Programa Analítico de Disciplina

EAM 630 - Aprendizado de Máquina: Princípios e Aplicações

Departamento de Engenharia Civil - Centro de Ciências Exatas e Tecnológicas	
Catálogo: 2024	
Número de créditos: 4 Carga horária semestral: 60h Carga horária semanal teórica: 1h Carga horária semanal prática: 3h	Semestres: I

Ementa
Introdução ao Aprendizado de Máquinas Processos de mineração de dados na descoberta de conhecimento Manipulação e visualização de dados por meio de liguanges de programação Avaliação de Desempenho do conhecimento obtido Análise de Agrupamento Análise Preditiva

Conteúdo

Unidade	T	P	To
1. **Introdução ao Aprendizado de Máquinas** 1. Abordagens do Aprendizado de Máquinas e Mineração de Dados aplicados á solução de problemas 2. Definições relacionado ao Dado, à informação e ao conhecimento obtido 3. Tipos de Dados (Estruturados, não estruturados, convenções) 4. Etapas de um processo de descoberta de conhecimento (Knowledge Discovery in Databases – KDD) 5. Tarefas de Mineração de Dados	2h	0h	2h
2. **Processos de mineração de dados na descoberta de conhecimento** 1. Aquisição, seleção, limpeza, normalização e transformação de dados	1h	3h	4h
3. **Manipulação e visualização de dados por meio de liguanges de programação** 1. Introdução à linguagem de programação - sintaxe e comandos 2. Análise Exploratória de Dados 1. Conceitos de Estatística Descritiva 2. Pré-Processamento de Dados	1h	3h	4h

A autenticidade deste documento pode ser conferida no site https://siadoc.ufv.br/validar-documento com o código: 6CQO.HQ13.PILY

4. **Avaliação de Desempenho do conhecimento obtido** 1. Avaliação para Dados Contínuos 1. Erro Médio Absoluto 2. RMSE 3. R^2 2. Avaliação para Dados Discretos 1. Matrix de Confusão 2. Acurácia 3. Sensibilidade e Especificidade 4. Curva ROC (Receiver Operating Characteristic) 3. Índice Kappa 4. Técnicas Probabilísticas 1. Validação Cruzada 2. Amostragem Bootstrap	1h	3h	4h
5. **Análise de Agrupamento** 1. Agrupamento Hierárquico 2. Agrupamento por Partição 1. PAM (Partition Around Medoids) 2. K-médias (K-means)	1h	6h	7h
6. **Análise Preditiva** 1. Tarefa de Classificação 1. K-vizinhos mais próximos 2. Árvores de Decisão 3. Random Forest 4. Naïve Bayes 5. Rede Neural Artificial 6. Suporte Vector Machine 2. Regressão Linear 3. Análise de Séries Temporais 4. Classificação de uso e cobertura da Terra	9h	27h	36h
7. **Implementação de scripts em Sistema de Informação Geográfica (SIG) para descoberta de conhecimento em dados Geoespaciais** 1. Integração do scritp R no Sistema de Informação Geográfica (Qgis)	0h	3h	3h
Total	**15h**	**45h**	**60h**

Teórica (T); Prática (P); Total (To);

A autenticidade deste documento pode ser conferida no site https://siadoc.ufv.br/validar-documento com o código: 6CQO.HQ13.PILY

EAM 630 - Aprendizado de Máquina: Princípios e Aplicações

Bibliografias básicas	
Descrição	**Exemplares**
Han, J.; Kamber, M.; Pei, J. Data mining : concepts and techniques. 3rd ed., Burlington, MA : Elsevier, 2012.	3
WITTEN, I. H.; FRANK, Eibe; HALL, Mark A. Data mining: practical machine learning tools and techniques. 3rd ed. Burlington, MA: Morgan Kaufmann Publishers, c2011. xxxiii, 629 p. ISBN 9780123748560.	6
Tan,Pang-Ning; Steinbach ,Michael; Karpatne, Anuj; Kumar, Vipin; Introduction to Data Mining (Second Edition). Disponível em https://www-users.cs.umn.edu/~kumar001/dmbook/index.php	0
OLIVEIRA, Paulo Felipe de; GUERRA, Saulo; MCDONNEL, Robert. Ciência de Dados com R – Introdução. Brasília: Editora IBPAD, 2018.	0
Ramasubramanian, Karthik ; Singh, Abhishek. Machine Learning Using R - With Time Series and Industry-Based - Use Cases in R. Apress, Berkeley, CA, https://doi.org/10.1007/978-1-4842-4215-5, 2019, Disponível em https://link.springer.com/content/pdf/10.1007%2F978-1-4842-4215-5.pdf	0
Tsai, K.-T. Machine Learning for Knowledge Discovery with R: Methodologies for Modeling, Inference, and Prediction (1st ed.). Chapman and Hall/CRC. https://doi.org/10.1201/9781003205685. 2021.	0

