

PLANO DE ENSINO

IDENTIFICAÇÃO		
HABILITAÇÃO:	Análise de Dados	C. H. TOTAL: 80 horas
PROFESSOR (A):	Gerson do Nascimento Silva (PhD Student, UnB)	

PLANEJAMENTO	
BASES TECNOLÓGICAS	
Ementa: Desenvolvimento de código-fonte (script) utilizando linguagem <i>python</i> para implementação, mostrando como:	
1. Uso da biblioteca “Seaborn”, PCA, ANOVA, Curva ROC, Hiper-parâmetros e Modelos de Regressão <ol style="list-style-type: none"> Reindexar <i>dataframes</i>; Substituir múltiplos valores usando biblioteca computacional <i>Pandas</i>; Plotar <i>dataframes</i> com biblioteca computacional <i>Seaborn</i>; Limpeza, estruturação e enriquecimento de dados brutos – “<i>Data Wrangling</i>”; Plotar gráfico de barras usando <i>dataframes Pandas</i>; Utilizar séries temporais em <i>Pandas</i>; Gerar séries temporais com <i>Pandas</i> e <i>Seaborn</i>; Gerar “<i>grouped bar</i>”; Determinar e plotar coeficientes de correlação; Reduzir dimensionalidade de matriz esparsa; Usar PCA (<i>Principal Component Analysis</i>) para reduzir dimensionalidade; Extrair características usando PCA; Usar “ANOVA <i>F-values</i>” para selecionar características; 	<ol style="list-style-type: none"> Usar “<i>Chi squared</i>” para selecionar características (<i>features</i>); Remover características altamente correlacionadas; Verificar acurácia de um modelo com validação cruzada; Verificar “pontuação AUC” de um modelo; Plotar curva de aprendizado de máquina; Plotar “curva ROC”; Usar o algoritmo “<i>Random Forest</i>”; Sintonizar hiper-parâmetros com algoritmo “<i>GridSearchCV</i>” e “<i>Random Search</i>”; Otimizar hiper-parâmetros de um modelo de regressão logística; Otimizar hiper-parâmetros de um modelo de árvore de decisão; Criar e otimizar “baseline” de um modelo de regressão linear; Criar e otimizar “baseline” de um modelo de regressão ridge; Criar e otimizar “baseline” de um modelo de regressão lasso e “<i>ElasticNet</i>”.
TEORIA	PRÁTICA
<p>-video aula: (<i>playlist</i> do PhD Francisco A. Rodrigues – USP)</p> <ul style="list-style-type: none"> “Estatísticas”; “Probabilidades”; “Redes Complexas”; “Processos Estocásticos”; “Ciência de Dados”; <p>-video aula: (<i>playlist</i> do Dr. Alexandre L. M. Levada – USP)</p> <ul style="list-style-type: none"> “Introdução a Teoria dos Grafos”; “Reconhecimento de Padrões”; 	<p>- Análise de dados:</p> <p>é um processo de inspeção, limpeza, transformação e modelagem de dados com o objetivo de descobrir informações úteis, informar conclusões e apoiar a tomada de decisões. A análise de dados tem múltiplas facetas e abordagens, abrangendo diversas técnicas sob uma variedade de nomes, e é usada em diferentes domínios dos negócios, ciências e ciências sociais. Desempenha um papel tornando a tomada de decisões mais científicas e ajudando no processo de operar com mais eficácia.</p>

ANÁLISE DA REALIDADE
Pré-Requisito
<ol style="list-style-type: none"> Lógica de Programação; Programação Orientada a Objetos;

3. Análise de Sistemas Orientada a Objeto;
4. Fundamentos e Modelagem de banco de Dados.

Necessidade da turma:

- Programação intermediário em *python* ou outra linguagem de programação que traga subsídios para entendimento dos conceitos computacionais.

PROJEÇÃO DE FINALIDADES

Objetivos Gerais (Competências)

Compreender os conceitos de IA e seus recursos e capacidade para implementar código-fonte reproduzível.

Objetivos Específicos (Habilidades)

- Capacidade para implementar código-fonte (uso de *python* ou outras tecnologias com mesma finalidade);
- Capacidade para implementar API'S para o pré-processamento de dados;
- Capacidade para implementar API'S para análise de dados;
- Integração do uso de API'S em manipulação de banco de dados;
- Ser capaz de implementar código-fonte reproduzível utilizando conceitos anteriores.

FORMAS DE MEDIAÇÃO

Procedimentos Metodológicos

- Elaborar modelos de implementação de código-fonte que valide as teorias, com uso de exemplos;
- Propor trabalhos práticos sobre o assunto;
- Corrigir os trabalhos práticos;
- Elaborar avaliação do conhecimento sobre o assunto ministrado;
- Uso de computadores para implementar exemplos;

AVALIAÇÃO*

Instrumento(s)	Data
<ul style="list-style-type: none"> • Avaliação única por meio do instrumento <i>Quiz</i>; • Criação de um projeto final "<i>End to End</i>" implementado em código-fonte reproduzível <i>python</i>. 	Ao fim do módulo estudado.

* No Exame Quiz não é permitido qualquer espécie de acréscimo de nota (trabalhos, participações e outros) estranhos à avaliação.

* Quiz é um jogo mental no qual os jogadores tentam responder corretamente a questões que lhes são colocadas. A palavra também é utilizada como sinônimo de avaliação de aquisição de conhecimentos ou capacidades em ambientes de aprendizagem.

BIBLIOGRAFIA

BIBLIOGRAFIA BÁSICA (Títulos, periódicos, etc)

Título/Periódico	Autor	Edição	Editora	Ano	Livro Texto
Álgebra Linear	José Luiz Boldrini [et al.]	3ª edição	Harbra Ltda	1980	-
Trilhas Python	Eduardo Pereira	1ª edição	Casa do Código	2018	-
Python para Desenvolvedores	Luiz Eduardo Borges	3ª edição	Novatec	2014	Sim

BIBLIOGRAFIA COMPLEMENTAR (Títulos, periódicos, etc)

Título/Periódico	Autor	Edição	Editora	Ano
Curso Intensivo de Python	Eric Matthes	1ª edição	Novatec	2017
Data Science from Scratch	Joel Grus	1ª edição	O'Reilly	2015
Introduction to Applied Linear Algebra	Stephen Boyd	1ª edição	Cambridge University Press	2018

Outros: