



RECENZJA

rozprawy doktorskiej mgra inż. Marcina Ochaba
pt.: „System wspomagania decyzji dla wczesnego wykrywania
dysplazji oskrzelowo-płucnej u wcześniaków”

Promotor rozprawy: prof. dr hab. inż. Wiesław Wajs

Podstawa formalna opracowania recenzji

Recenzję wykonano w odpowiedzi na uchwałę Rady Wydziału Elektroniki, Automatyki, Informatyki i Inżynierii Biomedycznej AGH (pismo Dziekana Wydziału WEAlIB-b/510-24-4/16 z dnia 27.01.2017 r.).

1. Rozwiązywany problem naukowy

Tematyka przedstawionej pracy doktorskiej dotyczy wspomagania decyzji medycznych odnośnie wcześniaków po pierwszym tygodniu życia, o skrajnie lub bardzo małej masie urodzeniowej, zagrożonych dysplazją oskrzelowo-płucną (BPD). Jest to interesujące zagadnienie o charakterze naukowym i dużym znaczeniu praktycznym. Rozwiązywane zadanie polegało na analizie możliwości wykorzystania w tym celu wybranych metod komputerowej predykcji stanu wcześniaka w oparciu o zestawy dostępnych parametrów oraz ocenie efektywności zaproponowanego podejścia. W badaniach uwzględniono regresję logistyczną (RL) i maszynę wektorów nośnych (SVM), oraz retrospektywne dane pozyskane u 109 małych pacjentów.

Celem pracy było opracowanie interaktywnego systemu wspomagania decyzji, mogącego działać w różnorodnych warunkach rzeczywistych, charakteryzujących się niedostępnością do niektórych z używanych parametrów, oferującego jednakże maksymalnie trafną w danych okolicznościach predykcję przyszłej diagnozy oraz dodatkowe sugestie działań mogących zwiększyć jej trafność.

Powyższy cel, odzwierciedlający ukierunkowanie i efekty przeprowadzonych badań, nie został jednak *explicite* zapisany przez Doktoranta. Paragraf „Cele pracy” (str. 21-22) pokazuje motywację pracy oraz zawiera zwięzły opis przeprowadzonych badań.

Teza pracy została przedstawiona na str. 23 w brzmieniu: „Możliwe jest zbudowanie systemu wspomagania decyzji dla predykcji dysplazji oskrzelowo-płucnej po pierwszym tygodniu życia wcześniaków w oparciu o metody regresji logistycznej oraz wektorów wspierających (SVM)”.

Tak postawiona teza jest odpowiednia dla pracy na stopień doktora, mimo to wydaje się, że została ona sformułowana zbyt ogólnikowo. Skoncentrowanie się w pracy jedynie na wykazaniu tezy w powyższym brzmieniu zaowocowałoby badaniami o mniejszym poziomie oryginalności, niż ma to miejsce. Wynika to z faktu, że metoda RL

była już z powodzeniem stosowana w predykcji dysplazji oskrzelowo-płucnej u wcześniaków, a jednocześnie znana jest wysoka skuteczność SVM w analogicznych zadaniach. W istocie praca koncentruje się na systemie wspomagania decyzji działającym optymalnie dla każdego podzbioru parametrów uwzględnionych w pracy – i to jest jej najistotniejszy walor.

2. Charakterystyka rozprawy

Praca ma charakter badawczy, z dobrze zarysowanymi podstawami teoretycznymi, a jej istotą jest synteza elastycznego systemu wspomagającego decyzje medyczne w oparciu o dane rzeczywiste oraz ocena efektywności jego działania, wykorzystująca złożone obliczenia komputerowe.

Rozprawa doktorska mgra inż. Marcina Ochaba, obejmująca 125 stron (poza str. tytułową i streszczeniami), składa się z 9 wyróżnionych części.

