|  |  |
| --- | --- |
| **PROGRAMA DA DISCIPLINA** | |
| **CURSO** | ENGENHARIA DE SOFTWARE |
| **DISCIPLINA** | Data Science & Statistical Computing |
| **CARGA HORÁRIA** | 80h |
| **DATAS** | 5 de fevereiro de 2024 a dezembro de 2024. |

|  |
| --- |
| **PROFESSOR:** Carlos André Marinho Vieira  **E-MAIL:** profcarlos.vieira@fiap.com.br |

|  |
| --- |
| **EMENTA** |
| **1º. Semestre:**   * Introdução à Ciência de Dados * Análise Exploratória e Estatística Descritiva * Distribuições de Probabilidade * Testes de Hipótese (Paramétricos e Não Paramétricos) * Correlação e Regressão Linear   **2º. Semestre:**   * Limpeza e Manipulação de Dados * Regressão Linear Aplicada * Algoritmos de Regressão * Algoritmos de Classificação * Algoritmos de Agrupamento |

|  |
| --- |
| **OBJETIVOS** |
| Fomentar a compreensão das bases e aplicações da Ciência de Dados, incluindo a identificação de problemas de negócios, o estudo aprofundado dos tipos de dados e a exploração de carreiras na área, capacitando os alunos a iniciar sua jornada no campo com uma base sólida;  Desenvolver habilidades em Análise Exploratória de Dados e Estatística Descritiva, ensinando os alunos a aplicar medidas de tendência central, variação e a construir gráficos relevantes, para uma interpretação eficaz dos dados;  Introduzir e aplicar conceitos de Distribuições de Probabilidade e Testes de Hipótese, permitindo aos alunos compreender e aplicar testes paramétricos e não paramétricos, além de explorar intervalos de confiança, para validar suposições sobre dados;  Ensinar a correlação e fundamentos de Regressão Linear, preparando os estudantes para analisar relações entre variáveis e aplicar regressão linear simples e múltipla, incluindo a realização de testes de hipótese em contextos de regressão;  Capacitar os alunos em técnicas de Limpeza e Manipulação de Dados, garantindo que estejam aptos a tratar dados incorretos, duplicados, categóricos e faltantes, além de aprender a lidar com outliers e realizar o escalonamento de variáveis;  Entender os principais conceitos e aplicações de modelos de Séries Temporais, introduzindo os alunos ao estudo de modelos de previsão e séries contidas nos diversos segmentos de estudos;  Explorar algoritmos de machine learning utilizados em técnicas de Regressão, Classificação e Agrupamento, ampliando as habilidades analíticas dos alunos para diversos tipos de problemas de dados;  Promover a aplicação prática de técnicas de Ciência de Dados em problemas do mundo real, incentivando a solução de desafios de negócios com a aplicação de técnicas aprendidas, preparando os alunos para contribuições significativas em ambientes profissionais. |

|  |
| --- |
| **METODOLOGIA DE AVALIAÇÃO** |
| **Checkpoints:** 3x no semestre, formatados pelo professor.  **Sprints:** 4 mentorias/ano, 2 entregas/semestre, challenge com empresa parceira.  **Global Solution:** Case de avaliação semestral contendo todas as disciplinas.  **Pesos:** 20% Checkpoints, 20% Sprints, 60% Global Solution. |

|  |
| --- |
| **BIBLIOGRAFIA BÁSICA** |
| BRUCE, Andrew; BRUCE, Peter. **Estatística prática para cientistas de dados.** Alta Books, 2019.  LARSON, Ron; FARBER, Betsy; CYRO. **Estatística aplicada**, 6 ed. Prentice Hall, 2016.  FAWCETT, Tom; PROVOST, Foster. **Data Science para Negócios**: O que você precisa saber sobre mineração de dados e pensamento analítico de dados. Alta Books Editora, 2016.  GRUS, J. **Data science do zero:** Primeiras regras com o Python. Alta Books: 2016. |
| **BIBLIOGRAFIA COMPLEMENTAR** |
| SHARDA, Ramesh; DELEN, Dursun; TURBAN, Efraim. **Business Intelligence e Análise de Dados para Gestão do Negócio**, 4d. Bookman Editora, 2019.  MCKINNEY, Wes. **Python para análise de dados:** Tratamento de dados com Pandas, NumPy e IPython. Novatec Editora, 2018.  HUFF, Darrell. **Como mentir com estatística.** Intrínseca, 2016.  KNAFLIC, Cole Nussbaumer. **Storytelling com dados:** um guia sobre visualização de dados para profissionais de negócios. Alta Books, 2017.  HUANG, Changquan; PETUKHINA, Alla. **Applied Time Series Analysis and Forecasting with Python.** Springer, 2022.  PEIXEIRO, Marco. **Time series forecasting in python.** Simon and Schuster, 2022. |