

## Implicações éticas e morais do uso de modelos de machine learning para diagnóstico de cancro

O uso de inteligência artificial na área da saúde apresenta grande potencial, mas também acarreta várias implicações morais e éticas. A área de machine learning (ML), em particular, pode ter um grande impacto em tarefas de diagnóstico de doenças em medicina humana, no entanto a sua integração na prática clínica é ainda limitada devido a questões relacionadas com a confidencialidade e privacidade dos doentes, discordância entre modelo e médico ou questões de responsabilidade, para enumerar apenas algumas [1]. O presente documento apresenta algumas das principais questões éticas e morais inerentes ao uso de modelos de ML no diagnóstico de cancro e como foram abordadas no contexto do desenvolvimento do nosso modelo.

### **Privacidade e confidencialidade dos dados**

A construção de modelos na área da saúde envolve a utilização de dados de pacientes reais que têm o direito à privacidade e confidencialidade. Assim, a utilização de tais dados implica a obtenção do consentimento informado dos pacientes. Além disso, os dados pessoais que permitem rastrear os dados até ao paciente devem ser eliminados tão exaustivamente quanto possível. No presente trabalho, os dados foram obtidos do TCIA, um serviço que desidentifica e hospeda um grande arquivo de imagens médicas de cancro acessíveis para download público. De acordo com a declaração de isenção de responsabilidade da fonte de dados: “Todos os esforços terão sido feitos para remover informações privadas de saúde das imagens, tanto pela fonte das imagens, como por processos de desidentificação automática testados pela UAMS, conforme exigido pela HIPAA [ 45CFR164.514(e)(4)], ou seja, salvaguardas adequadas para garantir que as informações de saúde protegidas não sejam usadas ou divulgadas de forma inadequada. No entanto, os princípios éticos ordenam que todos os utilizadores não façam qualquer tentativa de identificar indivíduos a partir de quaisquer elementos de dados e metadados que restem.” [2].

### **Enviesamento dos dados**

Os dados podem ser enviesados em relação a um determinado género, localização geográfica, religião, raça, estatuto social, etc., o que pode criar e/ou reforçar a discriminação. Assim, este enviesamento dos dados pode comprometer a capacidade de generalização do modelo desenvolvido e torná-lo inútil para aplicações da vida real. Técnicas de recolha e curadoria adequadas de dados devem reduzir o risco de enviesamento. No presente trabalho não temos como controlar esse tipo de problema, uma vez que o conjunto de dados obtido já foi desidentificado.

### **Transparência e explicabilidade do modelo**

A transparência e a explicabilidade de um modelo geram confiança no uso de ML na área da saúde, tanto por médicos quanto por pacientes [3]. O desenvolvimento de um modelo de ML requer a tomada de várias decisões ao longo do processo, o que, juntamente com a baixa interpretabilidade de alguns modelos (por exemplo, modelos de aprendizagem profunda), fazem do modelo uma “caixa preta” para a maioria dos médicos. Assim, qualquer manipulação de dados, engenharia de features, ajuste de hiperparâmetros do modelo e os critérios de avaliação devem ser cuidadosamente documentados para garantir transparência, credibilidade e replicabilidade. No presente trabalho, um grande esforço foi feito para garantir esses princípios de ética digital através da documentação exaustiva de todos os passos dados.

### **Responsabilidades do caso de discordância entre médicos e modelos de ML**

Um diagnóstico contraditório entre o modelo desenvolvido e o clínico levanta várias questões éticas. Os médicos devem seguir o seu diagnóstico quando discordam do resultado do modelo? De que forma esta discordância entra em conflito com os princípios éticos médicos? Estes conflitos entre o modelo e os médicos requerem diretrizes éticas e enquadramentos legais implementados pelas instituições para garantir que os interesses dos pacientes sejam protegidos e também para atribuir responsabilidades em caso de decisão errada. Actualmente, essas diretrizes ainda são escassas e implementadas caso a caso [4], no entanto, há uma necessidade premente de discutir e implementar estratégias para resolver questões de responsabilidade se se pretende evoluir e implementar de forma segura o uso de ML na área da saúde [5]. Estas questões éticas, embora fora do âmbito deste trabalho, foram consideradas relevantes para este documento.

## Referências

- [1] <https://www.cancerimagingarchive.net/disclaimer/>
- [2] Michael, R. (2019). Ethical Dimensions of Using Artificial Intelligence in Health Care. *AMA Journal of Ethics*, 21(2), E121-124. <https://doi.org/10.1001/amajethics.2019.121>
- [3] Pattarabanjird, T., McNamara, C. The clinicians' perspectives on machine learning. *Nat Cardiovasc Res* **1**, 189–190 (2022). <https://doi.org/10.1038/s44161-022-00033-9>
- [4] Alabi, R., Vartiainen, T., Elmusrati, M. (2020) Machine learning for prognosis of oral cancer: What are the ethical challenges? *Proceedings of the Conference on Technology Ethics 2020 - Tethics 2020*, 1-22. <https://urn.fi/URN:NBN:fi-fe2020111690467>
- [5] [Texts adopted - Framework of ethical aspects of artificial intelligence, robotics and related technologies - Tuesday, 20 October 2020 \(europa.eu\)](#) (13/10/2023)