Nienumerowany **Wstęp** w jasny sposób prezentuje aspekty medyczne rozwiązywanego problemu, z dobrze opartą o literaturę analizą „Czynników ryzyka BPD”. Dalej następuje systematyczny przegląd dotychczasowych podejść do predykcji BPD w pierwszej fazie życia wcześniaków, z właściwym przedstawieniem używanych w nich danych i metod oraz osiągniętych wyników. Ta część „Wstępu” daje bardzo dobre tło do prezentacji własnych osiągnięć Autora uzyskanych w ramach doktoratu. Zawarto w niej również omawiane już wcześniej „Cele pracy” i „Tezę pracy”.

Rozdział 1 w wyczerpujący sposób prezentuje dane rzeczywiste wykorzystane w badaniach. W pierwszej kolejności przedstawiono ich źródło, zawartość i uzasadniono wybór 14 parametrów zarejestrowanych u wszystkich 109 pacjentów. Parametry te precyzyjnie zdefiniowano i szczegółowo omówiono, dzieląc je na statyczne i dynamiczne. Dalej następuje analiza statystyczna zbioru parametrów, obejmująca testy: zgodności z rozkładem normalnym, różnic parametrów rozkładów cech w dwóch wyróżnionych grupach oraz współliniowości zmiennych losowych. Efektem prac omówionych w tym rozdziale jest wskazanie parametrów różnicujących oraz liniowo skorelowanych.

Rozdział 2 prezentuje szczegółowo matematyczne podstawy dwóch metod predykcji uwzględnionych w pracy, tj. regresji logistycznej (RL) i maszyny wektorów nośnych (SVM), oraz wartościowe porównanie mechanizmów ich działania. Materiał dotyczący RL (jak i własności parametrów oraz analizy ich współliniowości z rozdz. 1) przedstawiono w sposób podobny do pracy doktorskiej dra Pawła Stocha (AGH 2007), zrealizowanej w tym samym zespole badawczym.

Rozdział 3 zawiera omówienie wstępnych badań przeprowadzonych w celu pozyskania dodatkowej wiedzy potrzebnej do właściwego zaproponowania funkcjonalności i struktury opracowywanego modelu wspomagania decyzji. Zdefiniowano w nim podstawowe wskaźniki jakości predykcji (trafność ACC i jej odchylenie standardowe ACC_{dev} , czułość TPR oraz specyficzność diagnozy SPC), które dalej używane były do oceny zastosowanych modeli, jak i bezpośrednio jako element wspomagania decyzji. W tej części badań Doktorant szczególną uwagę poświęcił analizom (w tym porównawczym) wyników klasyfikacji dla wszystkich możliwych podzbiorów danych stosując RL oraz podstawową implementację SVM w Matlabie, oraz dużej ilości modeli z wykorzystaniem biblioteki LibSVM. Dodatkowo przeprowadzono też analizę zdjęć RTG klatki piersiowej wcześniaków, jako potencjalnego źródła wartościowych informacji diagnostycznych. Wyciągnięte wnioski stanowiły kluczowe przesłanki dla dalszych badań.

Rozdział 4, choć niewielki objętościowo, stanowi najważniejszą część pracy, gdyż bezpośrednio omawia badania uzupełniające i ostateczne działania, które doprowadziły do powstania bazy wiedzy (w formie relacyjnej bazy danych MySQL) opartej o uzyskane wyniki oraz implementację systemu wspomaganie decyzji (aplikacja webowa w PHP). Funkcjonalności tego systemu odzwierciedlają najważniejsze z wniosków wynikających z przeprowadzanych badań.

Rozdział 5 ilustruje działanie opracowanego systemu na przykładzie serii eksperymentów poszukiwania najlepszego klasyfikatora posługującego się modelami predykcji o rosnącej liczbie użytych parametrów. Uzyskane wyniki porównano z klasycznymi metodami selekcji cech (selekcji postępującej w przypadku RL i eliminacji wstecznej dla SVM), a następnie bardzo szczegółowo przeanalizowano.

Rozdział 6 zawiera najważniejsze wnioski wynikające z przeprowadzonych badań na temat komplementarności podejść z RL i SVM w rozwiązywanym problemie, jak i dotyczące znaczenia poszczególnych parametrów dla jakości diagnozy. Ponadto mgr M. Ochab przeprowadził w tym rozdziale szeroką dyskusję uzyskanych wyników w odniesieniu do wcześniejszych doniesień literaturowych oraz omówił możliwe kierunki dalszej rozbudowy systemu.

