



## IA Aplicada à Saúde

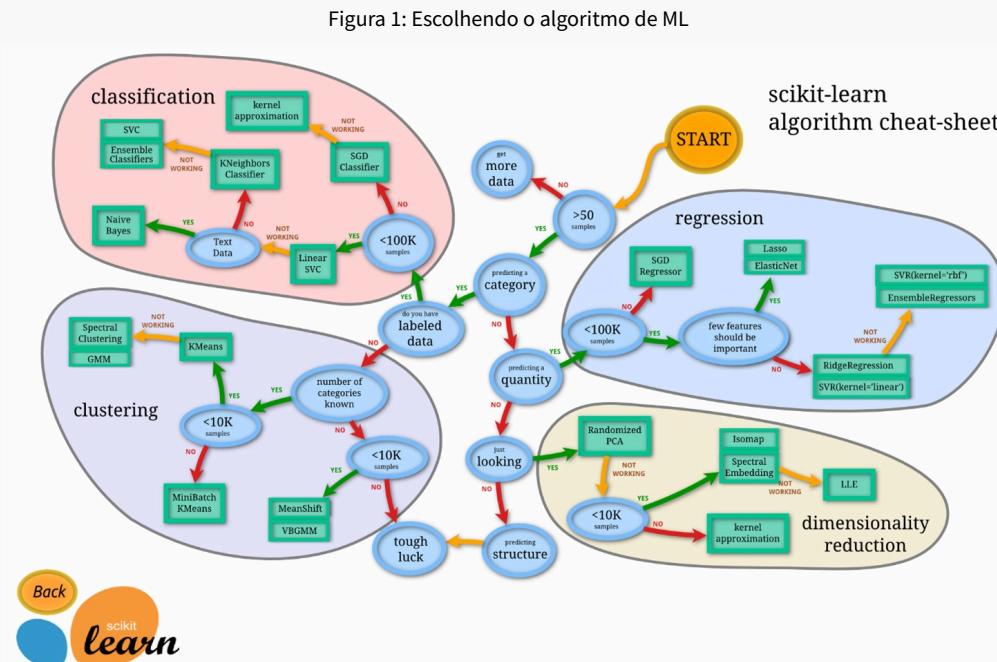
UNIDADE 04

Machine Learning Aplicado à Saúde

| Machine Learning Aplicado à Saúde

O Machine Learning (ML) representa grande parte das iniciativas de IA em Saúde, pois permite a descoberta de conhecimento oculto em meios aos dados (com a utilização de algoritmo de agrupamento) e também nos dá poder preditivo (através da classificação e regressão) que, dentre outras coisas, pode apoiar a decisão clínica e melhorar processos relacionados ao cuidado.

De maneira geral, quando desenvolvemos modelos de ML para área da saúde, seguimos as mesmas regras de utilização de algoritmos válidas para outros domínios. Apenas algumas atenções especiais se aplicam a determinadas tarefas específicas (ao longo da disciplina seremos nos depararemos com tais situações). Portanto, o fluxograma de decisão da biblioteca *scikit-learn* também pode ser aplicado para este contexto, conforme apresentado na Figura 1.



Fonte: scikit-learn. Disponível em: [https://scikit-learn.org/stable/tutorial/machine\\_learning\\_map/index.html](https://scikit-learn.org/stable/tutorial/machine_learning_map/index.html)

Na figura podemos ver os três principais tipos de algoritmos:

- **Classificação:** predição de valores categóricos (e.g., benigno/maligno)
- **Regressão:** predição de valores contínuos (e.g., estimar o custo de internação de um paciente)
- **Agrupamento:** descoberta de grupos de dados (e.g., segmentação de pacientes infectados com COVID-19 por grau de severidade)

Lembrando que a classificação e a regressão, fazem parte de grupo de algoritmos de **aprendizado supervisionado**, ou seja, precisam de dados rotulados/anotados para efetuar o treinamento do modelo. Já o agrupamento é integrante do grupo de **aprendizado não-supervisionado**, onde não há necessidade de uma padrão-ouro de anotações definidas, ou seja, não precisamos das “respostas” para treinar o modelo. Caso queira revisitar estes conceitos, indico a leitura a seguir.



