

Prof. dr hab. inż. Juliusz Kulikowski
Inst. Biocybernetyki i Inżynierii Biomedycznej
im. M. Nałęcz PAN
ul. Ks. Trojdena 4, 02-109 Warszawa

Warszawa, dnia 19.11.2018 r.

Recenzja
osiągnięcia naukowego i istotnej aktywności naukowej dr inż. Agnieszki Wosiak
w związku z postępowaniem w sprawie nadania w/w stopnia naukowego
doktora habilitowanego nauk technicznych

Niniejszą Recenzję przygotowano na podstawie pisma Dziekana Wydz. Fizyki Technicznej, Informatyki i Matematyki Stosowanej Politechniki Łódzkiej z dn. 01.10.2018 r. z powołaniem na pismo Centralnej Komisji ds. Stopni i Tytułów. do Dziekana Wydz. Fizyki Technicznej, Informatyki i Matematyki Stosowanej Politechniki Łódzkiej z dn. 7.09.2018 r. powołującego komisję habilitacyjną w przedmiotowej sprawie.

1. Ogólna sylwetka Kandydatki

Dr inż. Agnieszka Wosiak uzyskała stopień magistra inżyniera informatyki na Wydz. Fizyki Technicznej, Informatyki i Matematyki Stosowanej Politechniki Łódzkiej w 1998 r. oraz stopień doktora nauk technicznych w dziedzinie informatyki, tamże, w 2008 r. na podstawie rozprawy pt. „Inteligentny system analizy danych pomiarowych z urządzeń energetycznych”.

Kandydatka w latach 1998 – 2008 była zatrudniona w Instytucie Informatyki Politechniki Łódzkiej na stanowisku asystenta, a od 2009 r. do chwili obecnej – tamże – na stanowisku adiunkta.

2. Ocena osiągnięcia naukowego

Osiągnięcie naukowe Kandydatki pt.:

Techniki statystycznej i eksploracyjnej analizy danych medycznych

stanowiące podstawę ubiegania się o stopień naukowy doktora habilitowanego zostało (zgodnie z art. 16 ust. 2 ustawy z dnia 14 marca 2003 r. o stopniach naukowych i tytule naukowym oraz o stopniach i tytule w zakresie sztuki (Dz. U. 2016 poz. 882 ze zm. w Dz. U. 2016 poz. 1311.) udokumentowane przez:

a) Monografię naukową:

Wosiak A. (2018). „Techniki statystycznej i eksploracyjnej analizy danych medycznych”. AOW EXIT, Warszawa.

b) Zbiór artykułów naukowych:

1. **Agnieszka Wosiak**, Kinga Glinka, Danuta Zakrzewska (2017). „Multi-label classification methods for improving comorbidities identification”. Computers in Biology and Medicine. DOI: 10.1016/j.compbiomed.,2017.07.006.
2. Liliana Byczkowska-Lipińska, **Agnieszka Wosiak** (2017). „Instance Selection Techniques in Reduction of Data Streams Derived from Medical Devices”. Przegląd Elektrotechniczny. Vol. 93, 12/2017, ss. 115-118.
3. **Agnieszka Wosiak**, Sylwia Karbowski (2017). „Preprocessing compensation techniques for improved classification of imbalanced medical datasets”. IEEE Federated

