundamentais serão apresentados exemplos ilustrativos desses conceitos e resultados. De seguida são apresentados problemas que são resolvidos em laboratório usando o software R. Estes problemas são realizados em laboratório, pretendendo-se que os alunos participem ativamente na sua resolução. Deste modo os alunos adquirem as competências relativas ao software estatístico e também aos conceitos.

A avaliação consiste na realização de quatro trabalhos com relatório e/ou um teste final escrito. Os trabalhos que os alunos realizam para avaliação podem ser apresentados e discutidos na sala de aula.

3.3.7. Teaching methodologies (including assessment):

Lecture-lab classes are the adequate way to convey the course contents to students as, together with the explanation of the main concepts and results, illustrative examples are given. Then, some problems are proposed to students to be solved in a lab context, using the software R, and students are supposed to take active part in their resolution. In this way, students acquire the basic expertise in the application of the statistic software as well as the concepts.

The evaluation consists on the resolution of four practical problems and with written report and/or a written final test. The practical problems solved for the evaluation may be presented and discussed by the students in classroom.

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

A forte componente prática desta UC que faz uso intensivo de computação justifica que as aulas sejam teórico-práticas o que permite uma maior proximidade temporal entre a exposição dos resultados e conceitos e a sua aplicação através da implementação computacional dos correspondentes métodos e técnicas. A componente de avaliação em que os alunos resolvem problemas em grupo e depois apresentam na aula permite que todos os alunos possam aprender e discutir as diferentes especificidades de cada problema.

3.3.8. Evidence of the teaching methodologies coherence with the curricular unit's intended learning outcomes:

The strong practical component of this course that makes intensive use of computation justifies the lecture-lab classes which will allow a smaller time gap between the exposition of the concepts and results and their application by the computational implementation of the related methods and techniques. The evaluation component in which students solve the problems in group and then make an exposition to the class permit that all students learn and discuss the different specificities of each problem.

3.3.9. Bibliografia principal:

1. Davison, A.C., Hinkley, D.V., *Bootstrap Methods and their Application*, Cambridge University Press, 1997.
2. Gamerman, D., Lopes, H.F., *Stochastic Simulation for Bayesian Inference*, Chapman & Hall/CRC, 2006.
3. Gentle, J.E., *Random Number Generation and Monte Carlo Methods*, Springer-Verlag, 1998
4. McCullagh, P., Nelder, J.A., *Generalized Linear Models*, London: Chapman and Hall, 1983.
5. Rizzo, M., *Statistical computing with R*. CRC Press, 2007.
6. Robert, C.P., Casella, G., *Introducing Monte Carlo Methods with R*, Springer, 2010.
7. Ross, S.M., *Simulation*, 3rd Edition, Academic Press, 2002.

Mapa IV - Recuperação de Informação / Information Retrieval

3.3.1. Unidade curricular:

Recuperação de Informação / Information Retrieval

3.3.2. Docente responsável (preencher o nome completo) e respetivas horas de contacto na unidade curricular:

Nuno Manuel Robalo Correia (apenas Responsável, não tem horas de contacto)

3.3.3. Outros docentes e respetivas horas de contacto na unidade curricular:

João Miguel da Costa Magalhães (Regente) – T:28h; PL:28h

3.3.4. Objetivos de aprendizagem (conhecimentos, aptidões e competências a desenvolver pelos estudantes):

Conhecimentos

- *Compreender o conceito de relevância de informação*
- *Analisar dados Web e multimédia*
- *Compreender modelos de ordenação de informação*
- *Compreender protocolos de avaliação em recuperação de informação*

Aptidões

- *Implementar modelos de recuperação de informação*
- *Implementar módulos de um sistema de recuperação de informação*
- *Colocar em produção um sistema de recuperação de informação em larga-escala*
- *Desenhar um protocolo de avaliação e avaliar um motor de pesquisa*

Competências

- *Desenhar um sistema de recuperação de informação*
- *Selecionar as técnicas de RI adequadas a cada problema*
- *Capacidade de análise crítica dos resultados de avaliação*

3.3.4. Intended learning outcomes (knowledge, skills and competences to be developed by the students):**Knowledge**

- *Learn the concept of information relevance.*
- *Analyze Web and multimedia data.*
- *Learn how to rank information by relevance.*
- *Understand evaluation protocols.*

Know-how

- *Implement information retrieval models.*
- *Ability to adapt and improve components of a search engine.*
- *Deploy search engines with large-scale datasets.*
- *Design evaluation protocols and evaluate search engines.*

Soft-Skills

- *Select the right IR techniques for particular problems.*
- *Design information retrieval systems.*
- *Ability to do critical thinking about retrieval results.*

3.3.5. Conteúdos programáticos:

1. *Introdução*
2. *Pesquisa Web*
3. *Modelos básicos de recuperação de informação*
4. *Ordenação com base em hiperligações*
5. *Indexação*
6. *Recuperação probabilística de informação*
7. *Modelos de linguagem*
8. *Expansão de queries e feedback de utilizador*
9. *Protocolos de avaliação*
10. *Aprender a ordenar*
11. *Recuperação distribuída de informação*
12. *Indexação de dados de alta dimensão*

3.3.5. Syllabus:

1. *Introduction*
2. *Web Search*
3. *Basic retrieval models*
4. *Link-based ranking*
5. *Indexing*
6. *Probabilistic retrieval*
7. *Language models*
8. *User feedback and query expansion*
9. *Evaluation*
10. *Learning-to-rank*
11. *Distributed retrieval*
12. *High-dimensional data indexing*

3.3.6. Demonstração da coerência dos conteúdos programáticos com os objetivos de aprendizagem da unidade curricular:

A primeira parte do curso discute conceitos base e introduz técnicas e algoritmos elementares de Recuperação de Informação: o modelo de espaço vetorial, relevância, PageRank, indexação, modelos de linguagem e ordenação por aprendizagem.

A parte final do programa, aplica os conceitos base de RI à área de recuperação de informação multimédia e examina os requisitos específicos deste tipo de informação.

3.3.6. Evidence of the syllabus coherence with the curricular unit's intended learning outcomes:

The first part of the course discusses key concepts and introduces students to well-established techniques and algorithms in Information Retrieval: the vector space model, relevance, PageRank, indexing, language-models and learning to rank.

The final part of the course takes the initial concepts from classic IR and applies them to multimedia information and examines the special requirements of this data.

3.3.7. Metodologias de ensino (avaliação incluída):

Nas aulas teóricas é apresentada a matéria, com exemplos e discussão cuidada dos conceitos mais importantes. As aulas laboratoriais destinam-se à realização dos trabalhos de laboratório e o projeto final.

As aulas laboratoriais destinam-se à realização de 4 trabalhos laboratoriais, com duração de duas aulas cada, e do um projeto final que durará 5 semanas.

Será disponibilizada uma página Web da disciplina onde se mantém informação atualizada sobre o funcionamento da mesma. Os slides da matéria teórica e os guias de laboratório estarão disponíveis na página Web da disciplina.

A avaliação da disciplina é composta por 2 testes escritos individuais realizados a meio e no fim do semestre e trabalhos de grupo de laboratório: 4 mini-trabalhos quinzenais e um projeto final.

A nota mínima em cada um dos testes e trabalhos de laboratório é de 8 valores. A componente teórica vale 50%, cada mini-trabalho vale 5% e o trabalho final 30%.

3.3.7. Teaching methodologies (including assessment):

The course content is presented in the lectures with practical examples followed by a detailed discussion of the most important concepts. In the laboratory classes support will be given to students to the work assignments and final project.

A web page will be made available with updated information about the course. The teaching materials and laboratory guides will also be available on this web page.

The course grading is composed of two written tests and four laboratory assignments and a final project. The minimum grade of each componente is 8.0, the exam weights 50%, each work assignment is 5% and the final project is 30%.

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

O objetivo do curso é o de permitir os alunos compreenderem todos os aspetos de desenho e implementação de um motor de pesquisa. Os alunos irão dominar as principais técnicas de Recuperação de Informação: representação de informação, indexação, querying, e ranking por relevância. Este primeiro objetivo é tratado na primeira parte do programa e fortalecido na segunda parte onde conceitos fundamentais são estendidos a informação multimédia.

Os exercícios semanais permitem aos alunos dominar profundamente os conceitos apresentados e a estrutura do programa mantém os alunos atualizados com os últimos avanços na área e com os problemas atualmente tratados pela comunidade.

3.3.8. Evidence of the teaching methodologies coherence with the curricular unit's intended learning outcomes:

The goal of this course is to provide students with an understanding of all aspects of the design and implementation of search engines. Students will master fundamental concepts of Information Retrieval: information representation, indexing, querying, and ranking by relevance. This first objective is tackled by the first part of the course syllabus where fundamental concepts are introduced. The second part of the syllabus further strengthen these concepts by extending them to multimedia information.

The weekly assignments allow students understand deep concepts and syllabus structure keep students up-to-date with late-breaking research allowing them to pursue an actual research topic.

3.3.9. Bibliografia principal:

- C. D. Manning, P. Raghavan and H. Schütze, "Introduction to Information Retrieval", Cambridge University Press, 2008. <http://www-nlp.stanford.edu/IR-book/>

- S. Büttcher, C. L. A. Clarke, G. V. Cormack, "Information Retrieval: Implementing and Evaluating Search Engines", The MIT Press, 2010. <http://www.ir.uwaterloo.ca/book/>

Mapa IV - Estatística Multivariada / Multivariate Statistics

3.3.1. Unidade curricular:

Estatística Multivariada / Multivariate Statistics

3.3.2. Docente responsável (preencher o nome completo) e respetivas horas de contacto na unidade curricular:

Carlos Manuel Agra Coelho - TP 28h

3.3.3. Outros docentes e respetivas horas de contacto na unidade curricular:

Filipe José Gonçalves Pereira Marques - TP 14h

Miguel dos Santos Fonseca - TP 14h

3.3.4. Objetivos de aprendizagem (conhecimentos, aptidões e competências a desenvolver pelos estudantes):

No final desta unidade curricular o estudante terá adquirido conhecimentos, aptidões e competências que lhe permitam:

- *Conhecer e saber demonstrar resultados fundamentais*
- *Ser capaz de efetuar uma análise descritiva de um conjunto de dados multivariados*
- *Conhecer e utilizar as propriedades das principais distribuições em Estatística Multivariada*
- *Ser capaz de utilizar diferentes técnicas e métodos estatísticos na análise de dados multivariados*
- *Ser capaz de abordar desafios resultantes da análise de grandes volumes de dados*

3.3.4. Intended learning outcomes (knowledge, skills and competences to be developed by the students):

At the end of this course the student will have acquired knowledge and skills that will enable him to:

- *know and prove key results*
- *be able to perform a descriptive analysis of a set of multivariate data*
- *know and use the properties of the main distributions in Multivariate Statistics*
- *be able to use different techniques and statistical methods in the analysis of multivariate data.*
- *be able to address challenges arising from the analysis of large volumes of data*

3.3.5. Conteúdos programáticos:

- 1 - *Breves revisões e noções básicas de Álgebra Linear. Estatística descritiva multivariada.*
- 2 - *A distribuição Normal Multivariada. Estimadores de máxima verosimilhança e suas distribuições. A distribuição Wishart*
- 3 - *Inferência sobre vectores de médias*
 - 3.1 - *Testes com base numa amostra*
 - 3.2 - *Testes com base em duas amostras, amostras emparelhadas e amostras independentes*
 - 3.3 - *Testes com base em várias amostras*
- 4 - *Regressão Multivariada e Análise Canónica*
- 5 - *Análise de Componentes Principais e Análise Factorial*
- 6 - *Análise Discriminante e Classificação*
- 7 - *Análise de Clusters*

3.3.5. Syllabus:

- 1 – *Brief review of basic notions in Linear Algebra. Multivariate descriptive statistics*
- 2 – *The Multivariate Normal distribution. Maximum likelihood estimators and their distributions. The Wishart distribution*
- 3 – *Inference on mean vectors*
 - 3.1 – *One sample tests*
 - 3.2 – *Two sample tests – paired and independent samples*
 - 3.3 – *Tests based on more than two samples*
- 4 – *Multivariate Regression and Canonical Analysis*
- 5 – *Principal Component Analysis and Factorial Analysis*
- 6 – *Discriminant Analysis and Classification*
- 7 – *Cluster analysis*

3.3.6. Demonstração da coerência dos conteúdos programáticos com os objetivos de aprendizagem da unidade curricular:

Nos capítulos de 1 a 7 são explorados os seguintes temas: análise descritiva, distribuições multivariada Normal e Wishart, inferência sobre vetores de médias, regressão multivariada, análise canónica, análise de componentes principais, análise factorial, análise discriminante e análise de clusters. Estes temas abordam as principais técnicas e métodos estatísticos utilizados na análise de dados multivariados.

Nos materiais de apoio estão disponíveis slides de apoio a cada capítulo, exercícios com e sem resolução e ficheiros com dados que servem de base para a resolução dos mais diversos problemas. A bibliografia fornecida aos alunos serve de complemento aos materiais de apoio disponibilizados.

Durante as aulas serão dadas todas as indicações necessárias para a utilização correta do software R essencial para a resolução dos problemas propostos.

Muitos dos exercícios são baseados em problemas da vida real e em situações que envolvam grandes volumes de dados.

Os objetivos enunciados são assim abrangidos.

3.3.6. Evidence of the syllabus coherence with the curricular unit's intended learning outcomes:

In chapters 1 to 7 the following topics are explored: descriptive analysis, the multivariate Normal and Wishart distributions, inference about mean vectors, multivariate regression, canonical analysis, principal component analysis, factor analysis, discriminant analysis and cluster analysis. These topics cover the key techniques and statistical methods used in the analysis of multivariate data.

In the supporting materials are available: slides supporting each chapter, exercises with and without resolution and

files with data that are the basis for the resolution of the most diverse problems. The bibliography provided to the students complements the supporting materials available.

During the classes will be given all the necessary information for correct use of the R software essential to the resolution of the problems proposed.

Many of the exercises are based on real life problems and in situations involving large volumes of data.

The objectives are thus well covered.

3.3.7. Metodologias de ensino (avaliação incluída):

*As aulas são teóricas/práticas participadas, com exposição oral dos conceitos devidamente complementada com exemplos e resolução de problemas. Os alunos necessitam de assistir a um mínimo de 2/3 das aulas teórico/práticas lecionadas para obter frequência. A avaliação contínua é baseada em 3 testes, T1, T2 e T3 sendo a nota final (NF) atribuída através da seguinte fórmula $NF=0.25*T1+0.5*T2+0.25*T3$. Os alunos obtêm aprovação se NF for superior ou igual a 9.5. Se o aluno não obtiver aprovação através de avaliação contínua poderá ainda obtê-la no exame de recurso se tiver nota superior ou igual a 9.5.*

3.3.7. Teaching methodologies (including assessment):

*Classes are theoretical/practical with oral presentation of the concepts together with examples and problem solving. The students need to attend a minimum of two thirds of the classes to obtain frequency. The continuous evaluation is based on three tests, T1, T2 and T3 being the final note (NF) given by the following formula $NF=0.25*T1+0.5*T2+0.25*T3$. The students are approved if NF is greater than or equal to 9.5. If the students do not obtain approval through continuous evaluation they can still get approval in the resource examination with a score greater than or equal to 9.5.*

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

As componentes teóricas e práticas necessárias para atingir os objetivos de aprendizagem são ministradas nas aulas que já contemplam estas duas vertentes. Os alunos podem ainda recorrer ao horário de dúvidas para colmatar dúvidas que persistam. A componente teórica será aprofundada através da exposição oral dos conceitos apoiada em slides devidamente elaborados para o efeito e através da análise e discussão de diferentes exemplos. A vertente prática será desenvolvida através da resolução de problemas e da discussão dos mesmos. A avaliação contínua é assegurada com a realização de três provas escritas.

3.3.8. Evidence of the teaching methodologies coherence with the curricular unit's intended learning outcomes:

The theoretical and practical components necessary to achieve the learning objectives are administered in classes that already include these two components. Students can also use the doubts schedule to clarify persistent doubts. The theoretical component shall be further developed through the oral presentation of the concepts supported by slides, properly designed for that purpose, and through analysis and discussion of various examples. The practical component will be developed by solving problems and by discussing them. The continuous evaluation is provided with three written tests.

3.3.9. Bibliografia principal:

Livros base/ Base references:

Johnson, R. and Wichern, D. W. (2007), Applied Multivariate Statistical Analysis, 6th Edition, Prentice Hall, New Jersey

Morrison, D. F. (2004), Multivariate Statistical Methods, 4th Edition, Duxbury Press

Livros Complementares/ Complementary references:

Wilkinson, D. J (2014), Statistics for Big data, Lecture notes

(<https://www.staff.ncl.ac.uk/d.j.wilkinson/teaching/mas8381/notes14.pdf>)

Koch, I. (2013), Analysis of Multivariate and High-Dimensional Data, Cambridge Series in Statistical and Probabilistic

Mathematics, Cambridge University Press, New York

Flury, B. (1997), A First Course in Multivariate Statistics, Springer. New York

Rencher, A. C. (1998), Multivariate Statistical Inference and Applications, John Wiley & Sons

Rencher, A. C. (2002), Methods of Multivariate Analysis, John Wiley & Sons

Bibliografia Avançada/ Advanced references:

Anderson, T. W. (2003), An Introduction to Multivariate Statistical Analysis, 3rd ed., J. Wiley & Sons, New York

Mapa IV - Aprendizagem Automática / Machine Learning

3.3.1. Unidade curricular:

Aprendizagem Automática / Machine Learning

3.3.2. Docente responsável (preencher o nome completo) e respetivas horas de contacto na unidade curricular:

Pedro Manuel Corrêa Calvente Barahona (apenas Responsável não tem horas de contacto)

3.3.3. Outros docentes e respetivas horas de contacto na unidade curricular:

Ludwig Krippahl - T:28h

Susana Maria dos Santos Nascimento Martins de Almeida - PL:28h

3.3.4. Objetivos de aprendizagem (conhecimentos, aptidões e competências a desenvolver pelos estudantes):

No final desta unidade curricular o estudante terá adquirido conhecimentos, aptidões e competências que lhe permitam:

Compreender:

-Os paradigmas e desafios da área de Aprendizagem Automática: Aprendizagem Supervisionada e Aprendizagem Não-Supervisionada.

- Métodos fundamentais e suas aplicações na descoberta de conhecimento orientada aos dados.

- Métodos de selecção de modelos e atributos

- Vantagens e limitações dos métodos de Aprendizagem Automática estudados.

Ser capaz de:

- Aplicar, adaptar e implementar algoritmos de Aprendizagem Automática.

- Modelar experimentalmente dados reais usando esses algoritmos.

- Interpretar e avaliar resultados experimentais.

- Validar e comparar algoritmos de Aprendizagem Automática.

Conhecer:

- Os conceitos e a linguagem necessários ao aprofundamento autónomo dos conhecimentos de Aprendizagem Automática.

- Tipos de problemas passíveis de resolução por Aprendizagem Automática e os métodos mais adequados a cada caso.

3.3.4. Intended learning outcomes (knowledge, skills and competences to be developed by the students):

At the end of this unit, students will:

Understand:

- Paradigms and challenges of Machine Learning: supervised and unsupervised learning.

- Fundamental methods and their application in data oriented knowledge discovery

- Model and feature selection methods

- Advantages and limitations of Machine Learning methods.

Be able to:

- Apply, adapt and implement Machine Learning algorithms.

- Model real data experimentally using these algorithms.

- Interpret and evaluate experimental results in Machine Learning

- Validate and compare Machine Learning algorithms

Know:

- Concepts and language necessary for autonomous progression in Machine Learning.

- Problem types amenable to Machine Learning solutions and how to choose the best methods for each case.

3.3.5. Conteúdos programáticos:

1. Introdução à Aprendizagem Automática.

Paradigmas de Aprendizagem Automática: Aprendizagem Supervisionada, Aprendizagem Não-Supervisionada e Aprendizagem por Reforço.

2. Dados

2.1 Tópicos de normalização e visualização de dados

2.2 Redução de dimensionalidade e selecção de atributos

3. Aprendizagem Supervisionada

3.1 Regressão

3.2 Regressão logística

3.3 Redes Neurais

3.4 Máquinas de Suporte Vectorial

3.5 Modelos geradores

3.6 Aprendizagem Preguiçosa.

3.7 Validação, validação cruzada, selecção de modelos e comparação de métodos de classificação

3.8 Ensembles

4. Aprendizagem Não-Supervisionada

4.1 Métodos de Agrupamento por partição

4.2 Métodos de Agrupamento Probabilístico

4.3 Métodos de Agrupamento Difuso

4.4 Métodos de Agrupamento Hierárquico

4.5 Cadeias de Markov

4.5 Avaliação de métodos e de resultados de agrupamento

5 Outros tópicos de Aprendizagem Automática

3.3.5. Syllabus:

1. Introduction to Machine Learning*Machine Learning paradigms: Supervised Learning, Unsupervised Learning and Reinforcement Learning.***2. Data***2.1 Data normalization and visualization**2.2 Dimensionality reduction and feature selection***3. Supervised Learning***3.1 Regression**3.2 Logistic Regression**3.3 Artificial Neural Networks**3.4 Support Vector Machines**3.5 Generative models**3.6 Lazy learning**3.7 Validation, cross-validation, model selection and comparison**3.8 Ensemble methods***4. Unsupervised Learning***4.1 Partitional clustering**4.2 Probabilistic clustering**4.3 Partitional Fuzzy clustering**4.4 Hierarchical clustering**4.5 Markov chains**4.6 Clustering evaluation methods***5 Other machine learning topics, such as deep learning and semi-supervised learning****3.3.6. Demonstração da coerência dos conteúdos programáticos com os objetivos de aprendizagem da unidade curricular:**

O programa da disciplina, tanto na componente de Aprendizagem Supervisionada como na componente de Aprendizagem Não-Supervisionada, cobre todos os tópicos base e uma parte substancial dos tópicos considerados opcionais na sub-área 'Advanced Machine Learning' [elective] do Computer Science Curriculum 2013 da ACM (<http://cs2013.org/>).