Ostatnie dwie części rozprawy to **Bibliografia** zawierająca 87 pozycji i **Dodatek** prezentujący liczbowe przykłady predykcji BPD w oparciu o RL i SVM.

Praca napisana jest ładnym językiem polskim, na dobrym poziomie edycyjnym, z nielicznymi błędami interpunkcyjnymi. W niektórych częściach dominuje opis w 1. osobie, czego należałoby unikać w dysertacji doktorskiej.

3. Ocena merytoryczna rozprawy

Analiza stanu wiedzy

Doktorant dokonał prawidłowego rozeznania literaturowego w obszarze tematycznym pracy, co pozwoliło na właściwe określenie jej ram w stosunku do aktualnego stanu wiedzy. Efekt tych starań ilustruje bogaty wykaz wykorzystanej literatury, zawierający ok. 45% publikacji z ostatniej dekady. Starsze pozycje dotyczą głównie podstawowych zagadnień medycznych oraz tradycyjnych już metod statystycznych. Wyniki przeprowadzonych studiów literaturowych widać przede wszystkim we „Wstępie”, z zawłascza przy omawianiu „Znanych modeli predykcyjnych BPD”, oraz w dyskusji uzyskanych wyników zamieszczonej we „Wnioskach”.

Metody badawcze

Wykorzystane w pracy metody badawcze należą do obszaru analiz i wnioskowania statystycznego. Zostały one prawidłowo dobrane do rozwiązywanego problemu naukowego, jakim było opracowanie interaktywnego systemu wspomaganie decyzji medycznych. Do metod tych należały m.in.: testy zgodności (Shapiro-Wilka), zmienności (Manna-Whitney'a, chi-kwadrat, Fishera) i korelacji liniowej (współczynniki R^2 i VIF); klasyfikatory RL i SVM; techniki optymalizacji i badania klasyfikatorów (ROC, AUC, walidacja krzyżowa); czy synteza modeli statystycznych. Obejmują one podejścia zarówno tradycyjne, jaki i stosunkowo nowe (SVM, ROC z AUC). Wybór i umiejętne zastosowanie powyższych metod świadczy o wiedzy, dużych umiejętnościach i szerokim warsztacie badawczym Doktoranta.

Prezentacja i analiza wyników

Uzyskane wyniki zostały jasno i precyzyjnie opisane. Autor umiejętnie zilustrował zarówno trudniejsze zagadnienia teoretyczne, jak i osiągnięte wyniki własnych badań,

tworząc rysunki poglądowe, histogramy i przede wszystkim tabele (na str. 3 i 5 znajdują się spisy rysunków i tablic umieszczonych w pracy).

Analizy uzyskanych wyników (rozdz. 3.3, 3.5, 3.7), a zwłaszcza te zawarte w rozdz. 5.3 i 6.1, cechują się szerokim zakresem i dużym poziomem szczegółowości. Odczuwalnym niedostatkiem jest brak jasnych konkluzji o bardziej ogólnym charakterze poznawczym lub aplikacyjnym, wynikających z przeprowadzonych analiz.

