



RECENZJA

**rozprawy doktorskiej mgr inż. Darii Hemmerling,
pt.: „Wykorzystanie sygnału mowy
jako źródła informacji diagnostycznej, kontrolnej i prognostycznej
w wybranych problemach medycznych związanych z otolaryngologią”**

Promotor rozprawy: prof. dr hab. inż. Janusz Gajda
Promotor pomocniczy: dr inż. Andrzej Skalski

Podstawa formalna opracowania recenzji

Recenzję wykonano w odpowiedzi na uchwałę Rady Wydziału Elektroniki, Automatyki, Informatyki i Inżynierii Biomedycznej AGH (pismo Dziekana Wydziału WEAIIB-b/510-11-2/14 z dnia 30.10.2018 roku).

1. Rozwiązywany problem naukowy

Tematyka przedstawionej pracy doktorskiej dotyczy wspomagania decyzji medycznych związanych z dysfunkcją narządu głosu oraz wybranym schorzeniem neurologicznym – chorobą Parkinsona. W szczególności, przeprowadzono badania ukierunkowane na ustalenie czy sygnał mowy może być źródłem informacji diagnostycznej, kontrolnej i prognostycznej dla wybranych jednostek chorobowych opisywanych przez otolaryngologów i neurologów. Jest to interesujące zagadnienie o charakterze naukowym i dużym znaczeniu praktycznym. Rozwiązywane zadanie polegało na analizie możliwości wykorzystania w tym celu wybranych metod komputerowej klasyfikacji i predykcji stanu rozważanego systemu w oparciu o zestawy dostępnych parametrów, reprezentujących ilościowe cechy zarejestrowanego sygnału mowy, oraz na ocenie efektywności zaproponowanego podejścia. Na potrzeby prognozowania postępu choroby Parkinsona opracowano metodę estymacji punktacji UPDRS III z sygnału mowy, wykorzystywanej przez lekarzy do kwantyfikacji stopnia zaawansowania tej choroby u badanych pacjentów. W rozważaniach uwzględniono techniki statystycznej analizy danych oraz metody sztucznej inteligencji, włączając uczenie maszynowe. Eksperymenty prowadzono z wykorzystaniem danych pochodzących z bazy Saarbruecken Voice Database (SVD) – obszar analiz związanych z jednostkami chorobowymi zidentyfikowanymi przez otolaryngologię, a także nagrań głosu u osób z chorobą Parkinsona wykonanych przez Doktorantkę w Krakowskim Szpitalu Specjalistycznym im. Jana Pawła II.

Celem pracy było „przeprowadzenie badań zmierzających do ustalenia, w jakim stopniu zmiany w sygnale mowy, dające się wykryć i ilościowo określić na podstawie analiz fonetyczno-akustycznych tego sygnału, mogą być wykorzystane jako źródło informacji diagnostycznej, prognostycznej i kontrolnej przy rozwiązywaniu niektórych problemów medycznych związanych z wybranymi chorobami w otolaryngologii”.

Jabłoński

Doktorantka wskazała na powyższe, jako na główny cel pracy. Dodatkowo sformułowała cel o znaczeniu praktycznym (str. 12): „Uwagę skupiono na możliwościach wykorzystania wyników analizy sygnału mowy do stworzenia systemu, którego celami działania są:

- automatyczna detekcja chorób laryngologicznych,
- prognoza postępu choroby Parkinsona,
- estymacja punktacji w skali UPDRS opisującej nasilenie objawów choroby Parkinsona.”.