As várias metodologias estudadas são enquadradas com exemplos de aplicação.

Os trabalhos práticos e exercícios tutoriais permitem consolidar conhecimentos na implementação/adaptação dos algoritmos estudados a problemas reais. Será dada particular atenção à interpretação de resultados experimentais, avaliação dos métodos usados e comparação de modelos.

3.3.6. Evidence of the syllabus coherence with the curricular unit's intended learning outcomes:

Both the Supervised and Unsupervised Learning components of the syllabus cover all the core subjects and most of the elective subjects for Machine Learning in the ACM Computer Science Curricula 2013 (<http://cs2013.org/>).

The methods studied are applied to practical examples. Tutorial classes and practical assignments promote the consolidation of theoretical knowledge as well as practice in the implementation and adaptation of the algorithms to real life problems. Practical assignments also focus on the interpretation of experimental results and the evaluation of the methods used, including the comparison of different models.

3.3.7. Metodologias de ensino (avaliação incluída):

Nas aulas teóricas são lecionados os tópicos fundamentais da matéria, a complementar com a leitura da bibliografia aconselhada. Nas aulas teóricas será dedicado tempo para exposição da matéria e para dúvidas e discussão com os alunos.

As aulas laboratoriais destinam-se à orientação tutorial e realização dos trabalhos práticos sobre tópicos selecionados da matéria.

Será disponibilizada uma página Web da disciplina com informação atualizada sobre o seu funcionamento, incluindo os acetatos da matéria teórica.

Avaliação

A avaliação da disciplina tem uma componente teórica-prática (T) e uma laboratorial ou de projeto (P), cada uma com peso de 50% na nota final.

Para obter aprovação à disciplina é necessário: nota mínima de 10 valores em cada uma das componentes de avaliação, sendo a nota final a sua média ponderada.

A componente teórica-prática (T) consiste em dois testes, com igual peso para a nota desta componente, podendo ser substituída por um exame de recurso.

3.3.7. Teaching methodologies (including assessment):

Lectures will cover the fundamental topics of the subject matter, which the students should complement with the given bibliography. All lecture materials will be supplied for further study. Lectures will include some time questions and discussion of the subject matter.

Tutorial classes will be dedicated to exercises and guidance in the practical assignments, focusing on selected topics. Class schedules and materials will be supplied online.

Assessment:

The evaluation of this curricular unit is made by two components: theoretical/problems (T) and project (P). Each component contributes with 50% to the final grade.

To pass, the student must have: a grade of at least 10 points in each of the theoretical/problems and project components. The final grade is defined as the weighted average of the two components of evaluation.

The theoretical part consists of two written individual tests; alternatively, this component can be evaluated by a written exam.

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

As várias metodologias estudadas são enquadradas com exemplos de aplicação.

Os trabalhos práticos e exercícios tutoriais permitem consolidar conhecimentos na implementação/adaptação dos algoritmos estudados a problemas reais. Será dada particular atenção à interpretação de resultados experimentais, avaliação dos métodos usados e comparação de modelos.

As noções gerais e conceitos apresentados ao longo do curso serão avaliados nos dois testes ou no exame final.

3.3.8. Evidence of the teaching methodologies coherence with the curricular unit's intended learning outcomes:

The different techniques covered in this unit will be framed within practical problems to illustrate their application.

The tutorials and assignments will help consolidate knowledge and skills in the application and adaptation of the algorithms to real problems. Especial attention will be given to the interpretation of experimental results, evaluation and comparison of different models.

Understanding of the theoretical notions and concepts presented along this unit will be assessed in the written tests or final exam.

3.3.9. Bibliografia principal:

T. Mitchell. Machine Learning, McGraw-Hill, 1997.

C. M. Bishop. Pattern Recognition and Machine Learning, Springer, 2006.

E. Alpaydin. Introduction to Machine Learning, Second Edition, MIT Press, 2010.

T. Hastie, R. Tibshirani, J. Friedman. Elements of Statistical Learning, Second Edition, Springer, 2009.

Mapa IV - Sistemas para Processamento de Big Data / Systems for Big Data Processing

3.3.1. Unidade curricular:

Sistemas para Processamento de Big Data / Systems for Big Data Processing

3.3.2. Docente responsável (preencher o nome completo) e respetivas horas de contacto na unidade curricular:

Nuno Manuel Ribeiro Preguiça (Responsável e Regente) – T: 28h

3.3.3. Outros docentes e respetivas horas de contacto na unidade curricular:

Sérgio Marco Duarte – PL: 28h

3.3.4. Objetivos de aprendizagem (conhecimentos, aptidões e competências a desenvolver pelos estudantes):

Esta é uma UC obrigatória na área de sistemas para processamento de Big Data.

O UC visa apresentar uma perspetiva geral das soluções atuais para processamento de dados, focando-se no armazenamento e processamento distribuído dos dados. Na UC será dado particular ênfase aos modelos programação suportados e sua aplicação.

Conhecimento

- *Conhecer as diferentes facetas do processamento em Big Data*
- *Conhecer as principais classes de sistemas para armazenamento de grandes volumes de dados*
- *Conhecer os modelos de programação genéricos dominantes*
- *Conhecer soluções para domínios específicos*

Aplicação

- *Ser capaz de identificar qual o tipo de sistema apropriado para a resolução dum problema concreto*
- *Ser capaz de codificar (a solução para) um problema concreto usando o modelo de programação mais adequado*
- *Ser capaz de executar uma aplicação numa plataforma distribuída*

3.3.4. Intended learning outcomes (knowledge, skills and competences to be developed by the students):

This is a mandatory course on big data processing systems. The course provides a broad view of the current solutions for large-scale data processing. It focuses on distributed storage and data processing, with special emphasis on the

supported programming models and their application.

Knowledge

- *Know the different facets of processing large volumes of data.*
- *Know the main classes of systems for storage of large volumes of data*
- *Know the dominant programming models for Big Data*
- *Know solutions for specific problem domains*

Application

- *Be capable of identifying the best system class for solving a specific problem.*
- *Be capable of coding a specific problem solution in the most suitable programming model*
- *Be capable of executing a big data application in a distributed platform.*

3.3.5. Conteúdos programáticos:

1. Visão Geral

a. Motivação, Aplicações

b. Desafios

2. Modelos de programação

a. Batch vs. Incremental vs tempo-real

b. Dados estruturados vs. dados não-estruturados

c. Programação declarativa vs. generalista

3. Armazenamento de Dados

a. Sistemas de ficheiros distribuídos (e.g. HDFS)

b. Base de dados relacionais.

c. Bases de dados NoSQL (e.g. keyvalue stores, arquivos de documentos)

d. Integração de múltiplas fontes de dados (e.g. Hive)

4. Plataformas de processamento genéricas

a. Infraestrutura: contexto, propriedades e implicações

b. Modelo mapreduce and plataforma de suporte (e.g. Hadoop)

c. Plataformas de segunda geração (e.g. Pig, Spark)

5. Processamento em domínios específicos

a. Bibliotecas para aprendizagem automatic (e.g. Spark MLlib)

b. Plataformas para processamento em grafos (e.g. GraphX)

6. Introdução às plataformas para processamento em tempo-real

a. Fontes de dados (e.g. Flume, Kafka)

b. Modelos de dados: micro-batches vs. fluxos contínuos

c. Plataformas de processamento (e.g. Storm, Spark Streaming)

3.3.5. Syllabus:

1. Overview

a. Motivation, Applications

b. Challenges

2. Programming models

a. Batch vs. Incremental vs. Real-time

b. Structured data vs. Unstructured data

c. Declarative programming vs. General-purpose

3. Data storage

a. Distributed file systems (e.g. HDFS)

b. Relational databases

c. NoSQL databases (e.g. key-value stores, document stores)

d. Integration of multiple data sources (e.g. Hive)

4. Generic processing platforms

a. Infrastructure: context, properties and implications

b. Map-reduce model and supporting platform (e.g. Hadoop)

c. Second generation platforms (e.g. Pig, Spark)

5. Processing for specific domains

a. Machine learning libraries (e.g. Spark MLlib)

b. Platforms for graph processing (e.g. GraphX)

6. Introduction to real-time processing platforms

a. Data sources (e.g. Flume, Kafka)

b. Data models: micro-batch vs. continuous

c. Processing platforms (e.g. Storm, Spark Streaming)

3.3.6. Demonstração da coerência dos conteúdos programáticos com os objetivos de aprendizagem da unidade curricular:

Os conteúdos programáticos contribuem para alcançar os objetivos da unidade curricular como se explica de seguida.

O conhecimento das diferentes facetas do processamento em Big Data é abordado de forma geral no tópico 1 dos conteúdos programáticos, sendo as principais facetas abordadas de forma detalhada nos outros tópicos.

Relativamente ao conhecimento dos modelos de programação e das classes de sistemas de armazenamento, estes

são abordados nos tópicos 2 e 3 dos conteúdos programáticos.

Os tópicos 4, 5 e 6 contribuem para um conhecimento mais profundo das plataformas de processamento genéricas, para domínios específicos e para processamento em tempo real, incluindo conhecimentos do tipo de problemas que cada uma permite endereçar eficientemente, como se codificam e executam soluções nestas plataformas.

3.3.6. Evidence of the syllabus coherence with the curricular unit's intended learning outcomes:

The syllabus contribute to achieve the unit's objectives as explained next.

The different facets of Big Data processing are overviewed in syllabus topic 1, with the most important facets being detailed in the following topics.

Programming models and storage classes for Big Data are addressed in topics 2 and 3 of the syllabus.

Topics 4, 5 and 6 contribute to a deeper knowledge of generic processing platforms, as well as platforms for specific domains and processing of realtime data. These topics address the problems that can be efficiently handled by each type of platform, how to code and deploy solutions for such problems.

3.3.7. Metodologias de ensino (avaliação incluída):

Nas aulas teóricas expõe-se e discutem-se os tópicos do programa da UC, recorrendo a sistemas existentes para ilustrar as questões em destaque.

As aulas práticas fornecem aos alunos a experiência no desenvolvimento de soluções de processamento de dados em grande escala, usando uma seleção das plataformas existentes. São realizadas demonstrações, resolvidos exercícios e o acompanhamento dos dois pequenos trabalhos práticos a resolver durante o semestre.

A avaliação será baseada nos seguintes componentes: dois testes (25% cada); dois trabalhos de programação, resolvidos em grupo (25% cada).

3.3.7. Teaching methodologies (including assessment):

In the lectures, the topics that comprise the course syllabus are presented and discussed, using existing systems and platforms to highlight the issues and present concrete examples.

In labs, the students acquire experience on developing solutions for large-scale data processing problems, using a selection of current platforms and systems. Classes comprise demos, exercises and support for the two programming assignments.

Grading is based on the following components: two quizzes (25% each) and two team programming-assignments (25% each).

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

As metodologias de ensino têm como objetivo dar uma visão abrangente e crítica dos problemas e das soluções para processamento de dados em grande escala. Tal é conseguido através da exposição dos conceitos e técnicas fundamentais nas aulas teóricas. As aulas práticas visam a consolidação e a colocação em prática do conhecimento através de exemplos, exercícios e a realização dois trabalhos de programação.

3.3.8. Evidence of the teaching methodologies coherence with the curricular unit's intended learning outcomes:

The teaching methodology intend to provide students with broad and critical view of the problems and solutions for large-scale data processing. That is achieved by using lectures to present the key concepts, tools and techniques in data processing. Labs provide consolidation the acquired knowledge and its application using concrete examples, exercises and two programming assignments.

3.3.9. Bibliografia principal:

*Big Data Fundamentals: Concepts, Drivers, and Techniques
ISBN: 0134291077
Prentice Hall, 2016*

Mapa IV - Empreendedorismo / Entrepreneurship

3.3.1. Unidade curricular:

Empreendedorismo / Entrepreneurship

3.3.2. Docente responsável (preencher o nome completo) e respetivas horas de contacto na unidade curricular:

Virgílio António Cruz Machado - TP:28h

3.3.3. Outros docentes e respetivas horas de contacto na unidade curricular:

N/A

3.3.4. Objetivos de aprendizagem (conhecimentos, aptidões e competências a desenvolver pelos estudantes):

O curso pretende motivar os alunos para o empreendedorismo e para a necessidade da inovação tecnológica. O programa cobre vários tópicos que são importantes para a adoção de uma cultura aberta aos riscos suscitados em processos de criação de novos produtos ou atividades que exigem características empreendedoras.

No final desta unidade curricular, os alunos deverão ter desenvolvido um espírito empreendedor, uma atitude de trabalho em equipa e estar aptos a:

- 1) *Identificar ideias e oportunidades para empreenderem novos projetos;*
- 2) *Conhecer os aspetos técnicos e organizacionais inerentes ao lançamento dos projetos empreendedores;*
- 3) *Compreender os desafios de implementação dos projetos (ex: mercado, financiamento, gestão da equipa) e encontrar os meios para os ultrapassar;*
- 4) *Expôr a sua ideia e convencer os stakeholders.*

3.3.4. Intended learning outcomes (knowledge, skills and competences to be developed by the students):

This course is intended to motivate students for entrepreneurship and the need for technological innovation. It covers a list of topics and tools that are important for new venture creation as well as for the development of creative initiatives within existing enterprises. Students are expected to develop an entrepreneurship culture, including the following skills:

- 1) *To identify ideas and opportunities to launch new projects;*
- 2) *To get knowledge on how to deal with technical and organizational issues required to launch entrepreneurial projects;*
- 3) *To understand the project implementation challenges, namely venture capital and teamwork management, and find the right tools to implement it;*
- 4) *To show and explain ideas and to convince stakeholders.*

3.3.5. Conteúdos programáticos:

O empreendedorismo como estratégia de desenvolvimento pessoal e organizacional. Processos de criação de ideias. A proteção da propriedade intelectual: patentes e formalismos técnicos. A gestão de um projeto de empreendedorismo: planeamento; comunicação e motivação; liderança e gestão de equipas Marketing e inovação para o desenvolvimento de novos produtos e negócios. O plano de negócios e o estudo técnico-financeiro. Financiamento e Sistemas de Incentivos: formalidades e formalismos. A gestão do crescimento e o intra-empreendedorismo.

3.3.5. Syllabus:

Strategy for entrepreneurship. Ideation and processes for the creation of new ideas. Industrial property rights and protection: patents and technical formalities. Managing an entrepreneurial project: planning; communication and motivation; leadership and team work. Marketing and innovation for the development of new products and businesses. Business plan and entrepreneurial finance. System of Incentives for young entrepreneurs. Managing growth and intrapreneurship.

3.3.6. Demonstração da coerência dos conteúdos programáticos com os objetivos de aprendizagem da unidade curricular:

O conteúdo programático foi desenhado para incentivar o aluno ao empreendedorismo e à perceção e análise da envolvente em busca de oportunidades de negócio, de forma a que consiga aplicar os conhecimentos adquiridos:

- 1) *na transformação de conhecimento científico em ideias de negócio;*
- 2) *na criação, seleção e desenvolvimento de uma ideia para um novo produto ou serviço;*
- 3) *na elaboração de um plano de negócio e de um plano de marketing;*
- 4) *na exposição das suas ideias em curto tempo e em ambientes stressantes.*

3.3.6. Evidence of the syllabus coherence with the curricular unit's intended learning outcomes:

The syllabus was designed to encourage the student for entrepreneurship and for the perception and analysis of new business opportunities; with this program, the student may apply the knowledge provided:

- 1) *to transform scientific knowledge in business ideas;*
- 2) *to create, select and develop an idea for a new product or service;*
- 3) *to draw a business plan and a marketing plan;*
- 4) *to better explain and present its ideas in a short time and stressed environments.*

3.3.7. Metodologias de ensino (avaliação incluída):

Este curso será ministrado a alunos dos 4º ou 5º anos dos programas de Mestrado integrado e de 2º ciclo. O programa é dimensionado para decorrer entre o 1º e o 2º semestre, num período de 5 semanas, envolvendo um total de 45 horas presenciais (TP), organizadas em 15 sessões de 3 horas e exigindo um esforço global de 3 ECTS.

As aulas presenciais baseiam-se na exposição dos conteúdos do programa. Os estudantes serão solicitados a aplicar as competências adquiridas através da criação e desenvolvimento de uma ideia (produto ou negócio). As aulas integrarão alunos provenientes de diversos cursos com vista a promover a integração de conhecimento derivado de várias áreas científicas e envolverão professores e "mentores" com background diverso em engenharia, ciência, gestão e negócios.

A avaliação compreende a apresentação e defesa da ideia num elevator pitch e do respetivo relatório (realizado em grupo de 4-5 elementos). A apresentação contribuirá com 60% e o relatório com 40% para a nota final.

3.3.7. Teaching methodologies (including assessment):

This course is directed to students from the 4th and 5th years of the "Mestrado Integrado" (Integrated Master) and students from the 2nd cycle (Master). The program was designed for a duration of 5 weeks, with a total of 45 hours in class (15 sessions of 3 hours each) - 3 ECTS.

Classes are based in an exposition methodology. Students will be asked to apply their skills in the creation and development of an idea, regarding a new product or a new business. Classes integrate students from different study programs to promote the integration of knowledge derived from various scientific areas and involve academic staff and "mentors" with diverse background in engineering, science, management and business.

Students evaluation is based on the development and presentation of an idea/project in an elevator pitch, and its report. The work should be developed in teams of 4-5 members. The presentation should account for 60% of the final mark and the report 40%.

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

Considerando o tempo disponível (5 semanas), a metodologia de ensino preconiza que em cada semana sejam discutidos e trabalhados (em grupo) os temas apresentados, os quais tinham sido definidos nos objetivos de aprendizagem.

Na 1ª semana os temas a abordar estão relacionados com os aspetos estratégicos do empreendedorismo, a geração de ideias, a liderança e a gestão de equipas; como resultado os alunos deverão constituir e organizar as suas equipas para poderem definir o problema que se pretende resolver. Na 2ª semana, os temas apresentados permitirão que o aluno possa evoluir no seu projeto acrescentando opções de soluções ao problema identificado na semana interior e proceder à seleção de uma delas. Na 3ª semana, a abordagem ao mercado e às condições de comercialização viabilizarão a concretização do plano de marketing.

Na 4ª semana, abordar-se-ão os aspetos relacionados com a viabilidade financeira do projeto, possibilitando a realização do respetivo plano de negócio e do seu financiamento. Na última semana, abordar-se-á o processo de exposição da ideia aos potenciais interessados, tendo os alunos que realizar a apresentação e defesa do seu projeto num elevator pitch, perante um júri.

Neste sentido, a metodologia privilegia

- 1) a apresentação de casos práticos e de sucesso;*
- 2) a promoção de competências nos domínios comportamentais, nomeadamente, no que respeita ao desenvolvimento do sentido crítico, à defesa de ideias e argumentos baseados em dados técnico-científicos, à tolerância e capacidade de gestão de conflitos em situações adversas e stressantes.*
- 3) a participação dos alunos nos trabalhos colocados ao longo da unidade curricular e a sua apresentação.*

3.3.8. Evidence of the teaching methodologies coherence with the curricular unit's intended learning outcomes:

Considering the available time (5 weeks), the teaching methodology praises that, in each week, the subjects presented and defined in the learning objectives are discussed and worked (in groups). In the first week, the subjects introduced to students are related with entrepreneurial strategic issues, generation of ideas, leadership and work team management; as a result, the students will have to organize their teams to be able to define the problem. In the 2nd week, the subjects presented will allow the student to pursue its project; they have to consider different options for the problem identified in the previous week. In the 3rd week, the market related issues are approached, and the students are asked to build a marketing plan. In the 4th week, financial issues are addressed, making it possible to accomplish a business plan. In the last week, the process of how to expose the idea to potential stakeholders is addressed; the students are required to present and argue their project in an elevator pitch.

This methodology gives priority to:

- 1) the presentation of practical and successful cases;*
- 2) the promotion of soft skills, namely, in what concerns to the development of critical thinking, the defense of ideas and arguments based on technical-scientific data, to the tolerance and capacity of dealing with conflicts in adverse and stressful situations.*
- 3) the participation of the students in practical works and assessments and their presentation.*

3.3.9. Bibliografia principal:

Books

Burns, P. (2010). Entrepreneurship and Small Business: Start-up, Growth and Maturity, Palgrave Macmillan, 3rd Ed.

