IA Aplicada à Saúde

UNIDADE 07

Ética, Segurança e Privacidade de dados em saúde

Ética, Segurança e Privacidade de dados em saúde

Com o aumento na presença da IA em nosso dia-a-dia, diversas questões morais, éticas e legais são levantadas pela sociedade, e quando falamos de IA em saúde, todos estes aspectos são acentuados, pois as decisões clínicas podem ser uma questão de vida ou morte para um paciente.

Como vimos no decorrer da disciplina, grande parte dos algoritmos de IA em saúde baseiam-se em dados para o treinamento e execução de suas tarefas (i.e., *data-driven AI*). Além disso, entendemos todas as particularidades e complexidades ao lidarmos com dados contidos no prontuário eletrônico do paciente. Portanto, as questões de **ética, segurança e privacidade são essenciais**, por estarmos constantemente manipulando dados sensíveis e executando tarefas relacionadas ao bem-estar e saúde de seres humanos.

Mesmo que a IA possa ajudar no diagnóstico de doenças ou prever o risco de mortalidade, os humanos algum dia preferirão o conselho de uma IA a seu médico? À medida que a humanidade se acostuma a viver lado a lado com sistemas inteligentes, há uma série de obstáculos a superar.

Ética em IA

A ética define uma série de códigos morais de conduto que moldam a sociedade e suas atividades, e a moralidade refere-se aos princípios que diferenciam o comportamento certo do errado. **Mas, e quando falamos de ética na IA?** Neste caso estamos focando em privacidade, tomada de decisão e compartilhamento de dados, compondo três vertentes principais:

1. **Ética dos dados**: concentra-se na geração, coleta, uso, propriedade, segurança e transferência de dados.
2. **Ética da inteligência**: cobre a saída ou resultados da análise preditiva que os dados são usados para desenvolver.
3. **Ética das práticas**: refere-se à moralidade da inovação e sistemas para orientar questões emergentes.

A chamada **ética dos dados**, será coberta em mais detalhes na seção Segurança e Privacidade, já às restantes serão discutidas a seguir. O uso da IA na prática clínica tem um enorme potencial para transformá-la para melhor, mas também levanta desafios éticos relacionados à **segurança do paciente** e **transparência dos algoritmos**.

A **segurança do paciente** é um dos maiores desafios da IA na área da saúde, e mesmo sendo uma área relativamente atual, temos já diversos exemplos de problemas associados a este tema. Um deles é o sistema IBM Watson, que usa algoritmos de IA para avaliar as informações dos prontuários dos pacientes e ajudar os médicos a explorar as opções de tratamento do câncer para seus pacientes. No entanto, em 2018, em relatórios internos da empresa constavam informações sobre recomendações "inseguras e incorretas" dadas pelo sistema para tratamentos de câncer. O problema parece estar no treinamento do Watson, que ao invés de usar dados reais de pacientes, foi treinado apenas com alguns casos clínicos "sintéticos". Declarações dizem que os erros ocorreram apenas como parte do teste do sistema e, portanto, nenhuma recomendação de tratamento incorreta foi dada a um paciente real.

Exemplos como este acabam impactando negativamente a imagem da IA perante a comunidade médica, e também enfatizam a importância quando falamos de segurança e eficácia destes sistemas. **Mas como podemos garantir que as IAs sejam seguras e eficazes?** Para atingir todo o potencial da IA, as partes interessadas, especialmente os desenvolvedores de IA, precisam garantir a **confiabilidade e a validade dos conjuntos de dados e a transparência**.

Primeiramente, os **conjuntos de dados precisam ser confiáveis e válidos**. Quanto melhores forem os dados de treinamento (dados rotulados), melhor será o desempenho do IA. Além disso, os algoritmos geralmente precisam de refinamentos posteriores para gerar resultados precisos.

Em segundo lugar, **alguma transparência deve ser garantida**. Embora em um mundo ideal todos os dados e algoritmos estariam abertos para o público examinar, existem algumas questões legítimas relacionadas à proteção da propriedade intelectual e também ao não aumento do risco de segurança dos dados, que impedem um ambiente 100% transparente. Além disso, os desenvolvedores de IA devem ser suficientemente transparentes, por exemplo, sobre o tipo de dados usados e quaisquer deficiências do software (por exemplo, viés de dados). A transparência cria confiança entre as partes envolvidas no processo, principalmente dos médicos e pacientes, que são essenciais para uma implantação de IA bem-sucedida.

Ainda no tema da transparência, os **sistemas que utilizam abordagens "caixa-preta"** (que já debatemos quando falamos de IA explicável), pode levantar algumas preocupações, pois pode ser complexo determinar um algoritmo 100% transparente neste contexto, mesmo com as possibilidades de explicação dos modelos. Entretanto, talvez não haja necessidade de "abrir a caixa preta", pois em alguns casos os resultados positivos de estudos randomizados ou outras formas de teste sirvam como demonstração suficiente da segurança e eficácia dos IAs.

A IA tem o potencial de melhorar a saúde não apenas em ambientes de alta renda, mas de democratizar e globalizar a saúde. Porém, algoritmos treinados em dados de mundo real, serão tão confiáveis, eficazes e justos quanto os dados com os quais são treinados, ou seja, **se os dados refletem qualquer tipo de desigualdade ou viés, o sistema de IA pode aprender a replicar tais comportamentos**. Portanto, a IA apresenta o risco replicar preconceitos, podendo assim gerar discriminação. Isto posto, é vital que os desenvolvedores de IA estejam cientes desse risco e minimizem potenciais vieses em cada estágio do processo de desenvolvimento da tecnologia. Em particular, eles devem considerar o risco de vieses ao decidir quais tecnologias e procedimentos de ML eles desejam usar para treinar os algoritmos e quais conjuntos de dados, considerando sua qualidade e diversidade, eles desejam usar para a programação.

SAIBA MAIS

A seguir alguns exemplos reais de como a IA pode replicar preconceitos e dificultar a equidade entre gêneros, raças, idade, entre outros.

https://veja.abril.com.br/economia/discriminacao-o-desafio-da-inteligencia-artificial-em-processos-seletivos/

https://www.correiobraziliense.com.br/opiniao/2021/01/4902182-racismo-algoritmico-a-inteligencia-artificial-a-servico-da-discriminacao.html

https://inforchannel.com.br/2020/10/08/como-detectar-e-remover-o-preconceito-na-inteligencia-artificial/

Em saúde, onde informações relacionadas ao fenótipo e genótipo podem estar envolvidas, uma **IA com viés poderia, por exemplo, levar a diagnósticos falsos e tornar os tratamentos ineficazes para algumas subpopulações e, assim, comprometer sua segurança**. Por exemplo, imagine um software de suporte à decisão clínica baseado em IA que ajude os médicos a encontrar o melhor tratamento para pacientes com câncer de pele. No entanto, o algoritmo foi treinado predominantemente em pacientes brancos. Portanto, o software de IA provavelmente fornecerá recomendações menos precisas ou mesmo imprecisas para subpopulações para as quais os dados de treinamento eram insuficientes, como os negros.

Alguns desses vieses podem ser resolvidos devido ao aumento da disponibilidade de dados e tentativas de melhor coletar dados de populações minoritárias e especificar melhor para quais populações o algoritmo é ou não usado de forma apropriada.

Por fim, ainda nos resta uma última discussão. O corpo clínico deve informar ao paciente que uma IA está sendo aplicada aos seus dados, e consequentemente auxiliando na tomada de decisão? Até que ponto, por exemplo, um médico precisa revelar que não pode interpretar totalmente as recomendações de diagnóstico e tratamento da IA? Quanta transparência é necessária? Estas e outras questões ainda estão em aberto e um amplo debate ainda deve ser feito para que as devidas regulações sejam realizadas. Na próxima seção discutiremos acerca da privacidade dos dados do paciente, e como podemos utilizar dados sensíveis e sigilosos para desenvolver nossos algoritmos.

Segurança e Privacidade dos dados

Por utilizarmos dados sensíveis de pacientes, a segurança e privacidade são aspectos fundamentais de todo projeto de IA em saúde. O sigilo dos dados do paciente já é regulado há tempos, e, portanto, sempre foi tema de discussão, desde as primeiras aplicações de IA voltadas ao cuidado médico.

A manutenção e armazenamento dos dados do paciente em sistemas de prontuário eletrônico devem seguir rígidas normas de segurança, para evitar a perda, modificação acidental ou até mesmo vazamento dos dados. Como implementar camadas de segurança em sistemas de informação não fazem parte do escopo desta unidade de aprendizagem, portanto, iremos focar apenas em **como manter a privacidade dos dados em projeto de IA em saúde** (i.e., de-identificação dos dados) e dar uma visão geral de como a **Lei Geral de Proteção aos Dados (LGPD)** complementa a proteção aos dados dos pacientes.

Health Insurance Portability and Accountability Act (HIPAA)

Um dos primeiros países a complementar a proteção dos dados de saúde foram os Estados Unidos, que criaram a HIPAA, uma lei federal de 1996 que exigia a criação de padrões nacionais para proteger as informações confidenciais de saúde do paciente de serem divulgadas sem o consentimento ou conhecimento do paciente.

Entre outras definições, a HIPAA estabelece uma série de dados como ***Protected Health Information*** **(PHI)**, que são os dados de saúde em qualquer forma, incluindo registros físicos e eletrônicos. Portanto, o PHI inclui registros de saúde, resultados de exames laboratoriais e informações financeiras. Essencialmente, todas as informações de saúde são consideradas PHI quando incluem identificadores individuais. As informações demográficas também são consideradas PHI de acordo com as regras da HIPAA, assim como muitos identificadores comuns, como nomes de pacientes, números de previdência social, números de carteira de motorista, detalhes de seguro e datas de nascimento, quando estão vinculados a informações de saúde. Entre os principais identificadores que tornam as informações de saúde em PHI, estão:

- Nomes
- Datas, exceto ano
- Números de telefone
- Dados geográficos
- Números de FAX
- Números de previdência social
- Endereço de e-mail
- Números de registros médicos
- Números de conta
- Números de beneficiários de planos de saúde
- Números de certificado / licença
- Identificadores de veículos e números de série, incluindo placas
- URLs da web
- Identificadores de dispositivos e números de série
- Endereços de protocolo de Internet
- Fotos de rosto inteiro e imagens comparáveis
- Identificadores biométricos (ou seja, leitura de retina, impressões digitais)
- Qualquer número ou código de identificação único

Desde a aprovação da lei da HIPAA, é consenso entre pesquisadores da área de IA em saúde em todo mundo, que todos dados utilizados para treinamento de algoritmos devem passar por um **processo de de-identificação, ou anonimização**, para que todos dados identificáveis do paciente sejam removidos antes da utilização dos dados (ao final desta unidade vamos ver como construir um algoritmo de de-identificação).

Lei Geral de Proteção de Dados (LGPD)

Ao contrário dos Estados Unidos, o Brasil não tem uma lei específica para tratar dados da saúde, como a HIPAA, porém segue padrões similares de acesso à informação. Recentemente, entrou em vigor a LGPD, lei que determina como dados de cidadãos devem ser recolhidos e tratados. O objetivo principal da LGPD é proteger os dados das pessoas da coleta e utilização abusiva por parte das empresas, combatendo, por exemplo, o mercado de comercialização de dados pessoas para fins comerciais, sem que o usuário tenha consentido.

