

Machine Learning aplicado à Saúde

André Filipe de Moraes Batista, Alexandre Chiavegatto Filho

Faculdade de Saúde Pública da Universidade de São Paulo, São Paulo, SP,

Objetivos do tutorial

Apresentar o panorama atual do uso das técnicas de aprendizagem de máquina aplicadas na área de saúde, contextualizando os principais algoritmos utilizados. Além de cobrir os aspectos teóricos sobre o tema, o capítulo terá uma seção prática de desenvolvimento de um conjunto de algoritmos de aprendizado de máquina fazendo-se uso de um conjunto de dados da área da saúde.

- **Principais modelos aplicados na área da saúde** – apresentaremos principais conceitos e funcionamento dos algoritmos mais utilizados na área da saúde;
- **Interpretabilidade de modelos preditivos** – será abordada a capacidade de interpretação dos modelos preditivos e as ferramentas existentes. As demonstrações serão feitas na linguagem Python.

Público-alvo

Interessados em conhecer como algoritmos de aprendizagem de máquina estão sendo utilizados na área da saúde e que buscam compreender os princípios de seus funcionamentos.

Autores

Prof. Dr. André Filipe de Moraes Batista. Pós-doutorando na Faculdade de Saúde Pública da USP, com projeto na área de Machine Learning aplicado à Saúde. Professor do Curso de Data Science do INSPER. Doutor em Engenharia da Computação pela Escola Politécnica da Universidade de São Paulo (POLI-USP). É pesquisador do Laboratório de Big Data e Análise Preditiva em Saúde da Faculdade de Saúde Pública da USP.

Tópicos abordados

- **Panorama do uso de machine learning na área da saúde** – será apresentado um panorama contextualizado dos principais avanços da área de machine learning aplicado à saúde. Diversos trabalhos científicos serão apresentados, destacando-se iniciativas brasileiras na área;
- **Machine Learning** – apresentam-se os conceitos básicos de machine learning, o que é aprendizado por indução, destacando os tipos de modelos de aprendizagem (supervisionado, não supervisionado e por reforço). Além disso, será apresentado um fluxo de tarefas para o desenvolvimento de algoritmos de machine learning, para o qual serão detalhadas as atividades de pré-processamento. Além disso, será discutido o conceito de sobreajuste e como evitá-lo. Por fim, discutiremos as técnicas mais utilizadas para a mensuração do desempenho dos modelos;

Prof. Dr. Alexandre Chiavegatto Filho. Professor Livre-Docente da Faculdade de Saúde Pública da USP (FSP/USP) na área de Estatísticas de Saúde. Possui graduação em Economia na FEA/USP, doutorado direto em Saúde Pública na FSP/USP e pós-doutorado na Universidade de Harvard. Atualmente é o pesquisador principal de pesquisas financiadas pela FAPESP, CNPq e Fundação Lemann. Atuou como professor convidado (2016) e pesquisador visitante (2017) na Universidade de Harvard. É coordenador de cursos na FSP/USP sobre o uso do R para a análise de dados e sobre machine learning em saúde. É diretor do Laboratório de Big Data e Análise Preditiva em Saúde da Faculdade de Saúde Pública da USP.