

| <b>Disciplina</b>  | INTRODUÇÃO A CIÊNCIA DE DADOS  |                               | <b>Turma A</b>                      |                                |   |
|--|--|-------------------------------|-------------------------------------|--------------------------------|---|
| <b>Período</b>   | 2021.3   |                               | Código 219044                       |                                |   |
| <b>Departamento</b>                                      | Mecânica Aplicada e Computacional  |                               | <b>Professores</b> Leonardo Goliatt |                                |   |
| <b>Objetivos</b>   | <ol style="list-style-type: none"> <li>1. Identificar, obter e transformar um conjunto de dados para torná-lo adequado para a produção de dados estatísticos comunicados por escrito.</li> <li>2. Formular questões e hipóteses pertinentes baseados em dados.</li> <li>3. Recuperar automaticamente dados de fontes de informação</li> <li>4. Construir modelos de aprendizado de máquina com base em dados experimentais ou provenientes de simulações computacionais.</li> <li>5. Criar visualizações para comunicar a análise de dados.</li> </ol> |                               |                                     |                                |   |
| <b>Metodologia de ensino</b>                             | <ul style="list-style-type: none"> <li>- Videoaulas gravadas (assíncrono);</li> <li>- Estudo individual do material didático fornecido (apostila, slides, vídeos) e resolução de listas de exercícios (assíncrono);</li> <li>- Discussão sobre a teoria, exercícios e tira-dúvidas (síncrono);</li> <li>- Avaliações (assíncrono).</li> </ul>  |                               |                                     |                                |   |
| <b>Demandas de equipamentos e de conexão necessárias</b> | <ul style="list-style-type: none"> <li>- Smartphone ou tablet ou notebook ou desktop;</li> <li>- Conexão à internet suficiente para baixar o material e assistir aos vídeos disponibilizados;</li> </ul>   |                               |                                     |                                |   |
| <b>Programa</b>  |  |                               |                                     |                                |   |
| Video aula   | Data   | Quantidade de horas           |                                     | Assunto                        | Conteúdo                                      |
|  |  | Atividades ao vivo (síncrono) | Estudo individual                   |                                |   |
| 01   | 19/10/21   | 2                             |                                     | Princípios de Ciência de Dados | Processo de Ciência de Dados                  |
| 02   | 21/10/21   | 2                             |                                     |                                | Princípios FAIR                               |
| 03   | 26/10/21   | 2                             |                                     |                                | Obtenção, manipulação e limpeza de dados      |
| 04   | 28/10/21   | 2                             |                                     |                                | Obtenção, manipulação e limpeza de dados      |
| 05   | 02/11/21   | 2                             |                                     |                                | Análise exploratória de dados                 |
| 06   | 04/11/21   | 2                             |                                     |                                | Análise exploratória de dados                 |
| 07   | 09/11/21   | 2                             |                                     | Modelagem de Dados             | Recuperação e extração de informação          |
| 08   | 11/11/21   | 2                             |                                     |                                | Recuperação e extração de informação          |
| 09   | 16/11/21   | 2                             |                                     |                                | Aprendizado supervisionado: regressão         |
| 10   | 18/11/21   | 2                             |                                     |                                | Aprendizado supervisionado: regressão         |
| 11   | 23/11/21   | 2                             |                                     |                                | Aprendizado supervisionado: classificação     |
| 12   | 25/11/21   | 2                             |                                     |                                | Aprendizado supervisionado: classificação     |
| 13   | 30/11/21   | 2                             |                                     |                                | Aprendizado supervisionado: validação cruzada |
| 14   | 02/12/21   | 2                             |                                     |                                | Aprendizado supervisionado: validação cruzada |

|    |          |    |   |                             |  |
|----|----------|----|---|-----------------------------|--|
| 15 | 07/12/21 | 2  |   |                             | Aprendizado não supervisionado               |
| 16 | 09/12/21 | 2  |   |                             | Governança de dados                          |
| 17 | 14/12/21 | 1  | 1 | Projeto de Ciência de Dados | Projeto guiado: Classificação de textos      |
| 18 | 16/12/21 | 1  | 1 |                             | Projeto guiado: Previsão de séries temporais |
| 19 | 21/12/21 | 1  | 1 |                             | Projeto colaborativo: Raspagem de dados      |
| 20 | 11/01/22 | 2  | 1 |                             | Projeto guiado: Otimização de modelos ML     |
| 21 | 13/01/22 | 2  |   |                             | Introdução a Inteligência Artificial         |
| 22 | 20/01/22 | 1  | 1 |                             | Projeto Integrador de Ciência de Dados       |
|    |          | 40 | 5 |                             |  |
|    |          | 45 |   |                             |  |

|  |   |
|--|---|
| <b>Horário das aulas síncronas e atendimento</b> | <p>Aulas síncronas: 3ª feira, das 19h às 21h e 5ª feira, das 19h às 21h.</p> <p>O link das salas de aula será enviado à turma antes do início de cada atividade síncrona.</p> <p>Atendimento: pelo fórum do Google Classroom (preferencialmente) ou por e-mail, conforme demanda dos discentes.</p> |
|--|---|

|                     |   |
|---------------------|---|
| <b>Bibliografia</b> | <p>Referências Principais</p> <ol style="list-style-type: none"> <li>1. Géron, Aurélien. Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems. O'Reilly Media, 2019.</li> <li>2. Wes McKinney. 2012. Python for Data Analysis: Data Wrangling with Pandas, NumPy, and Ipython. O'Reilly Media, Inc.</li> </ol> <p>Bibliografia complementar</p> <ol style="list-style-type: none"> <li>1. Nina Zumel and John Mount. 2014. Practical Data Science with R (1st ed.). Manning Publications Co., Greenwich, CT, USA.</li> <li>2. Joel Grus. 2015. Data Science from Scratch: First Principles with Python. O'Reilly Media, Inc.</li> <li>3. Bart Baesens. 2014. Analytics in a Big Data World: The Essential Guide to Data Science and its Applications. John Wiley &amp; Sons</li> <li>4. Philipp K. Janert, 2010, Data Analysis with Open Source Tools. O'Reilly Media, Inc.</li> <li>5. McCallum, Q. Ethan. Bad data handbook: cleaning up the data so you can get back to work. O'Reilly Media, Inc., 2012.</li> </ol> |
| <b>Avaliação</b>    | <p>Avaliações <b>contínuas</b> com valor 100 pontos cada uma. A nota final (NF) é a média aritmética. Critério de aprovação: <math>NF \geq 60</math></p>  |