- Conference on Computer Science and Information Systems (FedCSIS). Annals of Computer Science and Information Systems. Vol. 11, s. 203-211. (Web of Science)
4. **Agnieszka Wosiak**, Danuta Zakrzewska (2017). „Unsupervised feature selection using reversed correlation for improved medical diagnosis”. IEEE International Conference on INnovations in Intelligent SysTems and Applications (INISTA). ss. 18-22
 5. Liliana Byczkowska-Lipińska, **Agnieszka Wosiak** (2016). „Fuzzy classification of medical data derived from diagnostic devices”. Przegląd Elektrotechniczny. Vol. 92(12). ss. 85-88.
 6. **Agnieszka Wosiak**, Agata Zamecznik, Katarzyna Niewiadomska-Jarosik. (2016). Supervised and unsupervised machine learning for improved identification of intrauterine growth restriction types. IEEE Federated Conference on Computer Science and Information Systems (FedCSIS), Annals of Computer Science and Information Systems. Vol. 8. ss. 323-329. (Web of Science).
 7. Kinga Glinka, **Agnieszka Wosiak**, Danuta Zakrzewska (2016). Improving Children Diagnostics by Efficient Multi-label Classification Method, Information Technologies in Medicine. Advances in Intelligent Systems and Computing. Vol. 471(1). ss. 253-266. (Web of Science).
 8. Byczkowska-Lipińska L., **Wosiak A.** Hybrid Classification of High-dimensional Biomedical Tumour Datasets, published by Springer in the book entitled Advanced and Intelligent Computations in Diagnosis and Control (Eds. Józef Korbicz and Marek Kowal) w: Advances in Intelligent Systems and Computing series – Proceedings of 12th International Conference on Diagnostics of Processes and Systems, (Web of Science).
 9. Byczkowska-Lipińska L., **Wosiak A.**: Classification of Medical Data Derived from Diagnostic Devices Using Ensembles of Classifiers, Przegląd Elektrotechniczny, Vol. 91 NR 12/2015, ss. 143 – 146,
 10. **Wosiak Agnieszka**, Zakrzewska Danuta: On Integrating Clustering and Statistical Analysis for Supporting Cardiovascular Disease Diagnosis, Proceedings of the 2015 Federated Conference on Computer Science and Information Systems, Annals of Computer Science and Information Systems ACSIS, Vol. 5, ss. 303–310, 2015, DOI: 10.15439/2015F151, FEDCSIS. (Web of Science)
 11. **Wosiak Agnieszka**, Zakrzewska Danuta: Feature Selection for Classification Incorporating Less Meaningful Attributes in Medical Diagnostics, Proceedings of the 2014 Federated Conference on Computer Science and Information Systems, Annals of Computer Science and Information Systems ACSIS, Vol. 2, ss. 235–240, 2014, FEDCSIS, (Web of Science)
 12. Byczkowska-Lipińska L., **Wosiak A.**: Multimedia Database Techniques for Medical Diagnosis Processes Support, published by Springer in the book entitled Intelligent Systems in Technical and Medical Diagnostics (Eds. Józef Korbicz and Marek Kowal) w: Advances in Intelligent Systems and Computing series, Proceedings of 11th International Conference on Diagnostics of Processes and Systems.
 13. **Agnieszka Wosiak**: On Statistical Analysis Methods Improving Epidemiological Studies, JOURNAL OF APPLIED COMPUTER SCIENCE, Vol. 21 No. 2 (2013), ss. 209-219
 14. Liliana BYCZKOWSKA-LIPIŃSKA, **Agnieszka WOSIAK**: Image signal metadata analysis in diagnosis of Alzheimer’s disease, Przegląd Elektrotechniczny, Vol. 88, NR 12b/2012.

Według oświadczeń Kandydatki, jej udział w pracach współautorskich wahał się od 60 do 90%, co potwierdzają też współautorzy w załączonych oświadczeniach. Spośród wyżej wymienionych artykułów 3 (poz. 4, 11 i 13) nie były punktowane, natomiast punktacja MNSWiT pozostałych 11 artykułów wynosiła od 10 (poz. 11 i 12) do 25 pkt. (poz. 1).

Wśród w/w pozycji 1 (poz. 1) została zamieszczona w czasopiśmie o profilu biomedycznym i szerszym zasięgu międzynarodowym, 2 (poz. 8 i 12) – jako rozdziały w – monografii o zasięgu międzynarodowym wyd. przez Wyd. Springer, 4 (poz. 3, 6, 10 i 11) – w krajowym czasopiśmie (organie PTI) o profilu informatycznym, specjalizującym się m. in. w publikowaniu materiałów pokonferencyjnych FedCSIS, 4 (poz. 2, 5, 9 i 14) – w krajowym czasopiśmie technicznym z dziedziny elektrotechniki, 1 (poz. 13) – w krajowym czasopiśmie o profilu informatycznym i raczej wąskim oddziaływaniu międzynarodowym, oraz 1 (poz. 4)

– w materiałach konferencji międzynarodowej. W sumie ranga w/wym. artykułów mierzona punktacją MNSWiT jest umiarkowana.

