|  |  |  |  |
| --- | --- | --- | --- |
| **Faculdade de Computação e Informática** | | | |
| **Tecnologia em Ciência de Dados** | | **Etapa: 3ª Etapa** | |
| **Componente curricular: Aprendizado de Máquina I** | | **Carga horária: EAD ( ) Teórica ( ) Prática** | |
|  | |  | |
| **Ementa:** *Conceito de aprendizado de máquina. Conceito de aprendizado supervisionado e não supervisionado. Estudo e aplicação dos principais modelos de aprendizado supervisionado para classificação e regressão de dados. Avaliação dos modelos de classificação por diferentes métricas.* | | | |
| **Objetivos** | | | |
| **Objetivos Conceituais** | **Objetivos Procedimentais e**  **Habilidades** | | **Objetivos Atitudinais e Valores** |
| * Conhecer os principais paradigmas do Aprendizado de Máquina. * Ter contato com aplicações da Aprendizagem de Máquina a problemas práticos. * Conhecer alguns dos principais frameworks de Aprendizagem de Máquina. | * Desenvolver a capacidade de   reconhecer oportunidades de  aplicação do Aprendizado de Máquina à problemas de pesquisa e desenvolvimento.   * Ser capaz de modelar um   através de métodos Aprendizado de Máquina para sua resolução.   * Estar apto a implementar   protótipos de soluções  de Aprendizado de Máquina. | | • Valorizar a interdisciplinaridade  do conhecimento científico.   * Estar atento para as tecnologias   de ponta e às oportunidades de  inovação.   * Estar atento para identificar   oportunidades de resolução de  problemas por meio do  Aprendizado de Máquina.   * Perceber o potencial de desenvolvimento de novos negócios a partir do Aprendizado de Máquina. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Conteúdo Programático** | | | | | | |
| **ASSUNTO** | | **LIVRO** | **PÁG / CAP** | | **BIBLIOTECA** | |
| **Aula 1**  **Introdução ao Aprendizado de Máquina**   * O que é o Aprendizado de Máquina e o novo paradigma de programação que ele traz * A importância e as aplicações do Aprendizado de Máquina * Alguns conceitos e termos básicos que envolvem o Aprendizado de Máquina | | * Larose, Chantal D.; Larose, Daniel T. **Data Science Using Python and R** Hoboken: Wiley, c2019. E-book (259 p.) (Wiley Series on Methods and Applications in Data Mining Ser.). ISBN 9781119526834 (electronic bk.). online em: **Biblioteca do Mackenzie.** * Brown, Sata. **Machine learning, explained**. April, 2021. Disponível em: <https://mitsloan.mit.edu/ideas-made-to-matter/machine-learning-explained>. Acesso em: 04 de Março de 2022. | *Cap1, 2*  *-* | | *Online*  *Online* | |
| **Aula 2**  **Aprendizado Supervisionado e Regressão Linear**   * O Aprendizado Supervisionado e as tarefas de Regressão e Classificação * Como aplicar modelos de Regressão Linear Simples e Múltipla na Predição de Quantidades * Como avaliar a eficiência dos modelos de Regressão Linear | * Jake VanderPlas. **Python Data Science Handbook**. O'Reilly Media, Inc. (2016). ISBN: 9781491912058. Disponível em: <https://jakevdp.github.io/PythonDataScienceHandbook/> Acesso: 04 de Março de 2022. * Kotu, Vijay; Deshpande, Balachandre **Data Science: concepts and practice**. 2nd ed. Cambridge, [England]: Morgan Kaufmann, c2019. E-book (570 p.) ISBN 9780128147627 (electronic bk.). online em: **Biblioteca do Mackenzie.**. | | *Cap 5:*  *What Is Machine Learning?*  *In Depth: Linear Regression*  *Cap 5, 5.1 Regression Methods* | | *Online*  *Online* | |