Apesar de não ser uma lei específica para dados de saúde, **o setor da saúde certamente é o setor que mais trata os dados pessoais considerados sensíveis pela LGPD**, portanto, organizações de saúde devem se atentar às suas definições. De maneira geral, as empresas somente poderão manter dados pessoais de usuários com prévia autorização, porém, na área da saúde já existe uma obrigatoriedade (definida por outras leis) de se registrar e armazenar as informações no prontuário do paciente.

> *Os dados dos pacientes podem ser utilizados para melhorias da organização de saúde, desde que respeitada a finalidade de trazer benefícios para o próprio paciente, visando a criação de soluções que melhorem algum aspecto relacionado ao atendimento da instituição.*

Um outro caso em que não é obrigatório o consentimento do uso dos dados é para a realização de estudos por órgão de pesquisa. De qualquer maneira, **comitês de ética em pesquisa pedem que os dados de pacientes sejam de-identificados para sua utilização no treinamento de algoritmos de IA**.

Enfim, a aplicação da LGPD para dados da saúde ainda conta com muitas áreas cinzas, ou seja, em que a definição da lei se sobrepõe a outras bases legais, por exemplo, o médico conta com uma base legal para armazenamento de vários dados sensíveis do paciente. Portanto, muitas discussões ainda ocorrerão no contexto da saúde para termos um consenso definitivo.

De-identificação ou anonimização dos dados

A construção de modelos de IA geralmente requer a construção de um padrão-ouro de dados anotados. No caso da saúde, eventualmente precisamos que uma equipe de especialistas rotule textos extraídos de prontuários de pacientes. Para que seja possível o compartilhamento destes dados, realizamos um processo de de-identificação destes dados, para que nenhum paciente seja identificado pela equipe que tem acesso aos dados.

Este processo geralmente é automatizado por algoritmos de anonimização, e a privacidade é garantida após um processo manual adicional para garantir o funcionamento correto do algoritmo de de-identificação. No vídeo a seguir, vamos ver na prática como construir um algoritmo de de-identificação simples.

Construindo um algoritmo de de-identificação de dados clínicos

Nesta videoaula vamos desenvolver um algoritmo de de-identificação utilizando técnicas de Processamento de Linguagem Natural e Machine Learning.

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].

NASCIMENTO, J. C.; MARQUES, M. L.; COSTA, N. H.; CASTELO BRANCO, M. I.; SANTOS, T. B.; SOUSA, R. V. INTELIGÊNCIA ARTIFICIAL NA SAÚDE E A PROTEÇÃO DE DADOS. Revista Brasileira de Direitos Fundamentais & Justiça, v. 14, n. 1, p. 207-230, 22 dez. 2020. Disponível em: <http://dfj.emnuvens.com.br/dfj/article/view/964>. Acesso em: 01 Jan. 2021.

MACHINE LEARNING AND AI FOR HEALTHCARE: BIG DATA FOR IMPROVED HEALTH OUTCOMES. New York, NY: Springer Science+Business Media, 2018.

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

RIGHTS (OCR), Office for Civil. Privacidade de Informação de Saúde e Norma de Seguranção HIPAA. [S. l.], 2020. Text. Disponível em: https://www.hhs.gov/ocr/get-help-in-other-languages/portuguese.html. Acesso em: 01 Jan. 2021.

LGPD NA SAÚDE: COMO A LEI DE PROTEÇÃO DE DADOS AFETA A ÁREA MÉDICA? [S. l.], 2020. Disponível em: https://telemedicinamorsch.com.br/blog/lgpd-na-saude. Acesso em: 01 Jan. 2021.

LGPD NOS ESTABELECIMENTOS DE SAÚDE: 10 DICAS PRÁTICAS. [S. l.], 2021. Disponível em: https://chcadvocacia.adv.br/blog/lgpd-nos-estabelecimentos-de-saude/. Acesso em: 01 Jan. 2021.