Tak postawione cele odwołują się do poczynionego przez Doktorantkę wkładu o charakterze poznawczym i aplikacyjnym, i są odpowiednie dla pracy na stopień doktora. Należy jednak zauważyć, że oryginalność tych przyczynków jest ograniczona dotychczasowymi wynikami uzyskanymi w omawianym zakresie przez innych autorów i zespoły badawcze, opublikowanymi w ostatnich latach w literaturze międzynarodowej (co koreluje z czasem realizacji opisanego w rozprawie projektu, a tym samym pozytywnie wskazuje na aktualność podjętej przez Kandydatkę tematyki). W przedstawionej do recenzji dysertacji realizacja zidentyfikowanych celów stanowi natomiast rodzaj uporządkowania w postaci logicznie spójnego ciągu czynności zrealizowanego z wykorzystaniem efektywnych narzędzi eksploracji danych eksperymentalnych (nie zestawionych dotychczas w opisanej konfiguracji dla tak postawionego zadania), co prowadzi do uzyskania dokładności i powtarzalności zaprojektowanego rozwiązania systemowego użytecznych dla zastosowań biomedycznych – tutaj: w otolaryngologii i neurologii.

Teza pracy nie została sformułowana w dysertacji *explicite*. Pomimo, iż nie określono w sposób jednoznaczny i formalny (np. w Ustawie) konieczności formułowania w dysertacji sekcji o takim właśnie znaczeniu, to biorąc pod uwagę kanon przygotowania prac na stopień naukowy, należy uznać ten fakt za zauważalny brak. Niemniej jednak, mając na względzie poczynione w rozprawie założenia, zdefiniowane cele, logiczną spójność przedstawionych rozważań oraz uzyskane wyniki, uzasadnionym wydaje się w tym względzie stwierdzenie: Wykorzystanie informacji zawartych w sygnale mowy człowieka umożliwia algorytmizację oraz opracowanie narzędzi digitalizacji, wsparcia i automatyzacji procesu podejmowania decyzji diagnostycznych oraz terapeutycznych w otolaryngologii i chorobach neurodegeneracyjnych.

2. Charakterystyka rozprawy

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

Rozprawa doktorska mgr inż. Darii Hemmerling, obejmująca 113 stron, składa się z dziewięciu wyróżnionych części.

Rozdział 1 jest wprowadzeniem do tematyki badań podjętej przez Doktorantkę. W pierwszej kolejności uzasadniono możliwość i korzyści wynikające z wdrożenia procedur technicznych do opisu rozważanego systemu medycznego w kontekście zadania klasyfikacji i regresji, co doprowadziło do postawionego w dysertacji celu – podrozdział 1.1. Dalej następuje systematyczny przegląd dotychczasowych podejść do automatyzacji detekcji i klasyfikacji dysfunkcji narządu głosu oraz w opisie i prognozie procesów neurodegeneracyjnych, ze szczególnym uwzględnieniem choroby Parkinsona. We właściwy sposób przedstawiono tutaj wykorzystywane dane i metody oraz osiągnięte wyniki. Ta część „Wprowadzenia” przynosi dobre tło do prezentacji własnych osiągnięć Autorki uzyskanych w

ramach doktoratu. W końcu następuje zwięzłe przedstawienie struktury pracy, co dodatkowo porządkuje i obrazuje koncepcję Doktorantki dotyczącą postawionego w dysertacji problemu badawczego.

Rozdział 2 przedstawia fizyczne (akustyczne) i medyczne aspekty generowania prawidłowego i zdeformowanego sygnału mowy człowieka, co jest istotne w kontekście zastosowania technicznych schematów jego parametryzacji, analizy i interpretacji uzyskanych obserwacji. Scharakteryzowano w szczególności mechanizmy patologii otolaryngologicznych rozważanych w recenzowanej dysertacji, tj. dysfonia hiperfunkcjonalna, porażenie nerwu krtaniowego wstecznego, ostre zapalenie krtani, zaburzenia spowodowane chorobą Parkinsona. W ostatniej części rozdziału drugiego usystematyzowano dotychczasową wiedzę na temat wykorzystywanych w praktyce medycznej skal klinicznych opisujących zaawansowanie choroby Parkinsona, co jest uzasadnione postawionym w pracy celem badawczym, odnoszącym się do prognozowania postępu choroby Parkinsona na podstawie sparametryzowanego sygnału mowy, a dalej translacji tej parametryzacji do punktacji w skali UPDRS.