| **Aula 3**  **Classificação: Regressão Logística**   * O Aprendizado Supervisionado para Classificação * A Métrica de Acuracidade * Dilema Bias-Variância | * Kotu, Vijay; Deshpande, Balachandre **Data Science: concepts and practice**. 2nd ed. Cambridge, [England]: Morgan Kaufmann, c2019. E-book (570 p.) ISBN 9780128147627 (electronic bk.). online em: **Biblioteca do Mackenzie.** * Larose, Chantal D.; Larose, Daniel T. **Data Science Using Python and R** Hoboken: Wiley, c2019. E-book (259 p.) (Wiley Series on Methods and Applications in Data Mining Ser.). ISBN 9781119526834 (electronic bk.). online em: **Biblioteca do Mackenzie.** | | *Cap 5, 5.2 Logistic Regression*  *Cap 13.* | | *Online*  *Online* | |
| **Aula 4**  **Classificação: Knn**   * Outras Métricas de Classificação: Matriz de Confusão, Precisão, Revocação, F1 * Como empregar essas métricas para comparar e selecionar melhores modelos * Modelos de K-Vizinhos mais Próximos para Classificação | * Larose, Chantal D.; Larose, Daniel T. **Data Science Using Python and R** Hoboken: Wiley, c2019. E-book (259 p.) (Wiley Series on Methods and Applications in Data Mining Ser.). ISBN 9781119526834 (electronic bk.). online em: **Biblioteca do Mackenzie.** * Jake VanderPlas. **Python Data Science Handbook**. O'Reilly Media, Inc. (2016). ISBN: 9781491912058. Disponível em: <https://jakevdp.github.io/PythonDataScienceHandbook/> Acesso: 04 de Março de 2022. * Kotu, Vijay; Deshpande, Balachandre **Data Science: concepts and practice**. 2nd ed. Cambridge, [England]: Morgan Kaufmann, c2019. E-book (570 p.) ISBN 9780128147627 (electronic bk.). online em: **Biblioteca do Mackenzie.** | | *Cap. 7*  *Cap. 5,*  *Introducing Scikit-Learn*  *Hyperparameters and Model Validation*  *Cap. 4.3* | | *Online*  *Online*  *Online* | |
|
| **Aula 5**  **Árvores de Decisão e Seleção de Atributos**   * Outras Métricas de Classificação: Matriz de Confusão, Precisão, Revocação, F1 * Como empregar essas métricas para comparar e selecionar melhores modelos * Modelos de K-Vizinhos mais Próximos para Classificação | * Larose, Chantal D.; Larose, Daniel T. **Data Science Using Python and R** Hoboken: Wiley, c2019. E-book (259 p.) (Wiley Series on Methods and Applications in Data Mining Ser.). ISBN 9781119526834 (electronic bk.). online em: **Biblioteca do Mackenzie.** * Jake VanderPlas. **Python Data Science Handbook**. O'Reilly Media, Inc. (2016). ISBN: 9781491912058. Disponível em: <https://jakevdp.github.io/PythonDataScienceHandbook/> Acesso: 04 de Março de 2022. | | *Cap. 6*  *Cap. 5,*  *In-Depth: Decision Trees and Random Forests* | | *Online*  *Online* | |
| **Aula 6**  **Validação Cruzada e GridSearch**   * O que são os conjuntos de Validação e Teste * Como e por que aplicar a Validação Cruzada dos modelos * Como empregar GridSearch do Scikit-Learn para selecionar os hiperparâmetros de um modelo | * **\_\_\_.** Scikit-learn. **3.2. Tuning the hyper-parameters of an estimator.** Disponível em: https://scikit-learn.org/stable/modules/grid\_search.html Acesso: 04 de Março de 2022. * Kotu, Vijay; Deshpande, Balachandre **Data Science: concepts and practice**. 2nd ed. Cambridge, [England]: Morgan Kaufmann, c2019. E-book (570 p.) ISBN 9780128147627 (electronic bk.). online em: **Biblioteca do Mackenzie.** | | -  *Cap 8* | | *Online*  *Online* | |