Kotler, P. (2011). Marketing Management, Prentice-Hall

Shriberg, A. & Shriberg (2010). Practicing Leadership: Principles and Applications, John Wiley & Sons, 4th Ed.

Spinelli, S. & Rob Adams (2012). New Venture Creation: Entrepreneurship for the 21st Century. McGraw-Hill, 9th Ed.

Byers, Thomas H., Dorf R. C., Nelson, A. (2010). Technology Ventures: From Idea to Enterprise, 3rd Ed., McGraw-Hill

Hirsh, R. D. (2009). International Entrepreneurship: Starting, Developing, and Managing a Global Venture, Sage Publications, Inc

Hirsh, R.D., Peters, M. P., Shepherd, D.A. Entrepreneurship, 7th Ed., McGraw-Hill, 2007

Journals

Entrepreneurship Theory and Practice

Mapa IV - Seminário / Seminar

3.3.1. Unidade curricular:

Seminário / Seminar

3.3.2. Docente responsável (preencher o nome completo) e respetivas horas de contacto na unidade curricular:

Pedro Manuel Corrêa Calvente Barahona /Jorge Orestes Lasbarrères Cerdeira - S:28h

3.3.3. Outros docentes e respetivas horas de contacto na unidade curricular:

*Convidados de Instituições Parceiras do Mestrado
Guests from Partner Institutions of the Master*

3.3.4. Objetivos de aprendizagem (conhecimentos, aptidões e competências a desenvolver pelos estudantes):

Saber: Informação e grandes volumes de dados existente em empresas e outras instituições.

Necessidades da sua transformação em conhecimento.

Saber Fazer: Explorar as oportunidades de extrair conhecimento dos dados existentes. Análise das várias tecnologias e ferramentas usadas para esse efeito.

Competências Complementares: Capacidades de apresentação oral e escrita. Interação com decisores para uma colaboração eficaz.

3.3.4. Intended learning outcomes (knowledge, skills and competences to be developed by the students):

Knowledge: Information and large volumes of data in existing companies and other institutions.

Needs of its transformation into knowledge.

Application: Explore opportunities to extract knowledge from existing data. Analysis of different technologies and tools used for this purpose.

Soft-skills: Oral and writing presentation skills. Interaction with decision-makers for effective collaboration.

3.3.5. Conteúdos programáticos:

1. Casos de Estudo apresentados por empresas e outras instituições.

2. Tipos de dados mantidos por empresas e sua dimensão.

3. Conhecimento implícito nos dados existentes.

4. Decisões que podem ser tomadas com o conhecimento extraído dos dados.

3.3.5. Syllabus:

1. Case studies submitted by companies and other institutions.

2. Types of data kept by companies and their size.

3. Implicit knowledge of existing data.

4. Decisions that may be improved from the knowledge extracted from the data.

3.3.6. Demonstração da coerência dos conteúdos programáticos com os objetivos de aprendizagem da unidade curricular:

Nesta unidade curricular expõem-se os estudantes a vários casos de estudo, em que as instituições (empresas, institutos públicos, centros de investigação, etc) obtêm e/ou administram grandes volumes de dados e pretendem retirar conhecimento e informação útil desses dados.

Os alunos deverão analisar estes casos de estudo, apresentados por responsáveis dessas instituições, de forma a analisar as possibilidades de extração desse conhecimento, bem como da sua utilização para o melhoramento das decisões a serem tomadas.

3.3.6. Evidence of the syllabus coherence with the curricular unit's intended learning outcomes:

In this course students are exposed to various case studies, in which several institutions (companies, public institutions, research centers, etc.) obtain and / or manage large volumes of data and want to take advantage from the knowledge and useful information extracted from these data.

Students should examine these case studies, presented by representatives of these institutions, in order to analyse the possibilities of extraction and use of this knowledge to improve the decisions to be taken.

3.3.7. Metodologias de ensino (avaliação incluída):

O ensino consiste na exposição de casos de estudo pelos responsáveis das instituições parceiras do curso, e na sua discussão com os alunos. Pretende-se que as apresentações identifiquem a informação mantida nessas instituições, nomeadamente em aspetos como o seu tipo, origem, volume, sistemas usados para a sua manutenção, incluindo a confidencialidade e segurança, e como ela é ou poderá ser utilizada para melhorar as decisões tomadas na atividade dessas instituições.

A avaliação é feita através de relatórios que os alunos elaboram a partir da informação recolhida nas apresentações e na sua análise posterior, e onde deverão identificar as oportunidades de utilização das metodologias e técnicas ensinadas no curso que mais adequadamente permitam abordar os problemas identificados e que conduzam a uma melhoria na atividade dessas instituições. A nota final é obtida pela avaliação dos relatórios pelos responsáveis da UC, com o apoio das responsáveis das instituições que fornecem os casos de estudo.

3.3.7. Teaching methodologies (including assessment):

Teaching consists of exposure of case studies by partner institutions in charge of the course, and its discussion with students. It is intended that the presentations identify several aspects of the information held in these institutions, such as its type, origin, volume, systems used for its maintenance, including confidentiality and security, and how the data can be used to improve decisions taken in the activity of these institutions.

The evaluation is done through reports that students prepare from information collected in the presentations and further analysis, and where they should identify opportunities for the use of the methodologies and techniques taught in the programme to adequately address the identified problems and lead to an improvement in the activity of these

institutions.

The final grade is obtained by evaluation of the reports by the management of UC, with the support of the representatives of the institutions that provide case studies.

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

Esta UC pretende expor os alunos a casos reais em que a análise de grandes volumes de dados pode trazer grandes vantagens na atividade de instituições de vários tipos, tais como empresas, institutos estatais ou centros de investigação.

A metodologia adoptada, baseada em apresentações por parte de responsáveis das instituições parceiras seguida do estudo dos casos conducente à adequada utilização da informação existente, garante que os alunos são de facto expostos a casos reais com toda a sua complexidade, e testa a sua capacidade para análise desses casos e apresentação de soluções / sugestões de uma utilização adequada da informação existente.

3.3.8. Evidence of the teaching methodologies coherence with the curricular unit's intended learning outcomes:

This course aims to expose students to real cases where analysis of large volumes of data can bring great advantages in the activity of various institutions, such as companies, government institutes and research centers.

The methodology, based on presentations by leaders of the partner institutions, followed by further analysis of the cases, ensures that students are in fact exposed to real cases with all their complexity, and tests their ability of analysis and presentation of solutions / suggestions for appropriate use of existing information.

3.3.9. Bibliografia principal:

Esta UC não tem materiais de estudo específicos, devendo ser utilizados os materiais das UCs nas quais se baseiem as metodologias e técnicas adequadas ao casos em estudo.

Adicionalmente pode ser obtida informação de casos de estudo em várias livros da área, por exemplo:

This course has no specific study material and should use materials from other courses of the programme addressing the methodologies and techniques adequate to the case studie.

Further information can be obtained from case studies in several books of the area, for example:

• *Bart Baesens , Analytics in a Big Data World: The Essential Guide to Data Science and its Applications, Wiley, ISBN: 978-1-118-89270-1, 2014*

• *Data Science and Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data EMC Education Services (Editor), Wiley, ISBN: 978-1-118-87613-8, 2015*

Mapa IV - Métodos Bayesianos / Bayesian Methods

3.3.1. Unidade curricular:

Métodos Bayesianos / Bayesian Methods

3.3.2. Docente responsável (preencher o nome completo) e respetivas horas de contacto na unidade curricular:

Isabel Cristina Maciel Natário (Responsável e Regente) – TP:28h

3.3.3. Outros docentes e respetivas horas de contacto na unidade curricular:

Luís Pedro Carneiro Ramos – TP:14h

Filipe José Gonçalves Pereira Marques – TP:14h

3.3.4. Objetivos de aprendizagem (conhecimentos, aptidões e competências a desenvolver pelos estudantes):

O enorme aumento das capacidades computacionais e a perícia dos modelos bayesianos na análise de estruturas de dados muito complexas, como as que surgem em áreas como a Inteligência Artificial, justificam a pertinência desta UC. O seu objetivo principal é introduzir a abordagem da estatística bayesiana. No final desta UC o estudante deve perceber os princípios que regem a inferência bayesiana, saber incorporar, em problemas variados, o conhecimento existente a priori e a sua correspondente incerteza numa distribuição de probabilidade, saber atualizar a distribuição a priori com os dados de forma a estimar analítica ou numericamente (através de métodos de programação intensiva como o MCMC, Markov Chain Monte Carlo) as distribuições de probabilidade a posteriori e preditiva resultantes, saber usar a modelação hierárquica para representar e analisar sistemas complexos, fazer seleção de modelos e previsão, usando os softwares R-project e OpenBUGS (executado pelo R).

3.3.4. Intended learning outcomes (knowledge, skills and competences to be developed by the students):

The huge increase of computational capabilities and the expertise of Bayesian models on the analysis of very complex data structures, as the ones that come up in areas as Artificial Intelligence, justify the relevance of this CU. Its main objective is to introduce the Bayesian statistics approach. By the end of this CU, a student must understand the principles that rules Bayesian inference, know how to incorporate, in various problems, the existing prior knowledge and its corresponding uncertainty in a probability distribution, know how to update the prior distribution with data to estimate analytically or numerically (through intensive programming methods as MCMC, Markov Chain Monte Carlo) the resulting posterior probability and predictive distributions, know how to use hierarchical modelling to represent

and analyze complex systems, how to make model selection and prediction, using the software R-project and OpenBUGS (run in R).

3.3.5. Conteúdos programáticos:

- 1 - O paradigma Bayesiano
- 2 - A distribuição a priori e métodos para a sua formulação
- 3 - A função verosimilhança, a distribuição a posteriori, as distribuições marginal e preditiva
- 4 - Inferência bayesiana
- 5 - Markov Chain Monte Carlo, MCMC
- 6 - Qualidade e seleção de modelos
- 7 - Modelos hierárquicos

3.3.5. Syllabus:

- 1 - The Bayesian paradigm
- 2 - The prior distribution and methods for its formulation
- 3 - The likelihood function, the posterior distribution, the marginal and predictive distributions
- 4 - Bayesian inference
- 5 - Markov Chain Monte Carlo, MCMC
- 6 - Model evaluation and selection
- 7 - Hierarchical models

3.3.6. Demonstração da coerência dos conteúdos programáticos com os objetivos de aprendizagem da unidade curricular:

Esta UC começa por introduzir no ponto 1 do conteúdo programático o paradigma bayesiano, contrastando-o com a abordagem clássica, explicando a sua capacidade de incluir nas análises outra informação além dos dados, apresentando a sua abordagem coerente de fazer inferência e selecionar modelos, cobrindo assim o objetivo de explicitar os princípios que regem a inferência bayesiana.

O ponto 2 é dedicado à quantificação da incerteza da informação existente antes da recolha dos dados na formulação da distribuição a priori, que aborda o objetivo seguinte proposto.

Os pontos 3, 4 e 5 estudam os ingredientes necessários ao objetivo de estimação das distribuições a posteriori, preditiva e marginal, quer de forma analítica, se possível, quer de forma numérica, recorrendo aos métodos de programação intensiva de MCMC (Markov Chain Monte Carlo), que permitem aplicar de forma eficiente os métodos bayesianos em modelos complexos, permitindo subseqüentes inferências estatísticas.

O ponto 6 refere-se ao objetivo da avaliação da qualidade e seleção de modelos através do estudo de metodologias adequadas, e o ponto 7 detalha os modelos hierárquicos, mais adequados à modelação de dados com estruturas complexas.

O uso dos softwares R-project e OpenBUGS (que corre a partir do R), acompanhará as aplicações das metodologias bayesianas aprendidas, de forma mais relevante quando for necessário recorrer aos métodos MCMC, implementados pelo OpenBUGS.

3.3.6. Evidence of the syllabus coherence with the curricular unit's intended learning outcomes:

This CU begins by introducing, in point 1 of the syllabus, the Bayesian paradigm, comparing it with the classical approach, explaining its capacity to include in the analyses information other than data, presenting its coherent approach to do inference and selecting models, thus covering the objective of explaining the principles that rule Bayesian inference.

Point 2 is dedicated to the quantification of uncertainty of the existing information before data collection, in the formulation of the prior distribution, addressing the following proposed objective.

Points 3, 4 and 5 study the necessary ingredients for the estimation objective of the posterior, predictive and marginal distributions, either in an analytic way, if possible, or using intensive programming methods of MCMC (Markov Chain Monte Carlo), which permits to apply in an efficient way the Bayesian methods in complex models, allowing subsequent statistical inferences.

Point 6 refers to the objective of quality assessment and model selection through the study of adequate methodologies, and point 7 details hierarchical models, more adequate for modelling data with complex structures.

Software R-project and OpenBUGS (that runs from R) are used for running the applications of the Bayesian methodologies learnt, in a more relevant way if MCMC methods are needed, implemented in OpenBUGS.

3.3.7. Metodologias de ensino (avaliação incluída):

Aulas teórico-práticas onde a par com a exposição dos conceitos e resultados fundamentais serão apresentados exemplos ilustrativos desses conceitos e resultados. De seguida são apresentados problemas que são resolvidos em laboratório usando os softwares R e OpenBUGS. Estes problemas são realizados em laboratório, pretendendo-se que os alunos participem ativamente na sua resolução. Deste modo os alunos adquirem as competências relativas não só à adequada implementação das metodologias ensinadas em cada situação concreta como também relativas aos softwares estatísticos.

A avaliação consiste na realização de dois trabalhos com relatório e/ou um teste final escrito. Os trabalhos que os alunos realizam para avaliação podem ser apresentados e discutidos na sala de aula.

3.3.7. Teaching methodologies (including assessment):

Lecture-lab classes are the adequate way to convey the course contents to students as, together with the explanation of the main concepts and results, illustrative examples are given. Afterwards, some problems are proposed to students to be solved in a lab using both software R and OpenBUGS, being students supposed to take active part in their resolution. Consequently, students acquire the basic expertise not only of the adequate implementation of the methodologies learned in each concrete situation but also of the application of the statistic software.

The evaluation consists on the resolution of two practical problems with a written report and/or a written final test. The practical problems solved for the evaluation may be presented and discussed by the students in classroom.

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

A forte componente prática desta UC que faz uso intensivo de computação justifica que as aulas sejam teórico-práticas o que permite uma maior proximidade temporal entre a exposição dos resultados e conceitos e a sua aplicação através da implementação computacional dos correspondentes métodos e técnicas. A componente de avaliação em que os alunos resolvem problemas em grupo e depois apresentam na aula permite que todos os alunos possam aprender e discutir as diferentes especificidades de cada problema.

3.3.8. Evidence of the teaching methodologies coherence with the curricular unit's intended learning outcomes:

The strong practical component of this course that makes intensive use of computation justifies the lecture-lab classes which will allow a smaller time gap between the exposition of the concepts and results and their application by the computational implementation of the related methods and techniques. The evaluation component in which students solve the problems in group and then make an exposition to the class permit that all students learn and discuss the different specificities of each problem.

3.3.9. Bibliografia principal:

1. Albert, J. (2009). *Bayesian Computation with R*. Springer.
2. Bernardo J.M. & Smith, A.F.M. (1994). *Bayesian theory*. Wiley.
3. Congdon P (2001). *Bayesian Statistical Modelling*. Wiley.
4. Cowles, M.K. (2013). *Applied Bayesian Statistics. With R and OpenBUGS Examples*. Springer.
5. Gamerman, D. & Lopes, H.F. (2006). *Markov chain Monte Carlo - stochastic simulation for Bayesian inference*. Chapman & Hall/CRC.
6. Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B. (2003). *Bayesian Data Analysis (2nd edition)*. Chapman and Hall / CRC, 2003.
7. Gilks, W.R., Richardson, S. and Spiegelhalter, D. (Edts.) (1996) *Markov chain Monte Carlo in Practice*. Chapman and Hall/CRC.
8. Lee, P.M. (2004). *Bayesian Statistics: An Introduction, 3rd edition*, Arnold.

Mapa IV - Decisão e Risco / Decision and Risk

3.3.1. Unidade curricular:

Decisão e Risco / Decision and Risk

3.3.2. Docente responsável (preencher o nome completo) e respetivas horas de contacto na unidade curricular:

Maria Isabel Azevedo Rodrigues Gomes - TP: 20h

3.3.3. Outros docentes e respetivas horas de contacto na unidade curricular:

Marta Faias - TP: 20h

Manuel Leote Tavares Inglês Esquível - TP:16h

3.3.4. Objetivos de aprendizagem (conhecimentos, aptidões e competências a desenvolver pelos estudantes):

Ao obter aprovação o aluno:

- *terá tomado contacto com a problemática da tomada de decisão com um único ou múltiplos decisores,*
- *conhecerá os principais modelos de apoio à decisão, com ou sem negociação, e de risco,*
- *compreenderá qual o tipo de dados que deverão ser recolhidos e tratados de modo a alimentarem os modelos anteriores,*
- *saberá identificar e construir os modelos adequados aos problemas que poderá enfrentar num contexto profissional futuro.*

3.3.4. Intended learning outcomes (knowledge, skills and competences to be developed by the students):

When approved the student will:

- *have had contact with the problematic of decision-making with a single or multiple decision makers,*
- *know the most important models of decision support, with or without negotiation, and risk,*
- *understand what kind of data should be collected and processed to support previous models,*

- be able to identify and build the adequate models for problems that he or she may face his/her future professional context.

3.3.5. Conteúdos programáticos:

I. Teoria de decisão

- O que é uma decisão?*
- Estruturação e análise do problema de decisão*
- Decisões sequenciais*
- Decisões com vários objetivos*

II. Teoria de jogos

- Modelação da interação estratégica*
- Jogos de estáticos com informação completa*
- Jogos dinâmicos com informação completa e incompleta*
- Jogos cooperativos*

III. Risco

- Introdução*
- Modelo fundamental*
- Principais medidas e fatores de risco*
- Outros modelos de risco*

3.3.5. Syllabus:

I. Decision theory

- what is a decision?*
- structure and analyses of decision-making problems*
- sequential decisions*
- decisions with several objectives*

II. Game theory

- modelling strategic interaction*
- static games of complete information*
- dynamic games of complete and incomplete information*
- cooperative game theory*

III. Risk

- definition*
- fundamental model of risk*
- factors and measures of risk*
- other models for risk*

3.3.6. Demonstração da coerência dos conteúdos programáticos com os objetivos de aprendizagem da unidade curricular:

A unidade encontra-se dividida em três módulos: teoria da decisão, teoria de jogos e risco. No primeiro módulo o aluno é introduzido no contexto da tomada de decisão, sendo-lhe apresentados diferentes modelos clássicos para problemas em contexto de risco e incerteza. Segue-se o módulo de teoria de jogos onde se estende a noção de preferência do decisor e se introduzem modelos de negociação e interação estratégica entre os intervenientes no processo de tomada de decisão. Por fim, serão estudados diferentes conceitos e modelos de risco, estendendo as noções apresentadas no módulo 1.

3.3.6. Evidence of the syllabus coherence with the curricular unit's intended learning outcomes:

The unit is divided into three modules: decision theory, game theory and risk. In the first module the student is introduced to the context of decision-making and learns about different classic models in presence of uncertainty and risk. Follows the game theory module which extends the notion of decision-maker's preferences and introduces negotiation models and strategic interaction among decision-making players. Finally, different concepts and models for risk will be studied, extending the concepts introduced in module 1.

3.3.7. Metodologias de ensino (avaliação incluída):

Apresentação em aula dos problemas, conceitos e resultados - método expositivo – sendo parte do tempo dedicado à discussão dos tópicos com os alunos e esclarecimento de dúvidas. Aulas de carácter prático onde o aluno, supervisionado pelo docente, resolverá exercícios que permitirão a aplicação dos tópicos abordados anteriormente. Alguns tópicos serão desenvolvidos em laboratório com recurso software adequado.

A avaliação será composta por três trabalhos com relatório e/ou um teste final escrito. Os trabalhos serão desenvolvidos em grupo e serão apresentados e discutidos em sala de aula.

3.3.7. Teaching methodologies (including assessment):

Classes are theoretical/practical with oral presentation of concepts, methodologies, and examples, complemented with problem solving (including software). Specific student difficulties will be addressed during classes or in individual sessions scheduled with the professor.

The evaluation consists on three group assignments. The written reports will be presented and discussed by the students in class room.

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

As várias metodologias estudadas são enquadradas com exemplos de aplicação.

Os trabalhos práticos e exercícios tutoriais, sob a supervisão do docente, permitem consolidar conhecimentos na implementação/adaptação dos modelos estudados a problemas com dados reais. Será dada particular atenção à interpretação de resultados computacionais dos modelos, avaliação dos métodos usados e comparação de modelos.

3.3.8. Evidence of the teaching methodologies coherence with the curricular unit's intended learning outcomes:

In classes, concepts and methodologies are presented and illustrated with examples, motivating students for the relevance of the corresponding study, and developing student's ability for identifying problems and selecting the most adequate methodology for the corresponding resolution.

Practical classes allow students to test these skills, under the supervision of a professor. Particular attention is given to the interpretation of computer model results, evaluation of the methods used and comparison of models.

