

Conteúdo Programático

Curso de Extensão em Ciência de Dados e Analytics

Disciplina: Introdução à Ciência dos Dados

C.H.: 5 horas

Professor: Alexandre Maciel

Ementa: Introdução a Ciência de Dados; Problemas de Negócios; Produção de Dados; Armazenamento de Dados; Análise de Dados (Descritiva e Preditiva); Visualização de Dados.

Disciplina: Programação para Ciência de Dados

C.H.: 15 horas

Professor: Byron Leite

Ementa: Fundamentos de Programação; Estruturas de Dados e Algoritmos; Programação Orientada a Objetos; Ambientes de Desenvolvimento; APIs e Bibliotecas.

Disciplina: Análise e Transformação de Dados

C.H.: 15 horas

Professor: Mailson Santos

Ementa: Conceitos da coleta e preparação de dados; Processo de descoberta de conhecimento; Análise de Fontes e tipos de dados; ETL: Aplicação de técnicas e ferramentas para coleta, limpeza, processamento, transformação e entrega de dados.

Disciplina: Banco de Dados Não Convencionais

C.H.: 15 horas

Professor: Andrêza Leite

Ementa: Introdução a bancos de dados não convencionais; Fundamentos em Big Data; Gerenciamento de dados volumosos e heterogêneos; Arquitetura para sistemas de dados escaláveis; Novos modelos de dados e persistência poliglota; Modelagem e Armazenamento de dados não convencionais (semiestruturado, grafo, NoSQL e geográfico); Desenvolvimento de Soluções com bancos de dados não convencionais.

Disciplina: Estatística Computacional

C.H.: 15 horas

Professor: Manoel Marinho

Ementa: Introdução a Estatística; Fases do processo de análise de dados; Tipos de variáveis e escalas de mensuração, Técnicas de amostragem; Análise Descritiva de Dados (Medidas de tendência central, Medidas de dispersão); Intervalos de Confiança; Testes de Hipóteses paramétricos e não-paramétricos; Análise de variância ANOVA, Modelos de Regressão Simples e Múltipla; Simulação de Monte Carlo e Particionamento do conjunto de dados (leave one out, validação cruzada, holdout, bootstrap).

Disciplina: Aprendizado de Máquina

C.H.: 30 horas

Professor: Carmelo Bastos

Ementa: Introdução à Inteligência artificial: relevância, aplicações e eixos de desenvolvimento; Paradigmas de aprendizado de máquina; Aprendizado não supervisionado; Tarefas de Mineração de Dados: Agrupamento e Identificação de Outliers; Algoritmos de agrupamento automático: K-médias, Fuzzy C-Means, GMM, Mean Shifting clustering, DBS-CAN, Métodos Hierárquicos; Métricas para Análise de Desempenho de algoritmos de agrupamento. Aprendizado supervisionado; Métodos: Máquinas de Vetores de Suporte, K-vizinhos Mais Próximos, Árvores de Decisão, Random Forest, e Redes Neurais Artificiais; Análise de desempenhos: Métricas, técnicas e processos.

Disciplina: Processamento de Linguagem Natural

C.H.: 15 horas

Professor: Cleyton Rodrigues

Ementa: Introdução a mineração de texto; Pré-processamento de texto: encodings, tokenização, stopwords, stemming; Similaridade em texto: similaridade de string, tfidf, vetor de espaço de modelos; Classificação de texto: criação de features, análise de sentimentos; Extração de informação; Recuperação de Informação; Modelagem de Tópicos; Chatbots.

Disciplina: Projeto de Ciência de Dados

C.H.: 10 horas

Professor: Alexandre Maciel

Ementa: Desenvolvimento de projeto de ciência de dados a partir de demandas reais do mercado, usando dados públicos ou privados (formalmente cedidos para uso acadêmico na disciplina). Aplicação das etapas da metodologia CRISP-DM (entendimento do negócio, entendimento dos dados, preparação dos dados, modelagem, avaliação dos resultados e implantação).