| **Aula 7**  **Seleção de Modelos**   * Como selecionar melhores modelos de Classificação e Regressão * Sobre Outros Estimadores importantes * Como empregar modelos de aprendizado supervisionado para a Classificação de Imagens de Dígitos e Séries Temporais | * Jake VanderPlas. **Python Data Science Handbook**. O'Reilly Media, Inc. (2016). ISBN: 9781491912058. Disponível em: <https://jakevdp.github.io/PythonDataScienceHandbook/> Acesso: 04 de Março de 2022. | | *Cap. 5,*  *In-Depth: Decision Trees and Random Forests* | | *Online* | |
| **Aula 8**  **Aprendizado de Máquina em R**   * Como empregar o R para criar modelos de Aprendizado de Máquina Supervisionado * Como resolver problemas de Classificação e Regressão * Como fazer seleções simples de Hiperparâmetros e Modelos em R | * Larose, Chantal D.; Larose, Daniel T. **Data Science Using Python and R** Hoboken: Wiley, c2019. E-book (259 p.) (Wiley Series on Methods and Applications in Data Mining Ser.). ISBN 9781119526834 (electronic bk.). online em: **Biblioteca do Mackenzie.** | | *Cap. 6-8 e 11* | | *Online* | |

|  |
| --- |
| **Metodologia** |
| *Curso EAD* |
| **Critério de Avaliação** |
| *Curso EAD* |
| **Bibliografia Básica** |
| BECKER, J. L. **Estatística básica: transformando dados em informação**. Porto Alegre: Bookman, 2015.  CASTRO, L. N.; FERRARI, D. G. **Introdução à mineração de dados: conceitos básicos, algoritmos e aplicações**. São Paulo Saraiva, 2016.  SILVA, L. A. **Introdução à mineração de dados: com aplicações em R**. Rio de Janeiro GEN LTC, 2016. |
| **Bibliografia Complementar** |
| DEVORE, J. L. **Probabilidade e estatística para Engenharia e Ciências**. 2ª. Ed. São Paulo. CENGAGE Learning. 2014.  GOLDSHIMIDT, R.; PASSOS, E.; BEZERRA, E. **Data Mining.** 2ª. Ed. Rio de Janeiro: LTC, 2015.  LIMA, I. **Inteligência Artificial.** Ed. Grupo Gen, 2014.  SILVA, F.M., et al. **Inteligência Artificial.** Porto Alegre: SAGAH, 2019.  SHARDA, R.; DELEN, D.; TURBAN, E. **Business Intelligence e Análise de Dados para Gestão do Negócio.** 4ª. Ed. Porto Alegre: Bookman, 2019. |
| **Bibliografia Adicional** |
| Larose, Chantal D.; Larose, Daniel T. **Data Science Using Python and R** Hoboken: Wiley, c2019. E-book (259 p.) (Wiley Series on Methods and Applications in Data Mining Ser.). ISBN 9781119526834 (electronic bk.). online em: **Biblioteca do Mackenzie.**  Jake VanderPlas. **Python Data Science Handbook**. O'Reilly Media, Inc. (2016). ISBN: 9781491912058. Disponível em: <https://jakevdp.github.io/PythonDataScienceHandbook/> Acesso: 04 de Março de 2022.  Kotu, Vijay; Deshpande, Balachandre **Data Science: concepts and practice**. 2nd ed. Cambridge, [England]: Morgan Kaufmann, c2019. E-book (570 p.) ISBN 9780128147627 (electronic bk.). online em: **Biblioteca do Mackenzie.**  Hadley Wickham, Garrett Grolemund. **R for Data Science: Import, Tidy, Transform, Visualize, and Model Data.** O'Reilly Media Inc. (2017). (alternativamente disponível em: <https://r4ds.had.co.nz/>). |