

Treinamento CIS - 1º Período

1. Conteúdos do Período

- a. Conceitos básicos de Data Science
- b. Python para DS
- c. Manipulação de base de dados
- d. Visualização de dados
- e. Pandas, numpy, seaborn e matplotlib
- f. Normalização de parâmetros
- g. Tipos de dados
- h. Pré-processamento de dados
- i. Redução de dimensionalidade
- j. Clustering
- k. PCA
- l. Datasets de Treinamento x Validação
- m. Regressão Linear
- n. Regressão Logística

2. Conteúdos de Apoio

- a. [Python Básico ao Intermediário](#) - Vídeo que ensina todo o básico de Python.
- b. [Fundamentos de Python para Análise de Dados](#) - Curso DataScience Academy, Módulos recomendados - 5, 8 e 9
- c. [Cursos do Kaggle](#) - Cursos recomendados para o primeiro período - Python, Data Visualization, Pandas e Data Cleaning
- d. [Playlist - Introdução à Data Science, Visualização de Dados e Machine Learning](#) - Playlist brasileira que explica dos conceitos básicos até as primeiras implementações dos conceitos citados.
- e. [Python DataScience HandBook](#) - Livro sobre Numpy, Pandas, Matplotlib, relativamente curto e direto.
- f. [Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow](#) - Livro completo: Para o primeiro período recomenda-se capítulos 1, 2, 8 e 9
- g. [Data Analysis with Dr Mike Pound](#) - Playlist com diversos conceitos de análise de dados avançada.
- h. [Tipos de Dados](#) - Artigo com os tipos de dados em modelos estatísticos.
- i. [Linear Regression StatQuest](#) - Playlist sobre regressão linear
- j. [Logistic Regression StatQuest](#) - Playlist sobre regressão logística
- k. [PCA StatQuest](#) - Vídeo explicando detalhadamente PCA
- l. [Machine Learning Sentdex](#) - Vídeos iniciais da playlist sobre Machine Learning.

3. Tarefas



- a. Procurar um dataset para fazer os trabalhos
- b. Fazer uma análise no dataset utilizando as ferramentas aprendidas no período
- c. No meio do período, haverá uma reunião com o monitor para consolidação das informações
- d. Demonstrar como pré processar e representar os diferentes tipos de dados
- e. Demonstrar a utilização de técnicas como one-hot encoding, redução de dimensionalidade e PCA, clustering;
- f. Aplicar regressão nos dados - dividir o dataset em treinamento e validação e aplicar regressão para prever um resultado. Testar e visualizar os resultados;
- g. As atividades descritas são apenas sugestões e podem ser modificadas;
- h. A entrega é individual e deverá ser colocada no seu GitHub pessoal