

Prof. Marek Wesolowski
Katedra i Zakład Chemii Analitycznej
Gdański Uniwersytet Medyczny

Recenzja

pracy doktorskiej mgr Anny Marii Badury pt.:
Wykorzystanie modeli sieci neuronowych do przewidywania aktywności
przeciwdrobnoustrojowej czwartorzędowych soli amoniowych,
wykonanej pod kierunkiem prof. dr hab. Adama Bucińskiego
w Katedrze Biofarmacji Wydziału Farmaceutycznego
Collegium Medicum w Bydgoszczy, UMK w Toruniu

Rozprawa doktorska mgr Anny Badury posiada układ typowy, na który składa się 7 rozdziałów, tj. Wstęp, Cel pracy, Materiały i metody, Wyniki, Podsumowanie, Piśmiennictwo i Oświadczenia autorów. Zasadniczym elementem rozprawy jest zbiór trzech artykułów powiązanych tematycznie i opublikowanych w międzynarodowych czasopismach. Treść tych artykułów stanowią wyniki badań nad zastosowaniem sztucznych sieci neuronowych do przewidywania aktywności przeciwdrobnoustrojowej, tj. przeciwbakteryjnej i przeciwgrzybiczej, czwartorzędowych soli amoniowych na podstawie danych uzyskanych metodami chemii obliczeniowej. Należy podkreślić, że podjęta przez mgr Annę Badurę problematyka badawcza jest nadzwyczaj aktualna i szczególnie istotna, biorąc pod uwagę fakt, że typową, niekorzystną cechą substancji leczniczych stosowanych przeciw drobnoustrojom jest możliwość wystąpienia lekooporności. Lekooporność redukuje skuteczność terapeutyczną preparatów, co skutkuje koniecznością poszukiwania nowych substancji czynnych, które będą w stanie zastąpić dotychczas stosowane preparaty i skutecznie zwalczać infekcje powodowane przez patogenne bakterie i grzyby. W związku z powyższym, realizowane przez mgr Annę Badurę prace nad zastosowaniem zaawansowanych, wielowymiarowych technik obliczeniowych, jakimi są sztuczne sieci neuronowe (ANN), w badaniach nad możliwością przewidywania aktywności przeciwdrobnoustrojowej czwartorzędowych soli amoniowych na podstawie ich struktury chemicznej, są wielce pożądane i wychodzą naprzeciw zapotrzebowaniu na nowe, skuteczne substancje przeciwdrobnoustrojowe.

Ponieważ podstawowym narzędziem badawczym wykorzystanym do realizacji celu pracy są sztuczne sieci neuronowe, mgr Anna Badura we Wstępie do doktoratu wprowadziła czytelnika w problematykę tych metod obliczeniowych. W zwięzłej formie przedstawiła rys historyczny sztucznych sieci neuronowych zwracając uwagę na algorytm propagacji wstecznej, który jako metoda uczenia sieci wielowarstwowych wyeliminował większość ograniczeń ANN i wraz z rozwojem technik komputerowych przyczynił się do istotnego wzrostu zainteresowania tymi metodami. Na uwagę zasługuje również zrozumiały sposób opisu budowy podstawowego elementu każdej sztucznej sieci, tj. sztucznego neuronu. Z uwagi na fakt, że funkcjonowanie każdego neuronu oraz tworzonej przez wszystkie neurony sztucznej sieci neuronowej zależy od rodzaju użytej funkcji aktywacji, Doktorantka przybliżyła również to zagadnienie wskazując na funkcje sigmoidalne z uwagi na prostszy sposób realizacji procesu uczenia.

Prowadzenie obliczeń za pomocą sztucznych sieci neuronowych wraz z uzyskaniem satysfakcjonujących wyników nie jest zadaniem prostym. Wymaga dużego doświadczenia oraz wiedzy teoretycznej. Dlatego na szczególną uwagę zasługują przedstawione w kolejnych podrozdziałach Wstępu zwięzłe opisy struktur sztucznych sieci neuronowych i sposobów uczenia sieci. Doktorantka wyczerpująco scharakteryzowała strukturę sieci jednokierunkowych i sieci rekurencyjnych, zwracając uwagę na ich istotną rolę jako elementu optymalizującego działanie ANN, a ponadto wyjaśniła zasadę uczenia sieci w przypadku uczenia nienadzorowanego oraz nadzorowanego. Omówiła także znaczenie zbiorów – uczącego, testowego i walidacyjnego oraz miarę błędu działania sieci w procesie uczenia. Wartościowy rozdział kończy charakterystyka zastosowania ANN ze szczególnym uwzględnieniem obszaru nauk medycznych.

