## Machine Learning in Medicine

Rahul C. Deo, MD, PhD

Uma conversa sobre o paper Circulation. 2015;132:1920-1930  $^{\rm 1}$ 

Eduardo Elias Ribeiro Junior Henrique Aparecido Laureano

Faculdade de Medicina Universide Federal de Minas Gerais - UFMG

13 de setembro de 2016

<sup>&</sup>lt;sup>1</sup>DOI: 10.1161/CIRCULATIONNAHA.115.001593.

#### Roteiro

- 1. Proposta
- 2. Aprendizado supervisionado aprendendo de florestas e árvores
- 3. C-Path: um patologista automatizado e a importância da extração de características
- 4. Atrator de metagenes em câncer e bake-offs em aprendizado de máquina
- ${\bf 5.}\;$  Aprendizado não supervisionado em HFpEF: Em direção a medicina precisa?

#### Disponibilização

#### https://stats-cwr.github.io/stats4med



#### stats4med: Statistics for Medicine

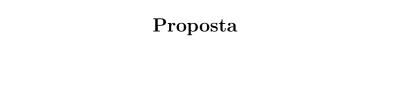
#### Statistics course with R for medicine students

Esta página web foi criada para armazenamento e disponibilização dos materiais elaborados para o curso de Estatística com R, ministrado para os alunos de graduação e pós-graduação em Medicina na Faculdade de Medicina da UFMG. Os materiais do curso são elaborados, essencialmente, como o pacote rmarkdown do R, cujo arquivos-fonte estão disponíveis em nosso repositório GitHub. As aulas podem ser acessadas pela barra de navegação ou ainda pela página Curso R.

Além de todo o material do curso também disponibilizamos aqui, os slides de eventuais seminários realizados para o grupo da Faculdade de Medicina. Esses materiais são listados no campo Seminars , da barra de navegação.

Atualizado em 16 de August de 2016.

© Copyright 2016 Ribeiro Jr., E. E.



#### Proposta do paper

Revisão das áreas médicas que podem ser beneficiadas com a utilização das abordagens de aprendizado e através de exemplos da literatura introduzir conceitos básicos em aprendizado de máquina:

▶ Apesar da enorme quantidade de conjuntos de dados da área médica e do uso dos algoritmos de aprendizagem adequados, pouquíssimos estudos tem significativa contribuição, o que causa um contraste com a significância obtida em várias outras áreas.

Portanto, outro objetivo do *paper* é encontrar razões para essa baixa relevância e como mudar esse cenário.

#### Divisões

#### Usualmente:

- ► Aprendizado supervisionado
  - Predição de uma saída ou alvo conhecido (interpretação de um ECG ou a detecção de nódulos num raio-x), i.e., o computador está aproximando o que um médico treinado já é capaz de fazer com alta precisão
- ► Aprendizado não supervisionado
  - Não existem saídas para serem preditas. O objetivo é encontrar naturalmente a ocorrência de padrões ou a realização de agrupamentos

#### Proposta:

- ▶ Técnicas utilizadas em situações cujas quais os médicos já se saem bem
- ► Técnicas utilizadas em situações cujas quais os médicos tem apenas sucesso limitado

## Algumas técnicas utilizadas

- ► Modelos estatísticos de regressão linear, como a regressão logística
  - Uma distribuição de probabilidade é assumida para a variável resposta condicionada as suas características
- ► Árvores de decisão
  - A cada ponto de quebra é feita uma partição com base nos valores de uma particular característica
- ► Redes neurais
  - Prediz uma resposta com base em representações transformadas das características
- ► K vizinhos mais próximos
  - Determina classes a partir dos valores das características dos k vizinhos mais próximos

# Aprendizado supervisionado - aprendendo de

florestas e árvores

C-Path: um patologista automatizado e a importância da extração de características

Atrator de metagenes em câncer e bake-offs em aprendizado de máquina

Aprendizado não supervisionado em HFpEF: Em direção a medicina precisa?