### SAIBA MAIS

Para entender mais sobre o assunto, realize a Leitura do capítulo 1, Introdução, do livro Inteligência artificial: uma abordagem de aprendizado de máquina.  [Inteligência Artificial: uma abordagem de aprendizado de máquina](#)

Nas seções a seguir, vamos discutir como podemos desenvolver um projeto de ML para Saúde, incluindo uma série de recomendações e boas práticas. Adicionalmente, vamos passar por diversos exemplos práticos que incluem código e dados para podermos trabalhar.

## | Etapas no desenvolvimento de um projeto de Machine Learning em Saúde

Os passos básicos para o desenvolvimento de um projeto de ML em saúde não se diferem muito de um projeto de ML para outras áreas, conforme podemos resumir na lista a seguir.

- **Concepção:** especificar o problema como uma tarefa de aprendizado de máquina (ter certeza deve ser resolvido com ML)
- **Preparação dos dados:** análise exploratória dos dados, insights e limpeza dos dados
- **Treinamento do modelo:** escolha dos métodos de aprendizado e otimização do treinamento
- **Avaliação:** determinar a performance do modelo
- **Implantação:** tornar o modelo disponível e monitorar seu uso e acurácia

Dentro destas etapas mais gerais, vale a pena nos atentarmos a algumas situações específicas da área da saúde, como por exemplo a sub-etapa de **coleta de dados**. Dependendo do tipo de dado necessário para tarefa, a obtenção dos dados pode não ser tão simples devido a problemas de **privacidade e segurança** (falaremos mais disso em uma unidade futura). Além desta questão, temos sempre que estar atentos a qualidade dos dados, pois em saúde alguns aspectos como **dados faltantes, inexatos** ou até mesmo **incorrectos** podem ser comuns no processo (conforme discutimos na segunda unidade). Portanto, a devida atenção deve ser dada na etapa de Preparação de dados.

Como qualquer área de estudo específica, a **experiência de especialistas** é primordial para atingir o sucesso em projetos de ML em saúde, pois eles detêm o conhecimento atrelado ao problema que está se tentando resolver. Por exemplo, em um modelo de previsão de hipertensão, o diagnóstico de diabetes estaria relacionado à probabilidade de hipertensão com base em dados e evidências médicas. Sendo assim, essa informação pode ser utilizada para melhorar a performance das categorizações.

Por ser uma área que pode ter impacto direto na saúde de pessoas, muitas questões éticas são levantadas, assim, a implantação de um modelo de ML em alguns momentos deve ser adiada, para que um processo de **disseminação e validação do projeto** seja realizado. Este processo pode ocorrer internamente, na própria organização, com apoio de funcionários, ou até mesmo em uma publicação científica submetida à um periódico da área. Alguns tópicos que podem ser discutidos são:

- Como lidar com falsos negativos, falsos positivos e a acurácia das predições?
- Quais possíveis implicações das predições feitas pelo modelo?
- Há implicações legais no uso dos dados do paciente?
- O projeto pode, direta ou indiretamente, afetar negativamente no cuidado a algum paciente?
- É necessária uma explicação da predição do modelo?

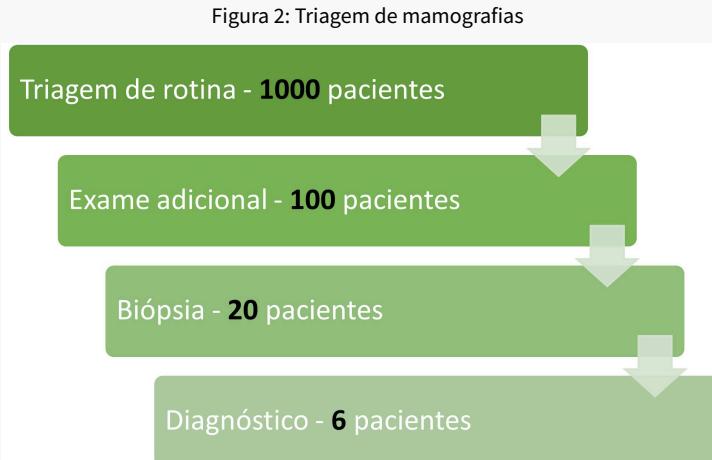
Quanto a questão dos falsos positivos e negativos, imagine o seguinte exemplo: um modelo preditivo que classifique o risco de câncer de mama. Um falso positivo significaria que a paciente é informada de que tem câncer de mama, mas isso deve ser identificado posteriormente no processo de tratamento. No entanto, um falso negativo significaria que uma paciente com câncer de mama não seria notificada - o que é potencialmente muito pior e custoso para o sistema de saúde.

