

健康診断データを用いた疾患予測における解釈可能なモデルの構築

大場 勇貴

健康診断は、検査による測定や問診を通して現在の健康状態を把握し、疾患の早期発見を目的とするものである。疾患の早期発見は、疾患の重症化を防ぐことができる。それにより、人々の健康寿命を伸ばすことができ、増え続ける医療費の削減にもつながる。そのため、より早期の段階で疾患の前兆をとらえるために医療データを用いた疾患の予測研究が行われている。それらの研究には、機械学習による手法が取り入れられており、様々な観点から研究が行われている。しかし、それらの研究では、予測性能の向上や新たな手法の提案に注力している。そのため、それらの予測モデルがなぜその予測結果を導いたのかといった解釈性については、大きく触れられていない。

本研究では、健康診断データを用いて疾患の進行について予測を行うモデルを構築し、また予測された理由が分かる解釈可能なモデルの構築を行うことを目的とする。

本研究では、まず機械学習手法である多層ニューラルネットワークによる予測モデルの構築を行った。予測対象は、健康診断が予防対象とする疾患の1つであり、生活習慣病の1つである糖尿病とした。次に、構築した予測モデルを解釈するために、モデルにおける属性の重要度を算出する手法である **Permutation importance** と、入力を摂動させることにより入力と出力の関係を解釈するための手法である **Sensitivity analysis** の2つの手法を用いた。

Sensitivity analysis では、受診者に生活習慣等を尋ねた質問票の質問項目に着目した。それぞれの質問項目の回答を摂動させ、回答を摂動させる前と摂動させた後のデータでそれぞれ予測を行い、それらの予測結果を比較して、質問項目の回答の変化と予測結果の変化が一致するか判断した。

実験の結果、関連研究で用いられている他の機械学習手法よりも、高い予測性能を持つ予測モデルを構築することができた。またデータセットに含まれる属性それぞれについて算出した **Permutation importance** により、構築した予測モデルが糖尿病の診断や判定に用いられる指標を重視しており、またそれら以外の属性も重視した予測を行っていることが分かった。**Sensitivity analysis** による質問項目の回答に対する摂動による予測結果の比較では、多くの質問項目において、回答の変化と予測結果の変化が一致した。一方で、回答の変化と予測結果の変化が一致しない質問項目もあった。

これらの回答の変化と予測結果の変化が一致しない原因が予測性能に起因しているのか、否かを検証するためには、生活習慣や健康状態の経年変化を考慮することができる複数年の健康診断データを入力とする予測モデルを構築する必要があると考えられる。

(指導教員 手塚太郎)