Oryginalność rozwiązania problemu naukowego

Postawiony problem naukowy – opracowanie efektywnego systemu wspomaganie decyzji medycznych dotyczących wcześniaków o skrajnie lub bardzo małej masie urodzeniowej i zagrożonych dysplazją oskrzelowo-płucną – został przez mgra M. Ochabą rozwiązany w sposób oryginalny. Stwierdzenie to wynika z kilku faktów. Po pierwsze, w celu predykcji diagnozy wykorzystana została SVM – metoda nie stosowana dotąd w tym konkretnym obszarze wspomaganie diagnostyki. Doktorant zaproponował też komplementarne wykorzystanie obu analizowanych metod (SVM i tradycyjnie stosowanej RL), co umożliwiłoby maksymalizację trafności predykcji w konkretnych sytuacjach. Najważniejszym jednak osiągnięciem przedłożonej pracy doktorskiej jest oparcie systemu wspomaganie decyzji o bazę wiedzy skonstruowaną dla wszystkich możliwych podzbiorów dostępnych parametrów (z analizowanej grupy 14) charakteryzujących stan dziecka zaraz po urodzeniu oraz umożliwienie lekarzowi interaktywnej współpracy z systemem w celu lepszego doboru (np. dodatkowego pomiaru) używanych parametrów i optymalnej w danych warunkach metody predykcji. Tym samym Kandydat dowiódł prawdziwości postawionej na wstępie tezy. Efekty swoich prac zawarł też w 5 publikacjach o zasięgu międzynarodowym (współautorstwo wraz z Promotorem): 1 artykule w czasopiśmie indeksowanym w *Web of Science* (IF = 1.521) oraz 4 referatach w materiałach pokonferencyjnych wydawnictwa Springer.

4. Uwagi krytyczne i zagadnienia do dyskusji

Mgr inż. Marcin Ochab, prowadząc badania naukowe i redagując rozprawę doktorską, nie ustrzegł się kilku niedociągnięć.

- 1) W pracy wykorzystano retrospektywne dane pochodzące od 109 noworodków. Biorąc pod uwagę, że głównym elementem badań były analizy statystyczne prowadzące do budowy bazy wiedzy, a następnie systemu wspomaganie decyzji, taka liczebność próby wydaje się zbyt mała. Wprawdzie prowadzenie pracy w wymienionych warunkach nie można uznać za błąd Doktoranta, lecz za obiektywne ograniczenie zewnętrzne, to jednak niewielka próba statystyczna (co w tym obszarze badań nie jest rzadkością, jak wynika z przeprowadzonej analizy literatury) zmniejsza znaczenie praktyczne uzyskanych wyników (widać to było szczególnie dobrze we wrażliwości SVM na selekcję parametrów podczas walidacji krzyżowej).
- 2) Wielokrotnie w pracy podawana jest liczba wszystkich podzbiorów 14 parametrów, które wykorzystano do budowy modeli predykcyjnych, jako równa 2^{14} . Ta znana z kombinatoryki zależność obejmuje jednak zbiór pusty (który z pewnością nie był badany), wtedy liczba analizowanych kombinacji to $2^{14} - 1$. Jeżeli jednak nie tworzono też modeli dla jednego parametru (co wydaje się wynikać z niektórych opisów), to liczba testowanych podzbiorów wynosiła $2^{14} - 15$.
- 3) Niektóre z testów statystycznych opisanych w rozdz. 1.3 dotyczą median, a nie wartości średnich. Należało zatem w odpowiednich tabelach umieścić także te wartości. Ponadto w niektórych miejscach zasugerowano, że testowane były wartości

średnie (np. pod Tab. 1.2). Jednym z wniosków analizy wyników pokazanych w Tab. 1.2 jest statystycznie istotna różnica rozkładów niektórych parametrów w badanych grupach. Ogólniejszym podejściem dla wykazania takiej różnicy byłoby zastosowanie nieparametrycznego testu Kołmogorowa-Smirnowa.

- 4) W tytule rozdz. 6.1. i innych miejscach pracy powtarza się termin „istotność parametrów” w sensie „znaczenie parametrów dla rozwiązywanego problemu”, co jest w tego typu pracy mylące, gdyż w statystyce określenie istotności parametrów wymaga przeprowadzenia odpowiednich testów.
- 5) W celu zilustrowania potencjału pracy należało z większą starannością dobrać zbliżone do rzeczywistości przykłady obliczeniowe pokazane w „Dodatku A”. Uzyskanie zupełnie sprzecznych prawdopodobieństw wystąpienia BPD za pomocą RL i SVM nie jest mocną stroną dysertacji.
- 6) We wzorze (1.3) brak jest kwadratów przy nawiasach w liczniku i mianowniku.
- 7) Analizowane parametry zostały zobrazowane m.in. za pomocą histogramów empirycznych. Informatywność tych rysunków zwiększyłaby się przy zastosowaniu wszędzie większej i równej liczby słupków (sugeruje się, aby ich liczba k była rzędu \sqrt{N} i jednocześnie $6 < k < 30$, czyli np. 10 w tej pracy).
- 8) Brak jest konsekwencji w stosowaniu niektórych oznaczeń. Np. w pracy pojawiają się zapisy AA i aa , PDA i pda , czy LOGIT zamiast RL.
- 9) Nie podano źródeł zależności (2.4) i (2.7).