Rozdział 3 stanowi najważniejszą część pracy, gdyż bezpośrednio omawia materiał badawczy, tj. zbiory danych rzeczywistych wykorzystywane do realizacji postawionych w dysertacji celów, oraz przyjętą metodologię badań. W pierwszej kolejności przedstawiono źródło sygnałów eksplorowanych na potrzeby opracowania procedury diagnostyki laryngologicznej w oparciu o sygnał mowy – baza Saarbruecken Voice Database (SVD) udostępniona przez Instytut Fonetyki w Saarland, Niemcy. Scharakteryzowano procedurę rejestracji danych oraz ich zawartość, ze wskazaniem na liczebność grup pacjentów dla kolejnych stanów systemu (trzy patologie oraz osoby zdrowe). Następnie wskazano na utworzenie własnych zasobów sygnałów (w Krakowskim Szpitalu Specjalistycznym im. Jana Pawła II w Krakowie, na Oddziale Neurologicznym z Pododdziałem Udarowym i Pododdziałem Rehabilitacji Neurologicznej), stanowiących nagrania głosu 27 pacjentów z chorobą Parkinsona. Techniczne warunki zapisu danych były analogiczne jak w przypadku bazy SVD, co uzasadnia założenie o spójności otrzymywanych wyników i dokonywanych interpretacji w ramach kolejnych etapów przyjętego programu badań. Paragraf 3.2 dysertacji omawia kwestię parametryzacji sygnału mowy, tłumacząc wybór sygnału akustycznego samogłosek o przedłużonej fonacji – [a, e, i, o, u] – jako podstawy diagnostyki medycznej w otolaryngologii i stanów neurodegeneracyjnych. Wskazano na listę trzydziestu miar opisujących periodyczność, zawartość szumów oraz nieliniowość procesu fonacji. Dalej, w dysertacji następują definicje kolejnych parametrów, kluczowych dla obiektywizacji wniosku medycznego na podstawie sygnału mowy człowieka jako nośnika informacji o stanie rozważanego systemu. W dwóch kolejnych podrozdziałach skupiono się na technicznych aspektach wydobywania informacji diagnostycznej (podrozdział 3.3) oraz kontrolnej i prognostycznej (podrozdział 3.4) z sygnału mowy człowieka. Na potrzeby procesu wniosku diagnostycznego zestawiono w schemacie algorytmicznym dwa zasadnicze etapy eksploracji danych: redukcję wektora danych wejściowych oraz operację klasyfikacji (binarnej oraz multiklasyfikacji), omawiając podstawy matematyczne i implementacji wybranych, dedykowanych do tego narzędzi analizy danych, tj. nieliniowa analiza składowych głównych, lasy losowe, sztuczne sieci neuronowe. W pracy zaproponowano, by zabiegi kontrolne i prognostyczne dotyczące neurodegeneracji w chorobie Parkinsona obserwować i wyrażać ilościowo w zredukowanym farmakologicznie (przez podanie leku lewodopy) horyzoncie czasu (trzy godziny) ewolucji objawów patologicznych. Jest to ciekawa propozycja w kontekście skrócenia czasu badania oraz eliminacji zniekształceń obserwacji wynikających z wahania stanu pacjenta. Założono

dwuetapowy schemat wnioskowania, w którym na podstawie sygnałów akustycznych oceniany jest najpierw stan neurologiczny pacjenta zgodnie ze skalą UPDRS część III, a następnie dokonywana jest predykcja jego stanu w 120 i 180 minucie od przyjęcia leku. Wskazano na metodę regresji wektorów nośnych (SVR, ang. *Support Vector Regression*) i sztuczne sieci neuronowe ze wsteczną propagacją błędów jako wybrane narzędzia, dedykowane do realizacji odpowiednio pierwszego i drugiego etapu wnioskowania ilościowego wyrażonego w jednostkach skali UPDRS III. Konsekwentnie, wprowadzono podstawy matematyczne SVR, omawiając z kolei projekt i implementację sztucznej sieci neuronowej w zadaniu prognozowania. Zaprezentowana w **Rozdziale 3** algorytmizacja procesu wnioskowania ilościowego na temat stanu (bieżącego i przyszłego) rozważanego obiektu medycznego warunkuje jakość i powtarzalność obserwacji relacjonowanych oraz interpretowanych w kolejnych częściach dysertacji.

