

Recenzja rozprawy doktorskiej lek. med. Konrada Stawiskiego p.t. “Głębokie sztuczne sieci neuronowe w integracji profilu ekspresji krążących i wewnątrzkomórkowych cząsteczek miRNA pacjentów z rakiem trzustki” na stopień dr. nauk medycznych.

Wprowadzenie do tematyki rozprawy

Rozwój metod sekwencjonowania nowej generacji (“rewolucja NGS”) doprowadził do dramatycznego zwiększenia ilości danych na temat poszczególnego pacjenta, w tym pacjenta z chorobą nowotworową. Wykorzystanie tych danych stwarza nadzieję na wczesne wykrywanie nowotworów poprzez badania przesiewowe profilu (sekwencji i stężeń) kwasów nukleinowych zawartych we krwi obwodowej, w tym mikroRNA (miRNA) – małych niekodujących stosunkowo stabilnych RNA, których wewnątrzkomórkowy profil ekspresji ulega zmianie w procesie transformacji nowotworowej. miRNA zawarte w komórkach, zarówno prawidłowych jak i nowotworowych, dostają się do osocza w wyniku śmierci komórek lub aktywnego wydzielania do pęcherzyków. Ze względu na zmienność ekspresji oraz stosunkowo małą specyficzność pojedynczych rodzajów miRNA, w diagnostyce nowotworów zwykle wykorzystuje się stężenia kilku rodzajów miRNA. Optymalny wybór takich miRNA (tzw. sygnatury) oraz sposób wykorzystania ich stężeń do klasyfikacji próbek stanowi obszar intensywnych badań.

W zarysowany powyżej nurt badań wpisuje się projekt doktorski lek. med. Konrada Stawiskiego, poświęcony możliwości wykorzystania profilu ekspresji miRNA we krwi obwodowej do diagnostyki raka trzustki. Perspektywa wykorzystania biomarkerów z krwi obwodowej do wczesnej diagnostyki raka trzustki ma szczególne znaczenie, ponieważ objawy raka trzustki są niespecyficzne i pojawiają się późno w historii przebiegu choroby. Wybór określonych miRNA i opracowanie opartego o ich stężenia algorytmu wnioskowania służącego do diagnostyki raka trzustki to powiązane zagadnienia z zakresu diagnostyki laboratoryjnej i uczenia maszynowego. Dlatego też, obok rewolucji NGS, projekt doktorski lek. med. Konrada Stawiskiego, bazuje na drugiej dokonującej się obecnie rewolucji – burzliwym rozwoju zastosowań sztucznych sieci neuronowych. Sztuczne sieci neuronowe to metoda uczenia maszynowego inspirowana biologicznymi sieciami neuronowymi, której główną ideą jest modelowanie zmiennej wynikowej za pomocą nieliniowej funkcji zastosowanej do kombinacji liniowej zmiennych wejściowych, przy czym proces ten może być powtarzany wielokrotnie przez kolejne warstwy sieci neuronowej, co wówczas określamy termiem “uczenie głębokie”, a realizujące je sieci mianem głębokich sieci neuronowych. Sieć neuronowa o zadanej strukturze jest poddawana uczeniu w oparciu o dane, czyli wartości zmiennych wejściowych (poziomy ekspresji miRNA) i wynikowych (binarna klasyfikacja próbki jako nowotworowej lub kontrolnej). Uczenie sieci neuronowej polega na modyfikacji wag połączeń pomiędzy neuronami, w taki sposób by zwiększyć poprawność klasyfikacji. Podczas gdy modyfikacja wag dokonywana jest w sposób automatyczny przez oprogramowanie służące do budowy i trenowania sieci, zagadnienie znalezienia optymalnej struktury sieci realizującej dane zadanie – w przypadku projektu doktorskiego jest to wykrywanie osób z rakiem trzustki – stanowi otwarty problem badawczy. *Jedną z zalet przedstawionej rozprawy doktorskiej jest zaproponowanie i implementacja metody systematycznego przeszukiwania przestrzeni struktur sieci neuronowych wraz z oceną ich skuteczności.*