Bardzo dobrze opracowany Wstęp wskazuje na dogłębne poznanie przez Doktorantkę zagadnień, którymi zajmowała się podczas wykonywania, a następnie publikowania wyników pracy doktorskiej. Trafny wybór sztucznych sieci neuronowych jako metody przewidywania aktywności przeciwdrobnoustrojowej soli amoniowych w oparciu o deskryptory opisujące ich strukturę chemiczną, znajduje także uzasadnienie w dokonanym przez Doktorantkę przeglądzie piśmiennictwa, wskazującym m.in. na dużą skuteczność ANN jako narzędzi do projektowania i poszukiwania nowych substancji farmakologicznie aktywnych, a ponadto, do przewidywania aktywności antyoksydacyjnej, przeciwbakteryjnej i przeciwgrzybiczej substancji organicznych. Istotne znaczenie ma również fakt, że sztuczne sieci nie wymagają programowania za pomocą odpowiednich algorytmów, lecz wykorzystują proces uczenia, tzn. sieć dostosowuje się do wykonania określonego zadania poprzez proces uczenia, korzystając z zestawu pobudzeń i odpowiadających im reakcji. Z tego powodu nie ma potrzeby formułowania skomplikowanych

hipotez, by zbudować model dla poszukiwania określonej zależności. Zastosowanie przez Doktorantkę sieci neuronowych pozwala także na redukcję kosztów syntez i czasu potrzebnego na ich wykonanie, a ponadto, spełnia założenia zielonej chemii.

Cel pracy doktorskiej został sformułowany jasno i logicznie, z wyszczególnieniem trzech szczepów drobnoustrojów (*Escherichia coli*, *Staphylococcus aureus*, *Candida albicans*), względem których mgr Anna Badura badała czwartorzędowe sole amoniowe pod kątem możliwości przewidywania wartości minimalnego stężenia hamującego (MIC) wzrost danego drobnoustroju (model regresyjny ANN) i sklasyfikowania badanych związków na aktywne lub nieaktywne w stosunku do danego drobnoustroju (model klasyfikacyjny ANN). Badała 140 nowo zsyntezowanych na Politechnice Poznańskiej chlorków 3,3'-($\alpha\omega$ -dioksaalkan)-bis-(1-alkilimidazoliowych), o różnej długości łącznika alkanowego i różnych podstawnikach alkilowych. Badane związki to substancje o potencjalnej aktywności przeciw drobnoustrojom, bakteriom i grzybom, których aktywność została scharakteryzowana przy użyciu wartości MIC i logarytmu z odwrotności tej wartości.

Do opisu struktury chemicznej czwartorzędowych soli amoniowych, Doktorantka wykorzystwała deskryptory molekularne. Ponieważ informacje, jakie niosą deskryptory zależą nie tylko od struktury cząsteczki, ale także od algorytmów użytych do ich obliczenia, stosując odpowiednie kryteria, mgr Anna Badura zajęła się wyselekcjonowaniem spośród ponad 3 tysięcy deskryptorów tylko tych, które mają istotny wpływ na przewidywaną aktywność przeciwdrobnoustrojową. Ostatecznie, stosując metodę prób i błędów, różne struktury sieci neuronowych i analizę wrażliwości do oceny istotności deskryptora, Doktorantka wybrała 38 deskryptorów molekularnych. Wprowadzając deskryptory do warstwy wejściowej, a wartości logarytmów z odwrotności MIC jako odpowiedź sieci, metodą prób i błędów, przy użyciu zbioru uczącego obejmującego 60% spośród 140 soli amoniowych i funkcji aktywacji, mgr Anna Badura zaprojektowała odpowiednie modele predykcyjne (regresyjne i klasyfikacyjne). Tak zoptymalizowane modele, różniące się liczbą neuronów w warstwie ukrytej, poddała procesowi uczenia stosując trzy najczęściej używane algorytmy uczenia, a zbiory – testowy i walidacyjny wykorzystwała do potwierdzenia poprawności działania zaproponowanych modeli predykcyjnych sztucznych sieci neuronowych.

Wykonany przez Doktorantkę zakres prac obliczeniowych związanych z realizacją celu badań, dokonanych przy użyciu specjalistycznych programów i korzystając z różnych baz danych, wraz z uzyskanymi wynikami, dyskusją oraz wnioskami został opublikowany w formie trzech artykułów naukowych:

1. Badura A., Krysiński J., Nowaczyk A., Buciński A. Application of artificial neural networks to prediction of new substances with antimicrobial activity against *Escherichia coli*. *Journal of Applied Microbiology*, 2021, 130, 40-49 (IF = 3,772; punkty ministerstwa = 70),
2. Badura A., Krysiński J., Nowaczyk A., Buciński A. Prediction of the antimicrobial activity of quaternary ammonium salts against *Staphylococcus aureus* using artificial neural networks. *Arabian Journal of Chemistry*, 2021, 14, 103233 (IF = 5,165; punkty ministerstwa = 70),
3. Badura A., Krysiński J., Nowaczyk A., Buciński A. Application of artificial neural networks to the prediction of antifungal activity of imidazole derivatives against *Candida albicans*. *Chemometrics and Intelligent Laboratory Systems*, 2022, 222, 104501 (IF = 3,491; punkty ministerstwa = 100).