Rozdział 4 ilustruje działanie opracowanego systemu procedur dedykowanych do diagnostyki oraz wspierania podejmowania decyzji medycznych na podstawie dostarczonych, obiektywnych informacji ilościowych dotyczących schorzeń otolaryngologicznych i neurodegeneracyjnych, uzyskanych z przetwarzania zarejestrowanych sygnałów mowy człowieka. Zapisano schematy blokowe procesów przetwarzania danych eksperymentalnych oraz zaraportowano szczegóły dotyczące parametryzacji procedur programowych używanych podczas obliczeń. Uzyskane wyniki bardzo szczegółowo przeanalizowano i porównano z opisami dostarczonymi przez lekarzy specjalistów.

Rozdział 5 podsumowuje czynności przeprowadzone przez mgr. inż. Darię Hemmerling, ukierunkowane na wykazanie możliwości wykorzystania sygnału mowy człowieka jako źródła informacji diagnostycznej, kontrolnej i prognostycznej w otolaryngologii oraz neurodegeneracyjnej chorobie Parkinsona. Doktorantka stwierdza osiągnięcie założonych celów poznawczych i aplikacyjnych. Podsumowanie zawiera także najważniejsze wnioski na temat komplementarności zaproponowanej metodologii, właściwości zaprojektowanych i zbudowanych struktur programowych, a także znaczenia poszczególnych sygnałów (samogłoska, intonacja) i parametrów dla jakości diagnozy. W końcowej części następuje wskazanie na możliwą ewolucję zaproponowanej w dysertacji metodologii i systemu diagnostyczno-kontrolno-prognostycznego, uwzględniającą także zwiększenie populacji analizowanych pacjentów, adaptację rozwiązania na platformy urządzeń mobilnych oraz włączenie do opisu jednostek chorobowych/podmiotów nie rozpatrywanych w recenzowanej pracy.

Ostatnie cztery (nienumerowane) części rozprawy, to **Bibliografia** zawierająca 121 pozycji, **Dodatek A**, **Dodatek B** i **Dodatek C** zawierające odpowiednio: macierze pomyłek klasyfikacji wybranych chorób laryngologicznych, informacje na temat parametryzacji sztucznych sieci neuronowych zbudowanych do analizy kolejnych wielkości wejściowych, zestawienie wyników predykcji przebiegu choroby Parkinsona dla kolejnych parametrów wejściowych, w teście po podaniu lewodopy.

Ponadto, dysertacja zawiera: Stronę tytułową, Oświadczenie o oryginalności i samodzielności pracy Autorki, Podziękowania, Spis treści, Deklarację na temat źródła finansowania badań i prowadzonej w ich ramach współpracy, Streszczenie w języku polskim i angielskim, a także Spis najważniejszych oznaczeń i skrótów.

Praca napisana jest ładnym językiem polskim, na dobrym poziomie edycyjnym. Stwierdzono pojedyncze błędy językowe (zwłaszcza związane z podziałem wyrazów, np. „nie-liniowe”, „polsko-języczne”, itp.), stylistyczne, błąd w konstrukcji zdania oraz błędy interpunkcyjne; przytaczane zależności w postaci wzorów należy uznać za elementy zdania i stosować dla nich interpunkcję. W opracowaniach naukowych należałoby także unikać

J. B. S.

niejednoznacznych stwierdzeń w formie przymiotnikowej, jak np. „Uzyskane wyniki należy uznać za obiecujące.”. Brak jest odwołania w tekście do Rys. 3.1, oraz do pozycji bibliografii nr [106] i [107]. Zanotowano pojedyncze przypadki kiedy występowanie obiektu typu rysunek poprzedzało w tekście jego wywołanie. W zapisie zmiennych nie stosowano rodzajów czcionek przyjętych jako kanon dla dokumentów naukowych i technicznych: zmienna skalarna – kursywa, wektory i macierze – czcionka pogrubiona.