3.3.9. Bibliografia principal:

Fudenberg F. and Tirole, J. (1995) Game Theory, The MIT Press, 1995.

Kelly A. (2003) Decision Making Using Game Theory: An Introduction for Managers, Cambridge University Press.

Minelli M., Chambers M. and Dhiraj A. (2013) Big Data, Big Analytics: Emerging Business Intelligence and Analytic Trends for Today's Businesses, Wiley, 1st edition.

Peterson M. (2009) An Introduction to Decision Theory, Cambridge University Press.

Siegel E., Davenport T. H. (2013) Predictive Analytics: The Power to Predict Who Will Click, Buy, Lie, or Die, Wiley, 1st edition.

Vose D. (2008) Risk Analysis: A Quantitative Guide. John Wiley & Sons.

Mapa IV - Visualização Interativa de Dados / Interactive Data Visualization**3.3.1. Unidade curricular:**

Visualização Interativa de Dados / Interactive Data Visualization

3.3.2. Docente responsável (preencher o nome completo) e respetivas horas de contacto na unidade curricular:

Nuno Manuel Robalo Correia (apenas Responsável, não tem horas de contacto)

3.3.3. Outros docentes e respetivas horas de contacto na unidade curricular:

João Carlos Gomes Moura Pires (Regente) – T: 18h ; PL: 18h

Fernando Pedro Reino da Silva Birra – T: 10h ; PL: 10h

3.3.4. Objetivos de aprendizagem (conhecimentos, aptidões e competências a desenvolver pelos estudantes):

Conhecimento

-O que é e quais os diferentes objetivos da VD.

-O papel da interação na VD e os padrões gerais de interação.

-O conceito de Variável Visual e as consequências práticas no desenho de Visualização interativa de Dados (VID).

-A classificação de tipos de dados para fins de visualização de dado e o seu impacto no desenho de sistemas VID.

- Para cada tipo de dados quais as técnicas de VD mais apropriadas.

-Quais as principais metodologias para comparação e avaliação de técnicas e sistemas de VD.

- Tendências atuais em VDI e o seu papel em sistemas mais gerais.

Aplicação

-Escolher as variáveis visuais e as técnicas de visualização mais apropriadas a um conjunto de dados e de objetivos.

-Usar um sistema VDI concreto para explorar e visualizar um ou mais conjuntos de dados.

-Desenhar e implementar um sistema VDI para uma classe de dados para um conjunto de objetivos de exploração e visualização.

-Desenhar e preparar uma bancada experimental para avaliar uma técnica VD.

3.3.4. Intended learning outcomes (knowledge, skills and competences to be developed by the students):

Knowledge

-What is Information Visualization, Data Visualization (DV) and the different purposes of DV.

-The role of interactivity in DV and the general interaction patterns.

-The concept of Visual Variable and the practical consequence in the design of Interactive Data Visualization (IDV).

-The classification of data for DV purposes and the impact on IDV.

-For each type of data the most relevant available techniques.

-The available approaches to Compare and Evaluate Visualization Techniques and Systems.

-The actual trends in IDV and their role in more general systems and applications.

Application

-Choose the visual variables and visualization techniques for a given data set and purposes.

-Use a given an existing IDV system to explore one or more data sets.

-Based on existing frameworks and platforms, design and build an IDV system appropriate for a class of data sets and a class of exploration and visualization tasks.

-Setup an experimental environment to evaluate a DV technique.

3.3.5. Conteúdos programáticos:

Introdução à Visualização de dados
O que é a visualização?
Relação entre visualização e outras áreas.
O processo de visualização.
Fundamentos de dados.
Perceção humana e processamento de informação.
Semiologia dos símbolos gráficos.
As variáveis visuais.
Técnicas de Visualização
Técnicas de visualização para dados geo-espaciais
Técnicas de visualização para dados temporais
Técnicas de visualização para dados multivariados
Técnicas de visualização para árvores, grafos e redes
Visualização de texto e de documentos
Conceitos e Técnicas de Interação
Operadores de interação, operandos e espaços
Espaço estruturado de visualização
Transformações de animação
Controle de interação
Desenho eficaz de visualização
Comparação e avaliação de técnicas de visualização
Sistemas de Visualização
Os sistemas baseados no tipo de dados
Sistemas baseados na análise de tipos
Análise de texto e visualização
Sistemas modernos e integrados de visualização
Toolkits
Direções de investigação em Visualização

3.3.5. Syllabus:

Introduction to Data Visualization
What Is Visualization?
Relationship between Visualization and Other Fields.
The Visualization Process.
Data Foundations.
Human Perception and Information Processing.
Semiology of Graphical Symbols.
The Visual Variables
Visualization Techniques
Visualization Techniques for Geospatial Data
Visualization Techniques for Time-Oriented Data
Visualization Techniques for Multivariate Data
Visualization Techniques for Trees, Graphs, and Networks
Text and Document Visualization
Interaction Concepts and Techniques
Interaction Operators, Operands and Spaces
Visualization Structure Space
Animating Transformations
Interaction Control
Designing Effective Visualizations
Comparing and Evaluating Visualization Techniques
Visualization Systems
Systems Based on Data Type
Systems Based on Analysis Type
Text Analysis and Visualization
Modern Integrated Visualization Systems
Toolkits
Research Directions in Visualization

3.3.6. Demonstração da coerência dos conteúdos programáticos com os objetivos de aprendizagem da unidade curricular:

O programa está desde logo bastante alinhado com os objetivos de “conhecimento” definidos. A primeira parte, “Introduction to Data Visualization” assegura que uma grande parte desses objetivos de “conhecimento” assim como as duas secções do programa: “Interaction Concepts and Techniques” e “Visualization Systems”.

A parte dos objetivos “Aplicação” são assegurados pela articulação das aulas teóricas em geral e em particular as relativas secção “Visualization Techniques”, e das aulas práticas onde os alunos vão tendo contato e aplicando sistemas concretos a problemas de visualização. Estes objetivos são ainda consolidados com o projeto realizado pelos alunos.

3.3.6. Evidence of the syllabus coherence with the curricular unit’s intended learning outcomes:

The syllabus is directed related and align with the knowledge objectives. The first part, “Introduction to Data Visualization” is a large contributor for those objectives. The “Interaction Concepts and Techniques” e “Visualization Systems” are also most aligned with the knowledge objectives.

The Application objectives are achieved with the contributions of many lectures, in particular the “Visualization Techniques”, and with the lab sessions where students apply concrete visualization systems to real data. The Application objectives are also reinforced with the project that has to be developed by the students.

3.3.7. Metodologias de ensino (avaliação incluída):

Nas aulas teóricas é apresentada a matéria, com exemplos e referências a sistemas relacionados. As aulas laboratoriais destinam-se à realização do projeto que acompanha a matéria apresentada nas aulas teóricas.

A avaliação da disciplina inclui cinco elementos: dois testes escritos individuais efetuados ao longo do semestre e 1 projeto, com várias fases (especificação, artigo com técnica(s) relacionada(s) e código/interface), que corresponde a um projeto de uma aplicação de interativa de visualização de dados.

3.3.7. Teaching methodologies (including assessment):

In the lectures the course content is presented, illustrated with application examples and references to related work. The laboratory classes are intended for specification, development and presentation of the project that deals with topics presented during the lectures.

The evaluation of the course consists of five elements: two mid-term written individual tests and one project, with several phases (specification, state of the art and code/interface) that together account for a project to develop a interactive data visualization application, using real data.

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

Nas aulas teóricas são expostos os princípios fundamentais da unidade curricular, acompanhados de exemplos e de referências para trabalho relacionado. Nas aulas práticas os alunos desenvolvem um projeto de dimensão média, que acompanha os tópicos da componente teórica.

Para além da bibliografia recomendada, os alunos têm acesso aos slides apresentados nas aulas teóricas e a artigos científicos, de forma a enquadrar os tópicos e o trabalho no contexto mais genérico dos atuais sistemas interativos de visualização.

Durante as aulas práticas, os alunos têm que trabalhar em grupos (preferencialmente de dois elementos) e aplicar os conhecimentos proporcionados pelas aulas teóricas. O projeto tem uma dimensão apreciável e vários momentos de verificação, de forma a ser semelhante a um trabalho de equipa real, no desenvolvimento de um sistema interativo de visualização de dados, com recurso a dados reais.

Os testes escritos individuais permitem a avaliação mais precisa dos conhecimentos adquiridos por cada aluno.

3.3.8. Evidence of the teaching methodologies coherence with the curricular unit’s intended learning outcomes:

The fundamental principles are presented during the lectures, illustrated with examples and references to related work. In the laboratory classes the students develop a project that implements the topics of the interactive data visualization systems.

In addition to the recommended readings, students have access to the slides presented in the lectures and related scientific, to contextualize the topics and the project in the context of current interactive data visualization systems.

During the laboratory classes, students have to work in groups (preferably with two elements) and apply the knowledge obtained from the lectures and other education materials. The project is relatively large with several deliverables to replicate as much as possible the situation of a real team developing an interactive data visualization system, using real data.

The individual written tests allow a more accurate assessment of knowledge acquired by each student.

3.3.9. Bibliografia principal:

- *Interactive Data Visualization: Foundations, Techniques, and Applications, Second Edition. Matthew O. Ward, Georges Grinstein, Daniel Keim, 2015, ISBN 9781482257373*
- *Visualization Analysis & Design, Tamara Munzner, 2015, ISBN: 9781466508910*

Mapa IV - Análise de Grandes Grafos / Large Graph Analytics

3.3.1. Unidade curricular:

Análise de Grandes Grafos / Large Graph Analytics

3.3.2. Docente responsável (preencher o nome completo) e respetivas horas de contacto na unidade curricular:

Jorge Orestes Lasbarrères Cerdeira - TP 40h

3.3.3. Outros docentes e respetivas horas de contacto na unidade curricular:

Maria Isabel Azevedo Rodrigues Gomes – TP 16h

3.3.4. Objetivos de aprendizagem (conhecimentos, aptidões e competências a desenvolver pelos estudantes):

No final desta unidade curricular o estudante terá adquirido conhecimentos, aptidões e competências que lhe permitam:

Identificar estruturas de conectividade, centralidade e comunidades em grafos grandes;

Prever processos que podem ocorrer em grafos grandes; e

Prever o crescimento de grandes grafos reais.

3.3.4. Intended learning outcomes (knowledge, skills and competences to be developed by the students):

At the end of this course the student will have acquired knowledge and skills that will enable him to:

Identify structures concerning connectivity, centrality and community within large graphs;

Predict processes that may occur over large graphs; and

Predict how real large graphs may grow.

3.3.5. Conteúdos programáticos:

1 Exemplos de grandes grafos reais;

2 Conceitos básicos de grafos;

3 Medidas topológicas (centralidade, comunidades, semelhança);

4 Análise da estrutura de grafos de grandes dimensões (componentes, caminhos mais curtos e efeito small-world, distribuição dos graus dos vértices, distribuição de medidas de centralidade);

5 Grafos aleatórios;

6 Processos em grandes grafo.

3.3.5. Syllabus:

1 Examples of large real graphs;

2 Basic graph concepts;

3 Topological measures (centrality, communities, similarity);

4 Large scale structure (components, shortest paths and small-world effect, vertices degree distribution, centrality measures distribution);

5 Random graphs;

6 Processes over large graphs.

3.3.6. Demonstração da coerência dos conteúdos programáticos com os objetivos de aprendizagem da unidade curricular:

O objetivo principal da UC é desenvolver competências que permitam estudar estruturas dos grandes grafos. As metodologias para atingir esse objetivo são apresentadas nos tópicos 3 e 4, após a introdução (no ponto 2) dos necessários conceitos básicos de grafos. No ponto 1 apresentam-se e destacam-se particularidades de grandes grafos que ocorrem em diversos contextos, e serve para motivar a matéria a tratar nesta UC. Os pontos 4 e 5 fornecem metodologias para prever a evolução de fenómenos em objectos representados por grandes grafos.

3.3.6. Evidence of the syllabus coherence with the curricular unit's intended learning outcomes:

The main goal of this course is to develop skills for studying structures of large graphs. The approaches to achieve this are covered by topics 3 and 4, after an introduction (point 2) to the necessary basics of graphs. In point 1, as a motivation for the course, large graphs that occur in a number of contexts are presented and particularities are pointed. Points 4 and 5 provide methodologies to predict the evolution of phenomena over objects that are represented as large graphs.

3.3.7. Metodologias de ensino (avaliação incluída):

As aulas são teóricas/práticas participadas, com exposição e discussão dos conceitos e metodologias devidamente complementada com exemplos e resoluções de problemas (incluindo software). Eventuais dúvidas poderão ser esclarecidas no decurso das aulas ou em sessões individuais no período de atendimento docente aos alunos.

Os alunos necessitam de assistir a um mínimo de 2/3 das aulas para obter frequência à UC.

Avaliação contínua: Dois testes a realizar durante o período lectivo, e elaboração de um relatório escrito, sobre um tópico a definir, e apresentação oral do trabalho. Considera-se aprovado o aluno com frequência e soma das classificações obtidas nos três elementos de avaliação ≥ 10 . A classificação final será o valor dessa soma.

Avaliação por exame de recurso: Só os alunos com frequência podem realizar exame de recurso. Considera-se aprovado o aluno que obtenha frequência e nota de exame final ≥ 10 valores.

3.3.7. Teaching methodologies (including assessment):

Classes are theoretical/practical consisting of exposition and discussions of concepts and methodologies complemented with examples and problems proposed for solving.

The student will be excluded of the evaluation if his presences in classes is less than 2/3 of the total number of classes.

The student can be evaluated by two test assessments, and a written and oral presentation of a study on a topic to be

proposed. Student will be approved if the results obtained in the three evaluation elements sum up (rounded) at least 10. The grade will be the rounded sum.

The student can also be approved by a final exam if the exam's grade is at least 10. The grade will be the one attained in the exam (rounded).

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

Os tópicos necessários para atingir os objetivos estabelecidos para a unidade curricular são ministrados nas aulas teórico-práticas, com o apoio do docente em horários de atendimento aos alunos. A aquisição destes conhecimentos é avaliada nas provas escritas (testes/exames) e numa apresentação oral. As componentes práticas necessárias para atingir os objetivos de aprendizagem são desenvolvidas em todas as formas de horas de contacto: nas aulas teórico-práticas através da análise e discussão de problemas-tipo. A avaliação destas competências é também realizada nas provas escritas e na prova oral. A frequência pretende assegurar que os alunos acompanham a matéria ao longo do semestre.

3.3.8. Evidence of the teaching methodologies coherence with the curricular unit's intended learning outcomes:

The topics required to achieve the objectives established for the course are taught in problem-solving sessions, with the support of teachers in in pre-established schedules reserved for the students' attendance. The acquisition of knowledge is assessed in written tests (tests / exams) and in an oral presentation. The practical components necessary to achieve the learning objectives are developed in all forms of contact: in classes through analysis and discussion-type problems. The assessment of these skills is also performed in the written tests and oral tests. The requirement for the presence of at 2/3 of classes is meant to ensure that students follow the matter throughout the semester.

3.3.9. Bibliografia principal:

Networks: An Introduction, Mark Newman, Oxford University Press, 2010

Networks, Crowds, and Markets: Reasoning about a highly connected World, David Easley and Jon Kleinberg, Cambridge University Press, 2010

Graph Analysis and Visualization: Discovering Business Opportunity in Linked Data, Richard Brath and David Jonker, Wiley, 2015

Big Data Analytics: From Strategic Planning to Enterprise Integration with Tools, Techniques, NoSQL, and Graph, David Loshin, Imprint: Morgan KaufmannPrint, Elsevier, 2013

Mapa IV - Processamento de Streams / Stream Processing

3.3.1. Unidade curricular:

Processamento de Streams / Stream Processing

3.3.2. Docente responsável (preencher o nome completo) e respetivas horas de contacto na unidade curricular:

José Júlio Alves Alferes - T :10;h PL:10h

3.3.3. Outros docentes e respetivas horas de contacto na unidade curricular:

Sérgio Marco Duarte - T 9h PL 9h

João Carlos Gomes Moura Pires - T 9h PL 9h

3.3.4. Objetivos de aprendizagem (conhecimentos, aptidões e competências a desenvolver pelos estudantes):

No final desta unidade curricular o estudante terá adquirido conhecimentos, aptidões e competências que lhe permitam:

*Entender os fundamentos e conhecer as principais linguagens e sistemas para processamento de streams de dados;
Construir aplicações que processem streams de dados em tempo real.*

3.3.4. Intended learning outcomes (knowledge, skills and competences to be developed by the students):

At the end of this course the student will have acquired knowledge and skills that will enable him to:

*Understand the fundamental and get acquainted with the main languages and systems for processing data streams;
Build application that process streams of data in real time.*

3.3.5. Conteúdos programáticos:

1. Sistemas distribuídos para processamento de streams

Modelos para processamento de streams: streams como sequências de mini-batches (e.g. Spark streaming);

processamento contínuo (e.g. Storm). Modelos gerais de programação. Problema de computações cíclicas. Aspectos de sistemas: distribuição, escalabilidade e tolerância a falhas.

2. Data Stream Management Systems (DSMS)

Modelos estruturas de dados para streams. Operações algébricas sobre streams e relações (continuous queries, agregados, janelas temporais). Linguagens de continuous queries. Extensões ao SQL e sistemas de bases de dados para lidar com streams.

3. Processamento de Eventos

Streams como sequências de eventos. Regras de produção, regras reativas e computação guiada por eventos. Redes de processamento de eventos, agentes e canais. Eventos complexos e derivados. Detecção de padrões de eventos. Linguagens e sistemas para processamento de eventos.

3.3.5. Syllabus:

1. Distributed Stream Processing Systems.

System models for stream processing: streams as sequences of mini-batches (e.g. Spark streaming); continuous processing (e.g. Storm). General-purpose programming models. The problem of cyclic computations. System aspects: distribution, scalability and fault-tolerance.

2. Data Stream Management Systems (DSMS).

Structured Data Models for Streams. Algebraic operators on stream and relations (continuous queries, aggregates and blocking, time windows). Continuous query languages. Languages and systems that extend SQL and database management systems to deal with data streams.

3. Complex Event Processing.

Streams as sequences of events. Production rules, reactive rules, and event-driven computing. Event processing networks, agents and channels. Complex and derived events. Detection of event patterns. Event-processing languages and systems.

3.3.6. Demonstração da coerência dos conteúdos programáticos com os objetivos de aprendizagem da unidade curricular:

Cada um dos grandes tópicos do programa fornece os fundamentos de uma abordagem a processamento de streams, e inclui exemplos de linguagens e sistemas que permitem aos estudantes consolidar os conceitos e construir pequenas aplicações. A combinação dos grandes tópicos permite uma visão sistémica que vai das componentes mais baixo nível, de conceção de sistemas computacionais, até à de mais alto nível, construídas sobre estes, com linguagens declarativas para processamento de dados de stream.

3.3.6. Evidence of the syllabus coherence with the curricular unit's intended learning outcomes:

Each of the syllabus main topics provides the fundamentals of an approach to process data streams, and includes examples of languages and system that allow the student to consolidate the concepts and build small applications. The combination of the main topics provides a systemic view ranging from the more low level components of computer systems, to high level components built on top of those, with declarative languages to process data streams.

3.3.7. Metodologias de ensino (avaliação incluída):

Nas aulas teóricas é feita a exposição e discussão dos conceitos e metodologias, sendo estas seguidas de aulas práticas laboratoriais com exemplos e resoluções de problemas e implementação de pequenas aplicações.

Ao longo do semestre os estudantes realizam dois testes escritos (ou, em alternativa, exame de recurso), sendo que a aprovação na UC está dependente de uma média superior a 10 valores na média desses dois testes.

Os estudantes têm ainda que elaborar um projeto de implementação de uma aplicação, completa e funcional, que processe streams e dados dimensão significativa. Exemplos de tais aplicações, que têm sido usados, são os challenges da principal conferência nesta área (DEBS).

A nota final é a média das notas do projetos e dos testes.

3.3.7. Teaching methodologies (including assessment):

Concepts and methodologies are exposed and discussed in lectures. In the labs, students are given examples, problems to solve, and also have to make the implementations of small implementations.

During the semester the students have two written midterms (or, alternatively, one final exam). Approval in the course is subject to approval in the midterms.

There is also a final project, which consists in the implementation of a functional application processing data streams with significant size. Examples of such application are the challenges of the main conference in this area (DEBS).

The final grade is the average of the midterms with the project.

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

Os tópicos necessários para atingir os objetivos estabelecidos para a unidade curricular são ministrados nas aulas teórico e práticas. A aquisição destes conhecimentos é avaliada nas provas escritas (testes/exames). A componente de projeto garante que no final os estudantes conseguem de facto construir aplicações nesta área.

3.3.8. Evidence of the teaching methodologies coherence with the curricular unit's intended learning outcomes:

The topics required to achieve the objectives established for the course are taught in the lectures and labs. The acquisition of knowledge is assessed in written midterms. The project component assures that students are indeed able to build applications in this area.

3.3.9. Bibliografia principal:

Opher Etzion and Peter Niblett. Event Processing in Action. Manning Publications, 2010.
Lukasz Golab and Tamer Özsu. Data Stream Management. Morgan and Claypool, 2010.
Several papers are provide for further reading.