Publikacja wyników prac będących tematem doktoratu w recenzowanych czasopismach o obiegu międzynarodowym świadczy o wysokim poziomie naukowym badań, poprawności merytorycznej i zainteresowaniu tym tematem. Wobec wartościowego dorobku publikacyjnego, zadanie Recenzenta staje się więc swego rodzaju formalnością.

Wyniki badań opublikowanych przez Doktorantkę doprowadziły do potwierdzenia przydatności sztucznych sieci neuronowych do rozwiązywania zadań prognostycznych i klasyfikacyjnych postawionych w celu pracy. Wskazały również, że wartościowych danych dostarcza analiza wrażliwości identyfikując deskryptory szczególnie istotne w poszukiwaniu nieliniowych relacji między aktywnością biologiczną a strukturą chemiczną opisującą badane cząsteczki.

Problem przewidywania za pomocą ANN aktywności przeciwbakteryjnej czwartorzędowych soli amoniowych w stosunku do bakterii *Escherichia coli*, przedstawia publikacja pierwsza. Doktorantka opracowała dwa modele sieci o architekturze jednokierunkowej, tzw. perceptron wielowarstwowy (MLP), składający się z warstw: wejściowej, ukrytej i wyjściowej, zawierających odpowiednio: 20-13-1 neuronów dla modelu regresyjnego i 20-8-2 neuronów w przypadku modelu klasyfikacyjnego. Stosując dwa algorytmy uczenia, BFGS 95 i BFGS 58, dla modeli regresyjnego i klasyfikacyjnego odpowiednio, mgr Anna Badura uzyskała bardzo dobrą jakość przewidywania wartości minimalnego stężenia hamującego, na poziomie 0,87-0,91 (współczynniki regresji liniowej R pomiędzy wartościami MIC przewidzianymi przez ANN, a wyznaczonymi doświadczalnie), dla badanych związków zgrupowanych w trzech zbiorach: uczącym, testowym i walidacyjnym. Druga sieć sklasyfikowała poprawnie aż 95% badanych związków, przypisując je do grup o niskiej lub wysokiej aktywności w stosunku do bakterii

Escherichia coli. Istotnym osiągnięciem jest także identyfikacja deskryptorów, na podstawie analizy wrażliwości, które w największym stopniu wpływają na zdolność do przewidywania (deskryptory VRD2 i VRp2) oraz do klasyfikacji (deskryptory Hypertens-80 i Vindex), przez opracowane modele sieci.

Publikacja druga traktuje o opracowaniu modeli ANN skutecznych jako narzędzia do przewidywania aktywności czwartorzędowych soli amoniowych w stosunku do bakterii *Staphylococcus aureus*. Podobnie, jak w przypadku szczepu *Escherichia coli*, Doktorantka zbudowała dwa modele sieci MLP zawierających odpowiednio: 20-8-1 neuronów dla modelu regresyjnego i 20-7-2 neuronów dla modelu klasyfikacyjnego. Stosując algorytmy uczenia, odpowiednio BFGS 18 i BFGS 6 uzyskała również bardzo dobrą jakość przewidywania wartości MIC (współczynniki R na poziomie 0,91-0,92) oraz klasyfikacji badanych związków do grup o niskiej lub wysokiej aktywności, na poziomie od 91,43% dla zbioru testowego do 100% dla zbioru walidacyjnego. Z danych zestawionych w Tabeli 8 (publikacja druga) wynika, że cztery deskryptory w przypadku modelu regresyjnego: Vindex, VRZ2, VRp2 i VRm2 oraz dwa w przypadku modelu klasyfikacyjnego: Hypertens-80 i GVWAI-80, wnoszą najwięcej informacji do relacji między właściwościami fizykochemicznymi a aktywnością biologiczną badanych soli amoniowych wobec szczepu *Staphylococcus aureus*.

Zdolność ANN do przewidywania aktywności przeciwgrzybiczej badanych związków wobec szczepu *Candida albicans* przedstawiono z kolei w publikacji trzeciej. Zaprojektowane przez mgr Annę Badurę sieci MLP o strukturze 20-7-1 i 20-13-2 dla modeli regresyjnego i klasyfikacyjnego, odznaczały się, tak jak w przypadku szczepów bakterii badanych w dwóch poprzednich publikacjach, dużą zdolnością przewidywania. Algorytmy uczenia, odpowiednio BFGS 93 i BFGS 32, umożliwiły uzyskanie wartości współczynników R w zakresie 0,88-0,91 dla modelu regresyjnego i poprawną klasyfikację badanych związków, od 88,57% dla zbioru testowego do 95,24% dla zbioru walidacyjnego. Analiza wrażliwości (Tabela 9, publikacja trzecia) wskazała, że deskryptory Me, Neoplastic-80 i Infective-80 (model regresyjny) oraz X0A i GG16 (model klasyfikacyjny) odgrywają największą rolę podczas przewidywania aktywności badanych soli amoniowych wobec drożdżaka *Candida albicans*.