3. Ocena merytoryczna rozprawy

Analiza stanu wiedzy

Doktorantka dokonała 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 blisko 60% publikacji z ostatniej dekady. Starsze pozycje dotyczą głównie podstawowych zagadnień medycznych oraz tradycyjnych już metod eksploracji danych eksperymentalnych. Wartościowe jest także przeprowadzone rozeznanie i ukazanie udziału polskich ośrodków badawczych w podejmowanym zakresie rozważań na tle międzynarodowych grup badawczych. Wyniki przeprowadzonych studiów literaturowych widać przede wszystkim we „Wprowadzeniu” oraz w części rozprawy poświęconej opisowi procesu generacji sygnału mowy i mowy zdeformowanej. Dla komplementarności przeprowadzonego rozeznania literaturowego warto byłoby jeszcze wskazać na dotychczasowe zaawansowanie prac związanych z oceną możliwości wykorzystania informacji zawartych w mowie człowieka do diagnozy chorób laryngologicznych i neurodegeneracyjnych w warunkach domowych, np. przy użyciu urządzeń mobilnych.

Metody badawcze

Wykorzystane w pracy metody badawcze należą do obszaru analiz i wnioskowania statystycznego oraz wykorzystującego metody sztucznej inteligencji, w tym uczenia maszynowego. Zostały one prawidłowo dobrane do rozwiązanego problemu naukowego, którym była ocena możliwości wykorzystania zmian w sygnale mowy człowieka jako źródła informacji diagnostycznej, kontrolnej i prognostycznej przy rozwiązywaniu wybranych problemów medycznych związanych z chorobami otolaryngologicznymi i neurodegeneracyjnymi. Do metod tych należały m.in. analiza składowych głównych, lasy losowe, regresja wektorów nośnych (SVR), czy narzędzia bazujące na procesie uczenia z klasy sztucznych sieci neuronowych. Obejmują one podejścia zarówno tradycyjne, jak i stosunkowo nowe (lasa losowe, SVR). Wybór i umiejętne zastosowanie powyższych metod i narzędzi świadczy o wiedzy, dużych umiejętnościach i szerokim warsztacie badawczym Doktorantki.

Prezentacja i analiza wyników

Uzyskane wyniki zostały jasno i precyzyjnie opisane. Autorka umiejętnie zilustrowała zarówno trudniejsze zagadnienia teoretyczne, jak i osiągnięte wyniki badań własnych, tworząc schematy blokowe, rysunki ideowe/poglądowe, wykresy i tabele. Dotyczą one zarówno parametryzacji wykorzystywanych narzędzi analizy danych eksperymentalnych, informacji na temat przebiegu operowania tych narzędzi (np. przebiegu procesu uczenia sztucznych sieci neuronowych), jak i obserwacji ilościowych odnoszących się do studiowanego systemu i zachodzących w nim procesów. Wszystkie one (należy przy tej okazji wskazać także na szczegółowość informacji zamieszczonych w trzech dodatkach)

Jolanta

umożliwiają zrekonstruowanie i powtórzenie eksperymentów zaproponowanych przez Doktorantkę w dysertacji.

Analizy uzyskanych wyników cechują się szerokim zakresem i dużym poziomem szczegółowości, natomiast odczuwalny jest brak ich szerszego odniesienia do obserwacji raportowanych przez inne zespoły badawcze w literaturze krajowej i międzynarodowej, m.in. bazujących na analogicznych lub innych sygnałach wejściowych (np. wypowiedane samogłoski i wyrazy czy struktury zdaniowe), zapisach danych zrealizowanych w innych warunkach aniżeli wyspecyfikowano w rozprawie, wykorzystujących alternatywne metody i narzędzia eksploracji danych oraz systemów, co streszczono we „Wprowadzeniu”. Wartościowe byłoby także podanie jasnych konkluzji o bardziej ogólnym charakterze poznawczym lub aplikacyjnym, wynikających z przeprowadzonych analiz.

