



MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DO PARANÁ
SETOR DE CIÊNCIAS SOCIAIS APLICADAS
Departamento de Ciência e Gestão da Informação

Ficha 2 (variável)

Disciplina: Tópicos em análise de dados		Código: SIN233
Professor responsável: José Marcelo A. P. Cestari (cestari@ufpr.br)		Período da oferta: 17/10/2022 a 25/02/2023. Vagas: 30.
Natureza: () Obrigatória (X) Optativa	(X) Semestral	() Anual () Modular
Pré-requisito: não possui.	Co-requisito: não possui.	
Modalidade: () Presencial () EaD (X) Parcialmente EAD: 20% (9h)		
Pré-requisito: não há	Co-requisito: não há	

C.H.Total: 45 horas.

C. H. Semanal: 3 horas.

Prática como Componente Curricular (PCC): 0.

Atividade Curricular de Extensão (ACE): 0.

Padrão (PD): 30 horas. Laboratório (LB): 15 horas. Campo (CP): 0.

Estágio (ES): 0. Orientada (OR): 0. Prática Específica (PE): 0. Estágio de Formação Pedagógica (EFP): 0.

Recesso entre 26/12/2022 e 14/01/2023 (inclusive).

Período do exame final: de 27/02 até 04/03/2023

Exame final da disciplina: 01/03/2023.

Feriados e outras datas previstas (conforme [PORTARIA ME Nº 14.817, DE 20 DE DEZEMBRO DE 2021](#)): 12/10, 28/10 (ponto facultativo), 02/11, 15/11 e 25/12.

Carnaval 2023 previsto para 20, 21 e 22/02/2023 (quarta-feira).

Justificativa da oferta 20% EaD: Portaria 280/ PROGRAD, de 13/12/2017 que "Aprova Ajuste Curricular no Curso de Gestão da Informação, aplicável à Resolução 64/13-CEPE" (0690698).

EMENTA

Programação matemática: programação linear, inteira e binária; análise de sensibilidade. Simulação: método de Monte Carlo. Uso de softwares em programação matemática e simulação.

PROGRAMA

- Introdução
 - Revisão de programação em Python e estatística.
 - Visão geral de aprendizado de máquina (machine learning)
- Método Monte Carlo
- Aplicação em Python para tratar, dentre outros, dos seguintes temas:
 - Estatística descritivas.

- o Plotagem.
- o Regressão linear, múltipla, logística e polinomial.
- o Árvore de decisão.
- o Agregações
- o K-means
- o Validação cruzada.

OBJETIVO GERAL

O aluno deverá conhecer e ser capaz de aplicar diferentes técnicas estatísticas e computacionais para análise de dados e apoio à tomada de decisão utilizando, dentre outras coisas, programação em Python.

OBJETIVOS ESPECÍFICOS

O aluno deverá ser capaz de:

1. Utilizar métodos na linguagem de programação Python relacionados à machine learning e estatística;
2. Utilizar métodos para plotagem de dados em gráficos
3. Utilizar métodos com árvores de decisão, regressões e demais correlatas.

PROCEDIMENTOS DIDÁTICOS

ATENÇÃO: apesar de não existir uma obrigatoriedade de pré-requisito para esta disciplina, recomenda-se que o discente já tenha feito a disciplina "PROGRAMAÇÃO DE COMPUTADORES PARA GESTÃO DA INFORMAÇÃO (SIN176)" e/ou que conheça um pouco da linguagem de Programação Python. Usaremos o Python cotidianamente na disciplina. Uma sugestão e alternativa é que o discente entre no site da Data Science Academy (DSA), cujo acesso é gratuito, e façam o curso "Python Fundamentos para Análise de Dados 3.0"- <https://www.datascienceacademy.com.br/cursosgratuitos> (acessado em 31/08/2022).

Aulas, teórico práticas, expositivo dialogadas em encontros presenciais nos quais a docente faz exposição do tema do programa, discute com os alunos os métodos e aplicações e propõe exercícios e atividades em grupo ou individuais, além das atividades avaliativas.

