

Algorithmic recourse para explicação de diagnósticos automáticos de Retinopatia Diabética



UFMG
UNIVERSIDADE FEDERAL
DE MINAS GERAIS

Leonardo Santos Miranda¹

¹Departamento de Ciência da Computação - UFMG

380 Words

Student ID:2020726313

Keywords: algorithmic recourse, retinopatia diabética, interpretabilidade

25 de janeiro de 2021

1 Objetivo

O objetivo desse projeto é explicar o diagnóstico automático de algoritmos de aprendizado de máquina para a doença retinopatia diabética. Para isso serão utilizadas técnicas de aprendizado profundo para o diagnóstico das imagens e a técnica *algorithmic recourse* para a explicação do diagnóstico [1]. *Algorithmic recourse* é uma técnica de explicação de predições que foca nas ações que um indivíduo pode realizar para reverter um resultado, possivelmente adverso. Assim, no contexto da aplicação médica de interesse, o objetivo desse projeto é entender o quão diferentes características podem ser mudadas para que a doença seja mitigada.

2 Problema, contexto e motivação

Retinopatia Diabética (RD) é uma complicação das diabetes muito predominante na qual os indivíduos portadores da doença sofrem danos nos vasos sanguíneos da retina, podendo causar cegueira se não tratada [2]. Diagnóstico manual tem dificuldade de encontrar 80% de sensibilidade que é recomendado, por exemplo, no Reino Unido [3].

Sistemas que fazem o uso de algoritmos de aprendizagem de máquina estão sendo cada vez mais utilizados em áreas mais sensíveis, como a medicina, onde uma tomada de decisão pode impactar diretamente a vida de um certo indivíduo [4]. Por este motivo, além dos modelos de aprendizagem de máquina precisarem serem cada vez mais robusto e preciso, valores socialmente relevantes como privacidade, responsabilidade e explicabilidade desempenham um papel importante para a adoção e impacto da tecnologia e também devem ser fortemente considerados, o que acaba sendo um dos maiores problemas destes modelos automatizados.

Através de intervenções podemos imaginar como o mundo seria caso diferentes decisões tivessem sido tomadas, por exemplo: "Seria aquele paciente sido curado se tivessemos aplicado o tratamento X ao invés do tratamento Y?" [5]. A técnica de *algorithmic recourse* possui a motivação em justamente tentar fornecer explicações e recomendações que são raramente consideradas por um modelo totalmente automatizado de aprendizagem de máquina,

3 Dados e técnicas

Para o projeto proposto, será utilizado a base de dados IDRiD (Indian Diabetic Retinopathy Image Dataset) [6]. Essa base de dados possui 516 imagens extraídas por especialistas da área da saúde e se subdividem nos 4 grau de tensidade da Retinopatia Diabética, além de

arquivos tabulares (.csv) para auxílio. A abordagem utilizada para trabalhar na base de dados será a aplicação da técnica *Algorithmic recourse* e inferência causal, buscando reverter um resultado possivelmente adverso e entender quais mudanças podem ocorrer através de possíveis interferências em suas características.

4 Referências

1. Karimi, A.-H., Barthe, G., Schölkopf, B. & Valera, I. A survey of algorithmic recourse: definitions, formulations, solutions, and prospects. *arXiv preprint arXiv:2010.04050* (2020).
2. Watkins, P. J. Retinopathy. *Bmj* **326**, 924–926 (2003).
3. Stolte, S. & Fang, R. A Survey on Medical Image Analysis in Diabetic Retinopathy. *Medical Image Analysis*, 101742 (2020).
4. Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., Van Der Laak, J. A., Van Ginneken, B. & Sánchez, C. I. A survey on deep learning in medical image analysis. *Medical image analysis* **42**, 60–88 (2017).
5. Prosperi, M., Guo, Y., Sperrin, M., Koopman, J. S., Min, J. S., He, X., Rich, S., Wang, M., Buchan, I. E. & Bian, J. Causal inference and counterfactual prediction in machine learning for actionable healthcare. *Nature Machine Intelligence* **2**, 369–375 (2020).
6. Meriaudeau, P. P. S. P. R. K. M. K. G. D. V. S. F. *Indian Diabetic Retinopathy Image Dataset (IDRiD)* 2018.