Mapa IV - Algoritmos para Redes Complexas / Algorithms for Complex Networks**3.3.1. Unidade curricular:**

Algoritmos para Redes Complexas / Algorithms for Complex Networks

3.3.2. Docente responsável (preencher o nome completo) e respetivas horas de contacto na unidade curricular:

Luís Manuel da Costa Cairés (apenas Responsável, não tem horas de contacto)

3.3.3. Outros docentes e respetivas horas de contacto na unidade curricular:

Margarida Paula Neves Mamede (Regente) – T:28h; PL:28h

3.3.4. Objetivos de aprendizagem (conhecimentos, aptidões e competências a desenvolver pelos estudantes):

Saber: Definir e identificar as noções fundamentais que caracterizam a estrutura de uma rede. Conhecer alguns algoritmos eficientes usados na análise de redes complexas. Conhecer algumas ferramentas de análise e de visualização de redes. Compreender alguns modelos comportamentais de redes.

Saber Fazer: Analisar a estrutura de uma rede concreta do mundo real, calculando as métricas apropriadas, comunidades, etc., e extrair informação/insights dos dados. Simular um processo de propagação numa rede. Selecionar, adaptar e usar ferramentas apropriadas de análise e de visualização de redes.

Competências Complementares: Capacidade de investigação e autonomia.

3.3.4. Intended learning outcomes (knowledge, skills and competences to be developed by the students):

Knowledge: Define and identify the fundamental notions that characterize the structure of a network. Know some efficient algorithms for analysing complex networks. Know some network analysis tools and some visualization tools. Understand some network behaviour models.

Application: Analyse the structure of a real-world network, computing the appropriate metrics, communities, etc., and extract information/insights from the data. Simulate a spreading process on a network. Choose, adapt, and use suitable network analysis and visualization tools.

Soft-skills: Skills in research and autonomy.

3.3.5. Conteúdos programáticos:

1. Aspectos de redes complexas. Exemplos de redes de diferentes domínios.

2. Estudo Empírico de Redes: Redes tecnológicas. Redes sociais. Redes de informação. Redes biológicas.

3. Grafos, Métricas e Algoritmos: Caminhos, componentes, distribuição do grau, métricas de centralidade, formação de agrupamentos, equilíbrio estrutural, homofilia. Algoritmos e estruturas de dados para computar caminhos, métricas e partições de grafos.

4. A Estrutura de Redes Reais Complexas: Componentes. A propriedade do mundo pequeno. Leis de potência e a propriedade da independência de escala. Distribuições de medidas de centralidade. Coeficientes de agrupamento.

5. Métodos e Algoritmos para Detecção de Comunidades.

6. Dinâmica de Redes: Percolação e resiliência de redes. Fenómenos de propagação.

3.3.5. Syllabus:

1. Aspects of complex networks. Examples of networks from different domains.

2. Empirical Study of Networks: Technological networks. Social networks. Networks of information. Biological networks.

3. Graphs, Metrics and Algorithms: Paths, components, degree distribution, centrality metrics, clustering, structural balance, homophily. Algorithms and data structures for computing paths, metrics and graph partitions.

4. The Structure of Real-World Networks: Components. The small-world property. Power laws and the scale-free property. Distributions of centrality measures. Clustering coefficients.

5. Community Detection Methods and Algorithms.

6. Network Dynamics: Percolation and network resilience. Spreading phenomena.

3.3.6. Demonstração da coerência dos conteúdos programáticos com os objetivos de aprendizagem da unidade curricular:

Os objetivos da unidade curricular podem ser sintetizados da seguinte forma:

1. Conhecer os conceitos chave da estrutura de redes complexas para extrair informação/insights dos dados: Os tópicos 1, 2, 3 e 4 dos conteúdos programáticos cobrem esta matéria com muitos exemplos, que ilustram e permitem compreender as definições estudadas.

2. Analisar a estrutura de uma rede do mundo real, calculando as métricas apropriadas, comunidades, etc.: Os principais algoritmos eficientes para analisar a estrutura de uma rede são estudados nos tópicos 3 e 5; as ferramentas de análise e de visualização de redes são introduzidas e usadas ao longo do semestre nas aulas de laboratório.

3. Simular um processo de propagação numa rede: Isto é estudado no tópico 6.

3.3.6. Evidence of the syllabus coherence with the curricular unit's intended learning outcomes:

The unit objectives can be summarized in the following way:

- 1. Know the key concepts of the structure of complex networks in order to extract information/insights from the data: Syllabus topics 1, 2, 3 and 4 cover this material with many examples, which illustrate and allow all studied definitions to be understood.*
- 2. Analyse the structure of a real-world network, computing the appropriate metrics, communities, etc.: The main efficient algorithms for analysing the structure of a network are studied in topics 3 and 5; network analysis and visualization tools are introduced and used throughout the semester in the lab sessions.*
- 3. Simulate a spreading process on a network: This is studied in topic 6.*

3.3.7. Metodologias de ensino (avaliação incluída):

O ensino consiste na exposição da matéria em aulas teóricas e na resolução de problemas em aulas práticas de laboratório. No laboratório, os alunos analisam e visualizam redes, implementam algoritmos e efetuam simulações em redes.

A avaliação é composta por um trabalho de grupo e dois testes. Há dois tipos de trabalhos. Os alunos (com formação em Informática) podem desenvolver software específico (por exemplo, para deteção de comunidades) para redes de larga escala. Os alunos (com outras formações académicas) podem efetuar a análise dos dados de uma rede, recorrendo apenas a ferramentas existentes. Em caso de reprovação, se a nota do trabalho não for inferior a 9, os alunos podem realizar um exame final. Os testes e o exame são com consulta.

Condição para obter aprovação: $NotaP \geq 9$ e $NotaT \geq 9$ e $NotaF = (NotaP + NotaT) / 2 \geq 10$, onde: $NotaP$ é a nota do trabalho; $NotaT$ é a média das notas dos testes ou a nota do exame; $NotaF$ é a nota final.

3.3.7. Teaching methodologies (including assessment):

There are two hours of lectures and a lab session each week. In the laboratory, students analyse and visualize networks, implement algorithms, and perform simulations on networks.

Assessment comprises one group programming project and two tests. There are two types of programming projects. Students (with a computer science background) can develop software for performing some network analysis (for instance, community detection) on large-scale networks. Students (with other academic backgrounds) can perform the analysis of a network dataset, making use of existent data analysis tools. In case of failure, if the project grade is not less than 9 (out of 20), students can do a final exam. The tests and the exam are open-book.

Condition to succeed: $Pgrade \geq 9$ and $Tgrade \geq 9$ and $Fgrade = (Pgrade + Tgrade) / 2 \geq 10$, where: $Pgrade$ is the project grade; $Tgrade$ is the mean of the test grades or the exam grade; $Fgrade$ is the final grade.

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

Nas aulas práticas, os alunos analisam e visualizam redes concretas, utilizando os conceitos, os algoritmos e os métodos dados nas aulas teóricas.

Nas componentes de avaliação teóricas (testes e exame), os alunos são confrontados com questões e problemas que permitem avaliar a compreensão dos conceitos e dos algoritmos envolvidos na unidade curricular, a capacidade para conceber a análise de uma rede concreta e a aptidão para extrair informação de uma rede.

No trabalho, os alunos escolhem a rede que pretendem estudar e efetuam um estudo teórico e experimental da sua estrutura, utilizando técnicas e métodos aprendidos. O relatório apresenta e classifica a rede, identifica os vários modelos teóricos adotados e descreve a sua aplicação, inclui a análise e a implementação dos respetivos algoritmos ou as ferramentas utilizadas e contém uma análise dos resultados obtidos experimentalmente. Ao realizar o trabalho, os alunos desenvolvem a aptidão para selecionar e aplicar as técnicas e as ferramentas apropriadas, a aptidão para conceber e avaliar estudos e a capacidade de comunicação escrita. Com este trabalho, também se pretende incentivar a autonomia e desenvolver a capacidade de investigação dos alunos.

3.3.8. Evidence of the teaching methodologies coherence with the curricular unit's intended learning outcomes:

In lab sessions, students analyse and visualize concrete networks, making use of the concepts, algorithms, and methods presented in the lectures.

The questions and problems posed in tests and exams allow the assessment of (1) the understanding of the concepts and algorithms involved in the unit, (2) the ability to design the analysis of a concrete network, and (3) the ability to extract information from a network.

In the programming project, students choose the network they want to study and perform a theoretical and experimental study of its structure, applying the techniques and methods learnt. The report presents and classifies the network, identifies the theoretical models adopted, describes their application, includes the analysis and the implementation of the corresponding algorithms or the tools used, and contains an analysis of the experimental results. By doing the programming project, students develop the ability to select and apply the appropriate techniques and tools, the ability to create and evaluate studies, and skills in writing communication. The project also aims at improving autonomy and research skills.

3.3.9. Bibliografia principal:

1. Mark E. J. Newman. *Networks: An Introduction*. Oxford University Press, 2010.
2. David Easley and Jon Kleinberg. *Networks, Crowds and Markets: Reasoning about a Highly Connected World*. Cambridge University Press, 2010.
Available on-line at <https://www.cs.cornell.edu/home/kleinber/networks-book/>
3. Albert-László Barabási. *Network Science*.
Available online at <http://barabasi.com/book/network-science>

Mapa IV - Tópicos de Otimização / Topics in Optimization

3.3.1. Unidade curricular:

Tópicos de Otimização / Topics in Optimization

3.3.2. Docente responsável (preencher o nome completo) e respetivas horas de contacto na unidade curricular:

Jorge Orestes Lasbarrères Cerdeira - TP 8h

3.3.3. Outros docentes e respetivas horas de contacto na unidade curricular:

Ana Luísa da Graça Batista Custódio – TP 8h
Isabel Cristina Silva Correia – TP 8h
Manuel Valdemar Cabral Vieira – TP 8h
Maria do Carmo Proença Caseiro Brás – TP 8h
Maria Isabel Azevedo Rodrigues Gomes – TP 8h
Paula Alexandra da Costa Amaral – TP 8h

3.3.4. Objetivos de aprendizagem (conhecimentos, aptidões e competências a desenvolver pelos estudantes):

No final desta unidade curricular o estudante terá adquirido conhecimentos, aptidões e competências que lhe permitem:

1. *Enquadrar um problema de otimização numa das subáreas deste domínio científico;*
2. *Identificar metodologias e algoritmos que permitam a resolução de problemas pertencentes às diferentes subáreas da otimização.*

3.3.4. Intended learning outcomes (knowledge, skills and competences to be developed by the students):

At the end of this course the student will have acquired knowledge and skills that will enable him to:

1. *Assign a given optimization problem to a subfield of this scientific domain;*
2. *Identify methodologies and algorithms for the resolution of problems belonging to the different subfields of optimization.*

3.3.5. Conteúdos programáticos:

Esta UC visa abordar fundamentos e aplicações em diversas áreas da otimização tais como:

1. *Programação linear;*
2. *Programação inteira;*
3. *Programação não linear;*
4. *Programação estocástica;*
5. *Programação multiobjetivo.*

3.3.5. Syllabus:

This course aims to address fundamentals and applications in different areas of optimization such as:

1. *Linear programming;*
2. *Integer programming;*
3. *Nonlinear programming;*
4. *Stochastic programming;*
5. *Multiobjective programming.*

3.3.6. Demonstração da coerência dos conteúdos programáticos com os objetivos de aprendizagem da unidade curricular:

Em qualquer um dos capítulos serão apresentados exemplos ilustrativos dos problemas pertencentes à respetiva subárea da otimização, permitindo ao estudante familiarizar-se com a classe de problemas em causa, cumprindo-se o objetivo 1 de aprendizagem.

Para cada classe de problemas serão estudadas as suas características principais, bem como algoritmos adequados à respetiva resolução. Cumpre-se assim o objetivo 2 de aprendizagem.

3.3.6. Evidence of the syllabus coherence with the curricular unit's intended learning outcomes:

Examples of problems belonging to each one of the optimization subareas will be provided, allowing the student to be familiar with the problem class under analysis. This is in accordance with learning outcome 1.

The main characteristics of each problem class, as well as algorithms for the corresponding solution, will be analyzed. This fulfils the learning outcome 2.

3.3.7. Metodologias de ensino (avaliação incluída):

As aulas são teóricas/práticas participadas, com exposição oral dos conceitos e metodologias devidamente complementada com exemplos e resoluções de problemas. Eventuais dúvidas poderão ser esclarecidas no decurso das aulas ou em sessões individuais marcadas com os professores.

Os alunos necessitam de assistir a um mínimo de 2/3 das aulas teóricas/práticas lecionadas para se submeterem a avaliação. A avaliação contínua é baseada em três testes. Se um aluno não obtiver aprovação através de avaliação contínua poderá vir a obtê-la num exame de recurso.

3.3.7. Teaching methodologies (including assessment):

Classes are theoretical/practical with oral presentation of concepts, methodologies, and examples, complemented with problem solving. Specific student difficulties will be addressed during classes or in individual sessions scheduled with the professor.

Students need to attend a minimum of two thirds of the classes in order to be evaluated. Continuous evaluation is based on three tests. If a student does not obtain approval through continuous evaluation he can try it in an additional examination.

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

A apresentação de conceitos e metodologias, apoiados com exemplos ilustrativos, pretende motivar os estudantes para a relevância dos tópicos estudados, desenvolvendo a sua capacidade de identificação de problemas e seleção da metodologia mais adequada para a respetiva resolução.

Os alunos terão oportunidade de testar estas capacidades nas aulas, com o apoio de um professor, ou em estudo individual, comparecendo eventualmente a sessões de atendimento individual, em caso de dificuldades.

A avaliação da unidade curricular incide sobre a aquisição das competências referidas.

3.3.8. Evidence of the teaching methodologies coherence with the curricular unit's intended learning outcomes:

In classes, concepts and methodologies are presented and illustrated with examples, motivating students for the relevance of the corresponding study, and developing student's ability for identifying problems and selecting the most adequate methodology for the corresponding resolution.

Additionally, classes will allow students to test these skills, under the supervision of a professor. Students can also test these skills by themselves, scheduling individual appointments with professors in case of difficulties.

The evaluation of the curricular unit focuses in the acquisition of the mentioned capabilities.

3.3.9. Bibliografia principal:

Introduction to Stochastic Programming, John R. Birge, François Louveaux, Springer Series in Operations Research and Financial Engineering, 2011

ISBN: 978-1-4614-0236-7 (Print) 978-1-4614-0237-4 (Online)

Multiobjective Linear and Integer Programming, Carlos Henggeler Antunes, Maria João Alves, João Climaco, Springer International Publishing, 2016.

ISBN: 978-3-319-28744-7 (Print) 978-3-319-28746-1 (Online)

Numerical Optimization, Jorge Nocedal, Stephen Wright, Springer Series in Operations Research and Financial Engineering, 2006

ISBN: 978-0-387-30303-1 (Print) 978-0-387-40065-5 (Online)

Operations Research: Applications and Algorithms, Wayne L. Winston, Wadsworth, Belmont, CA, 1994

ISBN: 978-0-534-38058-8

Mapa IV - Aprendizagem com Dados Não Estruturados / Learning from Unstructured Data

3.3.1. Unidade curricular:

Aprendizagem com Dados Não Estruturados / Learning from Unstructured Data

3.3.2. Docente responsável (preencher o nome completo) e respetivas horas de contacto na unidade curricular:

Pedro Manuel Corrêa Calvente Barahona (apenas Responsável não tem horas de contacto)

3.3.3. Outros docentes e respetivas horas de contacto na unidade curricular:

Rui Alberto Pimenta Rodrigues - T:14h; PL:14h

Ludwig Krippahl - T:14h; PL:14h

3.3.4. Objetivos de aprendizagem (conhecimentos, aptidões e competências a desenvolver pelos estudantes):

No final desta unidade curricular o estudante terá adquirido conhecimentos, aptidões e competências que lhe permitam:

Compreender:

-Os princípios básicos do 'deep learning':

extração automática das características ou representações, usadas para classificação ou regressão, através de múltiplas camadas.

métodos de optimização e regularização a usar nos modelos com grande número de parâmetros

-As diferenças entre os modelos apresentados: redes 'feed forward' com todas as conexões, redes de convolução e redes recorrentes.

-Problemas e técnicas para processar dados não estruturados

Ser capaz de:

-Seleccionar e implementar um modelo para resolver alguns problemas típicos e conseguir otimizá-lo para obter uma solução razoável

Conhecer:

Os problemas aos quais se tem aplicado o 'deep learning': reconhecimento de objectos em imagens, reconhecimento de voz, processamento de linguagem natural, sistemas de recomendação, etc.

3.3.4. Intended learning outcomes (knowledge, skills and competences to be developed by the students):

At the end of this unit, students will:

Understand:

-basic principles of deep learning.

automated extraction of features and representations for use in regression or classification in multi-layer models.

optimization and regularization methods applicable to models with a large number of parameters

- the different models presented: fully connected feed forward neural networks, convolution networks and recurrent networks.

- Problems and techniques for processing unstructured data.

Be able to:

-Select and implement models to solve some typical problems and optimize training to obtain a reasonable solution

Know:

Problems to which deep learning applies: object recognition in images, voice recognition, natural language processing, recommendation systems and others.

3.3.5. Conteúdos programáticos:

1 -Introdução: tipos de dados e problemas para resolver.

2 - Deep Feedforward Nets

3 - Otimização das redes Feedforward

4 - Regularizações das redes Feedforward

5 - Redes de convolução

6 - Redes recorrentes e problemas sequenciais

7 - Introdução aos modelos geradores e aplicação ao 'unsupervised learning'.

8 - Aprendizagem semi-supervisionada, representações e transference learning

9 - Mais alguns tipos de problemas (sistemas de recomendação, etc)

3.3.5. Syllabus:

1 - Introduction: data types and problems.

2 - Deep Feedforward networks

3 - Optimization of Feedforward networks.

4 - Regularization of Feedforward networks.

5 - Convolution networks.

6 - Recurrent networks and sequential problems.

7 - Introduction to generator models and their application to unsupervised learning.

8 - Semi-supervised learning, representation learning and transference learning

9 - Additional problems and examples.

3.3.6. Demonstração da coerência dos conteúdos programáticos com os objetivos de aprendizagem da unidade curricular:

1-O primeiro ponto dos conteúdos programáticos visa situar o aluno no tipo de problemas que irão ser abordados e para os quais são dirigidas as técnicas que serão apresentadas.

2-As redes 'feedforward' são os modelos básicos apresentados dos quais os outros derivam. A extração automática de características, a partir dos dados não processados é realizada nas camadas intermédias destas redes.

3- A otimização das redes é o passo necessário para ajustar o modelo aos dados de treino.

4- A regularização na otimização é a técnica usada para impedir um sobreajustamento dos parâmetros do modelo aos dados de treino e permitir desta forma que este se generalize a dados desconhecidos na fase de treino.

Os pontos 2, 3 e 4 contém as técnicas básicas a usar na resolução de um problema com a abordagem do deep learning.

5- As redes de convolução são usadas para problemas em que se procura uma certa invariância das características ou representações, invariância na translação espacial ou no tempo, por exemplo. Entre outras aplicações está o reconhecimento de objetos em imagem.

6- As redes recorrentes são usadas em problemas sequências. Por exemplo a interpretação de um texto pode ser visto como um problema que envolve uma sequência de palavras.

7- Modelos geradores são usados quer para atribuir uma probabilidade à ocorrência de um objeto ou conjunto de dados quer para gerar dados semelhantes aos encontrados num conjunto de treino. Uma das aplicações será a o preenchimento de entradas ausentes nos dados.

8- Muitas vezes é possível usar dados não classificados para melhorar a aprendizagem de dados classificados ou transferir representações aprendidas num domínio para facilitar a aprendizagem noutra domínio. A aprendizagem semi-supervisionada e a transferência podem recorrer a técnicas como, por exemplo, autoencoders ou outros modelos que aprendam representações latentes.

9- Um outro conjunto de problemas, que envolve uma combinação de técnicas ou mesmo técnicas que não serão abordadas neste curso, será introduzido de forma mais superficial.

3.3.6. Evidence of the syllabus coherence with the curricular unit's intended learning outcomes:

- 1- The first topic is aimed at introducing the type of problems and techniques that will be covered in this unit.
 - 2- Feedforward neural networks are the basic models from which the others are derived, with automatic feature extraction from unprocessed data occurring in the hidden layers of these networks.
 - 3- Network optimization is a necessary step for fitting the model to the training data.
 - 4- Regularization of the training process is a widely used technique to prevent overfitting of the model parameters to the training data, which degrades generalization to new data.
- Topics 2, 3 and 4 cover the basic techniques for solving problems with deep learning.
- 5- Convolution networks are used to model invariance in some features or representations, such as spacial or temporal invariance. Examples include object recognition in images.
 - 6- Recurrent networks are useful in problems with sequential data. For example, in text processing, since texts are sequences of words.
 - 7- Generator models are used both to assign probabilities to the occurrence of specific examples from a data set or to generate new examples from a similar probability distribution. This can be applied, for example, for reconstructing missing values from a data set.
 - 8- It is often possible to use unlabelled data to improve the classification of labelled data or to transfer learning and representations from one domain to another to facilitate the training of new models. Semi-supervised learning and transfer learning can resort to techniques such as autoencoders or other latent representation models.
 - 9- Other problems, involving combinations of the techniques covered in this unit or even other techniques, will be introduced in less detail to provide a broader overview of this area.

