Czym różnią się między sobą testy na koronawirusa?

4 czerwca 2020

Wyniki fałszywie pozytywne, fałszywie negatywne, czułość, specyficzność – to pojęcia oceniające skuteczność testów przeprowadzanych m.in. na obecność koronawirusa. Co oznaczają te sformułowania, wyjaśnia prof. Marek Majewski z Uniwersytetu Łódzkiego. Dla zobrazowania porównuje pacjentów do klientów banku, starających się o kredyt.

Jak tłumaczy naukowiec z Wydziału Matematyki i Informatyki UŁ w opracowaniu przekazanym PAP przez Uniwersytet Łódzki, pojęcia te są używane zarówno w diagnostyce medycznej, jak i w analizie danych, gdzie konstruuje się modele matematyczne.

„Wyobraźmy sobie zbiór danych zawierający informacje o historii kont bankowych klientów pewnego banku. Model klasyfikacyjny może nam w tym przypadku pomóc w odpowiedzi na pytanie, czy dany klient będzie (lub nie będzie) solidnie spłacać kredyt, o który wnioskuje – proponuje prof. Majewski. – Zupełnie podobna sytuacja jest w przypadku pewnych testów diagnostycznych. Mają one odpowiedzieć na pytanie, czy pacjent należy do jednej z dwóch kategorii: jest chory na pewną chorobę, czy też jest zdrowy.”

Naukowiec wyjaśnia, że jakość takich modeli ocenia się za pomocą tych samych technik i pojęć. Potrzeba do tego pewnego zbioru danych, np. stu klientów albo stu pacjentów, o których wiadomo, że będą (lub nie będą) spłacać kredyt albo czy są (lub nie są zakażeni). Taki zbiór nazywa się zbiorem testowym.

W przypadku klientów banku zbiór testowy może być zbudowany na podstawie danych o historii klientów. W przypadku medycyny – będzie to porównanie z pewnym (absolutnie wiarygodnym) testem referencyjnym (tzw. gold standard).

Inne testy ocenia się, porównując je z tym „złotym” testem”. Dla każdej osoby ze „złotego” zbioru przeprowadza się oceniany test i porównuje jego wynik go z danymi rzeczywistymi. Jak wylicza prof. Majewski, istnieją cztery możliwe wyniki takiego porównania.

Po pierwsze – wynik testu jest pozytywny, czyli wg testu pacjent jest zakażony, choć w rzeczywistości jest zdrowy. Taki przypadek nazywa się fałszywie pozytywnym (false positive). Po drugie – wynik testu jest pozytywny i w rzeczywistości pacjent jest zakażony. Wtedy wynik jest prawdziwie pozytywny (true positive). Po trzecie – wynik testu jest negatywny, czyli wg testu pacjent jest zdrowy, choć w rzeczywistości pacjent chory – wynik fałszywie negatywny (false negative). I po czwarte – wynik testu jest negatywny i w rzeczywistości pacjent jest zdrowy. To wynik prawdziwie negatywny (true negative).

Liczbę obserwacji, dla których otrzymano poszczególne wyniki, zestawia się zwykle w tabeli (tzw. macierzy pomyłek). Na podstawie macierzy pomyłek wprowadza się parametry danego testu. Prof. Majewski definiuje najważniejsze z nich.

Współczynnik dokładności ACC (accuracy rate) oznacza liczbę obserwacji sklasyfikowanych poprawnie podzieloną przez liczbę wszystkich obserwacji.

Błąd modelu (error rate) oznacza iloraz obserwacji fałszywie sklasyfikowanych, do liczby wszystkich obserwacji.

Czułość (sensitivity) – mierzy proporcję liczby poprawnych pozytywnych klasyfikacji względem liczby wszystkich (prawdziwie) pozytywnych przypadków. Czułość odpowiada na pytanie, jaką część wyników pozytywnych wykrywa test albo jakie jest prawdopodobieństwo, że test wykonany dla osoby chorej wykaże, że jest ona chora.

Specyficzność (specifity) testu to liczba prawdziwie negatywnych klasyfikacji względem wszystkich (prawdziwie) negatywnych przypadków. Specyficzność informuje, jaką część

wyników negatywnych wykrywa test albo jakie jest prawdopodobieństwo, że dla osoby zdrowej test wykaże, że osoba jest zdrowa.

Precyzja przewidywania pozytywnego (positive predictive value – PPV) mierzy proporcję prawdziwie pozytywnych klasyfikacji względem wszystkich pozytywnych klasyfikacji. Precyzja przewidywania pozytywnego odpowiada na pytanie, ile z pozytywnie sklasyfikowanych przypadków zostało dobrze sklasyfikowanych albo jeśli wynik testu jest pozytywny, to jakie jest prawdopodobieństwo, że badana osoba jest chora.

Precyzja przewidywania negatywnego (negative predictive value – NPV) to stosunek liczby przypadków prawdziwie negatywnie sklasyfikowanych do wszystkich negatywnych klasyfikacji. Wskaźnik odpowiada na pytanie: jeśli wynik testu jest negatywny, to jakie jest prawdopodobieństwo, że osoba badana jest zdrowa?

„W diagnostyce medycznej do interpretacji dwóch ostatnich wskaźników należy podchodzić ostrożnie ponieważ, jak wykazano, zależą on od tzw. chorobowości (czyli liczby osób aktualnie chorych, ang. prevelence). Przykładowo, wraz ze wzrostem chorobowości PPV również rośnie. Niemniej jednak w analizie danych nadal są one stosowane” – mówi prof. Majewski.

Naukowiec zaznacza, że testy, które dają bardzo wysoką dokładność, zwykle są drogie. Tańsze są na ogół tzw. testy przesiewowe. Jego zdaniem, sensownie jest używać takich testów przesiewowych, które mają wysoką precyzję przewidywania negatywnego. Dają one możliwie najmniej wyników fałszywie negatywnych.

Autorstwo: Karolina Duszczyk
Źródło: NaukawPolsce.PAP.pl