Monografia naukowa wym. w p. a) spełnia formalne standardy rozprawy habilitacyjnej. Zawiera ona zbiorcze przedstawienie najważniejszych wyników naukowych dr inż. Agnieszki Wosiak zawartych w publikacjach wym. w p. b), przedstawione na nieco szerszym tle problemów i metod analizy medycznych danych diagnostycznych. Pod względem formalnym monografia liczy 146 stron i składa się z 5 rozdziałów, Bibliografii liczącej 219 pozycji (w tym 13 pozycji sygnowanych nazwiskiem Kandydatki), Spisu rysunków (28 pozycji) i Spisu tablic (38 pozycji).

Dwa pierwsze rozdziały („Wstęp” i „Metody statystyczne w analizie danych medycznych”) mają charakter wprowadzający. Poza przedstawieniem ogólnej charakterystyki danych medycznych i najważniejszych metod ich statystycznej analizy w Rozdz. 2 przedstawiono też charakterystykę 5 przypadków ocen medycznych, z którymi związane zbiory danych są przedmiotem dokładniejszej analizy w następnych rozdziałach: ciśnienia krwi u dzieci hipotroficznym, zaburzeń rytmu serca oraz autonomicznego układu nerwowego u dzieci hipotroficznym, funkcji rozkurczowej i skurczowej lewej komory serca u dzieci z nadciśnieniem tętniczym pierwotnym, skuteczności i powikłań leczenia farmakologicznego przetrwałego przewodu tętniczego indometacyną lub ibuprofenem u noworodków ze skrajnie małą masą urodzeniową oraz skuteczności stosowania reżimu płynowego w leczeniu monosymptomatycznego moczenia nocnego u dzieci. W każdym z powyższych 5 przypadków przedstawiono tło medyczne przypadku, zastosowane metody badań statystycznych danych doświadczalnych, uzyskane wyniki analizy danych i wniosek końcowy. Te ostatnie można streścić jak następuje:

1. Techniki statystycznej analizy danych potwierdziły hipotezę o istnieniu zależności pomiędzy wewnątrzmacicznym zahamowaniem wzrostu płodu, a wystąpieniem u tych dzieci nadciśnienia tętniczego.

2. Komputerowa, statystyczna analiza danych pozwoliła na wykrycie istnienia zaburzeń autonomicznego układu nerwowego u dzieci hipotroficznym.(...) Zastosowanie technik statystycznej analizy danych w procesie oceny parametrów elektrokardiograficznych oraz zmienności rytmu serca stanowi użyteczną metodę wykrywania zaburzeń autonomicznego systemu nerwowego.

3. Zastosowanie technik komputerowej analizy statystycznej do oceny parametrów uzyskanych z wykorzystaniem metody doplera tkankowego u dzieci z nadciśnieniem pierwotnym umożliwiło wykrycie zaburzeń funkcji rozkurczowej i skurczowej lewej komory.

4. Skuteczność zamykania istotnego hemodynamicznie przetrwałego przewodu tętniczego u noworodków z wyjątkowo niską masą urodzeniową leczonych ibuprofenem lub indometacyną, w rutynowo stosowanych dawkach jest porównywalna. Częstość występowania ciężkich powikłań (NEC, perforacja przewodu pokarmowego, małopłytkowość i zgon) jest podobna w obu grupach pacjentów, natomiast zaburzenia funkcji nerek częściej występują w grupie dzieci leczonych indometacyną.

5. Stosowanie reżimu płynowego u dzieci z cierpiących na monosymptomatyczne moczenie nocne wykazuje stosunkowo niską (18%) i późną (3 miesiące) skuteczność. Analiza danych wykazała, że tylko u pacjentów o korzystnym profilu, związanym z okazjonalnymi początkowymi epizodami moczenia oraz z dużą maksymalną objętością pęcherza, warto stosować reżim płynowy przez dłuższy okres czasu (...). Dopiero w przypadku nieskuteczności tej terapii, należy przejść do leczenia farmakologicznego.

Wnioski te mają bezpośrednie praktyczne znaczenie przede wszystkim dla lekarzy pediatrów, ale ogólniej wskazują też na stwierdzoną skuteczność i przydatność komputerowo wspomaganą statystyczną analizę danych w różnych obszarach medycyny. Ważne jest przy

tym stwierdzenie, że „Każdy przypadek jest inny i stąd wynika konieczność indywidualnego podejścia do prowadzonej analizy, a dzięki temu możliwość uzyskania bardziej wartościowych rezultatów”.

