

**Nome da Disciplina.**

Aprendizado de Máquina na Saúde.

**Disciplina do PGC:**

Tópicos Avançados em Sistemas de Computação I, II, III ou IV.

**Disciplinas da graduação (eletiva):**

TCC00259, TCC00260, TCC00261 Tópicos em Inteligência Artificial I, II, III.

**Pré-requisitos para a graduação:**

Estrutura de dados ou Estrutura de Dados e seus Algoritmos

Probabilidade e estatística

**Carga horária:** 60h

**Nº de créditos:** 4

**Objetivo da disciplina.**

Discutir e compreender os conceitos, métodos e aplicações do aprendizado de máquina na saúde, e os desafios da análise de dados clínicos, considerando a sua heterogeneidade e os aspectos legais e éticos da saúde, e de proteção à privacidade dos pacientes.

**Ementa.**

A disciplina trata dos conceitos, métodos e aplicações do aprendizado de máquina na saúde, incluindo o uso de modelos supervisionados, não supervisionados, por reforço e de aprendizado profundo em aplicações voltadas a triagem, diagnóstico, tratamento, reabilitação e bem-estar dos pacientes. Inclui tópicos sobre a estratificação de risco, modelagem do progresso de doenças e aplicações para o auxílio ao diagnóstico baseado em imagens e ao prognóstico de doenças, considerando os aspectos legais e éticos da saúde, e a proteção à privacidade dos dados dos pacientes.

**Conteúdo programático.**

1. Introdução ao aprendizado de máquina para a saúde.
  - a. Desafios do aprendizado de máquina na saúde
  - b. Princípios gerais de modelagem e avaliação de desempenho de modelos de aprendizado de máquina.

2. Características dos dados clínicos.
  - a. Desafios da heterogeneidade dos dados clínicos.
  - b. Preparação dos dados, normalização e padronização.
  - c. Métodos de exploração de dados e visualização para análise de dados clínicos.
3. Aplicação de modelos de triagem e estratificação de riscos.
  - a. Modelos supervisionados como regressão linear e logística.
  - b. Árvores de decisão e florestas aleatórias.
4. Modelos de sobrevivência.
  - a. Modelos paramétricos, não paramétricos e semi-paramétricos, e aplicações na saúde.
5. Inferência causal.
  - a. Redes bayesianas.
6. Noções sobre diagnóstico diferencial.
7. Aplicações do aprendizado profundo na saúde.
8. Aplicações do aprendizado por reforço na saúde.
9. Discussão sobre ética, privacidade de dados.
  - a. Desafios éticos na utilização de dados clínicos em aprendizado de máquina.
  - b. Noções sobre justiça algorítmica.
  - c. Medidas de privacidade de dados e regulamentação.

#### **Bibliografia básica:**

BERNER, Eta S., Clinical Decision Support System: Theory and Practice, Springer-Verlag New York, Second Edition, 2007.

KORB, K. B.; NICHOLSON, A. E. Bayesian artificial intelligence. Clayton, Victoria, Australia: Chapman & Hall/CRC, 2004.

PANESAR, Arjun, Machine Learning and AI for Healthcare: Big Data for Improved Health Outcomes, Apress, 2019.

PEARL, Judea. Causal inference in statistics: An overview. Statistics surveys, v. 3, p. 96-146, 2009.

SHORTLIFFE, E. H., CIMINO, J. E., Biomedical Informatics: Computer Applications in Health Care and Biomedicine, Forth Edition, Springer, 2014.

WANG, Ping; LI, Yan; REDDY, Chandan K. Machine learning for survival analysis: A survey. ACM Computing Surveys (CSUR), v. 51, n. 6, p. 1-36, 2019.

#### **Bibliografia complementar:**

AGRESTI, Alan, An Introduction to Categorical Data Analysis, Wiley Series in Probability and Statistics, 3rd edition, 2019.

BENSON, T.; GRIEVE, G. Principles of Health Interoperability. 4o ed. Springer, 2016.

BRAUNSTEIN, Mark L., Health Informatics on FHIR: How HL7's New API is Transforming Healthcare, 2018

CRAWLEY, Michael J. The R book. John Wiley & Sons, 2012.

GEISS, Linda S. et al. Prevalence and incidence trends for diagnosed diabetes among adults aged 20 to 79 years, United States, 1980-2012. *Jama*, v. 312, n. 12, p. 1218-1226, 2014.