Sieci neuronowe stanowią modele o dużej liczbie parametrów, dlatego ich wykorzystanie wiąże się z ryzykiem nadmiernego dopasowania, czyli dopasowania do przypadkowych koincydencji pomiędzy wartościami zmiennych w konkretnym zbiorze danych. Na szczęście, istnieją sposoby aby ograniczyć nadmierne dopasowanie, które klasycznie opierają się na losowym podziale całości danych na trzy rozłączne podzbiory: pierwszy podzbiór – określany jako zbiór treningowy, w oparciu o który zachodzi uczenie sieci (modyfikacja wag); drugi podzbiór – na którym, równoległe do zbioru treningowego, dokonywana jest ocena skuteczności działania sieci dla aktualnego zestawu jej wag; oraz trzeci podzbiór – służący do oceny najlepszej sieci wybranej w oparciu o dwa poprzednie podzbiory. Należy zauważyć, że w opisanym w

rozprawie doktorskiej zastosowaniu diagnostycznym parametrem stworzonego końcowego klasyfikatora jest również wybór sygnatury, czyli genów miRNA stanowiących zmienne wejściowe, pomimo iż wybór ten został dokonany przed rozpoczęciem treningu głębokiej sieci neuronowej.

Możliwość wykorzystania stężeń miRNA we krwi do wykrywania nowotworów jest komplikowana przez fakt, że na stężenie danego miRNA w surowicy wpływa jego uwalnianie nie tylko przez komórki guza, ale również przez pozostałe komórki organizmu. Dlatego dla sieci neuronowej łatwiejszym zadaniem powinno być rozpoznanie nowotworu na podstawie wewnątrzkomórkowych stężeń miRNA mierzonych w tkance, niż na podstawie stężeń tych samych miRNA w surowicy. Transfer wiedzy pomiędzy sztucznymi sieciami neuronowymi realizującymi pokrewne zadania, w szczególności od sieci realizującej zadanie prostsze do sieci realizującej zadanie bardziej złożone, stanowi obszar aktywnych badań, dlatego kolejnym innowacyjnym aspektem ocenianej pracy doktorskiej jest zbadanie zagadnienia nauczania przenoszonego sieci neuronowych w zastosowaniu do diagnostyki raka trzustki.

Szczegółowa charakterystyka formy i zawartości rozprawy

Przedstawiona do recenzji rozprawa doktorska liczy wraz z załącznikami 217 stron, na które składają się liczący 25 stron Wstęp, przedstawiony syntetycznie "Cel pracy i hipoteza badawcza", liczący 28 stron rozdział "Materiały i metody", najdłuższy, liczący 78 stron rozdział "Wyniki", przedstawiona na 16 stronach Dyskusja oraz wypunktowane na 1 stronie wnioski. W rozprawie zacytowano około 250 pozycji zawartego na 22 stronach (nienumerowanego) piśmiennictwa. Wynik rozprawy przedstawione są w postaci 42 tabel oraz zilustrowane na 62 rycinach. Zwraca uwagę bardzo wysoka jakość edytorska i graficzna całości rozprawy, w której znalazłem jedynie 2 błędy typograficzne: powtórzenie słowa "metabolitów" na str. 12; oraz kilkakrotne, m.in. w tytule rozdziału 6.2, użycie określenia "grupa badawcza" zamiast "grupa badana".

Niezwykle jasno napisany Wstęp przystępnie wprowadza do tematyki doktoratu, dokumentując jednocześnie gruntowną wiedzę Doktoranta, zarówno kliniczno-laboratoryjną, jak i dotyczącą uczenia maszynowego. Rozdział "Materiały i metody" szczegółowo opisuje sposób wykonania pracy, zaś rozdział "Wyniki" wyczerpująco przedstawia jej rezultaty. Dyskusja zawiera ocenę wyników rozprawy z punktu widzenia uczenia maszynowego oraz pod kątem ich użyteczności klinicznej.