Serão utilizados como recursos: sala de aula no SA, laboratório de computadores, sala no ambiente virtual de aprendizagem – AVA, mais especificamente o UFPR Virtual; o Microsoft Office 365 disponível institucionalmente, o site da Data Science Academy como suporte ao conteúdo e o software Anaconda Python para programação em Python. Para as aulas não presenciais o material didático será elaborado observando utilização de linguagem dialógica, apresentação e formato próprios para o ensino à distância. Como suporte ao ensino será utilizada a infraestrutura da UFPR tais como: software (Office 365) e AVA. Em atendimento à Resolução 72/10-CEPE, serão observados os seguintes aspectos:

1. Sistema de comunicação – encontros presenciais, AVA (UFPR Virtual), correios eletrônicos trocados entre coordenação e alunos e professor e alunos;
2. Modelo de tutoria a distância e presencial – não está prevista a tutoria nesta disciplina. As dúvidas serão esclarecidas pela docente e, quando existir, pelos alunos do programa de Monitoria ou alunos de PPGs em estágios docência;
3. Material didático específico – o material didático será elaborado observando utilização de linguagem dialógica, apresentação e formato próprios para o ensino presencial;
4. Infraestrutura de suporte tecnológico, científico e instrumental à disciplina – o curso conta com Ambiente Virtual de Aprendizagem específico e pacote Office 365, incluindo sala no Teams para eventuais esclarecimentos e interação;
5. Previsão de período de ambientação dos recursos tecnológicos a serem utilizados pelos discentes: os alunos recebem capacitação no Ambiente de Aprendizagem Virtual no primeiro semestre do Curso e esta é uma disciplina de sexto período. Assume-se que os alunos já tenham domínio no uso das tecnologias.
6. As atividades podem ser presenciais ou remotas, conforme definição em sala de aula e rubricas.

IMPORTANTE:

A) Cursos DSA. Há uma série de materiais e cursos gratuitos no portal DSA. Alguns deles são "Inteligência Artificial Fundamentos 2.0" e "Introdução à Ciência de Dados 3.0" link: <https://www.datascienceacademy.com.br/cursosgratuitos> (visitado em 17/08/2022).

B) Minha Biblioteca (Livros on-line - UFPR). Link aqui: <https://minhabiblioteca.ufpr.br/biblioteca/> (visitado em 17/08/2022).

C) Instalação de ferramentas. Dentro das possibilidades, para melhor acompanhar a disciplina usando recursos próprios (computador de casa ou notebook - se for o caso) providenciar a instalação das seguintes ferramentas:

- **Anaconda Python (Python 3.7 - Anaconda 2020.02 (Windows)). Obrigatório.**
 - o Link direto: https://repo.anaconda.com/archive/Anaconda3-2020.02-Windows-x86_64.exe
 - o Link para todo repositório de versões: <https://repo.anaconda.com/archive/>
 - Procurem pela versão de 2020-03-11.
- **Jupyter Notebook :** <http://jupyter.org/>
- **Sublime Text 3**(editor de textos - opcional). É um editor de textos 'normal' mas que tem algumas facilidades para o Python. Link: <https://www.sublimetext.com/3>
- **PyCharm Community Edition (Windows) - opcional:** <https://www.jetbrains.com/pt-br/pycharm/download/#section=windows>

Os links acima foram visitados no dia 31/08/2022. As ferramentas necessárias estarão instaladas no laboratório pertinente para a disciplina.

D) Informações gerais.

Site oficial: python.org.

Usar o Python 3. NUNCA iremos usar a Versão 2 do Python.

Não iremos utilizar o interpretador padrão da linguagem Python (lembrando que Python é uma linguagem interpretada e não compilada). Nós iremos usar o **Anaconda Python** junto com o editor Jupyter Notebook como interpretador (já vem com muitos pacotes prontos).

Como alternativa há algumas opções para escrever código Python on-line, uma delas é o Google Colab (<https://colab.research.google.com/>)

ATIVIDADES NÃO PRESENCIAIS

Sistema de Comunicação: Plataforma Moodle da UFPR, com acesso em www.ufprvirtual.ufpr.br.

Tutoria/Monitoria: As atividades no Moodle terão a supervisão do professor da disciplina, com apoio de monitor do Programa Institucional de Monitoria (PIM) quando este for o caso. A tutoria/monitoria prevê a orientação no cumprimento das tarefas estabelecidas para cada estratégia, verificação de prazos, resposta de dúvidas e identificação de necessidades e problemas na condução da disciplina.

Material didático: No Moodle está disponível o Guia Didático da disciplina, onde consta o cronograma e demais orientações gerais das práticas e atividades. Além deste documento, na plataforma também estão disponíveis demais materiais pertinentes à disciplina, incluindo mas não limitados à arquivos para leitura (e-books ou arquivos do tipo pdf ou links para acesso a artigos em periódicos); e vídeos (link para visualização de material audiovisual em plataformas como Youtube).