3.3.7. Metodologias de ensino (avaliação incluída):

- As aulas teóricas cobrem os tópicos fundamentais da matéria, dedicando tempo para exposição da matéria e para dúvidas e discussão com os alunos.
- As aulas laboratoriais destinam-se à orientação tutorial e realização dos trabalhos práticos.
- A avaliação da disciplina é composta por uma componente teórica-prática (T) e uma componente laboratorial ou de projeto (P), tendo cada componente um peso de 50% na nota final.
- A componente teórica-prática (T) consiste em dois testes que contarão em partes iguais para a nota desta componente. Em alternativa, esta componente pode ser realizada por exame de recurso.

3.3.7. Teaching methodologies (including assessment):

- Lectures will cover the fundamental topics of the subject matter, include some time for questions and discussion of the subject matter.
- Tutorial classes will be dedicated to exercises and guidance in the practical assignments, focusing on selected topics.
- The evaluation of this curricular unit is made by two components: theoretical/problems (T) and project (P). Each component contributes with 50% to the final grade.
- The theoretical part consists of two written individual tests; alternatively, this component can be evaluated by a written exam.

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

- As várias metodologias estudadas são enquadradas com exemplos de aplicação.
- Os trabalhos práticos e exercícios tutoriais permitem consolidar conhecimentos na implementação/adaptação dos algoritmos estudados a problemas reais. Será dada particular atenção à interpretação de resultados experimentais, avaliação dos métodos usados e comparação de modelos.
- As noções gerais e conceitos apresentados ao longo do curso serão avaliados nos dois testes ou no exame final.

3.3.8. Evidence of the teaching methodologies coherence with the curricular unit's intended learning outcomes:

- The different techniques covered in this unit will be framed within practical problems to illustrate their application.
- The tutorials and assignments will help consolidate knowledge and skills in the application and adaptation of the algorithms to real problems. Especial attention will be given to the interpretation of experimental results, evaluation and comparison of different models.
- Understanding of the theoretical notions and concepts presented along this unit will be assessed in the written tests or final exam.

3.3.9. Bibliografia principal:

- Ian Goodfellow, Yoshua Bengio, Aaron Courville, *Deep Learning*, MIT Press
- Yann LeCun, Yoshua Bengio and Geoffrey Hinton, *Deep learning*, *Nature*, Vol. 521, pp 436-444, 2015
- Li Deng and Dong Yu, *Deep Learning Methods and Applications*, *Foundations and Trends in Signal Processing*, Volume 7 Issues 3-4 (2013), pp 97-387, ISSN: 1932-8346
- <https://www.microsoft.com/en-us/research/wp-content/uploads/2016/02/DeepLearning-NowPublishing-Vol7-SIG-039.pdf>

Mapa IV - Prospecção de Dados / Data Analytics and Text Mining**3.3.1. Unidade curricular:**

Prospecção de Dados / Data Analytics and Text Mining

3.3.2. Docente responsável (preencher o nome completo) e respetivas horas de contacto na unidade curricular:

Pedro Manuel Calvente Barahona (apenas Responsável, não tem horas de contacto)

3.3.3. Outros docentes e respetivas horas de contacto na unidade curricular:

Joaquim Francisco Ferreira da Silva – T:14h; PL:14h

Susana Maria dos Santos Nascimento Martins de Almeida -- T:14h; PL:14h

3.3.4. Objetivos de aprendizagem (conhecimentos, aptidões e competências a desenvolver pelos estudantes):

Esta UC aborda princípios, métodos e recomendações práticas para extrair padrões relevantes a partir de dados estruturados e não estruturados (numéricos textuais) numa perspectiva de interface entre Ciência da Computação e Estatística.

Conhecimentos:

- *Pré-processamento de dados*
- *Métodos de redução de dimensionalidade*
- *Algoritmos de agrupamento para dados entidade-atributo e dados em grafo*
- *Métodos computacionais de validação de agrupamento*
- *Métodos de Extração de Informação Relevante*
- *Análise simbólica e análise estatística de textos*
- *Descritores e classificação de Documentos*
- *Casos de estudo em Data Analytics e Text Mining*

Aptidões:

- *Implementar/adaptar algoritmos de Data Analytics e Text Mining*
- *Modelar experimentalmente dados reais*
- *Interpretar e avaliar resultados experimentais*

Competências:

- *Avaliar a adequação dos métodos a dados e casos de estudo*
- *Avaliar criticamente os resultados obtidos*
- *Autonomia em aplicar e aprofundar os conhecimentos adquiridos*

3.3.4. Intended learning outcomes (knowledge, skills and competences to be developed by the students):

This course addresses principles, methods and practical recommendations for extracting meaningful patterns from structured and unstructured data (numeric and textual data), from a perspective at the interface between Computer Science and Statistics.

Knowledge

- *Data pre-processing*
- *Dimensionality reduction methods*
- *Clustering algorithms for entity-attribute data and graph-data*
- *Computational validation methods*
- *Relevant Information extraction methods*
- *Symbolic analysis and statistical analysis of text*
- *Document descriptors and document Classification*
- *Case studies in Data Analytics and Text Mining*

Skills

- *Implement and adapt Data Analytics and Text Mining algorithms*
- *Model real data experimentally*
- *Interpret and evaluate experimental results*

Competences

- *Evaluate the suitability of each method to case studies*
- *Critical evaluation of experimental results*
- *Autonomy and self-reliance in the application and furthering studies*

3.3.5. Conteúdos programáticos:

Introdução

Data Analytics: Exemplos de aplicação e perspectivas

Text Mining

Prospecção de Informação estruturada ou não estruturada

Data Analytics

- *Compreender os dados*

Sumarização, correlação e visualização

Verificação da estrutura dos dados

- *Pré-processamento de dados*
- *Redução de Dimensionalidade*

*Análise de Componentes Principais
Decomposição em valores singulares*

- *Modelação Descritiva*

Algoritmos de agrupamento por pontos representativos

Algoritmos de agrupamento espectral

- *Validação e Interpretação de Modelos Descritivos*
- *Casos de Estudo de Data Analytics*

Text Mining

- *Extratores de Expressões Relevantes*

- *Análise Simbólica e análise estatística de textos*

Tokenization, Stemming e etiquetagem morfológica

Distribuição das palavras em texto

Desambiguação semântica

- *Descritores de Documentos*

Extração de palavras-chave explícitas e implícitas. Sumarização

- *Classificação supervisionada vs não-supervisionada de Documentos*
- *Casos de estudo em Text Mining*

3.3.5. Syllabus:

Introduction

Data Analytics

What is data? Examples of data analytic tasks and perspectives

Text Mining

Mining of structured or unstructured data

Data Analytics

- *Data Understanding*

Summarization, correlation and visualization

Verification of structure in data

- *Data Pre-processing*

- *Dimensionality Reduction*

Principal Component Analysis

Singular Value Decomposition

- *Descriptive Modeling*

Representative-based clustering algorithms

Spectral clustering

- *Validation and Interpretation of Descriptive Modeling*

- *Data Analytics Case Studies*

Text Mining

- *Extractors of Relevant Expressions*

- *Symbolic and Statistical Analysis of texts*

Tokenization, Stemming and Part-Of-Speech Tagging

Word distribution in texts

Word Sense Disambiguation

- *Document Descriptors*

Extraction of explicit and Implicit Keywords. Summarization

- *Supervised vs unsupervised Document Classification*

- *Text Mining Case Studies*

3.3.6. Demonstração da coerência dos conteúdos programáticos com os objetivos de aprendizagem da unidade curricular:

São apresentadas e discutidas várias abordagens de estado da arte em data analytics e text mining de acordo com a bibliografia adotada. As várias metodologias estudadas são enquadradas com exemplos de aplicação e casos de estudo.

Os trabalhos práticos e exercícios tutoriais permitem consolidar conhecimentos na implementação/adaptação dos algoritmos estudados a problemas reais orientados aos dados.

A programação de algoritmos e de protocolos experimentais de validação orientados aos dados oferece experiência na modelação e computacional de dados.

3.3.6. Evidence of the syllabus coherence with the curricular unit's intended learning outcomes:

Several state of the art models and methods of data analytics and text mining are presented and discussed according to the adopted text books. The methods studied are applied to practical examples and case studies.

Tutorial classes and practical assignments promote the consolidation of theoretical knowledge as well as practice in the implementation and adaptation of the algorithms to data-driven real world problems. Practical assignments also focus on the interpretation of experimental results and on the evaluation of the methods used.

The programming of algorithms and data-driven experimental protocols provide experience on computational data modeling.

3.3.7. Metodologias de ensino (avaliação incluída):

Nas aulas teóricas são lecionados os tópicos fundamentais da matéria, os quais deverão ser complementados com a leitura da bibliografia adotada.

As aulas laboratoriais destinam-se à orientação tutorial e realização dos trabalhos práticos.

Em todas as aulas são disponibilizados os acetatos da matéria teórica.

A avaliação desta UC é composta por uma componente teórica-prática (T) e uma componente laboratorial ou de projeto (P), tendo cada componente um peso de 50% na nota final.

Para obter aprovação à disciplina um estudante necessita de: nota mínima de 10 valores em cada uma das componentes de avaliação de cada módulo (I e II). A nota final é a média das duas componentes de avaliação.

A componente teórica-prática (T) consiste em dois testes, que contarão em partes iguais para a nota desta componente. Em alternativa, esta componente pode ser realizada por exame de recurso.

A componente de projeto consiste na realização de dois projetos de programação acompanhados de relatórios escritos.

3.3.7. Teaching methodologies (including assessment):

Lectures will cover the fundamental topics of the subject matter, which the students should complement with the adopted bibliography. All lecture materials will be supplied for further study.

Tutorial classes will be dedicated to exercises and guidance in the practical assignments, focusing on selected topics.

The evaluation of this CU is made by two components: theoretical/problems (T) and project (P). Each component contributes with 50% to the final grade.

To pass, the student must have: a grade of at least 10 points (out of 20 points) in each of the theoretical/problems and project components, in each module (I and II). The final grade is the average of the two components of evaluation.

The theoretical part consists of two written individual tests; alternatively, this component can be evaluated by a written exam.

The project component is evaluated by a set of assignments, and two programming projects accompanied by written reports.

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

Os objetivos desta unidade curricular são:

a) a aquisição de conhecimentos teóricos para a compreensão de várias aproximações de Data Analytics e Text Mining, das metodologias e dos princípios da sua aplicação;

b) a aptidão para implementar estes algoritmos e, assim, resolver problemas práticos;

c) a competência para avaliar a adequação de cada método a casos de estudo e para avaliar criticamente os resultados obtidos.

O objetivo a) será atingido pela exposição da matéria, discussão e esclarecimento de dúvidas nas aulas teóricas, onde se abordará o fundamento teórico desta unidade curricular. Esta matéria será avaliada em testes individuais, onde os alunos têm um tempo limitado para mostrar o seu grau de assimilação e compreensão dos assuntos.

Os objetivos b) e c) serão atingidos principalmente pela componente prática da UC, se bem que sempre assente nas aulas teóricas, por meio das aulas práticas e do trabalho autónomo do aluno. As aulas práticas darão aos alunos oportunidade para aplicar os conhecimentos a casos concretos, experimentar as diferentes abordagens e examinar, em detalhe, os vários algoritmos em condições realistas. A componente prática será avaliada em dois trabalhos de grupo, avaliação esta que focará não só a implementação e aplicação prática dos algoritmos mas também a capacidade crítica de avaliação dos métodos escolhidos e dos resultados. Esta componente prática contribuirá também para a autonomia dos alunos.

3.3.8. Evidence of the teaching methodologies coherence with the curricular unit's intended learning outcomes:

The course objectives are:

a) the acquisition of theoretical knowledge to allow the understanding of different approximations of Data Analytics and Text Mining, methodologies and their principles of application;

b) the skill to implement these algorithms and, thus, solve practical problems;

c) the competence to evaluate the suitability of each method to case studies and for a critical evaluation of the results.

Objective a) will be reached mainly through the theoretical lectures, which include exposition of the subject matter and discussion. These lectures will cover the theoretical foundation of the course and this knowledge will be evaluated in individual written tests, where the students will have limited time to show their success in assimilating and understanding the subjects.

Objectives b) and c) will be reached mainly through practice, during the tutorial classes and the student's own work. Tutorial classes will give the students the opportunity to apply their knowledge to specific cases, experiment with different approaches and examine, in detail, the different algorithms in realistic conditions. This practical component will be evaluated in two group assignments, focusing on the implementation and practical application of the algorithms but also on the critical assessment of the methods and results. This component will contribute to give the student the autonomy and self-reliance needed for future applications and further studies in this field.

3.3.9. Bibliografia principal:

1. Module I

Larose, D.T., Larose, C. D. (2015). Data Mining and Predictive Analytics, 2nd Edition, Wiley.

Zaki, M. J., Meira Jr., W. (2014). Data Mining and Analysis, Fundamental Concepts and Algorithms, Cambridge

University Press

Mirkin, B. (2011). *Core Concepts in Data Analysis: Summarization, Correlation, Visualization. Undergraduate Topics for Computer Science Series*, Springer, London.

2. Module II

Weiss, S.M., Indurkha, N., Zhang, T., Damerou, F. (2005). *Predictive Methods for Analyzing Unstructured Information*, Springer.

Outras publicações científicas de acordo com cada tópico/ other scientific publications according to each topic.

Mapa IV - Dissertação / Dissertation

3.3.1. Unidade curricular:

Dissertação / Dissertation

3.3.2. Docente responsável (preencher o nome completo) e respetivas horas de contacto na unidade curricular:

Pedro Manuel Corrêa Calvente Barahona /Jorge Orestes Lasbarrères Cerdeira - OT:40h

3.3.3. Outros docentes e respetivas horas de contacto na unidade curricular:

Todos os docentes envolvidos na proposta do ciclo de estudos - OT:40h

3.3.4. Objetivos de aprendizagem (conhecimentos, aptidões e competências a desenvolver pelos estudantes):

O trabalho conducente à dissertação de Mestrado em Análise e Engenharia de Big Data enquadra-se nos termos definidos na alínea b) do Artigo 20º do Decreto Lei 74/2006, e consiste num trabalho individual de investigação e/ou desenvolvimento que explora os conhecimentos adquiridos ao longo do curso, original e especialmente concebido para esse fim.

O trabalho de dissertação pode basear-se na elaboração de uma tese de natureza científica ou no desenvolvimento de um projeto avançado de engenharia de concepção, de preferência inspirado ou tendo como base um caso de estudo de uma instituição parceira do mestrado, podendo combinar as duas vertentes, e sendo avaliado de acordo com parâmetros uniformes.

3.3.4. Intended learning outcomes (knowledge, skills and competences to be developed by the students):

The dissertation work for Master in Big Data Analytics and Engineering conforms with the requirements of item b) Art 20º of DL 74/2006, and consists of an individual research and / or development work which explores the knowledge acquired in the whole program, original, and specifically identified for that purpose.

The dissertation work may be based on the elaboration of a research-oriented thesis of scientific nature, or in the development of an advanced conception-oriented engineering project, preferably based on a case study presented by a partner institution of the MSc programme, possibly combining the two orientations, and evaluated according to common criteria.

3.3.5. Conteúdos programáticos:

Desenvolvimento de trabalho conducente a elaboração de dissertação de Mestrado em Análise e Engenharia De Big Data.

3.3.5. Syllabus:

Development of the work leading to the elaboration of a master dissertation in Big Data Analytics and Engineering.

3.3.6. Demonstração da coerência dos conteúdos programáticos com os objetivos de aprendizagem da unidade curricular:

Os conteúdos são naturalmente coerentes com os objetivos, dada a sua natureza especial da unidade curricular.

3.3.6. Evidence of the syllabus coherence with the curricular unit's intended learning outcomes:

Given the special of the curricular unit, the syllabus is naturally coherent with the unit's objectives nature

3.3.7. Metodologias de ensino (avaliação incluída):

O trabalho conducente à dissertação de mestrado tem a duração de dois semestres. No 1º semestre, o estudante desenvolve o trabalho introdutório do tema da sua dissertação e algum trabalho de desenvolvimento preliminar, apresentando um relatório que inclui um plano do trabalho a realizar para concluir a Dissertação do Mestrado, durante o segundo semestre. Esse relatório é discutido e avaliado numa sessão pública por um júri que o deverá validar e eventualmente apresentar sugestões relevantes.

No final do 2º semestre, o trabalho desenvolvido será formalmente apresentado na Dissertação de Mestrado, defendida em provas públicas perante o júri, que a avalia e lhe atribui uma nota final, de acordo com o regulamento geral da FCT e da UNL.

3.3.7. Teaching methodologies (including assessment):

The work leading to the master dissertation develops in two consecutive semesters. In the first semester the student develops preliminary work for the dissertation theme and some initial development work, and presents a written report, including a plan of the work needed to conclude the master dissertation work. This report is discussed and evaluated in a public session by a jury that validates and possibly proposes suggestions for improving the future work. At the end of the 2nd semester, the work developed is formally reported in the Master Dissertation, which is defended by the candidate in a public exam before a jury who assigns the final grading, according to the general regulations of FCT and UNL.

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

Tendo em conta a natureza especial desta unidade curricular, decorre que existe coerência adequada entre as metodologias de ensino e os objetivos de aprendizagem da mesma. Acresce que a metodologia de ensino adotada, estruturada em duas fases, com um ponto de aferição intermédio após a fase de preparação, tem-se revelado um fator promotor do aumento de qualidade do processo de desenvolvimento das dissertações, na opinião de todos os envolvidos (estudantes, docentes, membros de júri, e parceiros exteriores).

O trabalho de dissertação, independentemente da sua natureza de introdução à investigação ou de projeto avançado de engenharia de concepção, poderá desenvolver-se quer em contexto académico, quer em contexto de colaboração académico-empresarial, em qualquer dos casos, sob coordenação científica geral da comissão científica de curso e orientação formal de um membro do corpo docente do curso.

Tipicamente, serão avaliados os seguintes parâmetros:

1. *Qualidade da análise do estado da arte e da sua adequação aos objetivos do trabalho;*
2. *Qualidade do trabalho realizado;*
3. *Extensão do trabalho realizado;*
4. *Qualidade da análise crítica da contribuição e dos resultados obtidos;*
5. *Qualidade da apresentação oral e da argumentação;*
6. *Qualidade do Relatório;*
7. *Apreciação Geral.*

3.3.8. Evidence of the teaching methodologies coherence with the curricular unit's intended learning outcomes:

Having in mind the special nature of this curricular unit, it clearly results that there is adequate coherence between the teaching methodologies and the learning outcomes. Additionally, we notice that the teaching methodology, structured in two phases, with an interim assessment point after the preparation first phase, has revealed itself as a factor that promotes a quality increase of the dissertation supervision process, in the opinion of all the stakeholders (students, faculty members, jury members, and external collaborators).

The dissertation work, independently of its orientation towards research or advanced engineering project, may be developed either in academic context or in some collaboration with an external partner institution (such as a firm), in any case under the scientific coordination of the program's scientific committee and formal supervision of a faculty member.

Typically, the following parameters are taken into account in the final assessment:

1. *Quality of the state-of-the-art review and its suitability to the work objectives;*
2. *Quality of the performed work;*
3. *Extension of the performed work;*
4. *Quality of the critical analysis of the contributions and results obtained;*
5. *Quality of the oral presentation and argumentation;*
6. *Quality of the written document;*
7. *Global appreciation.*

3.3.9. Bibliografia principal:

Depende de cada trabalho de dissertação / Depends on each specific dissertation work.

