

Nome primário: Tópicos Avançados em Ciência da Computação I, II, III, IV.

Nome secundário: Introdução ao Aprendizado de Máquina para Saúde.

Disciplinas da graduação (eletiva): TCC367 Introdução ao Aprendizado de Máquina para Saúde.

Pré-requisitos para a graduação:

Estrutura de Dados OU Estrutura de Dados e seus Algoritmos OU Probabilidade e Estatística

Carga horária: 60h

Nº de créditos: 4

Ementa

Conceitos, métodos e aplicações do aprendizado de máquina e da inteligência artificial em saúde, integrando perspectivas da computação, da medicina, da epidemiologia e da informática em saúde. Estudo do ciclo de vida de soluções de IA em saúde: definição do problema clínico, natureza e qualidade dos dados, interoperabilidade, modelagem, validação, interpretação, avaliação clínica, implementação, monitoramento e governança. Exploração de modelos supervisionados, não supervisionados, de sobrevivência, inferência causal, aprendizado profundo, NLP biomédico, modelos fundacionais e multimodais, aplicados a triagem, diagnóstico, prognóstico, apoio à decisão clínica, medicina de precisão e saúde pública. Discussão crítica de limitações de dados observacionais, vieses, equidade, privacidade, LGPD, explicabilidade, usabilidade, segurança, regulação de software como dispositivo médico e diretrizes de relato para modelos preditivos e intervenções de IA em saúde.

Objetivo

Capacitar estudantes a compreender, desenvolver, avaliar criticamente e comunicar soluções de aprendizado de máquina para problemas de saúde, considerando desempenho técnico, validade clínica, riscos, viabilidade de implementação, aspectos éticos, regulatórios e impacto no cuidado ao paciente.

Conteúdo Programático

1. **Fundamentos de IA e aprendizado de máquina em saúde:** Evolução histórica: sistemas especialistas, informática cognitiva, aprendizado estatístico, deep learning, modelos fundacionais e IA generativa; diferenças entre predição, diagnóstico, prognóstico, triagem, recomendação e decisão clínica.
2. **Dados clínicos, qualidade e interoperabilidade:** EHR, dados administrativos, imagens, sinais, notas clínicas, dados ômicos e determinantes sociais; qualidade, missingness, viés de seleção, leakage, governança de dados; padrões HL7 FHIR, OMOP, LOINC, SNOMED CT e terminologias clínicas.
3. **Desenho de estudos e ciclo de vida de modelos:** Formulação do problema, definição de população, desfecho, janela temporal, baseline, particionamento temporal, validação interna/externa, calibração, curvas ROC/PR, curvas de decisão e análise de subgrupos.
4. **Aprendizado supervisionado e não supervisionado:** Classificação, regressão, árvores, ensembles, regressão logística, SVM, redes neurais rasas, clustering, redução de dimensionalidade e fenotipagem computacional.
5. **Aprendizado profundo para dados biomédicos:** Redes convolucionais e transformers para imagens médicas, sinais e séries temporais; transferência de aprendizado; cuidados com balanceamento, aumento de dados e avaliação por paciente/instituição.

6. **NLP biomédico, LLMs e modelos multimodais:** Extração de informação de notas clínicas, embeddings, transformers, sumarização, codificação clínica, recuperação aumentada por busca, avaliação de alucinação, segurança, privacidade e uso responsável de LLMs.
7. **Sobrevivência, inferência causal e evidência do mundo real:** Modelos de sobrevivência, riscos competitivos, censura, emulação de ensaios alvo, confundimento, DAGs, propensity score, inferência causal e limites de dados observacionais.
8. **Sistemas de apoio à decisão clínica e implementação:** CDS, integração ao fluxo de trabalho, usabilidade, fadiga de alertas, interação humano-IA, desenho centrado no usuário, segurança do paciente e avaliação prospectiva.
9. **Explicabilidade, equidade, privacidade e segurança:** Interpretabilidade local/global, saliency, SHAP/LIME, explicações clinicamente úteis, vies algorítmico, equidade, LGPD, privacidade, anonimização, segurança e robustez.
10. **Regulação, relato científico e monitoramento pós-implantação:** SaMD, RDC ANVISA 657/2022, FDA, plano de mudanças predeterminadas, TRIPOD+AI, CONSORT-AI, SPIRIT-AI, DECIDE-AI, auditoria, drift, MLOps/MLHops e governança contínua.

