

CURSO	Licenciatura em Informática		
UNIDADE CURRICULAR	Modelos e Simulações Computacionais	Obrigatória	X
		Opcional	
ÁREA CIENTÍFICA	Engenharia Informática		

Ano: 3º	Semestre: 2º	ECTS: 6	Horas de Contacto teórico práticas: 60h
---------	--------------	---------	---

OBJETIVOS DA APRENDIZAGEM

Esta unidade curricular tem como finalidade proporcionar aos alunos uma iniciação às técnicas de aprendizagem automática (machine learning), usando a linguagem R. Para concluir com sucesso esta unidade curricular, os alunos deverão demonstrar possuir os seguintes conhecimentos e capacidades:

1. Compreender os conceitos fundamentais sobre Big Data e Machine Learning;
2. Compreender a estrutura, instalar um IDE para a linguagem e saber utilizar as suas funcionalidades básicas;
3. Compreender e saber utilizar algoritmos de supervised learning;
4. Compreender e saber utilizar algoritmos de unsupervised learning.

PROGRAMA

1. Introdução à aprendizagem automática (Machine Learning)
 - 1.1. Origem e evolução do conceito
 - 1.2. Big Data e Machine Learning
 - 1.3. Aplicações atuais
 - 1.4. Supervised versus Unsupervised learning
 - 1.5. Classificação de dados de input
 - 1.6. Modelos preditivos e descritivos
2. Introdução à linguagem R
 - 2.1. Origens e desenvolvimento
 - 2.2. Ambientes de desenvolvimento
 - 2.3. Estrutura da linguagem
 - 2.4. Suporte de funções científicas
 - 2.5. Exercícios práticos de familiarização com a linguagem
3. Técnicas de classificação com a técnica nearest neighbor
 - 3.1. O algoritmo K-NN
 - 3.2. Preparação de dados para K-NN
 - 3.3. Projeto prático de aplicação de K-NN
4. Aprendizagem probabilística
 - 4.1. Conceito de probabilidade

- 4.2. Conceito de independência e de probabilidade condicionada
- 4.3. Fórmula de Bayes
- 4.4. O algoritmo naïve Bayes
- 4.5. Projeto prático de aplicação do algoritmo naïve Bayes
- 5. Árvores de decisão e regras de classificação
 - 5.1. Conceito e utilização de árvores de decisão
 - 5.2. Algoritmo C5.0
 - 5.3. Projeto prático utilizando C5.0
 - 5.4. Conceito e utilização de regras de classificação
 - 5.5. Algoritmo 1R
 - 5.6. Algoritmo RIPPER
 - 5.7. Projeto prático utilizando regras de classificação
- 6. Métodos de regressão
 - 6.1. Conceito e modelos de regressão
 - 6.2. Regressão linear simples
 - 6.3. Estimacão dos mínimos quadrados
 - 6.4. Correlacão
 - 6.5. Regressão linear múltipla
 - 6.6. Projeto prático usando regressão linear
 - 6.7. Árvores de regressão
 - 6.8. Projeto prático usando árvores de regressão
- 7. Redes neuronais e máquinas de suporte vetorial
 - 7.1. Conceito e utilização de redes neuronais
 - 7.2. Projeto prático com redes neuronais
 - 7.3. Conceito e utilização de máquinas de suporte vetorial (SVM-Support Vector Machines)
 - 7.4. Projeto prático com SVM
- 8. Regras de associaão (association rules)
 - 8.1. Descoberta de padrões
 - 8.2. Algoritmo Apriori
 - 8.3. Projeto prático
- 9. Clustering
 - 9.1. Conceito de clustering
 - 9.2. Algoritmo de clustering k-means
 - 9.3. Projeto com utilização de algoritmo de clustering
- 10. Projeto prático final

DEMONSTRAÇÃO DE COERÊNCIA ENTRE CONTEÚDOS PROGRAMÁTICOS E RESULTADOS DA APRENDIZAGEM

O objetivo 1 é concretizado através dos pontos 1.1 a 1.6. Os pontos 2.1 a 2.5 permitem realizar o objetivo 2. O objetivo 3 é concretizado através dos pontos 3.1 a 3.3, 4.1 a 4.5, 5.1 a 5.7, 6.1 a

6.8 e 7.1 a 7.4. Os pontos 8.1 a 8.3 e 9.1 a 9.3 permitem concretizar o objetivo 4. O ponto 10 é concebido para consolidar os objetivos 3 e 4.

METODOLOGIA DE ENSINO E AVALIAÇÃO

Estão previstas 60 horas de contato. O tempo total de trabalho do aluno corresponde a 162 horas. A metodologia baseia-se em exposições teóricas para apresentação dos conceitos científicos seguidas de aplicações práticas de utilização dos diversos algoritmos de aprendizagem automática.

De acordo com o Regulamento de Funcionamento do ISTECS a avaliação é efetuada através de um exame escrito individual e obrigatório. Na classificação final, poderão ser considerados elementos de avaliação contínua, tais como testes, trabalhos individuais ou em grupo, assim como a participação nas aulas presenciais e em recursos de aprendizagem proporcionados por sistemas de e-learning.

DEMONSTRAÇÃO DE COERÊNCIA ENTRE METODOLOGIAS DE ENSINO E RESULTADOS DE APRENDIZAGEM

A participação dos alunos na implementação dos algoritmos utilizando software permite, não apenas a compreensão teórica dos conceitos como a capacidade para os aplicar na prática, alcançando assim os resultados de aprendizagem definidos como objetivos desta unidade curricular.

BIBLIOGRAFIA

Fundamental:

Lantz, Brett. Machine Learning with R Second Edition (2015) Packt Publishing
Lesmeister, Cory. Mastering Machine Learning with R (2015). Packt Publishing

Complementar:

Harrington, Peter. Machine Learning in Action (2012). Manning
Yu-Wei, Chiu Machine Learning with R Cookbook (2015). Packt Publishing

INTERNET:

Acesso a publicações da especialidade, gratuitamente, através da rede SPRINGER:
<https://link.springer.com/>