Oryginalność rozwiązania problemu naukowego

Postawiony problem naukowy – ustalenie w jakim stopniu zmiany w sygnale mowy, dające się wykryć i ilościowo opisać na podstawie analiz fonetyczno-akustycznych tego sygnału, mogą być wykorzystane jako źródło informacji diagnostycznej, kontrolnej i prognostycznej w otolaryngologii i chorobach neurodegeneracyjnych – został przez mgr inż. Darię Hemmerling rozwiązany w sposób oryginalny. Stwierdzenie to wynika z kilku faktów. Po pierwsze, w celach ustalenia diagnozy, kontroli i prognozy zestawiono ze sobą, w postaci spójnego systemu analizy danych, metody i narzędzia nie stosowane dotąd w tym konkretnym obszarze automatyzacji i wspomagania diagnostyki wybranych obiektów biomedycznych. Za szczególnie oryginalne należy uznać tutaj wykorzystanie regresji wektorów nośnych do oceny stanu neurologicznego pacjentów zgodnie ze skalą UPDRS część III na podstawie sygnałów akustycznych, a także sprzężenie zabiegów farmakologicznych ze wskazanymi w rozprawie metodami eksploracji danych i systemów na potrzeby wnioskowania o przebiegu ewolucji neurodegeneracji charakterystycznej dla choroby Parkinsona. Uzyskane w ten sposób rzutowanie obserwacji na trzygodzinny odcinek czasu pozwala wydatnie skrócić czas badania i do podjęcia decyzji diagnostyczno-terapeutycznej, a także zautomatyzować procedurę pomiaru w postawionym zadaniu. Proponując komplementarne wykorzystanie wybranych metod, możliwa jest maksymalizacja dokładności oraz powtarzalności zadań klasyfikacji (zwłaszcza multiklasyfikacji) i regresji. W tym kontekście, najważniejszym osiągnięciem przedłożonej pracy doktorskiej jest opracowana metodologia digitalizacji, wsparcia i automatyzacji procesu podejmowania decyzji diagnostycznych oraz terapeutycznych w otolaryngologii i chorobach neurodegeneracyjnych, a także narzędzie w postaci zbudowanego systemu implementującego procedury przetwarzania sygnału mowy człowieka na zdefiniowane powyżej potrzeby. Efekty swoich prac Kandydatka zawarła też w dziesięciu publikacjach (dwie samodzielnie, współautorstwo z Promotorem i Promotorem pomocniczym) o zasięgu międzynarodowym (stan na styczeń 2019 r., wg bazy Scopus): cztery artykuły w czasopismach aktualnie indeksowanych w Web of Science (w tym dwa z IF; IF = 2.115, IF = 1,694), pięć w materiałach pokonferencyjnych towarzystwa IEEE (2), EUSIPCO (1), INTERSPEECH (1), TSP (1) oraz jeden rozdział w książce wydawnictwa Springer. Ponadto, Kandydatka jest współautorką dwóch artykułów opublikowanych w czasopismach (w tym jedno zestawione na liście MNiSW – lista B).

Należy zauważyć, iż zdefiniowane w rozprawie pole badawcze pozostaje otwarte na dalsze studia o charakterze podstawowym i stosowanym.