W tej części monografii poważniejszych błędów lub usterek nie zauważyłem poza jedną niekonsekwencją: w p. 2.7 poświęconym ocenie istotności statystycznej stwierdzono: „Warto zatem w badaniach wskazać dokładną wartość p , zamiast oznaczeń $p < 0,05$ lub $p > 0,05$ ”; jednakże w wielu dalszych częściach monografii używa się właśnie nierównościowego określania współczynnika istotności statystycznej p .

Rozdział 3 monografii jest poświęcony technikom nadzorowanego i nienadzorowanego uczenia maszynowego w klasyfikacji danych medycznych. W pierwszej części zawiera on krótki przegląd faktów znanych na ten temat z literatury.

W podrozdziale 3.4 poświęconym „algorytmom klasyfikacyjnym” można dopatrzeć się kilku nieścisłości nie podważających jednak ogólnie prawidłowego doboru i sposobu przedstawienia treści. Są to:

- Brak wyraźnego rozróżnienia pojęć *klasyfikacja* i *rozpoznawanie*. Z klasyfikowaniem mamy do czynienia m.in. przy tworzeniu zbioru uczącego z dużej liczby zarejestrowanych przypadków lub obiektów. Lekarz, który używając tego zbioru posługuje się algorytmem SVM lub k-NN do postawienia diagnozy konkretnemu pacjentowi, nie dokonuje *klasyfikacji*, lecz *rozpoznania*.

- W podrozdziale 3.4.1 przytoczono definicję: „Formalnie drzewo określa się jako dowolny, spójny, skierowany graf acykliczny”. Definicja nie jest prawidłowa, gdyż na przykład struktura: $A \rightarrow B \leftarrow C \rightarrow D$ spełnia definicję, ale nie jest drzewem, zawiera bowiem dwa korzenie (A i C) i dwa liście (B i D). Prawidłową definicję drzewa skierowanego (zwanego tam „drzewem rozplywu”) można znaleźć np. w pracy: J.L. Kulikowski. „Zarys teorii grafów”, PWN, Warszawa, 1986, str.150.

- Wyrażenia (3.5) i (3.7) są formalnie niepoprawne i niezrozumiałe. Nie wiadomo, jaką rolę spełniają w nich separatory pionowe | i jak należy rozumieć w tych wzorach znaki równości =. Np. we wzorze (3.5) pierwszy znak równości oznacza równość zbiorów, ale drugi – równość zbioru i liczby arytmetycznej 1. Jeśli natomiast człon środkowy między tymi znakami równości rozumieć jako liczbę kardynalną zbioru, to nie można stawiać znaku jej równości ze zbiorem P , bo były by to różne obiekty matematyczne.

- W opisie *maszyny wektorów nośnych* (p. 3.4.3) i klasyfikatora *k-najbliższych sąsiadów* (p.3.4.4) zabrakło ważnego zastrzeżenia, że w obu tych przypadkach zbiory obserwacji muszą być elementami przestrzeni metrycznych. Ten warunek może nie być spełniony w przypadku, gdy obserwacje (jak to często ma miejsce właśnie w przypadku danych medycznych) złożone są z różnych typów danych. Problem ten został wprawdzie zauważony w dalszej części monografii (p. 4.5.6), gdzie w przypadku mieszanych typów danych proponuje się sprowadzenie ich do jednego typu, co może być jednak zabiegiem sztucznym, naginającym rzeczywistość do założeń formalnych.