4. Descrição e fundamentação dos recursos docentes do ciclo de estudos

4.1 Descrição e fundamentação dos recursos docentes do ciclo de estudos

4.1.2 Equipa docente do ciclo de estudos

4.1.2. Equipa docente do ciclo de estudos / Teaching staff of the study programme

Nome / Name	Grau / Degree	Área científica / Scientific Area	Regime de tempo / Employment link	Informação/ Information
Nuno Manuel Ribeiro Preguiça	Doutor	Informática	100	Ficha submetida
Nuno Manuel Robalo Correia	Doutor	Engenharia Informática e Electrotécnica	100	Ficha submetida
Sérgio Marco Duarte	Doutor	Informática	100	Ficha submetida
João Miguel da Costa Magalhães	Doutor	Computer Science	100	Ficha submetida

Hervé Miguel Cordeiro Paulino	Doutor	Informática	100	Ficha submetida
Susana Maria dos Santos Nascimento Martins de Almeida	Doutor	Informática	100	Ficha submetida
Ludwig Krippahl	Doutor	Bioquímica Estrutural	100	Ficha submetida
João Carlos Gomes Moura Pires	Doutor	Informática / Inteligência Artificial	100	Ficha submetida
Joaquim Francisco Ferreira da Silva	Doutor	Informática	100	Ficha submetida
Pedro Manuel Corrêa Calvente Barahona	Doutor	Computer Science	100	Ficha submetida
Margarida Paula Neves Mamede	Doutor	Informática	100	Ficha submetida
Jorge Orestes Lasbarrères Cerdeira	Doutor	Matemática	100	Ficha submetida
Isabel Cristina Maciel Natário	Doutor	Estatística e Investigação Operacional - ramo Probabilidades e Estatística	100	Ficha submetida
Filipe José Gonçalves Pereira Marques	Doutor	Matemática - Especialidade de Estatística	100	Ficha submetida
Pedro José dos Santos Palhinhas Mota	Doutor	Matemática- Especialidade Estatística	100	Ficha submetida
Luís Pedro Carneiro Ramos	Doutor	Matemática	100	Ficha submetida
Rui Alberto Pimenta Rodrigues	Doutor	Matemática, equações diferenciais	100	Ficha submetida
Manuel Leote Tavares Inglês Esquível	Doutor	Matemática/Processos Estocásticos	100	Ficha submetida
Maria Isabel Azevedo Rodrigues Gomes	Doutor	Engenharia e Gestão Industrial	100	Ficha submetida
Miguel dos Santos Fonseca	Doutor	Matemática	20	Ficha submetida
Fernando Pedro Reino da Silva Birra	Doutor	Informática	100	Ficha submetida
Ana Luísa da Graça Batista Custódio	Doutor	Matemática	100	Ficha submetida
Isabel Cristina Silva Correia	Doutor	Estatística e Investigação Operacional	100	Ficha submetida
Manuel Valdemar Cabral Vieira	Doutor	Matemática Aplicada	100	Ficha submetida
Maria do Carmo Proença Caseiro Brás	Doutor	Matemática na especialidade de Investigação Operacional	100	Ficha submetida
Carlos Manuel Agra Coelho	Doutor	Estatística/Bioestatística	100	Ficha submetida
Paula Alexandra da Costa Amaral	Doutor	Matemática	100	Ficha submetida
Marta Cristina Vieira Faias Mateus	Doutor	Economia	100	Ficha submetida
Luís Manuel Marques da Costa Caires	Doutor	Informática	100	Ficha submetida
José Júlio Alves Alferes	Doutor	Informática	100	Ficha submetida
Carlos Augusto Isaac Piló Viegas Damásio	Doutor	Informática	100	Ficha submetida
Virgílio António da Cruz Machado	Doutor	Computer Integrated Manufacturing	100	Ficha submetida
João Alexandre Carvalho Pinheiro Leite	Doutor	Informática	100	Ficha submetida
(33 Items)			3220	

<sem resposta>

4.2. Dados percentuais dos recursos docentes do ciclo de estudos

4.2.1. Corpo docente próprio do ciclo de estudos

4.2.1. Corpo docente próprio do ciclo de estudos / Full time teaching staff

Corpo docente próprio / Full time teaching staff	ETI / FTE	Percentagem* / Percentage*
Nº de docentes do ciclo de estudos em tempo integral na instituição / No. of full time teachers:	32	99.4

4.2.2. Corpo docente do ciclo de estudos academicamente qualificado

4.2.2. Corpo docente do ciclo de estudos academicamente qualificado / Academically qualified teaching staff

Corpo docente academicamente qualificado / Academically qualified teaching staff	ETI / FTE	Percentagem* / Percentage*
Docentes do ciclo de estudos com o grau de doutor (ETI) / Teaching staff with a PhD (FTE):	32.2	100

4.2.3. Corpo docente do ciclo de estudos especializado

4.2.3. Corpo docente do ciclo de estudos especializado / Specialized teaching staff

Corpo docente especializado / Specialized teaching staff	ETI / FTE	Percentagem* / Percentage*
Docentes do ciclo de estudos com o grau de doutor especializados nas áreas fundamentais do ciclo de estudos (ETI) / Teaching staff with a PhD, specialized in the main areas of the study programme (FTE):	28.2	87.6
Especialistas, não doutorados, de reconhecida experiência e competência profissional nas áreas fundamentais do ciclo de estudos (ETI) / Specialists, without a PhD, of recognized professional experience and competence, in the main areas of the study programme (FTE):	0	0

4.2.4. Estabilidade do corpo docente e dinâmica de formação

4.2.4. Estabilidade do corpo docente e dinâmica de formação / Teaching staff stability and training dynamics

Estabilidade e dinâmica de formação / Stability and training dynamics	ETI / FTE	Percentagem* / Percentage*
Docentes do ciclo de estudos em tempo integral com uma ligação à instituição por um período superior a três anos / Full time teaching staff with a link to the institution for a period over three years:	32	99.4
Docentes do ciclo de estudos inscritos em programas de doutoramento há mais de um ano (ETI) / Teaching staff registered in a doctoral programme for more than one year (FTE):	0	0

4.3. Procedimento de avaliação do desempenho

4.3. Procedimento de avaliação do desempenho do pessoal docente e medidas para a sua permanente atualização:

A FCT tem um Regulamento de Avaliação de Desempenho dos Docentes (RAD) (Despacho 13109/2012, publicado em DR, 2ª Série, n.º 193, de 4 de outubro), que se rege pelos princípios de universalidade e obrigatoriedade, imparcialidade e objetividade, equidade, confidencialidade e direito ao contraditório.

De acordo com o referido RAD, todos os docentes são avaliados em períodos trienais, com monitorização anual, nas vertentes de:

- Docência (e.g. diversidade de unidades curriculares lecionadas; resultados dos questionários aos estudantes; disponibilização de material pedagógico; orientação de dissertações de mestrado e de teses de doutoramento; participação em júris);*
- Investigação científica, desenvolvimento e inovação (e.g. coordenação e participação em projetos de investigação e direção de unidades de investigação; publicação de artigos e livros; comunicações em congressos científicos; participação em órgãos de revistas científicas; registo de patentes; participação em comissões, organizações ou redes científicas);*
- Tarefas administrativas e de gestão académica;*
- Extensão universitária, divulgação científica e prestação de serviços à comunidade (e.g. prémios e distinções públicas; transferência de tecnologia; serviços prestados a outras entidades).*

Da avaliação em cada vertente, resulta uma avaliação global no triénio expressa numa menção final de Excelente, Muito Bom, Bom ou Insuficiente.

A avaliação de cada docente é feita por dois avaliadores (um escolhido pelo próprio docente, e outro pelo presidente do departamento), com o contributo do presidente de departamento. Todo o processo é coordenado por um conselho eleito para esse efeito. O Conselho Científico e o Conselho Pedagógico são obrigatoriamente ouvidos sobre os resultados finais agregados do processo de avaliação. O Diretor atua como entidade de recurso, e os resultados finais são homologados pelo Reitor.

Os resultados da avaliação têm consequências no posicionamento remuneratório dos docentes, contratação por tempo indeterminado e renovações de contratos, e são tidos em conta na prioridade de concessão de licenças sabáticas, fixação do trabalho docente e obtenção de apoios extraordinários para coordenação ou dinamização de atividades.

A FCT concluiu o processo de avaliação de todos os seus docentes no triénio 2010-2012, estando agora a ser finalizado o processo de avaliação relativo ao triénio 2013-2015.

Tal como preconizado no próprio RAD, o Conselho Científico propôs, após debate na Faculdade, uma revisão do RAD, com vista à implementação de melhorias. As alterações ao RAD deverão ser publicadas em DR muito em breve.

4.3. Teaching staff performance evaluation procedures and measures for its permanent updating:

The school has an official Performance Assessment Regulation for the academic staff (Despacho 13109/2012, published in DR, 2.ª série, n.º 193, in October 4), governed by the principles of universality, impartiality, fairness, confidentiality, and right to adversarial.

By this regulation, all members of the academic staff are evaluated triennially, with observation every year, in the following aspects:

- Teaching (e.g. diversity of courses taught, students' satisfaction inquiries, teaching materials, MSc and PhD supervision, participation in academic juries);*
- Research (e.g., coordination and participation in research projects, coordination of research units, publication of scientific articles and books, conference papers, editorial boards of scientific journals and programme committees, patents);*
- Administrative and academic duties;*
- Extension activities, dissemination and services to the community (e.g., academic honours and awards, technology*

transfer, consultancy and other services to the community).

The final global evaluation for the 3years period results from the evaluation in each of the 4 subjects above, and is expressed in a grade of Excellent, Very Good, Good or Poor.

The evaluation of each professor is done by two evaluators (one chosen by the professor, and one by the corresponding head of department), and also has the contribution of the head of department. A commission elected for this purpose coordinates the whole process. The Scientific and Pedagogical Boards are consulted about the final aggregated results. The Director acts as appeal instance, and the Rector approves the final results of the evaluation. The results of the evaluation have an effect in the remuneration of the academic staff, in tenure, and in renovation of contracts of professors. They are also taken into account when authorising sabbatical leaves, in distribution of teaching load, or in the attribution of grants.

The evaluation process of the 2010-2012 period is already concluded, and the one for 2013-2015 is in an advanced state, to be concluded within 2016.

As advocated by the regulation itself, the evaluation process and results for 2010-2012 have been assessed, in order to introduce changes to be applied in the next evaluation period. This assessment was made by the Scientific Board, after debate in the School, and slight changes in the regulation were approved and will be published in DR very soon.

5. Descrição e fundamentação de outros recursos humanos e materiais

5.1. Pessoal não docente afeto ao ciclo de estudos:

Participam em atividades de suporte à gestão do ciclo de estudos o seguinte pessoal não docente:

- *Susana Pereira – Coordenadora Técnica / Especialista Informática Grau II Nível II (DI)*
- *Anabela Fonseca Nunes da Silva Duarte - Assistente Técnica (DI)*
- *Sandra Rosa Rafael Rainha - Assistente Técnica (DI)*
- *Maria da Graça Nobre dos Santos – Assistente Técnica (DM)*
- *Maria Libânia Patrício Gaspar – Assistente Técnica (DM)*
- *Maria Deolinda da Conceição Teixeira Mata – Assistente Operacional (DM)*

Apoio à coordenação e disseminação do curso, apoio aos docentes, atendimento dos estudantes.

- *Bruno Machado - Técnico de Informática.*

Apoio à gestão dos laboratórios de ensino, da redes de computadores, dos serviços de impressão, e da manutenção de informação na web, incluindo atividades de ligação com a Divisão de Informática, que também disponibiliza recursos humanos sempre que necessário.

5.1. Non teaching staff allocated to the study programme:

The following non-teaching staff participate in management support activities of the programme:

- *Susana Pereira - Technical Coordinator / Specialist Computer Grade II Level II (DI)*
- *Anabela Fonseca Nunes da Silva Duarte - Technical Assistant (DI)*
- *Sandra Rosa Rafael Queen - Technical Assistant (DI)*
- *Maria da Graça Nobre dos Santos - Technical Assistant (DM)*
- *Maria Libânia Patrick Gaspar - Technical Assistant (DM)*
- *Maria Deolinda da Conceição Teixeira Mata - Operational Assistant (DM)*

Support for the coordination and dissemination of course, support for teachers, attendance of students.

- *Bruno Machado - Computer Technician.*

Support for the management of teaching laboratories, computer networks, print services, and maintenance of information on the web, including liaison activities with the Information Technology Division, which also provides human resources whenever necessary.

5.2. Instalações físicas afetas e/ou utilizadas pelo ciclo de estudos (espaços letivos, bibliotecas, laboratórios, salas de computadores, etc.):

As instalações utilizadas pelo ciclo de estudos são comuns aos Departamentos de Informática e de Matemática.

Os Departamentos envolvidos dispõem de quantidade suficiente de: Gabinetes docentes e investigadores; Salas de reuniões com data show de acesso livre; Sala de reuniões departamental com data show e vídeo conferência; Salas de seminários com data show e vídeo conferência; Gabinetes e salas para pessoal administrativo e técnico; Salas servidores e de infraestrutura informática (área total 100m²); Laboratórios de Investigação; Laboratórios de Ensino Informática com data show incluindo laboratórios específicos de redes e projetos de alunos; Salas alunos de utilização livre com computador ou mesa de trabalho; Infraestrutura de controle de acessos suportada em cartões magnéticos; Câmaras de videovigilância.

5.2. Facilities allocated to and/or used by the study programme (teaching spaces, libraries, laboratories, computer rooms, etc.):

The facilities used for the study cycle are common to the Departments of Computer Science and Mathematics. The departments involved have enough of: Teachers and researchers offices; Meeting rooms with free data show access; Departmental meeting rooms with data projector and video conferencing; Seminar rooms with data projectors and video conferencing; Offices and rooms for administrative and technical staff; Rooms for servers and computer infrastructures; Research Laboratories; Computer teaching laboratories with data show (including specialised laboratory for networks and student projects); Rooms for free use of students with computer or desk; Access control infrastructure supported on magnetic cards; Video surveillance cameras.

5.3. Indicação dos principais equipamentos e materiais afetos e/ou utilizados pelo ciclo de estudos (equipamentos didáticos e científicos, materiais e TICs):

Os equipamentos utilizados pelo ciclo de estudos são comuns aos Departamentos de Informática e de Matemática. Os Departamentos envolvidos dispõem de quantidade suficiente de: Equipamento data show (fixo); Equipamento data show (portátil); Computadores (para alunos); Computadores (docentes); Computadores (serviços técnicos e administrativos); Computadores (serviços centrais); Computadores (CPUs) agregados em cluster (alunos); Computadores (CPUs) agregados em cluster (investigação); Infraestruturas de rede sem fios (APs); Infraestruturas de rede com fios (switches); Firewalls; Impressoras para docentes, serviços e pós-graduação; Fotocopiadora; Equipamento de videoconferência; 1 Sistema de som; Televisor plasma; Computador portáteis tipo Tablet; Monitores TFT 22" (para ligação de portáteis), entre outros.

5.3. Indication of the main equipment and materials allocated to and/or used by the study programme (didactic and scientific equipments, materials and ICTs):

The equipment's used for the study cycle are common to the Departments of Computer Science and Mathematics. The departments involved have enough: data show equipment (fixed); data show equipment (portable); Computers (for students); Computers (teachers); Computers (technical and administrative services); Computers (central services); Computers (CPUs) clusters clustered (students); Computers (CPUs) clusters clustered (research); wireless network infrastructure (APs); network infrastructure Wired (switches); firewalls; Printers for teachers, and post-graduation; Photocopying; video conferencing equipment; 1 sound system; TV plasma; Portable Tablet Computer type; TFT monitors 22 " (for laptop connection), among others.

6. Atividades de formação e investigação

Mapa VI - 6.1. Centro(s) de investigação, na área do ciclo de estudos, em que os docentes desenvolvem a su. Atividade científica

6.1. Mapa VI Centro(s) de investigação, na área do ciclo de estudos, em que os docentes desenvolvem a sua atividade científica / Research Centre(s) in the area of the study programme, where the teachers develop their scientific activities

Centro de Investigação / Research Centre	Classificação (FCT) / Mark (FCT)	IES / Institution	Observações / Observations
NOVA Laboratory for Computer Science and Informatics	Excelente / Excellent	FCT NOVA	
Centro de Matemática e Aplicações (CMA)	Muito Bom / Very Good	FCT NOVA	

Perguntas 6.2 e 6.3

6.2. Mapa resumo de publicações científicas do corpo docente do ciclo de estudos, na área predominante do ciclo de estudos, em revistas internacionais com revisão por pares, nos últimos cinco anos (referenciação em formato APA):

<http://a3es.pt/si/iportal.php/cv/scientific-publication/formId/4d8b51d4-ff22-e670-b327-57d124ad688a>

6.3. Lista dos principais projetos e/ou parcerias nacionais e internacionais em que se integram a. Atividades científicas, tecnológicas, culturais e artísticas desenvolvidas na área do ciclo de estudos:

O Dep. de Informática e LINCS NOVA estão envolvidos em muitas redes e projetos de investigação, com financiamento nacional (FCT / MCT, QREN, ADI), da União Europeia (ERC, Strep, IP, NoE, COST, AsiaLink), e das parcerias Carnegie Mellon e UT Austin (que incluem o Departamento de Matemática), e pelo Mestrado Europeu em Lógica Computacional. Estamos atualmente envolvidos em dezenas de projetos financiados, alguns por empresas privadas, no valor de 1,7 M € desde 2015. Vários estudantes de doutoramento são coordenados com instituições como Carnegie Mellon, UT Austin, IMT Lucca, e o Instituto Max Planck. Muitas colaborações nacionais e internacionais, académicos e industriais estão em curso (INRIA, Google, Microsoft Research, IBM, Imperial College, Dresden U, ESA, OutSystems, ZON, RTP, Novabase, Edisoft).

Ambos os departamentos de Informática e Matemática colaboraram com muitas escolas nacionais (IST, U. Lisboa, U. Minho, U. Porto, U. Aveiro, U. Coimbra, U. Évora e muitos outros).

6.3. List of the main projects and/or national and international partnerships, integrating the scientific, technological, cultural and artistic activities developed in the area of the study programme:

The Department of Informatics and NOVA LINCS are involved in many research networks and projects, funded nationally (FCT/MCT, QREN, ADI), or by the European Union (ERC, Strep, IP, NoE, COST, AsiaLink), by the Carnegie Mellon and UT Austin partnerships (also including the Department of Mathematics), and by the European Master in

Computational Logic. We are currently involved in tens of funded projects, some by private firms, worth >1.7M€ since 2015. Several of our PhD students are co-supervised with institutions such as Carnegie Mellon, UT Austin, IMT Lucca, and Max Planck Institute. Many national and international, academic and industrial collaborations are on going (INRIA, Google, Microsoft Research, IBM, Imperial College, Dresden U, ESA, OutSystems, ZON, RTP, Novabase, Edisoft). Both departments of Informatics/CS and Mathematics collaborate with many national schools (IST, U. Lisboa, U. Minho, U. Porto, U. Aveiro, U. Coimbra, U. Évora and many other).

7. Atividades de desenvolvimento tecnológico e artísticas, prestação de serviços à comunidade e formação avançada

7.1. Descreva esta. Atividades e se a sua oferta corresponde às necessidades do mercado, à missão e aos objetivos da instituição:

Várias atividades dos departamentos e unidades de investigação envolvidos, muitas vezes desenvolvidos em colaboração com empresas nacionais e internacionais, conduzem a transferência de tecnologia e de conhecimento, produzindo inovação e criação de valor, em alinhamento com a missão da FCT NOVA. Essas atividades abordam vários aspetos e necessidades da sociedade da informação, incluindo exemplos de segurança na web, televisão imersiva, gestão de emergências, cuidados de saúde e sistemas de saúde pública, vigilância dos oceanos, apoio ao desenvolvimento artístico, deteção de comportamento padrão de utilizadores, análise de rede de energia, e bioinformática.

Estes são campos onde a engenharia e ciência de dados está sendo aproveitada na sociedade. Recentemente, a NOVA assinou um acordo de colaboração na área de Big Data com a Mello Saúde e os departamentos (DI e DM) têm várias colaborações no mesmo tema (p.ex, com a Altran, Deloitte, o Banco de Portugal, IMA, Saude24, IHMT).

7.1. Describe these activities and if they correspond to the market needs and to the mission and objectives of the institution:

Several activities of the involved Departments and research units lead to technology and knowledge transfer, often developed in collaboration with national and international firms, producing innovation and creating value, in alignment with FCT NOVA mission. Activities address multiple aspects and needs of information society, examples include web application security, immersive television, emergency management, healthcare and public health systems, ocean monitoring, artistic development support, user behaviour pattern detection, energy grid analysis, bioinformatics to list just a few.

These are examples of fields where big data science and engineering is being leveraged nowadays in society. As further examples, NOVA has recently signed a collaboration agreement in the area of Big Data with Mello Saúde and DI/DM FCT NOVA has several collaborations in the same theme (e.g., with Altran, Deloitte, Banco de Portugal, IMA, Saude24, IHMT).

8. Enquadramento na rede de formação nacional da área (ensino superior público)

8.1. Avaliação da empregabilidade dos graduados por ciclos de estudos similares com base nos dados do Ministério que tutela o emprego:

É difícil reportar indicadores específicos para a área focada neste curso de mestrado, nem tais dados estão disponíveis nas bases de dados do Ministério da Economia, mas em termos de grandes áreas, é bem sabido que a empregabilidade de profissionais com competência na área de Engenharia Informática é de essencialmente 100%, havendo cada vez mais necessidade de especialistas, o mesmo sucedendo com a Matemática.

No contexto da agenda Portugal Digital “Estratégia e Plano de Ação para a Empregabilidade Digital”, alinhada com as prioridades da Agenda Digital para a Europa e da Estratégia Europa 2020, é reconhecido claramente que se pretende nos próximos anos obter (citamos) um “Acréscimo significativo de recursos humanos com licenciaturas, mestrados e doutoramentos em Ciências da Computação e Engenharia Informática, acompanhando a evolução de competências exigidas pela dinâmica do setor a nível internacional (vide casos do Big Data e Analytics)”.

8.1. Evaluation of the graduates' employability based on Ministry responsible for employment data:

It is difficult to report specific indicators for the area focused in this MSc course, as no such data are available on the of the Ministry of Economy databases, but in terms of broad areas, it is well known that the employability of professionals with expertise in Computer Engineering area is essentially 100%, with an increasing need of specialists, the same applying to mathematics.