Autor uzyskał w laboratorium, a następnie poddał analizie dane z 3 niezależnych kohort pacjentów, oznaczonych: BOSTON, GERMANY i LODZ, do których będę się dalej odnosił za pomocą pierwszych liter nazwy: B, G i L. Kohorta B liczyła 60 pacjentów, kohorta G – 150 pacjentów, kohorta L – 122 pacjentów, łącznie 322. Każda kohorta obejmowała pacjentów z rakiem trzustki oraz pacjentów kontrolnych bez raka trzustki, którymi mogły być zarówno osoby zdrowe jak i pacjenci z przewlekłym lub ostrym zapaleniem trzustki. Połączone kohorty L+B poddano analizie ekspresji wszystkich miRNA metodą miRNA-seq, a uzyskane dane poddano standardowej analizie eksploracyjnej za pomocą metod uczenia nienadzorowanego obejmujących analizę składowych głównych, analizę skupień, identyfikację miRNA o różnicowej ekspresji. Następnie, dane miRNA-seq przeanalizowano nadzorowanymi metodami uczenia maszynowego. W tym celu dane miRNA-seq (próbki) podzielono losowo (w proporcjach 60-20-20) na zbiór treningowy, testowy i walidacyjny. Następnie, za pomocą stworzonego przez Doktoranta autorskiego oprogramowania *OmicSelector*, dokonano wyłonienia sygnatur miRNA, których ekspresja we krwi obwodowej najlepiej rozdziela pacjentów z rakiem od osób kontrolnych. Założono, że sygnatura użyteczna w realiach klinicznych może liczyć do 10 miRNA, dlatego wyboru najlepszej sygnatury dokonano spośród sygnatur o długości do 10 miRNA. Następnie, z wykorzystaniem najlepszej sygnatury, która okazała się mieć długość 10 miRNA, przeprowadzono uczenie głębokich sieci neuronowych w celu wyłonienia sieci zapewniających najwyższą dokładność klasyfikacji, czyli odróżnienia próbek nowotworowych od kontrolnych. Co warto podkreślić, poszukiwania, uczenie

i wybór najlepszego klasyfikatora objęły sieci o różnej strukturze, zbudowane przez systematyczne przeszukiwanie przestrzeni hiperparametrów określającej strukturę sieci i sposób jej uczenia.

W drugim etapie badań, dla połączonych kohort L+G, dokonano metodą qPCR pomiaru ekspresji 10 miRNA zawartych we wcześniej wybranej sygnaturze, uzyskując zbiór danych określony jako walidacyjny. Dla tego zbioru przeprowadzono tę samą co poprzednio standardową eksplorację danych oraz przeprowadzono na nim procedurę uczenia głębokich sieci neuronowych, w sposób identyczny i dla tej samej sygnatury jak wcześniej w oparciu o dane miRNA-seq dla kohorty B+L.

W trzecim etapie badań wykorzystano tkankowe profile ekspresji miRNA zawartych w wybranej wcześniej sygnaturze w próbkach raka trzustki oraz prawidłowych narządów pobrane z referencyjnej bazy TCGA. Również te dane zostały poddane tej samej standardowej analizie eksploracyjnej oraz przeprowadzono na nich taką samą procedurę uczenia głębokich sieci neuronowych, z tą samą wcześniej wybraną sygnaturą. W fazie integracji, dane miRNA-seq z krwi, qPCR z krwi oraz miRNA-seq z tkanek poddano analizie zgodności ekspresji oraz dokonano oceny efektywności nauczania przenoszonego, polegającego na przeniesieniu części struktury i wag najlepszej sieci uzyskanej dla danych wewnątrzkomórkowych i użyciu ich jako punktu startowego dla uczenia na danych z krwi, cały czas dla tej samej wcześniej wybranej sygnatury. Skuteczność klasyfikacji za pomocą sieci uzyskanych w wyniku nauczania przenoszonego porównano ze skutecznością sieci uzyskanych bez przenoszenia. Ponadto, dla każdego zbioru danych i wariantu analizy zbadano wpływ wartości hiperparametrów na dokładność klasyfikacji.

Dla każdego zbioru danych i wariantu analizy, skuteczność najlepszego modelu diagnostycznego oceniano obliczając dokładność, czułość i swoistość na zbiorze treningowym, testowym i walidacyjnym oraz dodatkowo wykonując analizę krzywej ROC. Wartości uzyskane na podzbiorach walidacyjnych porównano z dostępnymi w literaturze wartościami dla klasyfikacji za pomocą wcześniej znanego biomarkera białkowego raka trzustki (CA19-9). Dodatkowo, dla kohorty L+G, dla której dostępne były informacje o czasie przeżycia, oceniono wartości prognostyczne miRNA wchodzących w skład sygnatury i porównano je z wartością dla CA19-9 w tej samej kohorcie. Uzyskane wyniki jednoznacznie wskazują na użyteczność miRNA wchodzących w skład wyłonionej sygnatury w wykrywaniu pacjentów z rakiem trzustki oraz na pozytywny wpływ uczenia przenoszonego na parametry klasyfikatora, potwierdzając tym samym hipotezę roboczą projektu. Ważnym elementem rozprawy jest zawarta w Dyskusji ocena użyteczności klinicznej uzyskanych wyników.