Outra situação que devemos estar atentos é o **engajamento da equipe médica** com o projeto, pois sem ele, o desenvolvimento, e principalmente a implantação do mesmo, fica ainda mais complexa. Tente sempre estabelecer um propósito comum ao iniciar o projeto, deixando claros todos os benefícios que ele pode trazer, oferecendo recompensas à equipe se necessário. Deve-se sempre incluir a equipe no planejamento e trata-los como parceiros na tomada de decisão.

O aspecto da “explicabilidade” é especialmente importante na saúde, pois os profissionais devem entender a resposta do modelo para interagir e confiar no mesmo. Esta questão será aprofundada em uma unidade futura.

### | Um exemplo de aplicação na triagem de pacientes e detecção de câncer de mama

A triagem pacientes através da análise de exames de imagem é papel do radiologista, que verifica os exames e dá seu parecer. Porém, este é um processo custoso, por demandar de um trabalho minucioso do profissional. Por exemplo, um centro hospitalar recebe as imagens de mamografias de 1000 pacientes, que devem ser verificadas pela equipe de radiologia. Dentre estes, 100 pacientes são chamados para realização de um exame adicional. As imagens então são analisadas novamente, para que seja realizada a biópsia em 20 pacientes. E finalmente, 6 destes pacientes são diagnosticados com câncer (este fluxo está resumido na Figura 2).



Fonte: Adaptado das notas de aula de Andrew Beck. Disponível em: [https://mlhc19mit.github.io/slides/lecture13\\_2.pdf](https://mlhc19mit.github.io/slides/lecture13_2.pdf)

Neste contexto, mais de 99% dos pacientes não são diagnosticados com câncer, e o fluxo de trabalho da equipe de radiologia é enorme. Este é um caso em que podemos considerar a utilização de um **modelo de detecção de câncer em imagens médicas**, para realizar a **triagem automática** dos pacientes que não tem diagnóstico de câncer, reduzindo assim os falsos positivos e melhorando a eficiência do processo. Assim, os radiologistas podem ignorar a verificação de mamografias que estejam abaixo de um limite determinado de probabilidade de câncer.

Mesmo com vantagens de utilização visíveis, este tipo de abordagem ainda enfrenta certa resistência, justamente por envolver o risco, mesmo que mínimo, de o algoritmo errar, e fatalmente não diagnosticar o câncer de um paciente. A pandemia do COVID-19, por exemplo, acelerou a utilização deste e outros tipos de tecnologia pelo mundo, pois devido às restrições de funcionamento e distanciamento social, algumas instituições tiveram funcionamento limitado, necessitando assim de novas maneiras para maximizar o fluxo de trabalho (como é o caso deste projeto de pesquisa)

A utilização de algoritmos de ML para detecção de doenças em exames de imagem é algo muito utilizado atualmente. Alguns exemplos práticos deste e de outros tipos de abordagem, serão listados na próxima seção, incluindo um algoritmo para detecção de câncer de mama utilizando SVMs.



## SAIBA MAIS

Neste vídeo vários exemplos de aplicações de ML em saúde são citados.

Fonte: Machine Learning nas Ciências da Saúde - Professor Ives Cavalcante Passos, MD, PhD.

## Exemplos práticos com código e dados

Um dos grandes gargalos na área de estudo de IA em saúde está justamente na dificuldade de acesso aos dados. Devido a isto, esta seção tem o objetivo de apresentar alguns exemplos de soluções de ML em saúde disponíveis para que possamos praticar todos conceitos aprendidos e nos familiarizar com as características deste tipo de projeto.



