

LOB1266 - Introdução à Ciência de Dados e Aprendizado de Máquinas

Introduction to Data Science and Machine Learning

Créditos-aula: 4

Créditos-trabalho: 0

Carga horária: 60 h

Ativação: 01/01/2024

Departamento: Ciências Básicas e Ambientais

Curso (semestre ideal): EA (8)

Objetivos

Utilização de ferramentas e bibliotecas computacionais para resolver problemas básicos de ciências de dados e aprendizado de máquinas

1) Introdução a ciência de dados: mercado profissional e a Indústria 4.0; 2) Ferramental e bibliotecas Python: Jupyter Notebook, Colab Google, NumPy e Pandas; 3) Raspagem web e uso de arquivos, bibliotecas Scikit-Learn, Matplotlib e Seaborn; 4) Descoberta e preparação de dados: descoberta de conhecimento em bancos de dados, análise, mineração, preparação e pré-processamento de dados; 5) Aprendizado de máquina: algoritmos e técnicas, ética e privacidade; 6) Predição: K-Vizinhos mais próximos, árvores de decisão e avaliação de modelos preditivos; 7) Desenvolvimento de um projeto de predição.

Docente(s) Responsável(eis)

Programa resumido

Introdução a ciência de dados; Ferramental e bibliotecas Python; Descoberta e preparação de dados; Aprendizado de máquina; predição, K-Vizinhos mais próximos e árvores de decisão; Desenvolvimento de um projeto de predição

1) Introdução a ciência de dados: mercado profissional e a Indústria 4.0; 2) Ferramental e bibliotecas Python: Jupyter Notebook, Colab Google, NumPy e Pandas; 3) Raspagem web e uso de arquivos, bibliotecas Scikit-Learn, Matplotlib e Seaborn; 4) Descoberta e preparação de dados: descoberta de conhecimento em bancos de dados, análise, mineração, preparação e pré-processamento de dados; 5) Aprendizado de máquina: algoritmos e técnicas, ética e privacidade; 6) Predição: K-Vizinhos mais próximos, árvores de decisão e avaliação de modelos preditivos; 7) Desenvolvimento de um projeto de predição.

Programa

1) Introdução a ciência de dados: mercado profissional e a Indústria 4.0; 2) Ferramental e bibliotecas Python: Jupyter Notebook, Colab Google, NumPy e Pandas; 3) Raspagem web e uso de arquivos, bibliotecas Scikit-Learn, Matplotlib e Seaborn; 4) Descoberta e preparação de dados: descoberta de conhecimento em bancos de dados, análise, mineração, preparação e pré-processamento de dados; 5) Aprendizado de máquina: algoritmos e técnicas, ética e privacidade; 6) Predição: K-Vizinhos mais próximos, árvores de decisão e avaliação de modelos preditivos; 7) Desenvolvimento de um projeto de predição.

1) *Introduction to data science: professional market and Industry 4.0*; 2) *Python tools and libraries: Jupyter Notebook, Colab Google, NumPy and Pandas*; 3) *Web Scraping and use of files, Scikit-Learn, Matplotlib and Seaborn*; 4) *Data discovery and preparation, Knowledge Discovery in Databases (KDD), data analysis, mining, preparation and pre-processing*; 5) *Machine learning: algorithms and techniques, ethics and privacy*; 6) *Prediction: K-Nearest Neighbors (KNN), Decision Trees and evaluation of predictive models*; 7) *Development of a prediction project*.

Avaliação

Método: Aulas teóricas e práticas nos laboratórios de informática. Serão aplicadas pelo menos duas avaliações denominadas P1 e P2. A P2 deverá englobar toda a matéria ministrada ao longo do semestre, abrangendo todos os tópicos previstos na ementa. As avaliações podem ser: seminários, trabalhos, projetos ou outra forma de avaliação definida pelo professor.

Critério: Média ponderada das avaliações (M) igual ou superior a 5,0 pontos.

Norma de recuperação: A recuperação será composta por uma única prova (RC) englobando toda a matéria ministrada ao longo do semestre. A média final, para os alunos em recuperação, será calculada com base na relação: $MF = (M + RC) / 2$

Bibliografia

A) FACELI, Katti et al. *Inteligência artificial: uma abordagem de aprendizado de máquina*. Rio de Janeiro: GEN, 2021. B) GRUS, J. *Data Science do Zero*. 2ª. ed. Rio de Janeiro: Alta Books, 2021. C) VanderPlas, J. *Python Data Science Handbook*. Sebastapol, CA: O'Reilly, 2016. D) BANIN, S. L. *Python 3: conceitos e Aplicações: uma abordagem didática*. São Paulo: Saraiva, 2018. E) BARI, A.; CHAOUCHI, M.; JUNG, T. *Análise preditiva para leigos*. 2a. ed. Rio de Janeiro: Alta Books, 2019. F) Manuais online das ferramentas utilizadas.

Requisitos

LOB1056 - Introdução aos Métodos Numéricos e Computacionais (Requisito fraco)