Bibliografias complementares	
Descrição	**Exemplares**
About Quick R. Disponível em https://www.statmethods.net/index.html	0
Ciência de Dados em R, Curso-R, (e-book), Disponível em https://livro.curso-r.com/index.html	0
Golschimidt, R.; Passos, E. Data Mining; Um guia prático. São Paulo: Elsevier; 2005.	0
Kumar, V.; Steinbach, M.; Tan, P. Introdução ao Data Mining; Mineração de dados. Ed. Ciência Moderna, 2009.	0
Landeiro et al. Introdução ao uso do programa R. 2011, Disponível em https://cran.r-project.org/doc/contrib/Landeiro-Introducao.pdf	0
M. Kuhn and K. Johnson. Applied Predictive Modeling, https://doi.org/10.1007/978-1-4614-6849-3, Springer, New York, NY, 2013.	0
Pinheiro, C. A. R. Inteligência analítica : mineração de dados e descoberta de conhecimento. Rio de Janeiro : Ciência Moderna, 2008.	0
R Bloggers - R news and tutorials contributed by hundreds of R bloggers. https://www.r-bloggers.com/	0
Silva, L.A.; Peres, S.M.; Boscarioli, C. Introdução à Mineração de Dados – Com Aplicações em R. Rio de Janeiro, Ed. Elsevier, 2016.	0

A autenticidade deste documento pode ser conferida no site https://siadoc.ufv.br/validar-documento com o código: 6CQO.HQ13.PILY

SPRINGERLINK (ONLINE SERVICE). Data Mining and Decision Support: Integration and Collaboration. 1st ed. 2003. XXIII, 275 p (The Springer International Series in Engineering and Computer Science, 0893-3405 ; 745). ISBN 9781461502869.	0
Tan, P.; Steimbach, M.; Kumatr, V. Introduction to data mining. Rio de Janeiro : Ciência Moderna, 2009.	0
Witten, I. H.; Frank, E.; Hall, M. A. Data mining : practical machine learning tools and techniques. 3rd ed. Burlington, MA : Morgan Kaufmann, 2011.	0

Pontos de controle		
Campo	**Anterior**	**Atual**
Conteúdo	Há alterações no conteúdo da disciplina	

A autenticidade deste documento pode ser conferida no site https://siadoc.ufv.br/validar-documento com o código: 6CQO.HQ13.PILY

Syllabus

EAM 630 - Machine Learning: Principles and applications

Departamento de Engenharia Civil - Centro de Ciências Exatas e Tecnológicas	
Catalog: 2024	
Number of credits: 4 Total hours: 60h Weekly workload - Theoretical: 1h Weekly workload - Practical: 3h	Period: I

Content

Introduction to Machine Learning
Data mining in Knowledge dicovery process
Data manipulation and visualization through programming languages
Performance Assessment of the knowledge
Cluster Analysis
Predictive Analytics

Course program

Unit	T	P	To
1. **Introduction to Machine Learning** 1. Machine Learning and Data Mining approaches applied to problem solving 2. Definitions related to Data, information and knowledge obtained 3. Data Types (Structured, unstructured, conventions) 4. Knowledge Discovery in Databases – KDD 5. Data Mining Tasks	2h	0h	2h
2. **Data mining in Knowledge dicovery process** 1. Data acquisition, selection, cleaning, normalization and transformation	1h	3h	4h
3. **Data manipulation and visualization through programming languages** 1. Introduction to programming language - syntax and commands 2. Exploratory Data Analysis 1. Descriptive Statistics 2. Data Pre-Processing	1h	3h	4h
4. **Performance Assessment of the knowledge** 1. Assessment for Continuous Data 1. Mean Absolute Error 2. RMSE 3. R^2 2. Evaluation for Discrete Data 1. Confusion Matrix 2. Accuracy 3. Sensitivity and Specificity 4. ROC curve (Receiver Operating Characteristic)	1h	3h	4h

A autenticidade deste documento pode ser conferida no site https://siadoc.ufv.br/validar-documento com o código: 6CQO.HQ13.PILY