Uwagi i pytania do Doktoranta

Podczas lektury rozprawy nasunęły mi się następujące uwagi:

- 1) We Wstępie, na str. 25 Autor pisze: "Zjawisko przeuczenia (nadmiernego dopasowania) [oznacza że] model wychwytał niezależny od biologii błąd techniczny pomiaru lub metody", podczas gdy nadmierne dopasowanie może zachodzić także do niezależnej od klasy docelowej zmienności biologicznej.
- 2) Określanie zbioru danych uzyskanych metodą qPCR jako "zbiór walidacyjny" jest niefortunne, w sytuacji gdy ta sama nazwa jest wykorzystywana jako określenie jednego z trzech podzbiorów wykorzystywanych w uczeniu maszynowym. Dodatkowo, w tym drugim znaczeniu, nazwy: "zbiór testowy" i "zbiór walidacyjny", użyte zostały odwrotnie niż zwykle w literaturze, w tym w klasycznym podręczniku "The Elements of Statistical Learning" autorstwa Hastie, Tibishirani i Friedman.
- 3) Wyboru najlepszej sygnatury dokonano w oparciu o meta-indeks, czyli średnią harmoniczną dokładności 10 metod klasyfikacji dla podzbiorów: treningowego, testowego i walidacyjnego zbioru miRNA-seq z krwi. Wykorzystanie całego zbioru danych na etapie wyboru sygnatury stwarza problem na etapie oceny właściwości diagnostycznych wykorzystującego tę sygnaturę końcowego klasyfikatora w oparciu o wyniki uzyskane na podzbiórze walidacyjnym. Problem ten

opisano wcześniej, m.in. w publikacji: Ambroise and McLachlan, 2002, PMID: 11983868 oraz na podanym poniżej forum dyskusyjnym: <https://stats.stackexchange.com/questions/64825/should-feature-selection-be-performed-only-on-training-data-or-all-data>.

W rozprawie przedstawiono własności diagnostyczne optymalnych sieci dla każdego z 4 wariantów analizy (miRNA-seq krew, qPCR krew, miRNA-seq tkanka, miRNA-seq krew przeniesienie z tkanki), jednak nie pokazano ani nie omówiono struktur i parametrów tych sieci. Mam wobec tego do Doktoranta następujące pytania:

- 1) Czy struktury optymalnych sieci były podobne, czy też różniły się wyraźnie pomiędzy wariantami analizy? (Być może da się je czytelnie zwizualizować).
- 2) Czy miRNA o istotnej zmianie ekspresji w danym zbiorze danych mają wyższe wagi połączeń niż pozostałe miRNA?
- 3) Czy optymalne sieci korzystały ze standaryzowanych danych?

Powyższe uwagi i pytania nie umniejszają mojej bardzo wysokiej oceny rozprawy doktorskiej lek. med. Konrada Stawiskiego. Na podkreślenie zasługuje także fakt, że Doktorant zdobył finansowanie swoich badań w postaci grantu NCN PRELUDIUM, oraz że wytworzone przez Doktoranta oprogramowanie zostało umieszczone w publicznym repozytorium kodu źródłowego.

Podsumowanie

W podsumowaniu stwierdzam, że przedstawiona rozprawa stanowi oryginalne rozwiązanie problemu naukowego i wykazuje ogólną wiedzę kandydata w dyscyplinie doktoratu oraz umiejętność samodzielnego prowadzenia przez niego pracy naukowej, a zatem odpowiada warunkom określonym w Art. 13 ust. 1 Ustawy z dnia 14 marca 2003 r. o stopniach naukowych i tytule naukowym oraz o stopniach i tytule w zakresie sztuki (Dz. U. poz. 1852) z późn. zmianami. Wnioskuje zatem do Rady Nauk Medycznych Uniwersytetu Medycznego w Łodzi o dopuszczenie lek. med. Konrada Stawiskiego do dalszych etapów przewodu doktorskiego. Jednocześnie, ze względu na wysoki poziom naukowy ocenianej rozprawy wnioskuje o jej wyróżnienie.

Dr hab. Michał Dąbrowski



Warszawa, 16 sierpnia 2021 r.