## EXEMPLO

Detecção de câncer de mama com SVMs

Predição de diabetes utilizando Redes Neurais e otimização de parâmetros com GridSearch

Predição de doenças cardiovasculares utilizando Redes Neurais

Triagem de transtorno do espectro autista infantil utilizando Redes Neurais

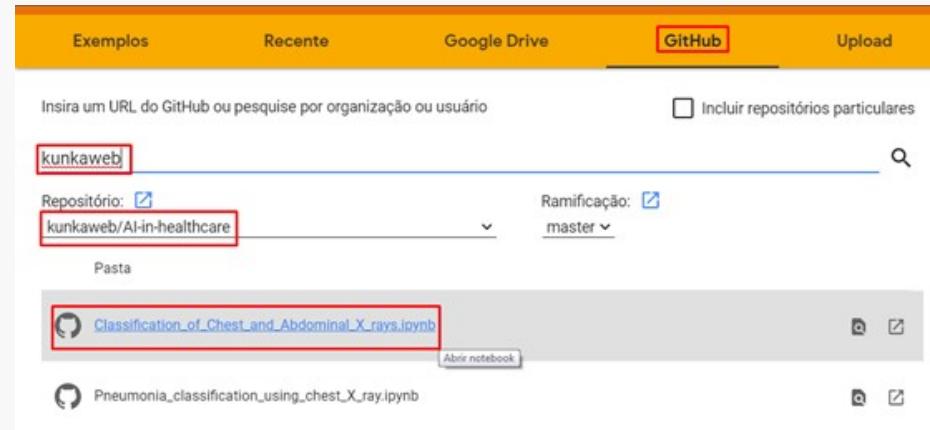
Classificação das radiografias de tórax e abdômen

Classificação da pneumonia usando radiografia de tórax

Você pode optar por executar todos os exemplos em sua própria máquina, utilizando o **Jupyter Notebook**, ou ainda abrir diretamente via **Google Colab**, conforme Figura 3. Fique atento se o código que você executar necessita que seja feito o upload dos dados para seu ambiente.

**Figura 3: Abrir notebook do Github diretamente no Google Colab**

Ao acessar o menu Arquivo > Abrir notebook você pode seguir os passos ilustrados na Figura.



Os exemplos apresentados podem ser separados em dois grandes grupos, os que utilizam dados estruturados para efetuar alguma predição (i.e., predição de diabetes, doenças cardiovasculares e autismo) ou os pertencentes a área de classificação de imagens médicas (i.e., detecção de câncer e manipulação de radiografias). Assista o vídeo a seguir, onde executaremos um dos exemplos apresentados acima.

## Classificando imagens médicas utilizando Redes Neurais e Transfer Learning

Nesta videoaula veremos de maneira prática como ocorre o treinamento de um algoritmo que faz a classificação de exames de imagem.

## Classificando imagens médicas utilizando redes neurais e transfer learning



### Conclusão

---

A partir dos aprendizados desta Unidade, nós:

- Revisitamos classificações de algoritmos de ML;
- Entendemos algumas particularidades no desenvolvimento de projetos de ML em saúde;
- Executamos exemplos práticos de aplicação de ML em dados da saúde.

### Referências Bibliográficas

---

COLICCHIO, T. K. Introdução à informática em saúde: fundamentos, aplicações e lições aprendidas com a informatização do sistema de saúde americano. Porto Alegre: Artmed, 2020. [Minha Biblioteca].

ARTIFICIAL INTELLIGENCE IN HEALTHCARE. [S. l.]: Elsevier, 2020. E-book. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/C20180040979>. Acesso em: 01 jan. 2021.

FACELI, K. Inteligência artificial: uma abordagem de aprendizado de máquina. Rio de Janeiro: Grupo Gen - LTC, 2021. E-book. [Minha Biblioteca].

PANESAR, Arjun. Machine Learning and AI for Healthcare: Big Data for Improved Health Outcomes. Berkeley, CA: Apress, 2019. E-book. Disponível em: <http://link.springer.com/10.1007/978-1-4842-3799-1>. Acesso em: 01 jan. 2021.



© PUCPR - Todos os direitos reservados.