4. Uwagi krytyczne i zagadnienia do dyskusji

Mgr inż. Daria Hemmerling, prowadząc badania naukowe i redagując rozprawę doktorską, nie ustrzegła się kilku niedociągnięć:

- 1) Prócz niezmiennego dążenia do zwiększania dokładności i powtarzalności wyników biopomiarów, współczesna inżynieria biomedyczna wyznacza tendencje dla urządzeń dedykowanych do oceny właściwości systemów, uwzględniające m.in. minimalizowanie inwazyjności, czasu trwania testów i udziału pacjenta w realizacji procedury, personalizację metod, minimalizowanie rozmiarów narzędzi pomiarowych oraz ich adaptowalność do zastosowań domowych i mobilnych, itd. Pomimo, iż jednym z celów prac w recenzowanej dysertacji była budowa systemu dedykowanego do pomiarów na potrzeby otolaryngologii i neurologii, nie poczyniono jednoznacznych założeń co do funkcjonalnych cech projektowanego narzędzia, ani też nie skomentowano takich kwestii, także w relacji do prac raportowanych w innych doniesieniach międzynarodowych i krajowych.
- 2) W zadaniach klasyfikacyjnych operujących na danych rzeczywistych, nierzadko stajemy przed potrzebą operowania na zbiorach, gdzie liczba próbek reprezentujących jedną z klas jest znacznie mniejsza aniżeli innej. Jest to częsta cecha medycznych zbiorów danych, a co w recenzowanej rozprawie dotyczy, np. liczebności grup pacjentów z bazy SVD (patrz Tab. 3.1). Nie wskazano w rozprawie na związek między niezbilansowaniem danych a dokładnością multiklasyfikacji w przypadku zastosowania wybranych narzędzi analizy.
- 3) Zagadnienie maksymalizacji dokładności pomiaru i wnioskowania dla ograniczonej liczby informacji wejściowych stanowi otwarty wątek poznawczy, a mała ilość wejściowych informacji (danych) na temat systemu/procesu jest zarazem jednym z argumentów przemawiających za rozwojem i wykorzystaniem w analizie metod sztucznej inteligencji. Należy wziąć tutaj pod uwagę również fakt, że w takim reżymie analiz niezbędne jest operowanie w schemacie uczenie-walidacja-testowanie, z adekwatnymi zbiorami próbek. W rozprawie przyjęto proporcje 80-10-10 dla takiego schematu i posiadanych wektorów danych. Biorąc pod uwagę liczebność zbiorów dla zadania multiklasyfikacji (Tab. 3.1) i/czy prognozowania (27 przykładów), należy liczyć się z niedokładnościami takiego właśnie pochodzenia, co widoczne jest w raportowanych i komentowanych przez Doktorantkę wynikach. Dodatkowo, prognozowanie stanów następnych na podstawie informacji o trzech/czterech poprzednich stanach układu (jak w ocenie ewolucji choroby Parkinsona) może być obciążone znaczącą medycznie niedokładnością, szczególnie w przypadku obserwowania dużej dynamiki w zachowaniu systemu (pacjent 18 i pacjent 27; Rys 4.13 oraz komentarze w tekście rozprawy). Wprawdzie prowadzenie pracy w wymienionych warunkach nie można uznać za błąd Doktorantki, lecz za obiektywne ograniczenie zewnętrzne, to jednak należy się liczyć ze zmniejszonym znaczeniem praktycznym uzyskanych wyników i potrzebą dalszych prac w tym zakresie, do czego Kandydatka odniosła się w „Podsumowaniu”.
- 4) Na stronie 51 (rozdział 3.4) stwierdzono: „W kolejnym etapie prac, przeprowadzono predykcję stanu pacjenta w 120 i 180 minucie od przyjęcia leku.”. Tymczasem w dalszej części pracy (rozdział 4.3, str. 73 i 75) wyspecyfikowano: „Pierwszą rejestrację przeprowadzono, kiedy pacjent znajdował się w stanie *off*, kolejną po 30 minutach od spożycia leków podanych przez lekarza, następnie po 60 i 120 minutach od pierwszego pomiaru. Na etapie uczenia dane wyjściowe sieci neuronowej stanowiły wyniki analizy akustycznej wykonanej w oparciu o nagrania głosu wykonane w 180 minucie od zażycia leków.”.
- 5) W badaniu stopnia neurodegeneracji mózgu w chorobie Parkinsona z wykorzystaniem pobudzenia farmakologicznego i techniki prognozowania danych w postaci szeregów czasowych, do realizacji zadania prognozowania posłużono się jednokierunkowymi sztucznymi sieciami neuronowymi. Biorąc pod uwagę fizyczną naturę przebiegu procesów zakodowanych w zarejestrowanych próbkach sygnałów należy oczekiwać, iż

