The importance of interpretability and visualization in machine learning for applications in medicine and health care

(Vellido, 2020)

Vellido, A. (2020). The importance of interpretability and visualization in machine learning for applications in medicine and health care. *Neural Computing and Applications*, *32*(24), 18069–18083. https://doi.org/10.1007/s00521-019-04051-w

Abstrato

Num curto período de tempo, muitas áreas da ciência fizeram uma transição brusca para métodos dependentes de dados. Em algumas este processo foi possibilitado pelos avanços simultâneos na aquisição de dados e pelo desenvolvimento de redes tecnologias de sistemas. Esta nova situação é particularmente clara nas ciências da vida, onde a superabundância de dados desencadeou uma fluxo de novas metodologias de gestão e análise de dados. Isto pode ser visto como um cenário perfeito para a utilização de aprendizagem de máquinas e técnicas de inteligência computacional para resolver problemas em que a análise de dados mais tradicional abordagens podem ter dificuldades. Mas, este cenário também coloca alguns desafios sérios. Um deles é a interpretabilidade de modelos e explicabilidade, especialmente para modelos complexos não lineares.

Em algumas áreas tais como medicina e cuidados de saúde, não abordar tal desafio pode limitar seriamente as hipóteses de adopção, na prática real, de sistemas baseados em computador que dependam de aprendizagem de máquinas e métodos de inteligência computacional para análise de dados. Neste documento, reflectimos sobre investigações recentes sobre a interpretabilidade e explicabilidade dos métodos de aprendizagem de máquinas e discutimos o seu impacto na medicina e cuidados de saúde. Prestamos atenção específica a uma das formas em que a interpretabilidade e a explicabilidade neste contexto podem ser abordada, que é através da visualização de dados e modelos. Argumentamos que, para além de melhorar a interpretabilidade do modelo como um objetivo em si, precisamos de integrar os peritos médicos na concepção de estratégias de interpretação de análise de dados. Caso contrário, é pouco provável que a aprendizagem de máquinas se torne parte da prática clínica e de cuidados de saúde de rotina.

4 Aprendizagem mecânica interpretável na saúde cuidados e medicina: uma chave para a adopção

Tal como mencionado na introdução, estamos a assistir a um transição radical e extremamente rápida para os dados dependência nas ciências da vida. Este processo está a colocar muita pressão sobre o desenvolvimento de estratégias inovadoras para gestão de dados biológicos, curadoria e, em última análise, análise.

Estes desafios também foram reconhecidos nos domínios particulares dos cuidados de saúde e da medicina, nos quais podem ter tido comparativamente menos repercussões devido ao facto de, na maioria dos casos, não gerarem dados nas grandes quantidades que se estão a tornar comuns, por exemplo, na bioinformática. Mesmo aqui, porém, a potencial complexidade e heterogeneidade dos dados médicos implica que "'ainda não é possível criar um modelo abrangente capazes de considerar todos os aspectos dos sistemas de cuidados de saúde" [34].

Dada a superabundância deste tipo de informação, a utilização de técnicas mineiras para a extração automatizada do conhecimento pode parecer uma ideia impecável. Tem-se argumentado, no entanto, que, a menos que devidamente concebidos e implementados, estes métodos poderiam levar a uma redução de competências entre os peritos médicos. A pressão exercida sobre os médicos para que façam uso dos EHR numa base rotineira pode também levar a relatórios de conteúdo empobrecido devido a restrições de tempo e a um aumento da dificuldade de encontrar um equilíbrio entre um envolvimento pessoal adequado com os pacientes e o cumprimento das directrizes de utilização dos EHR [36].