W Rozdz. 3 na szczególną uwagę zasługuje p. 3.9 poświęcony problematyce oceny i wyboru algorytmów uczących się dla konkretnych zastosowań medycznych. Jako kryterium oceny przyjęto: dokładność (ACC), precyzję (PREC), czułość (SENS) i pole powierzchni pod wykresem krzywej ROC (AUROC), a więc wskaźniki jakości używane w odniesieniu do decyzji binarnych. W badaniach rozróżniano przypadki symetrycznego i asymetrycznego zahamowanie wzrostu płodu (IUGR1 i IUGR2). Eksperymenty przeprowadzono w środowisku programowym WEKA Open Source Data Mining Tool. Ocenie porównawczej poddano 8 typów algorytmów klasyfikacyjnych, a następnie – 4 typy zespołów klasyfikatorów. Wyniki ocen przedstawiono w postaci tabelarycznej stwierdzając, że najwyższą dokładność wykazały algorytmy regresji logistycznej (80%) oraz stochastycznego

gradientu prostego (79%), a w klasie zespołów klasyfikatorów – zespoły uzyskane metodą boostingu z bazowym klasyfikatorem stochastycznego gradientu prostego (82,35%), najniższe zaś (76,45%) – zespoły oparte na metodzie lasów losowych. W podsumowaniu tej części pracy stwierdzono, że „w kontekście identyfikacji postaci IUGR żadna z technik nie wyróżniła się znacząco na tle pozostałych metod. Niemniej jednak warto podkreślić, że zastosowanie klasyfikatorów złożonych pozwoliło na uzyskanie lepszych rezultatów niż użycie pojedynczych klasyfikatorów oraz technik grupowania.”

Są to stwierdzenia interesujące, lecz nie można ich przenosić na podobne klasyfikatory stosowane do innych zbiorów danych.

Rozdz. 4 monografii poświęcono ważnemu problemowi selekcji cech obiektów poddawanych klasyfikacji i rozpoznawaniu, wciąż aktualnemu w przypadku rozszerzania tych zadań na nowe obszary zastosowań. W obszernej części wprowadzającej przedstawiono m.in. klasyfikację metod automatycznej selekcji cech z ogólnym podziałem na metody „filtrujące”, „powłoki” i „wbudowane”. Nazwy tych metod nie są zbyt udane, nie odzwierciedlają one bowiem dobrze ich istoty. Opisy metody filtrującej i powłoki zilustrowano prostymi schematami i algorytmami w pseudokodzie. W opisach tych występują drobne nieścisłości:

- W p. 4.5.1 w opisie metody selekcji cech uwzględniającej statystykę χ^2 parokrotnie pojawia się termin „wskaznik niespójności”, który nie jest jasno zdefiniowany, wskutek czego opis staje się niejasny.

- W p.4.5.2 stwierdzono: „im więcej przypadków należy do jednej z klas, tym większa jest wartość całkowitej entropii zbioru”, co jest oczywiście błędne i nie ma potwierdzenia w dalszej części wywodu.

Są to usterki drobne, nie podważające istotnie wartości tej części monografii, jaką jest próba usystematyzowania i charakterystyka wybranych ośmiu metod selekcji cech dla potrzeb komputerowo wspomaganey diagnostyki medycznej. W konkluzji Autorka stwierdza, iż „Źródła literaturowe opisujące poszczególne algorytmy nie pozwalają na wyciągnięcie wniosków umożliwiających ustanowienie rankingu rozwiązań. Nie jest możliwe również stwierdzenie, która metoda zachowa się najlepiej w określonym przypadku”. W związku z tym w ostatniej części rozdziału rozważono dwa szczególne, ale ważne problemy medyczne: selekcji cech genetycznych (genów) dla potrzeb wczesnego rozpoznawania zmian nowotworowych oraz selekcji cech dla potrzeb wczesnego wykrywania nadciśnienia tętniczego u dzieci. W pierwszym z tych problemów do badań użyto danych z 5 ogólnodostępnych zbiorów danych genetycznych, wstępnie usuwając z nich geny (tj. cechy) metabolizmu podstawowego i pozostawiając geny specyficzne tkankowo, istotne dla ewentualnego rozpoznawania zmian nowotworowych. Badania objęły siedem (spośród wcześniej omówionych), opartych na filtracji metod automatycznej selekcji cech oraz sześć technik klasyfikacyjnych. Wyniki eksperymentu zawarte w Tabelach 4.3 i 4,4 wskazują na wysoką skuteczność zbadanych algorytmów selekcji cech: redukcję ich liczny o od 93.5% do 99.7% przy zachowaniu wysokiego poziomu jakości klasyfikacji

W drugim problemie wzięto pod uwagę pierwotny zbiór 42 cech o niskiej informatywności i przy pomocy algorytmów J48 oraz AdaBoost dokonano selekcji cech: 9 w pierwszym i 6 w drugim przypadku; jest rzeczą interesującą, że tylko 3 cechy wystąpiły jednocześnie w obu podzbiorach.