3. Kappa Index 4. Probabilistic Techniques 1. Cross-Validation 2. Bootstrap Sampling			
5. **Cluster Analysis** 1. Hierarchical Clustering 2. Partitioning Clustering 1. PAM (Partition Around Medoids) 2. K-means	1h	6h	7h
6. **Predictive Analytics** 1. Classification Task 1. K-Nearest Neighbors (KNN) 2. Decision Trees 3. Random Forest 4. Naïve Bayes 5. Artificial Neural Network (ANN) 6. Support Vector Machine (SVM) 2. Linear Regression 3. Time Series Analysis 4. Land Use and Land Cover Classification	9h	27h	36h
7. **Implementation of scripts in Geographic Information System (GIS) for Knowledge Discovery in Spatial Databases** 1. Implementation of scripts in R for the Geographic Information System (QGIS)	0h	3h	3h
Total	**15h**	**45h**	**60h**

Theoretical (T); Practical (P); Total (To);

A autenticidade deste documento pode ser conferida no site https://siadoc.ufv.br/validar-documento com o código: 6CQO.HQ13.PILY

EAM 630 - Machine Learning: Principles and applications

Fundamental references	
Description	**Copies**
Han, J.; Kamber, M.; Pei, J. Data mining : concepts and techniques. 3rd ed., Burlington, MA : Elsevier, 2012.	3
WITTEN, I. H.; FRANK, Eibe; HALL, Mark A. Data mining: practical machine learning tools and techniques. 3rd ed. Burlington, MA: Morgan Kaufmann Publishers, c2011. xxxiii, 629 p. ISBN 9780123748560.	6
Tan,Pang-Ning; Steinbach ,Michael; Karpatne, Anuj; Kumar, Vipin; Introduction to Data Mining (Second Edition). Disponível em https://www-users.cs.umn.edu/~kumar001/dmbook/index.php	0
OLIVEIRA, Paulo Felipe de; GUERRA, Saulo; MCDONNEL, Robert. Ciência de Dados com R – Introdução. Brasília: Editora IBPAD, 2018.	0
Ramasubramanian, Karthik ; Singh, Abhishek. Machine Learning Using R - With Time Series and Industry-Based - Use Cases in R. Apress, Berkeley, CA, https://doi.org/10.1007/978-1-4842-4215-5, 2019, Disponível em https://link.springer.com/content/pdf/10.1007%2F978-1-4842-4215-5.pdf	0
Tsai, K.-T. Machine Learning for Knowledge Discovery with R: Methodologies for Modeling, Inference, and Prediction (1st ed.). Chapman and Hall/CRC. https://doi.org/10.1201/9781003205685. 2021.	0

Complementary references	
Description	**Copies**
About Quick R. Disponível em https://www.statmethods.net/index.html	0
Ciência de Dados em R, Curso-R, (e-book), Disponível em https://livro.curso-r.com/index.html	0
Golschimidt, R.; Passos, E. Data Mining; Um guia prático. São Paulo: Elsevier; 2005.	0
Kumar, V.; Steinbach, M.; Tan, P. Introdução ao Data Mining; Mineração de dados. Ed. Ciência Moderna, 2009.	0
Landeiro et al. Introdução ao uso do programa R. 2011, Disponível em https://cran.r-project.org/doc/contrib/Landeiro-Introducao.pdf	0
M. Kuhn and K. Johnson. Applied Predictive Modeling, https://doi.org/10.1007/978-1-4614-6849-3, Springer, New York, NY, 2013.	0
Pinheiro, C. A. R. Inteligência analítica : mineração de dados e descoberta de conhecimento. Rio de Janeiro : Ciência Moderna, 2008.	0
R Bloggers - R news and tutorials contributed by hundreds of R bloggers. https://www.r-bloggers.com/	0
Silva, L.A.; Peres, S.M.; Boscarioli, C. Introdução à Mineração de Dados – Com Aplicações em R. Rio de Janeiro, Ed. Elsevier, 2016.	0
SPRINGERLINK (ONLINE SERVICE). Data Mining and Decision Support: Integration and	0

A autenticidade deste documento pode ser conferida no site https://siadoc.ufv.br/validar-documento com o código: 6CQO.HQ13.PILY

Collaboration. 1st ed. 2003. XXIII, 275 p (The Springer International Series in Engineering and Computer Science, 0893-3405 ; 745). ISBN 9781461502869.	
Tan, P.; Steimbach, M.; Kumatr, V. Introduction to data mining. Rio de Janeiro : Ciência Moderna, 2009.	0
Witten, I. H.; Frank, E.; Hall, M. A. Data mining : practical machine learning tools and techniques. 3rd ed. Burlington, MA : Morgan Kaufmann, 2011.	0

Pontos de controle		
Campo	**Anterior**	**Atual**
Conteúdo	Há alterações no conteúdo da disciplina	

A autenticidade deste documento pode ser conferida no site https://siadoc.ufv.br/validar-documento com o código: 6CQO.HQ13.PILY