Suporte técnico: O Setor de Ciências Sociais Aplicadas dispõe de laboratório de Informática para os estudantes que necessitarem de acesso a computador e à Internet. O mesmo fica no 1º andar do edifício do Setor, no campus Jardim Botânico. Suporte técnico também é oferecido em relação ao UFPR Virtual pela Coordenação de Integração de Políticas de Educação a Distância (CIPEAD/PROGRAD). Eventuais problemas no Moodle deverão ser comunicados sempre que verificados ao monitor da disciplina ou, então, ao administrador da plataforma, em <https://ufprvirtual.ufpr.br/course/view.php?id=8860>.

Ambientação: o curso de Gestão da Informação prevê a ambientação de seus estudantes no AVA durante o 1º ano, 1o período, a partir da oferta de duas disciplinas 100% EaD, a partir das quais todos os alunos são orientados e experimentam o Moodle e suas funcionalidades. Caso persistam dúvidas sobre a utilização da plataforma, recomenda-se recorrer aos tutoriais disponíveis em <https://ufprvirtual.ufpr.br/course/view.php?id=8737#section-0>.

Frequência: Nos dias previstos para atividades de EaD a frequência é computada a partir da entrega das atividades. A não conclusão das mesmas no prazo é considerada falta.

FORMAS DE AVALIAÇÃO E CONTROLE DE FREQUÊNCIA

A aprovação na disciplina ocorrerá conforme a Resolução nº 37/97-CEPE. Ela dependerá do resultado das avaliações realizadas ao longo do período letivo (atividades), segundo o plano de ensino e cronograma divulgado no início do semestre, sendo o resultado global expresso de zero a cem. Será aprovado por média a e o discente que alcançar, no total do período letivo, frequência mínima de 75% da carga horária inerente à disciplina e obtiver, no mínimo, grau numérico 70 de média aritmética no conjunto de provas e outras tarefas propostas. Discente que não obtiver a média prevista deverá prestar exame final, desde que alcance a frequência mínima exigida e média não inferior a 40. No exame final a aprovação na disciplina dependerá da obtenção de grau numérico igual ou superior a 50 na média aritmética entre o grau do exame final e a média do conjunto das avaliações realizadas.

Os critérios para avaliação da aprendizagem envolverão a participação ativa nas atividades propostas; a posição crítica em relação às leituras e debates; o envolvimento no processo de ensino-aprendizagem; a evolução do aprendizado ao longo da disciplina; e o desenvolvimento de habilidades e competências em relação ao tema abordado.

As etapas avaliativas poderão ser assíncronas e/ou presenciais. Os critérios das atividades e avaliações estão definidos no Guia Didático da disciplina e serão publicados no ambiente virtual. As duas notas principais para composição da média final são:

1. **primeira nota:** média aritmética das atividades (AT) realizadas individualmente ou em equipe (conforme solicitado);
2. **segunda nota:** média aritmética das notas das avaliações (AV) individuais;

A nota final da disciplina será composta resumidamente conforme a seguinte equação:

$$\text{Nota} = \frac{\frac{\sum_{i=1}^x AT_i}{x} + \frac{\sum_{i=1}^n AV_n}{n}}{2}$$

Onde i é um contador incremental que começa valendo 1, x e n são a quantidade de atividades e avaliações (respectivamente) da disciplina. AT é sigla para Atividade e AV é sigla para Avaliação.

De maneira geral as atividades e avaliações irão considerar os seguintes aspectos:

- profundidade na abordagem do tema;
- qualidade do levantamento bibliográfico;
- contextualização do tema na área de interesse;
- explicação do funcionamento do algoritmo em questão;
- exemplificações práticas e contextualizadas do tema;
- análises dos resultados obtidos;
- criatividade da apresentação;
- uso de normas para elaboração de trabalhos de pesquisa (normas ABNT);
- desempenho individual e coletivo dos membros da equipe durante a apresentação.

As frequências serão controladas por chamada em sala de aula e/ou por entrega das atividades (no caso dos encontros remotos). O limite de ausências permitido é 25% da carga horária da disciplina.

Para mais detalhes sobre as datas previstas para as avaliações favor consultar o cronograma dentro do Guia Didático. Eventuais necessidades de ajustes nas datas, atividades e afins serão combinadas em acordo com a turma.

Exame final da disciplina: 01/03/2023.

BIBLIOGRAFIA BÁSICA (mínimo 03 títulos)

COLIN, Emerson Carlos. Pesquisa Operacional: 170 aplicações em estratégia, finanças, logística, produção, marketing e vendas. Rio de Janeiro: LTC, 2007.

LACHTERMACHER, Gerson. Pesquisa operacional na tomada de decisões. 3. ed São Paulo: Pearson Prentice Hall, 2009.

MOORE, Jeffrey H; WEATHERFORD, Larry R. Tomada de decisão em administração com planilhas eletrônicas. 6. ed Porto Alegre: Bookman, 2005.