W wyniku eksperymentu stwierdzono, że dla użytych kombinacji algorytmów klasyfikacyjnych z metodami grupowania wystąpiła ponad 70-procentowa zgodność z grupami początkowymi, wskazanymi przez ekspertów medycznych

W obu typach zadań uzyskane wyniki można ocenić jako interesujące z medycznego i ważne z informatycznego punktu widzenia.

Rozdział 5 monografii porusza dwa aspekty rozszerzające problematykę klasyfikacji danych medycznych o klasyfikację niezbalansowaną i wieloaspektową. Oba przypadki dotyczą ważnych sytuacji występujących w diagnostyce medycznej: rozpoznawania rzadkich schorzeń, które w zbiorach

danych są reprezentowane przez relatywnie niską liczbę zarejestrowanych przypadków oraz rozpoznawania kilku jednocześnie występujących schorzeń.

W rozdziale przedstawiono wyniki analizy technik równoważenia klas dla zbiorów danych niezbalansowanych. W tym celu porównano skuteczność algorytmów: drzew decyzyjnych, naiwnego klasyfikatora Bayesa, k -najbliższych sąsiadów i maszyny wektorów nośnych, uczonych na niezbalansowanych i zbalansowanych zbiorach danych. Klasyfikatory zastosowano do rozpoznawania 8 jednostek chorobowych, przy czym jako kryterium skuteczności została użyta czułość i dokładność rozpoznawania. Wyniki badań przedstawiono w postaci tablic numerycznych i wykresów. Wyniki badań nie we wszystkich przypadkach wykazały korzystny wpływ balansowania zbiorów danych na skuteczność klasyfikacji. Może to świadczyć nie tylko o niedoskonałości zastosowanych metod balansowania zbiorów danych, jak i potwierdzać fakt, iż niedostatek informacji o klasach zdominowanych nie da się zastąpić zabiegami formalnymi. Nie mniej jednak, w rozprawie została wskazana droga postępowania przy badaniu powyższego, ważnego problemu.

W części Rozdziału 5 poświęconej klasyfikacji wieloaspektowej przedstawiono podział metod takiej klasyfikacji na metody transformujące problem (sprowadzające go do problemu klasyfikacji jednoaspektowej), metody adaptujące algorytm klasyfikacji oraz metody łączone. Dalsze rozważania poświęcono metodom transformującym problem, wyodrębniając 7 rodzajów takich metod. Ważnym elementem tej części rozdziału jest propozycja rozszerzenia wskaźników (niezbyt szczęśliwie nazywanych tu *metrykami*): *dokładności*, *precyzji*, *czułości* oraz *miary F1*, a także *straty Hamminga* i *straty 0-1* na klasyfikatory wieloaspektowe (p. 5.3.2). Badaniu poddano dane z 4 zbiorów danych kardiologicznych, a wyniki przedstawiono w postaci tablic numerycznych i wykresów. Pozwoliły one Autorce na stwierdzenie, iż „...zastosowanie podejścia wieloperspektywicznego realizowanego za pomocą klasyfikacji wieloetykietowej jest efektywne i może być stosowane w diagnostyce medycznej do wykrywania chorób współistniejących.”

Choć stwierdzenie to jest oparte tylko na przykładach komputerowej analizy danych kardiologicznych, jest ono ważne także w odniesieniu do innych działów medycyny.

Podsumowując stwierdzam, że monografia dr inż. Agnieszki Wosiak pt. „Techniki statystycznej i eksploracyjnej analizy danych medycznych” zawierająca:

- opis autorskich algorytmów uczenia maszynowego uwzględniających statystyczne metody analizy danych medycznych, ich implementacji i analizę wyników przeprowadzonych badań eksperymentalnych;
- opracowanie nowych metod selekcji cech, zmniejszających wymiarowość danych, poprawiającej ich czytelność i przydatność dla wnioskowania medycznego;
- propozycje nowych rozwiązań umożliwiających analizę danych medycznych opartą na niezbalansowanych zbiorach danych oraz wieloetykietową klasyfikację,

zawiera tym samym bogaty oryginalny wkład naukowy dr inż. Agnieszki Wosiak do inżynierii biomedycznej, a także do informatyki (w tym szczególnie do techniki uczenia maszynowego). Monografia o tak skomponowanej treści nie ma dotychczas równoważnego jej odpowiednika w polskiej literaturze naukowej. Na podkreślenie zasługuje fakt, iż opisane w monografii eksperymenty zostały wykonane w bliskiej współpracy dr inż. Agnieszki Wosiak z zespołami lekarskimi. Z oświadczeń współautorów publikacji, na których oparte są odnośne rozdziały monografii, wynika, że ich współudział polegał najczęściej na konsultacji medycznej. **Wysoka wartość merytoryczna monografii przy relatywnie niskim poziomie usterek redakcyjnych w pełni zasługuje na uznanie jej za pracę spełniającą wymogi rozprawy habilitacyjnej.**

3. Dorobek naukowy poza głównym osiągnięciem naukowym

Przed uzyskaniem stopnia doktora dorobek dr inż. Agnieszki Wosiak obejmował prace związane z metodologią diagnostyki stanu urządzeń energetycznych na podstawie danych pomiarowych i z zastosowaniem metod statystycznej i eksploracyjnej analizy danych. Na podstawie opracowanych algorytmów powstał autorski system analizy danych pomiarowych, który został wdrożony na stacji transformatorowej Piotrków Trybunalski, a wyniki zostały zawarte w referatach konferencyjnych i publikacjach w czasopiśmie naukowo-technicznym o zasięgu krajowym.

Po uzyskaniu stopnia doktora dr inż. Agnieszka Wosiak kontynuowała prace nad budową algorytmów wspomagających diagnostykę urządzeń technicznych. Dotyczyły one optymalizacji sposobu przechowywania dużej ilości danych pomiarowych oraz rozwinięcia metodologii o algorytmy dotyczące diagnostyki układu chłodzenia transformatora. Wyniki tych prac zostały zawarte w dwóch publikacjach, w tym – jednej o zasięgu międzynarodowym. Ponadto Kandydatka uczestniczyła w pracach badawczych dotyczących sieci teleinformatycznych; wyniki tych prac zostały przedstawione w dwóch współautorskich publikacjach o zasięgu krajowym.

Ważną część dorobku naukowego dr inż. Agnieszka Wosiak stanowią też prace poświęcone architekturze oraz algorytmom związanym z przetwarzaniem danych obrazowych i z łączeniem ich z informacjami pochodzącymi z innych badań. Wyniki tych prac zostały zawarte w 6 innych publikacjach o zasięgu krajowym lub międzynarodowym.

Efektom współpracy dr inż. Agnieszki Wosiak ze środowiskiem lekarskim jest opracowany przez nią **system analizy danych medycznych ADaMed**, wdrożony (i nadal rozwijany) w Uniwersyteckim Szpitalu Klinicznym nr 4 im. M. Konopnickiej w Łodzi.

Ogólnie, dorobek dr inż. Agnieszki Wosiak niewchodzący w zakres pracy habilitacyjnej jest udokumentowany 23 publikacjami, głównie współautorskimi, umieszczonymi jednak w wydawnictwach o ograniczonym zasięgu.

Wskaźniki metryczne dorobku publikacyjnego dr inż. Agnieszka Wosiak przedstawiają się następująco:

- Sumaryczny impact factor według listy Journal Citation Reports (JCR), zgodnie z rokiem opublikowania: **8,704**;
- Liczba cytowań publikacji według bazy Web of Science (WoS) bez autocytowań: **10**;
- Liczba cytowań publikacji według bazy Scopus: **37**;
- Liczba cytowań publikacji według bazy Google Scholar: **65**;
- Indeks Hirscha według bazy Web of Science: **3**;
- Indeks Hirscha według bazy Scopus: **4**;
- Indeks Hirscha według bazy Google Scholar (stan na dzień 25 kwietnia 2018): **5**.