główny

stan następny systemu jest konsekwencją stanów układu w kilku wcześniejszych punktach czasu. Zatem przyjęta w rozprawie metodologia prowadzi do aproksymacji wynikającej z niedoskonałości modelu predykcyjnego w postaci jednokierunkowej sztucznej sieci neuronowej. Ponadto, jak pokazują to wyniki estymacji punktacji (metodą ϵ -SVR) w skali UPDRS III na etapie kontrolnym (Tab. 4.13), należy się liczyć z dodatkowym obciążeniem prognozowanych wyjść, wyrażonych w wartościach (skali) interpretowanych przez środowiska medyczne.

- 6) W raportach o charakterze naukowym i technicznym przyjęło się stosować regułę uzasadniania kolejnych etapów/czynności w prowadzonym rozumowaniu i/czy budowaniu narzędzi. Autorka operuje na wybranych narzędziach analizy danych, przy czym nie podano w rozprawie jednoznacznej argumentacji przemawiającej za słusznością dokonanych wyborów. Biorąc pod uwagę także stan dotychczasowej wiedzy w studiowanym obszarze zagadnieniowym, uzasadnione byłoby również odniesienie dokonanych obserwacji do wyników badań prowadzonych przez inne zespoły pracujące w omawianym zakresie tematycznym (np. wykorzystujące alternatywne narzędzia ekstrakcji informacji z sygnału mowy człowieka, co streszczono we „Wprowadzeniu”).

Recenzent zwraca się z prośbą, aby w czasie obrony Doktorantka odniosła się do poniższych kwestii:

- Jaką tezę postawiono dla zdefiniowanego, przeprowadzonego i zreferowanego w dysertacji zakresu studiów?

- Na stronach 47-50 rozprawy opisano budowę systemu uczącego się, wykorzystującego sztuczne sieci neuronowe, dedykowanego do rozwiązania zadania klasyfikacji.

Czy i jakie inne algorytmy – wymagające i/lub niewymagające modelowania – rozważano na potrzeby rozwiązania postawionego w rozprawie celu? Co przesądziło o dokonanych wyborach?

- Na potrzeby predykcji przyjęto strategię budowy 33 modeli sztucznych sieci neuronowych, oddzielnie dla każdego parametru akustycznego. Czy rozważano eksperyment, w którym do predykcji stanu przyszłego wykorzystuje się informacje zawarte we wzajemnych relacjach pomiędzy parametrami akustycznymi jako wartościami wejściowymi?

- Proszę przeprowadzić dyskusję uzyskanych w dysertacji wyników (dokładność, złożoność i efektywność opracowanych narzędzi, implikacje dla funkcjonalności systemu pomiarowego) w odniesieniu do obserwacji dotyczących analogicznych eksperymentów, raportowanych w doniesieniach literaturowych.

5. Wnioski końcowe

Zawartość dysertacji pokazuje, że cel pracy został osiągnięty. Przyjęta metodologia, zakres prac oraz uzyskane wyniki świadczą o pracowitości i umiejętności samodzielnego rozwiązywania problemów naukowo-technicznych przez Doktorantkę.

Reasumując uważam, że praca doktorska pt. „Wykorzystanie sygnału mowy jako źródła informacji diagnostycznej, kontrolnej i prognostycznej w wybranych problemach medycznych związanych z otolaryngologią” spełnia wymagania, jakie stawia rozprawom doktorskim – w dyscyplinie biocybernetyka i inżynieria biomedyczna – 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. Darii Hemmerling do publicznej obrony przedłożonej pracy.

Inżynier Jankowski