Podsumowując materiał faktograficzny zestawiony w trzech opublikowanych pracach mogę z pełnym przekonaniem stwierdzić, że wszystkie prace badawcze zostały bardzo dobrze zaplanowane, wykonane, a uzyskane wyniki zinterpretowane i opublikowane w wiodących wydawnictwach zagranicznych przypisanych do dyscypliny nauki farmaceutyczne. Łączna wartość współczynnika IF dla tych prac wynosi 12,428, natomiast punktacja ministerstwa =

240. We wszystkich artykułach Doktorantka jest pierwszym i korespondencyjnym autorem. Zgodnie z oświadczeniami współautorów, jest również twórcą koncepcji badań, opracowała metodologię badań, zastosowała techniki obliczeniowe i utworzyła modele predykcyjne do analizy danych doświadczalnych. Ponadto, przygotowała manuskrypty do druku i odpowiadała na uwagi recenzentów. Świadczy to o Jej wiodącym udziale w uzyskaniu satysfakcjonujących wyników obliczeń chemometrycznych. Znajduje to też odzwierciedlenie w logicznie i zwięźle ujętym Podsumowaniu i wnioskach z przeprowadzonych badań, które w rozprawie doktorskiej przedstawiła w formie czterech punktów.

Całości rozprawy doktorskiej dopełnia zamieszczone na końcu pracy Piśmiennictwo. Obejmuje ono 84 pozycje, w zdecydowanej większości profesjonalne, angielskojęzyczne prace i monografie. Należy podkreślić aktualność zebranej literatury, z wyłączeniem kilkunastu prac, wszystkie pozostałe artykuły zostały opublikowane po roku 2000, a z tego prawie dwie trzecie prac pochodzi z ostatniego dziesięciolecia. Cytowane piśmiennictwo przekonuje, że przed rozpoczęciem badań mgr Anna Badura należycie rozpoznała temat, którym zajmowała się w trakcie doktoratu. Potwierdza to również dobre przygotowanie teoretyczne Doktorantki do poprawnej interpretacji uzyskanych wyników i ich krytycznej dyskusji. Do usterek, jakie zauważyłem w piśmiennictwie należy dwukrotne cytowanie tych samych prac; pozycje 53 i 76, 54 i 58 oraz brak konsekwencji w zapisie tytułów czasopism. Doktorantka podawała zamiennie – pełne nazwy, skróty czasopism lub nazwy czasopism pisane małymi literami.

Pod względem redakcyjnym rozprawa doktorska została opracowana bardzo starannie. Układ pracy jest przejrzysty i logiczny, a rozmieszczenie tabel i rycin prawidłowe. Nie stwierdziłem błędów literowych, jedynie w Tabeli 1 brak jest kodu 11 (i wzoru podstawnika R), natomiast na stronach 46 i 58 należy poprawić numerację podrozdziałów.

W świetle wyżej przedstawionej, wysokiej oceny rozprawy autorstwa mgr Anny Marii Badury pt. *Wykorzystanie modeli sieci neuronowych do przewidywania aktywności przeciwdrobnoustrojowej czwartorzędowych soli amoniowych*, stwierdzam, że praca doktorska spełnia warunki formalne i merytoryczne stawiane rozprawom doktorskim w ustawie o stopniach naukowych i tytule naukowym oraz o stopniach i tytule w zakresie sztuki z dnia 14 marca 2003 roku (Dz.U. Nr 65, poz. 595, z późn. zm.), a także wydanych na jej podstawie rozporządzeń. W związku z powyższym wnoszę do Rady Dyscypliny Nauki Farmaceutyczne Collegium Medicum w Bydgoszczy, UMK w Toruniu, wniosek o dopuszczenie mgr Anny Marii Badury do dalszych etapów przewodu doktorskiego.

Biorąc pod uwagę wkład pracy wniesiony przez Doktorantkę w trakcie konsekwentnej realizacji celu badań, innowacyjny charakter podjętej problematyki naukowej i opublikowanie wszystkich wyników badań w międzynarodowych czasopismach o wysokiej renomie naukowej, wnioskuję o wyróżnienie rozprawy doktorskiej mgr Anny Marii Badury.

Gdańsk, dnia 23 czerwca 2022 r.

Gdański Uniwersytet Medyczny
Katedra i Zakład Chemii Analitycznej



Prof. dr hab. Marek Wesolowski