Publikacje te, choć nisko punktowane na liście MNSWiT. świadczą o tym, że **dr inż. Agnieszka Wosiak stara się swoją wiedzę i kwalifikacje w dziedzinie statystycznych metod analizy danych stosować w praktyce zarówno w służbie zdrowia, jak i w wybranych dziedzinach techniki.** W dorobku dr inż. Agnieszki Wosiak **nie ma natomiast patentów krajowych lub zagranicznych**, co jednak nie dziwi, gdyż algorytmy i programy komputerowe nie podlegają patentowaniu.

Dr inż. Agnieszka Wosiak uczestniczyła jako specjalista w realizacji **5 międzynarodowych lub krajowych projektów badawczych**, w tym w 3 przypadkach w charakterze kierownika zespołu badawczego.

Dr inż. Agnieszka Wosiak wygłosiła referaty na 21 konferencjach naukowych w kraju i na 2 konferencjach naukowych za granicą (Praga, Cape Town), uczestniczyła też w organizacji 2 konferencji naukowych o zasięgu międzynarodowym. W sumie jednak – **aktywność naukowa lub naukowo-organizacyjna Dr inż. Agnieszki Wosiak ograniczała się głównie do działalności na terenie kraju.** Kandydatka nie odbyła żadnego stażu w zagranicznym ośrodku naukowym lub akademickim, nie brała udziału w wykonywaniu ekspertyz lub innych opracowań na zamówienie, nie brała udziału w zespołach eksperckich lub konkursowych, nie recenzowała też projektów międzynarodowych lub krajowych, natomiast recenzowała manuskrypty publikacji dla 5 wydawnictw, w tym 3 zagranicznych.

Kandydatka jest członkiem 3 krajowych organizacji naukowych, natomiast w pracach komitetów redakcyjnych lub radach naukowych czasopism nie uczestniczy.

4. Dorobek dydaktyczny

Dorobek dydaktyczny Dr inż. Agnieszki Wosiak na terenie macierzystej uczelni obejmuje liczne **wykłady z dziedziny informatyki** dla słuchaczy studium doktoranckiego oraz studentów I i II stopnia, a także zajęcia laboratoryjne dla studentów I i II stopnia; niektóre z tych zajęć Kandydatka prowadziła w jęz. angielskim. Podobne zajęcia Kandydatka prowadziła też na terenie innych uczelni (we Włocławku i w Łodzi).

Dr inż. Agnieszka Wosiak sprawowała **kierownictwo i opiekę** nad kilkudziesięciu dyplomowymi pracami inżynierskimi i magisterskimi, zarówno w Politechnice Łódzkiej, jak i w Państwowej Wyższej Szkole Zawodowej we Włocławku.

Kandydatka sprawowała też funkcję **promotora pomocniczego** w odniesieniu do 2 doktorantów.

Na terenie Politechniki Łódzkiej Dr inż. Agnieszka Wosiak sprawowała szereg **funkcji dydaktyczno-organizacyjnych**: jako opiekun roku studiów, członek Komisji Rekrutacyjnej, Komisji Oceny Jakości Kształcenia itp.

W sumie – **osiągnięcia dydaktyczne Dr inż. Agnieszki Wosiak oceniam jako wysokie.**

5. Podsumowanie opinii

Biorąc pod uwagę dorobek naukowy Dr inż. Agnieszki Wosiak stanowiący **istotny wkład do informatyki**, w tym zwłaszcza – do jej zastosowań w medycynie w postaci opracowanych oryginalnych metod statystycznej analizy danych medycznych, szczególnie pod kątem redukcji danych w zbiorach uczących klasyfikatorów oraz powiązania tych metod z konkretnymi problemami medycznymi, a także – elementy dorobku naukowego dotyczące analizy danych dla potrzeb diagnostyki wybranych urządzeń technicznych, biorąc też dodatkowo pod uwagę duży dorobek dydaktyczny Kandydatki, pomimo zastrzeżeń dotyczących niskich wskaźników punktacji publikacji, **uważam, że spełnione są najważniejsze kryteria, określone w Dz. U. z 2011 r. poz. 1165, uzasadniające kontynuowanie postępowania w sprawie nadanie Dr inż. Agnieszce Wosiak stopnia naukowego doktora habilitowanego w dziedzinie nauk technicznych, dyscyplinie informatyka i wnoszę o tę kontynuację.**


(Juliusz L. Kulikowski)