Recenzent zwraca się też z prośbą, aby w czasie obrony Doktorant odniósł się do poniższych kwestii:

- Na początku rozdz. 1 bezosobowo wspomniano o pracach pozwalających na bezpośredni eksport wyników badań z monitora medycznego do bazy danych. Czy prace te wykonane zostały przez Kandydata? Jeżeli tak, to warto było ten wartościowy dla pracy aspekt techniczny podkreślić.
- Jednym z elementów badań statystycznych przedstawionych w rozdz. 1.3 jest analiza współliniowości parametrów (Tab. 1.3 i 1.4). Typowo analizę taką prowadzi się na etapie selekcji cech używanych następnie w zadaniu klasyfikacji. Dlaczego, pomimo wykazania współliniowości pomiędzy niektórymi parametrami, w dalszej części badań nadal analizowano wszystkie 14 cech? Co może być powodem, że niektóre z optymalnych modeli wykorzystywały parametry silnie skorelowane liniowo? (Tab. 3.2, 3.4, 4.1)
- Do oceny modeli predykcji użyto m.in. średnią trafność diagnozy ACC i jej odchylenie standardowe ACC_{dev} , zauważono też stosunkowo duże rozrzuty predykcji, zwłaszcza dla SVM. Czy w takiej sytuacji – jako wskaźnika jakości modelu – nie lepiej było użyć wartości oczekiwanej (kwadratu) różnicy między predykcją komputerową a rzeczywistą diagnozą, czyli błędu średniokwadratowego $MSE = (100 - ACC)^2 + ACC_{dev}^2$?
- W rozdz. 3.7 stwierdzono, że „oba typy klasyfikatorów mają dobre zdolności do generalizacji i nie następuje tzw. efekt przeuczenia”. Czy uzyskane wyniki, a zwłaszcza ACC_{dev} często dwukrotnie większe dla SVM niż dla RL, uprawniają do takiego wniosku?

UKP

- Do ostatnich stron zasadniczego tekstu rozprawy brakowało informacji na temat implementacji bardzo przydanych funkcjonalności proponowanego systemu wspomagania decyzji, a mianowicie informowania lekarza o prawdopodobieństwie przyszłego zdiagnozowania dysplazji oskrzelowo-płucnej w oparciu o wybrany model optymalny oraz o opcji automatycznego uczenia się systemu w danym ośrodku na nowych danych pochodzących ze zdiagnozowanych przypadków. Jakie są plany i możliwości (w tym dostępne metody) uzupełnienia systemu o takie moduły?

5. Wnioski końcowe

Zawartość dysertacji pokazuje, że cel pracy został osiągnięty, a postawiona na wstępie teza udowodniona. Przyjęta metodologia, zakres prac oraz uzyskane wyniki świadczą o pracowitości i umiejętności samodzielnego rozwiązywania problemów naukowo-technicznych przez Doktoranta.

Reasumując uważam, że praca doktorska pt. *„System wspomagania decyzji dla wczesnego wykrywania dysplazji oskrzelowo-płucnej u wcześniaków”* **spełnia wymagania**, jakie stawia rozprawom doktorskim *Ustawa o stopniach naukowych i tytule naukowym oraz stopniach i tytule w zakresie sztuki* z dnia 14 marca 2003 r. (z późniejszymi zmianami), w związku z czym stawiam wniosek o dopuszczenie mgr inż. Marcina Ochaba do publicznej obrony przedłożonej pracy.