In the context of the Agenda Digital Portugal “Strategy and Action Plan for Digital Employability”, in line with the priorities of the Digital Agenda for Europe and the Europe 2020 strategy, it is clearly recognized that in the years ahead it is intended to get (quote) a 'Surcharge significant human resources with undergraduate degrees, masters degrees and doctorates and Computer Science and Computer Engineering, following the evolution of skills required by the dynamics of the international sector (see cases of Big Data and Analytics). ”

8.2. Avaliação da capacidade de atrair estudantes baseada nos dados de acesso (DGES):

Decorre da procura de formação avançada na área geral das TIC que existe uma muito boa capacidade de atração de estudantes. Por exemplo, nos últimos anos, as ofertas de primeiro ciclo e mestrados integrados nas áreas que o curso requiere têm sido totalmente preenchidas na 1ª Fase de acesso.

8.2. Evaluation of the capability to attract students based on access data (DGES):

From the advanced training demand in the general area of ICT arises that there is a very good capacity to attract new students. For example, in recent years, the offers for first cycle and integrated master's degrees in the areas relevant for the course have been completely filled in the 1st round of access.

8.3. Lista de eventuais parcerias com outras instituições da região que lecionam ciclos de estudos similares:

Nesta fase de lançamento, o curso será, do ponto de vista de responsabilidade de ensino, assegurado exclusivamente pelo Departamento de Informática e Departamento de Matemática da FCT NOVA, podendo no futuro procurar-se parcerias com outras instituições de ensino superior, em áreas que venham a ser identificadas. Será colocada prioridade, dadas as características do curso, a parcerias com entidades externas utilizadores do conhecimento, como empresas e instituições públicas que providenciem casos de estudo e test-beds, suportando a oferta de colaborações de R&D e coorientação de teses de mestrado.

8.3. List of eventual partnerships with other institutions in the region teaching similar study programmes:

In this launching stage phase, and from the point of view of teaching responsibility, the course will be of exclusive responsibility of the departments of Informatics/CS and Mathematics of FCT NOVA, although in the future there partnerships might be sought with other institutions of higher education from areas to be then identified. Given the programme characteristics, priority is being given to partnerships with external entities that are users of data-driven knowledge, such as companies and public institutions that provide case studies and test-beds, supporting the supply of R & D collaborations and co-supervision of master theses.

9. Fundamentação do número de créditos ECTS do ciclo de estudos

9.1. Fundamentação do número total de créditos ECTS e da duração do ciclo de estudos, com base no determinado nos artigos 8.º ou 9.º (1.º ciclo), 18.º (2.º ciclo), 19.º (mestrado integrado) e 31.º (3.º ciclo) do Decreto-Lei n.º 74/2006, de 24 de Março:

De acordo com o Artigo 18º do Decreto-Lei no 74/2006, e tratando-se de um ciclo de estudos de 2º ciclo, com 2 anos (4 semestres), foi atribuído ao mesmo um total de 120 ECTS para a obtenção do grau de Mestre. Ao trabalho de dissertação de mestrado é atribuído um total de 60 ECTS.

9.1. Justification of the total number of ECTS credits and of the duration of the study programme, based on articles no.8 or 9 (1st cycle), 18 (2nd cycle), 19 (integrated master) and 31 (3rd cycle) of Decreto-Lei no. 74/2006, March 24th:

According to Article 18º of Decreto-Lei no 74/2006, being the case of an 1st cycle with 2 year full duration (4 terms), we assign it a total of 120 ECTS for obtaining the Master degree. To the master dissertation work is assigned a total of 60 ECTS.

9.2. Metodologia utilizada no cálculo dos créditos ECTS das unidades curriculares:

Para cálculo dos créditos ECTS das unidades curriculares foi utilizado como indicador a equiparação de 1 unidade ECTS a 28 horas de trabalho do estudante. O esforço do estudante nas várias componentes de atividade de cada unidade curricular será continuamente aferido e quando necessário reajustado pelos docentes e comissão científica do curso, de forma informada por inquéritos periódicos na FCT NOVA, desde o início do processo de Bolonha. Na presente proposta foi analisada a creditação das UC, tendo sido preocupação equilibrar o esforço do estudante entre os vários semestres do curso. De acordo com as recomendações gerais da FCT NOVA, todas unidades curriculares são medidas em múltiplos de 3 ECTS, por questões de modularidade e flexibilidade.

9.2. Methodology used for the calculation of the ECTS credits of the curricular units:

Our calculation of ECTS credits of curricular units was based on the correspondence of 1 ECTS unit to 28 working hours. The student work load in the various components of activity of each curricular unit will be continuously monitored and adjusted whenever considered necessary by the academic staff and program scientific committee, informed by periodic student surveys at FCT NOVA, in particular after the installation of the Bologna process. When preparing this proposal, the ECTS for all curricular units was again reconsidered, with a special concern to better balance the student workload along the several terms of the program. According to the general recommendations of FCT NOVA, all curricular units are measured in multiples of 3 ECTS units, for the sake of modularity and flexibility.

9.3. Forma como os docentes foram consultados sobre a metodologia de cálculo do número de créditos ECTS das unidades curriculares:

O esforço do estudante nas várias componentes de atividade de cada unidade curricular do curso de Mestrado em Análise e Engenharia de Big Data (MAEBD) que em alguns casos resultam de adaptação de UC existentes, será continuamente aferido e reajustado pelos docentes sob a coordenação das Comissões Científicas, de forma informada por inquéritos na FCT NOVA, desde o início do processo de Bolonha.

Estes inquéritos periódicos auscultam os estudantes sobre o número de horas que consomem nas várias atividades, e informam o processo de ajuste de créditos ECTS. Durante o processo de elaboração das fichas das UC incluídas nesta proposta, os docentes estiveram novamente ativamente envolvidos e auscultados sobre o método de cálculo das unidades ECTS, tendo também sido promovidas reuniões com os estudantes.

9.3. Process used to consult the teaching staff about the methodology for calculating the number of ECTS credits of the curricular units:

The student load in the various components of activity of each curricular unit of the Master in Big Data Analytics and Engineering (MAEBD), that in some cases resulted from updating UCs existing ones will be continuously monitored and adjusted by the academic staff under the supervision of unit coordinators and the program's scientific committees. This process has also been informed by the periodic student surveys carried out at FCT NOVA. These surveys consult students about their workload in several academic activities and courses, and are taken into account in the ECTS unit calculation processes. During the process of elaboration of this proposal and of the CU descriptions included, the academic staff and steering committees were again actively involved and consulted about the calculation method, and meetings with students were also organized.

10. Comparação com ciclos de estudos de referência no espaço europeu

10.1. Exemplos de ciclos de estudos existentes em instituições de referência do Espaço Europeu de Ensino Superior com duração e estrutura semelhantes à proposta:

A área de Big Data / Data Science / Data Analytics tem recebido um interesse crescente no Ensino Superior, nomeadamente em países europeus e norte-americanos, devido ao enorme impacto da mesma. Tratando-se de uma oferta inovadora, não é ainda possível referir ofertas similares a nível nacional. Com um perfil semelhante ao agora proposto, podem ser referidas várias ofertas a nível europeu e norte-americano, e.g 41 mestrados no Reino Unido, 14 em França, 8 na Alemanha e 6 em Itália, bem como 53 nos Estados Unidos e 5 no Canadá. A título de exemplo, o Mestrado em Data Science do University College London (1º departamento de informática no UK de acordo com a avaliação UK REF 2014) tem uma estrutura com três unidades curriculares obrigatórias e várias opcionais das áreas de Informática e de Matemática, incluindo um projeto/dissertação com peso de 1/3 no curso.

10.1. Examples of study programmes with similar duration and structure offered by reference institutions of the European Higher Education Area:

The area of Big Data / Data Science / Data Analytics is receiving an increased interest in higher education, particularly in European and North American countries, due to its huge impact. Being an innovative offer, it is not yet possible to refer to similar offerings at a national level. With a similar profile to that now proposed, several offers can be listed worldwide, namely 41 master's degrees in the United Kingdom, 14 in France, 8 in Germany and 6 in Italy as well as 53 in the United States and 5 in Canada. For example, the Master in Data Science at University College London (1st computer department in the UK according to the assessment UK REF 2014) has a structure with three compulsory courses and various optional areas of Computing and Mathematics, including a project / dissertation weighing 1/3 on the course.

10.2. Comparação com objetivos de aprendizagem de ciclos de estudos análogos existentes em instituições de referência do Espaço Europeu de Ensino Superior:

Os objetivos de aprendizagem dos vários cursos referidos variam entre aquelas de cariz mais fundamental e os que privilegiam as tecnologias e ferramentas de aplicação. As matérias que são lecionadas estão em linha com as do mestrado proposto. A título de exemplo, o Mestrado em Data Science do University College London tem uma estrutura com três unidades curriculares obrigatórias e várias opcionais, das áreas de informática e matemática (e.g. arquitetura de sistemas computacionais para grandes volumes de dados, modelação e visualização de dados multimédia, aprendizagem automática e datamining, inferência estatística e otimização, modelos de risco e decisão). O curso inclui ainda um projeto/dissertação com peso de 1/3 no curso. Neste curso, como na generalidade dos cursos desta área, é realçada a importância da utilização de casos de estudo reais pelo que utilizam parcerias com instituições, nomeadamente no projeto/dissertação obrigatória. Conclui-se que a oferta aqui proposta está alinhada com a prática internacional nomeadamente no espaço europeu.

10.2. Comparison with the intended learning outcomes of similar study programmes offered by reference institutions of the European Higher Education Area:

The learning objectives of several such courses range from those more fundamental nature and those who prefer the technology and application tools. The topics that are taught are in line with the proposed master course. For example, the Master in Data Science at University College London has a structure with three compulsory courses and various optional courses in areas of computer science and mathematics (e.g. architecture of computer systems for large volumes of data, modelling and multimedia data visualisation, machine learning and data mining, statistical inference and optimization, decision and risk models). The programme also includes a project / dissertation weighing 1/3 of the whole programme. In this programme, as in most programmes in this area, the importance of using real case studies is highlighted by using partnerships with institutions, namely in the mandatory project / dissertation mandatory. In conclusion, the offer proposed here is well in line with international practice particularly in Europe.

11. Estágios e/ou Formação em Serviço

11.1. e 11.2 Locais de estágio e/ou formação em serviço (quando aplicável)

Mapa VII - Protocolos de Cooperação

Mapa VII - Protocolos de Cooperação

11.1.1. Entidade onde os estudantes completam a sua formação:

<sem resposta>

11.1.2. Protocolo (PDF, máx. 150kB):

<sem resposta>

Mapa VIII. Plano de distribuição dos estudantes

11.2. Mapa VIII. Plano de distribuição dos estudantes pelos locais de estágio e/ou formação em serviço demonstrando a adequação dos recursos disponíveis.(PDF, máx. 100kB).

<sem resposta>

11.3. Recursos próprios da Instituição para acompanhamento efetivo dos seus estudantes nos estágios e/ou formação em serviço.

11.3. Recursos próprios da Instituição para o acompanhamento efetivo dos seus estudantes nos estágios e/ou formação em serviço:

N/A

11.3. Resources of the Institution to effectively follow its students during the in-service training periods:

N/A

11.4. Orientadores cooperantes

Mapa IX. Normas para a avaliação e seleção dos elementos das instituições de estágio e/ou formação em serviço responsáveis por acompanhar os estudantes

11.4.1 Mapa IX. Mecanismos de avaliação e seleção dos orientadores cooperantes de estágio e/ou formação em serviço, negociados entre a Instituição de ensino superior e as instituições de estágio e/ou formação em serviço (PDF, máx. 100kB):

<sem resposta>

Mapa X. Orientadores cooperantes de estágio e/ou formação em serviço (obrigatório para ciclo de estudos com estágio obrigatório por Lei)

11.4.2. Mapa X. Orientadores cooperantes de estágio e/ou formação em serviço (obrigatório para ciclo de estudos com estágio obrigatório por Lei) / External supervisors responsible for following the students' activities (mandatory for study programmes with in-service training mandatory by law)

Nome / Name	Instituição ou estabelecimento a que pertence / Institution	Categoria Profissional / Professional Title	Habilitação Profissional (1)/ Professional qualifications (1)	Nº de anos de serviço / Nº of working years
----------------	--	--	--	--

<sem resposta>

12. Análise SWOT do ciclo de estudos

12.1. Pontos fortes:

O curso oferece uma sólida formação científica e um currículo atualizado, diversificado e abrangente, numa área de enorme atualidade e impacto para a qual existe lacuna na oferta formativa em Portugal. O curso tem elevada empregabilidade, todos os docentes (tempo integral) são membros de centros de investigação com classificação "Excelente" e "Muito Bom", e existem excelentes infraestruturas de ensino e investigação. O equilíbrio entre as áreas

de Informática e Matemática assim como os temas cobertos fornecem ao diplomado uma sólida preparação prática e teórica, flexibilidade de perfil, e competências para abordar problemas concretos na área. Os diplomados adquirem as competências adequadas à sua inserção no mercado de trabalho como profissionais altamente qualificados na análise e engenharia de dados de grandes dimensões. Boa ligação com empresas/empregadores, incluindo projetos de R&D na área, facilitam o envolvimento de estudantes em estágios e dissertações em ambiente académico-empresarial.

12.1. Strengths:

The programme provides a solid scientific background and an updated, diversified and comprehensive curriculum in an area of great relevance and impact for which there is gap in the training offer in Portugal. Moreover, it offers high employability, all lecturers are full-time members of research centres rated as "Excellent" and "Very Good". Excellent teaching and research infrastructure.

The balance between the areas of Informatics / CS and Mathematics as well as the topics covered provide the graduate a solid theoretical and practical preparation, profile flexibility, and expertise to address specific problems in the area. The graduates acquire the skills appropriate to entering the labour market as highly qualified professionals in the analysis and engineering of large data volumes. Good connection with companies / employers, including R & D projects in the area, facilitate the involvement of students in "internships" and dissertations in an academic-business environment.

12.2. Pontos fracos:

O MAEBD poderá ter alguma dificuldade na captação de candidatos oriundos de Mestrados Integrados, que é o caso dos perfis de muitas formações superiores em Engenharia. No entanto, este aspeto poderá ser mitigado através de ações de recrutamento e disseminação adequadas.

12.2. Weaknesses:

The MAEBD may have some difficulty in attracting candidates coming from Integrated Masters, which is the profile of many higher education programmes in Engineering. However, this aspect can be mitigated by means of adequate recruitment and dissemination actions.

12.3. Oportunidades:

Vive-se atualmente um paradigma derivado das enormes quantidades de informação em todos os domínios, existindo falta de profissionais com formação específica e de nível adequado para manipular e analisar enormes quantidades de dados no contexto em que ocorrem. Estima-se (UE) que até 2020 a necessidade profissionais na área cresça mais de 160% na Europa. O Mestrado que aqui se propõe vem contribuir para responder a esta lacuna.

Muito possivelmente outras candidaturas a ciclos de curso nesta temática irão ser propostas num futuro próximo. Não há vantagem em protelar a criação deste curso, até porque estão reunidas todas as condições (qualidade e experiência do corpo docente, meios laboratoriais, parcerias) para o seu arranque no ano letivo 2017/2018.

12.3. Opportunities:

There is a current paradigm derived from the enormous amounts of information in all areas, and there is lack of professionals with specific training and adequate to handle and analyse huge amounts of data in the context in which they occur. It is estimated (EU) that by 2020 the need for professionals in the area will grow more than 160% in Europe. The Master being proposed will help to address this gap.

Quite possibly, other master programmes on this topic will be proposed in the near future. There is no advantage in delaying the creation of this programme, since all the conditions are well met (quality and experience of the teaching staff, laboratory facilities, partnerships) for its start in the academic year 2017/2018.

12.4. Constrangimentos:

A atual situação económica nacional e internacional, a par da diminuição da taxa de natalidade no país, reflete-se na restrição à procura de formação a nível do ensino superior. Limitações orçamentais conduzem a limitações de recursos humanos e materiais, e impedem certas ações que permitiriam assegurar de forma mais sustentada a execução do curso.

12.4. Threats:

The current national and international economic situation, along with the falling birth rate in the country, is reflected in the restriction on the demand for training at higher education level. Budgetary constraints lead to limitations of human and material resources, and prevent certain actions that could ensure more sustained manner the implementation of the course.

12.5. CONCLUSÕES:

O Mestrado em Big Data Analytics e Engenharia (MAEBD) visa a educação de analistas, líderes de desenvolvimento do projeto e especialistas em inovação no campo emergente da Ciência de Dados e Engenharia, que hoje em dia vem criando novas oportunidades para a gestão de processos e otimização de produtos em todos os campos de atividade. Com base numa uma colaboração equilibrada entre os departamentos de Informática e de Matemática da FCT NOVA, o Mestrado em Análise e Engenharia de Big Data (MAEBD) aqui proposto reúne competências computacionais e matemáticas fundamentais na área em rápida expansão de Big Data e cobrindo a Ciência de Dados, Data Warehousing e Data Mining, bem como Engenharia de Dados e tópicos de infraestruturas de computação. A estrutura global do currículo do MAEBD acomoda uma variedade de perfis de pós-graduação entre Ciência de Dados e Engenharia de Dados, garantindo ao mesmo tempo um forte núcleo de princípios fundamentais de ordem científica, de engenharia e de tecnologia.

O programa fornece uma base sólida através de um currículo muito atualizado numa área técnica e profissional de extrema atualidade e impacto, com equilíbrio de cursos de Ciência da Computação, Informática e Matemática, seminários pelos parceiros externos, empreendedorismo e cursos opcionais em outros temas complementares, nomeadamente uma unidade curricular em qualquer das disciplinas científicas oferecidas pela FCT NOVA, acomodando assim, um nível adequado de interdisciplinaridade, adequado para um campo amplamente apoiado como Big Science dados e Engenharia.

Existe atualmente uma falta de ofertas semelhantes em Portugal na área de Big Data, enquanto há uma alta procura por profissionais qualificados a nível de Mestrado em todo o mundo. Os departamentos proponentes estão bem posicionados e qualificados para iniciar e desenvolver o programa. Todos os membros do corpo docente são membros integrados das unidades de investigação envolvidas, classificadas como "Excelente" (NOVA LINCS) e "Muito Bom" (CMA) pela mais recente avaliação da FC&T, e que incluem grupos com pesquisa ativa em temas relacionados de Big Data. São disponibilizadas excelentes infraestruturas de ensino e pesquisa.

Pode-se assim aproveitar as muitas conexões existentes e parcerias com empresas e empregadores, incluindo projetos de I&D neste domínio, o que facilitará a integração dos estudantes em estágios e o desenvolvimento de teses de mestrado no âmbito de colaborações entre professores e empresas externas ou instituições públicas. A empregabilidade dos profissionais da área é extremamente alta e crescente em todo o mundo.

O MAEBD que se propõe oferece uma resposta eficaz e atempada para a necessidade crescente de profissionais sofisticados em Análise e Engenharia de Big Data, compatível com os mais altos padrões encontrados em programas similares na Europa e nos EUA.

12.5. CONCLUSIONS:

The Master in Big Data Analytics and Engineering (MAEBD) targets the education of analysts, project development leaders and innovation specialists in the emergent field of Data Science and Engineering, which nowadays creates novel opportunities for process management and product optimisation in every field of activity. Building on an evenly balanced collaboration between the departments of Computer Science and Mathematics of FCT NOVA, the Master in Analysis and Engineering of Big Data (MAEBD) here proposed brings together fundamental computational and mathematical skills in the rapidly expanding area of Big Data and covers Data Science, Data Warehousing and Data Mining as well as Data Engineering and Computing Infrastructures topics. The global structure of the MAEBD curriculum accommodates a variety of graduate profiles between Data Science and Data Engineering, while ensuring a strong core on key scientific, engineering and technology principles.

The programme provides a solid background through a very up-to-date curriculum in a technical and professional area of extreme timeliness and impact, and balances the presence of Computer Science, Informatics, and Mathematics courses, seminars by external partners, entrepreneurship, and elective courses in other complementary themes, namely one course in any of the scientific disciplines offered by FCT NOVA, thus accommodating a suitable level of interdisciplinary, adequate for a widely supporting field as Big Data Science and Engineering.

There is presently a lack of similar offers in Portugal in the area of Big Data, while there is a very high demand for qualified professionals at the Master level worldwide. The proponent academic Departments are well positioned and qualified to launch and develop the program. All members of the teaching staff integrated members of the involved research units, graded "Excellent" (NOVA LINCS) and "Very Good" (CMA) by the most recent FC&T evaluation exercise, and which host groups actively research on Big Data related topics. We also provide excellent teaching and research infrastructures.

We will leverage on the many existing connections and partnerships with companies and employers, including R&D projects in the field, which will facilitate the integration of students in internships and the development of master theses in the context of collaborations between faculty and external firms or public institutions. The employability of professionals in the area is extremely high and increasing worldwide.

The proposed MAEBD offers an effective and timely response to the increasing need for sophisticated professionals in Analysis and Engineering of Big Data, compatible with the highest standards found in similar programmes in Europe and the USA.