BIBLIOGRAFIA COMPLEMENTAR (mínimo 05 títulos)

ANDRADE, Eduardo Leopoldino de. Introdução à pesquisa operacional: métodos e modelos para a análise de decisão. 2, ed Rio de Janeiro: LTC, 2000.

CASTRUCCI, Plínio de Lauro. Modelos computacionais para gestão. Barueri: Manole, 2005.

NUTT, Paul; WILSON, David C. Handbook of decision making. Chichester UK: J. Wiley, 2010.

RAGSDALE, Cliff T. Modelagem e análise de decisão. São Paulo: Cengage Learning, 2010.

TARAPANOFF, Kira. Análise da informação para tomada de decisão: desafios e soluções. Curitiba: InterSaberes, 2015.

OUTRAS BIBLIOGRAFIAS

BURKOV, Andriy. **The hundred-page machine learning book**. Disponível em: <http://themlbook.com/wiki/doku.php>. Acesso em 17/08/2022.

DEISENROTH, Marc Peter; FAISAL, A. Aldo e ONG, Cheng Soon. **Mathematics for machine learning**. Disponível em: <https://mml-book.github.io>. Acesso em: 17/08/2022.

GOOGLE. **Machine learning crash course**. Disponível em: <https://developers.google.com/machine-learning/crash-course/>. Acesso em 17/08/2022.

POOLE, David L.; MACKWORTH, Alan K. **Artificial intelligence: foundations of computational agents**. 2ª ed. Disponível em: <https://artint.info/2e/html/ArtInt2e.html>. Acesso em 17/08/2022.

SHALEV-SHWARTZ, Shai; BEN-DAVID, Shai. **Understanding machine learning: From Theory to Algorithms**. Disponível em: <https://www.cse.huji.ac.il/~shais/UnderstandingMachineLearning/> Acesso em 17/08/2022.

SUTTON, Richard S.; BARTO, Andrew G. **Reinforcement learning: an introduction**. 2ª ed. Disponível em: <http://incompleteideas.net/book/RLbook2018.pdf>. Acesso em 17/08/2022.

Vídeos

RAMESH, Raj. **What is artificial intelligence?** In 5 minutes. Disponível em <https://www.youtube.com/watch?v=2ePf9rue1Ao&t=155s> . Acesso em 17/08/2022.

Ventury city. **Amazon's city of the future**. Disponível em <https://www.youtube.com/watch?v=ZjraUKJRIt8>. Acesso em 17/08/2022.

MUELLER, J. P. Começando a programar em Python para leigos 2Ed. Rio de Janeiro: Editora Alta Books, 2020. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9786555202298/>. Acesso em 12/08/2022.

BANIN, S. L. Python 3 - Conceitos e Aplicações - Uma abordagem didática. São Paulo: Editora Saraiva, 2018. 9788536530253. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9788536530253/>. Acesso em 12/08/2022.

PERKOVIC, L. Introdução à Computação Usando Python - Um Foco no Desenvolvimento de Aplicações. Rio de Janeiro: Grupo GEN, 2016. 9788521630937. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9788521630937/>. Acesso em 12/08/2022.

RIBEIRO, J.A. Introdução à Programação e aos Algoritmos. Rio de Janeiro: Grupo GEN, 2019. 9788521636410. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9788521636410/>. Acesso em 12/08/2022.

VIEIRA, M.J.; ANDRADE, S.C.; DA, S.P.F.; DUTRA, F.R. Raciocínio algorítmico. Porto Alegre: Grupo A, 2020. 9786581492915. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9786581492915/>. Acesso em 12/08/2022.

CORMEN, T. Desmistificando Algoritmos. Rio de Janeiro: Grupo GEN, 2013. 9788595153929. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9788595153929/>. Acesso em 12/08/2022.

PIVA JR, D. Algoritmos e Programação de Computadores. Rio de Janeiro: Grupo GEN, 2019. 9788595150508. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9788595150508/>. Acesso em 12/08/2022.

Data Science Academy (DSA). Python Fundamentos Para Análise de Dados 3.0. Disponível em <https://www.datascienceacademy.com.br/>. Acesso em 12/08/2022.

Demais materiais serão disponibilizados no UFPR Virtual conforme a necessidade



Documento assinado eletronicamente por **RODRIGO EDUARDO BOTELHO FRANCISCO, CHEFE DO DEPARTAMENTO DE CIENCIA E GESTAO DA INFORMACAO - SA**, em 21/09/2022, às 15:40, conforme art. 1º, III, "b", da Lei 11.419/2006.



A autenticidade do documento pode ser conferida [aqui](#) informando o código verificador **4795576** e o código CRC **55B2D74D**.