Metodologia

Aulas integradas com seminários onde os alunos analisam estudos de caso reais e desenvolvem um miniprojeto prático utilizando bases como MIMIC-IV/MIETIC, PhysioNet, imagens públicas ou bases sintéticas quando houver restrição ética.

Avaliação

Participação em aula e discussões
Trabalhos práticos e relatórios
Projeto final e apresentação

Bibliografia básica

SHORTLIFFE, Edward H.; CIMINO, James J. (eds.). Biomedical Informatics: Computer Applications in Health Care and Biomedicine. 5. ed. Cham: Springer, 2021.

BENSON, Tim; GRIEVE, Grahame. Principles of Health Interoperability: FHIR, HL7 and SNOMED CT. 4. ed. Cham: Springer, 2021.

JAMES, Gareth; WITTEN, Daniela; HASTIE, Trevor; TIBSHIRANI, Robert; TAYLOR, Jonathan. An Introduction to Statistical Learning: with Applications in Python. Cham: Springer, 2023.

RAJKOMAR, Alvin; DEAN, Jeffrey; KOHANE, Isaac. Machine Learning in Medicine. New England Journal of Medicine, v. 380, n. 14, p. 1347-1358, 2019.

HERNÁN, Miguel A.; ROBINS, James M. Causal Inference: What If. Boca Raton: Chapman & Hall/CRC, 2020.

COLLINS, Gary S. et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ, v. 385, e078378, 2024.

WORLD HEALTH ORGANIZATION. Ethics and governance of artificial intelligence for health: WHO guidance. Geneva: WHO, 2021.

Bibliografia complementar:

AGÊNCIA NACIONAL DE VIGILÂNCIA SANITÁRIA (ANVISA). Resolução da Diretoria Colegiada - RDC nº 657, de 24 de março de 2022. Dispõe sobre a regularização de software como dispositivo médico (SaMD). Brasília: ANVISA, 2022.

AGÊNCIA NACIONAL DE VIGILÂNCIA SANITÁRIA (ANVISA). Software como Dispositivo Médico: Perguntas e Respostas - RDC 657/2022. Brasília: ANVISA, 2022.

FOOD AND DRUG ADMINISTRATION (FDA). Marketing Submission Recommendations for a Predetermined Change Control Plan for Artificial Intelligence-Enabled Device Software Functions. Silver Spring: FDA, 2025.

LIU, Xiaoxuan et al. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. *Nature Medicine*, v. 26, p. 1364-1374, 2020.

CRUZ RIVERA, Samantha et al. Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension. *Nature Medicine*, v. 26, p. 1351-1363, 2020.

VASEY, Baptiste et al. Reporting guideline for the early-stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. *Nature Medicine*, v. 28, p. 924-933, 2022.

WORLD HEALTH ORGANIZATION. Ethics and governance of artificial intelligence for health: guidance on large multi-modal models. Geneva: WHO, 2024.

TOPOL, Eric J. *Deep Medicine: How Artificial Intelligence Can Make Healthcare Human Again*. New York: Basic Books, 2019.

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

BEAM, Andrew L.; KOHANE, Isaac S. Big Data and Machine Learning in Health Care. *JAMA*, v. 319, n. 13, p. 1317-1318, 2018.

AGRESTI, Alan. *An Introduction to Categorical Data Analysis*. 3. ed. Hoboken: Wiley, 2019.

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

BERNER, Eta S. *Clinical Decision Support Systems: Theory and Practice*. 2. ed. New York: Springer, 2007.

Horário

Sextas, das 14h às 18h.