

Głębokie sztuczne sieci neuronowe w integracji profilu ekspresji krążących i wewnątrzkomórkowych cząsteczek miRNA pacjentów z rakiem trzustki.

Rak trzustki jest jednym z nowotworów o najwyższej śmiertelności. Według Krajowego Rejestru Nowotworów, w Polsce co roku tę diagnozę słyszy ponad 4 tysiące osób i tylko około 200 z nich przeżywa 5 lat od tego momentu. Wysoka śmiertelność tej choroby wynika z faktu, że objawy raka trzustki pojawiają się zwykle bardzo późno i są niespecyficzne. Ponadto, z powodu wielu wad oznaczenia CA19-9, można stwierdzić, że nie istnieje rzetelny test diagnostyczny z krwi w kierunku raka trzustki.

Przez ostatnie lata uwagę naukowców przyciągały specjalne cząsteczki kwasu rybonukleinowego, których obecność we krwi nadal pozostaje zagadką, a których unikalna stabilność sprawia, że przypisuje im się cechy idealnych biomarkerów. Te cząsteczki, zwane mikroRNA (miRNA), wewnątrzkomórkowo odpowiadają za hamowanie procesu tworzenia białek w komórkach, regulując i odzwierciedlając tym samym stan komórki. Ich obecność na zewnątrz komórek, np. we krwi jako tzw. krążących miRNA, nadal pozostaje nie do końca wyjaśniona. Niektórzy badacze udowadniają, że krążące miRNA we krwi mogą przekazywać informacje i regulować stan innych komórek niż te z których się wywodzą. Inni zaś uważają, że krążące miRNA stanowią jedynie wypuszczone z komórek produkty uboczne, które są kompletnie niepotrzebne i nie pełnią żadnych funkcji. Niezależnie od tego, która z tych teorii jest prawdziwa, w przypadku chorób nowotworowych można spodziewać się innego profilu miRNA we krwi niż u osób zdrowych. Jednakże wydaje się także oczywiste, że taka mieszanka miRNA uwolniona z różnych komórek nie może być łatwa w interpretacji.

Klasyczne metody statystyczne w wielowymiarowych analizach profilu miRNA obciążone są błędem związanym z wielokrotną liczbą porównań. Dlatego też, w rozwiązaniu tego problemu coraz bardziej pomocne są techniki sztucznej inteligencji (data mining), których ideą *per se* jest poszukiwanie wzorów i wiedzy w wielkich zbiorach danych. Pomimo tego, zwłaszcza w przypadku miRNA, modele tworzone z wykorzystaniem tych technik często ulegają zjawisku przeuczenia, które sprawia, że modele te są wyjątkowo nieskuteczne na innych zbiorach danych (nowych grupach pacjentów), a zatem nie mogą być generalizowane.

Ostatnio nastąpił przełom w rozwoju technik sztucznej inteligencji poprzez rozwinięcie technik modelowania głębokimi sieciami neuronowymi (deep learning). Rozwiązania te zrewolucjonizowały np. rozpoznawanie mowy czy widzenie komputerowe i mogłyby zostać z powodzeniem wykorzystane także w takich zadaniach jak rozpoznawanie pacjentów z i bez raka trzustki na podstawie profilu krążących miRNA. Główną zaletą tych metod jest automatyczne poszukiwanie ukrytych cech (deep features) w zbiorze danych, które pozwala na lepsze rozróżnienie specyficznych cech dla problemu badawczego. Oprócz tego, rozwijające się aktywnie badania naukowe w dziedzinie sztucznej inteligencji pokazują, że techniki te znacznie lepiej radzą sobie z tzw. nauczaniem przeniesionym (transfer learning). W nauczaniu przeniesionym próbuje się otworzyć jedną z unikalnych i najbardziej wartościowych cech ludzkiej inteligencji, która pozwala na przeniesienie wiedzy uzyskanej w jednym zdaniu jako wiedzę, która później ulega tylko dalszej szybszej modyfikacji w nowych zadaniach.

W tym projekcie zamierzamy, zatem, zaprojektować i sprawdzić nowatorskie podejście do identyfikacji krążących miRNA jako biomarkerów raka trzustki poprzez zaproponowanie techniki integracji wcześniej uzyskanej wiedzy z analizy wewnątrzkomórkowych miRNA. Według naszej hipotezy wykorzystanie najnowocześniejszych technik sztucznej inteligencji wraz z technikami nauczania przeniesionego z analizy wewnątrzkomórkowych miRNA mogłoby zapewnić długo oczekiwany przełom. W tym celu najpierw zrekrutujemy grupę 44 pacjentów z rakiem trzustki, którzy rozszerzą zebraną już przez nas kohortę pacjentów, a których materiał zostanie następnie poddany sekwencjonowaniu nowej generacji miRNA w surowicy. Ostatecznie, poprzez sprawdzenie 3 zaproponowanych przez nas metod integracji mamy nadzieję dostarczyć dowody pozwalające na przypuszczenie, że nasza technika zapewni innowacyjną skuteczną integrację profili miRNA z różnych źródeł. Pozytywny wynik projektu pozwoliłby znacznie rozszerzyć i przełomowo przyspieszyć poszukiwanie biomarkerów w profilu krążących miRNA raka trzustki i innych chorób onkologicznych.