

Warszawa, 7 grudnia 2018

Prof. dr hab. Leon Bobrowski  
Wydział Informatyki Politechniki Białostockiej

**Recenzja osiągnięcia naukowego i istotnej aktywności naukowej dr inż. Agnieszki Wosiak z Wydziału Fizyki Technicznej, Informatyki i Matematyki Stosowanej Politechniki Łódzkiej w postępowaniu habilitacyjnym prowadzonym na tym Wydziale**

Podstawą do przedłożenia tej recenzji jest pismo z dnia 7 września 2018 r. Centralnej Komisji do Spraw Stopni i Tytułów powołującą Komisję w celu przeprowadzenia postępowania habilitacyjnego dr inż. Agnieszki Wosiak oraz pismo z dnia 1 października 2018 r. prof. dr hab. Piotra Liczberskiego, dziekana Wydziału Fizyki Technicznej, Informatyki i Matematyki Stosowanej Politechniki Łódzkiej.

Pani Agnieszka Wosiak uzyskała w 2008 roku stopień doktora nauk technicznych w dyscyplinie informatyka na Wydziale Fizyki Technicznej, Informatyki i Matematyki Stosowanej na podstawie rozprawy pt.: *Inteligentny system analizy danych pomiarowych z urządzeń energetycznych*.

Dr inż. Agnieszka Wosiak jest obecnie adiunktem w Instytucie Informatyki Politechniki Łódzkiej.

Wskazany przez habilitantkę osiągnięciem (wynikającym 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. Nr 65, poz. 595, z później zm., art. 16) jest: *Techniki statystycznej i eksploracyjnej analizy danych medycznych*. Wskazane osiągnięcie obejmuje monografię pt. *Techniki statystycznej i eksploracyjnej analizy danych medycznych* wydanej w 2018 r. przez wydawnictwo EXIT w Warszawie, której habilitantka jest samodzielną autorką. Prezentowane osiągnięcie naukowe obejmuje również 6 publikacji w czasopismach z listy Ministerstwa Nauki i Szkolnictwa Wyższego oraz 7 opublikowanych w ramach konferencji indeksowanych w bazie *Web of Science*. Dwa artykuły konferencyjne z powyżej wzmiankowanych zostało opublikowanych w serii Springer: *Advances In Intelligent Systems and Computing*. Trzy artykuły z wymienionych w ramach osiągnięcia naukowego zostało opublikowanych w *Przeglądzie Elektrotechnicznym*.

W prezentowanym osiągnięciu naukowym pierwszą autorką sześciu publikacji jest pani prof. dr hab. inż. Liliana Byczkowska – Lipińska z Wydziału Fizyki Technicznej, Informatyki i Matematyki Stosowanej Politechniki Łódzkiej.

Aktualny dorobek publikacyjny dr inż. Agnieszki Wosiak jest moim zdaniem skromny. Liczba cytowań wg bazy *Web of Science* wyniosła 12. Indeks Hirscha wg bazy *Web of Science* ma wartość 2.

W cykl publikacji wskazany przez habilitantkę jako osiągnięcie naukowe wchodzi monografia o objętości 149 stron pt. *Techniki statystycznej i eksploracyjnej analizy danych medycznych*. [H1.] Monografia ta została podzielona na pięć rozdziałów o następujących tytułach:

1. *Wstęp*
2. *Metody statystyczne w analizie danych medycznych*
3. *Techniki nadzorowanego i nienadzorowanego uczenia maszynowego w klasyfikacji danych medycznych*
4. *Selekcja Cech*
5. *Aspekty rozszerzające problematykę klasyfikacji danych medycznych*

Dr inż. Agnieszka Wosiak sformułowała w *Autoreferacie* cel naukowy swojej pracy w następujący sposób:

*"Celem naukowym zrealizowanych badań było opracowanie efektywnych algorytmów, bazujących na metodach z obszaru statystycznej i eksploracyjnej analizy danych medycznych, do rozwiązywania problemów z zakresu medycyny. Właściwie opracowana metodologia i dobór odpowiednich metod komputerowej analizy danych stanowią kluczowe czynniki procesu badawczego. Ogromna liczba gromadzonych danych powoduje, że rozważenie przez człowieka wszystkich czynników staje się niemożliwe. Wskazane jest zastosowanie zaawansowanych narzędzi analitycznych, które do tej pory dotyczyły przede wszystkim procesów biznesowych lub przemysłowych."*

*Autoreferat s. 5: "W ramach przeprowadzonych prac opracowane zostały algorytmy uwzględniające zastosowanie wybranych metod statystycznych:*

- *dobór testów parametrycznych i nieparametrycznych*
- *ocenę związków poprzez analizę korelacyjną i regresyjną*

*Podrozdziały 2.8.1 – 2.8.5 monografii stanowią podsumowanie wybranych autorskich badań będących wkładem w badania kliniczne i diagnostykę medyczną."*

*Autoreferat s. 7: "Rozszerzenie metodologii prowadzenia statystycznej analizy danych stanowiło opracowanie autorskiego systemu komputerowego wspomagającego badania epidemiologiczne poprzez statystyczną analizę danych."*

Habilitantka deklaruje najbardziej znaczące wyniki swoich prac na stronie 17 - tej *Autoreferatu*. Wśród tych deklaracji znaleźć można poniższe zagadnienia:

- *opracowanie autorskich algorytmów uwzględniających metody statystyczne do analizy danych medycznych oraz zaimplementowanie tych algorytmów w autorskim systemie komputerowym...*
- *opracowanie architektury systemu komputerowego do pozyskiwania metadanych z obrazów medycznych...*
- *opracowanie usprawnienia detekcji nowych czynników wpływających na wcześniejsze stwierdzenie diagnozy poprzez opracowanie hybrydowego rozwiązania, uwzględniającego połączenie technik analizy skupień z technikami statystycznej analizy danych...*
- *opracowanie nowego algorytmu selekcji cech, polegającego na wyborze atrybutów o mniejszej informatywności...*
- *opracowanie nowej metody selekcji cech, uwzględniającej ich odwrócony porządek względem wzajemnych zależności korelacyjnych...*
- *opracowanie nowej metody klasyfikacji wieloetykietowej...*

Poszczególne osiągnięcia wymienione przez habilitantkę w *Autoreferacie* jako najbardziej znaczące budzą wątpliwości merytoryczne z powodu ich ogólnikowości i małej precyzji sformułowań. Przykładem tego typu ogólnikowości jest już pierwsze sformułowanie: *"opracowanie autorskich algorytmów uwzględniających metody statystyczne do analizy danych medycznych"*. Na czym polega deklarowana przez habilitantkę *"autorskość"* tych algorytmów? Algorytmy dotyczące *doboru testów parametrycznych i nieparametrycznych* czy też *oceny związków poprzez analizę korelacyjną i regresyjną* (*Autoreferat* s. 5) trudno jest klasyfikować jako autorskie.

W zestawieniu najbardziej znaczących wyników, habilitantka odnosi się dwukrotnie do problemu selekcji cech: *"opracowanie nowego algorytmu selekcji cech "* oraz *"opracowanie nowej metody selekcji cech"*. Habilitantka deklaruje *opracowanie nowego algorytmu selekcji cech, polegającego na wyborze atrybutów o mniejszej informatywności"*. Tematyka selekcji cech jest aktualnie rozwijana i stosowana w wielu ośrodkach badawczych na świecie w tym w ośrodkach medycznych. Istotą rozwijanych współcześnie metod selekcji cech jest wydzielanie takich pojedynczych cech lub ich zestawów, które charakteryzują się największą informatywnością. Na przykład, poszukuje się takich genów lub ich zestawów, które wyróżniają się dużą siłą w reprezentowaniu lub w różnicowaniu chorób. Habilitantka zdaje się nie posiadać aktualnej wiedzy na temat opracowanych i skutecznie stosowanych metod selekcji cech oraz ich różnorodności. Podstawową techniką i punktem odniesienia dla różnych metod selekcji cech pozostaje *LASSO (Least Absolute Shrinkage and Selection Operator)*.

Prezentowane w ramach habilitacji wyniki dotyczą metod wspomagania diagnostyki medycznej. Wśród prezentowanych wyników znalazły się nowe techniki związane z klasyfikacją niezbalansowanych zbiorów danych oraz z klasyfikacją wieloetykiętową. Zgodnie z deklaracją habilitantki w *Autoreferacie* (s. 4): *"Opracowane algorytmy zostały zaimplementowane w nowych systemach komputerowych wdrożonych do pracy w szpitalu"*.

Wyniki osiągnięte przez habilitantkę zostały opisane w artykułach naukowych jak również zebrane w monografii pt. *"Techniki statystycznej i eksploracyjnej analizy danych medycznych"*, *EXIT*, Warszawa 2018. Zgodnie z deklaracją habilitantki w *Autoreferacie* (s. 4): *"Podrozdziały 2.8.1-2.8.5 monografii stanowią podsumowanie wybranych, autorskich badań będących wkładem w badania kliniczne oraz diagnostykę medyczną"*. Stwierdzenie habilitantki o wkładzie w badania kliniczne budzi jednak wątpliwości.

*Autoreferat* zawiera również opis pozostałych osiągnięć naukowo-badawczych habilitantki nie wchodzących w zakres wskazanego osiągnięcia naukowego. Zgodnie z tym opisem:

*"Przed uzyskaniem stopnia doktora moje badania koncentrowały się na opracowaniu metodologii 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. "* (s. 18)

*"Po uzyskaniu stopnia doktora, początkowo w dalszym ciągu rozwijane były prace nad budową algorytmów wspomagających diagnostykę urządzeń technicznych. W szczególności dotyczyły one optymalizacji sposobu przechowywania znacznej ilości danych pomiarowych*

oraz rozwinięcia metodologii o algorytmy związane z diagnostyką układu chłodzenia transformatora." (s. 19)

Habilitantka podsumowuje na stronach 16-tej i 17-tej *Autoreferatu* najbardziej znaczące wyniki prac, wchodzące w zakres osiągnięcia naukowego. W zestawieniu tym pewne zdziwienie i wątpliwości budzi stwierdzenie: *"opracowanie metody automatycznej selekcji cech, uwzględniającej ich odwrócony porządek względem wzajemnych zależności korelacyjnych, a w rezultacie wyłonienie podzbioru cech o największym zróżnicowaniu – zastosowanie cech wyznaczonych przez algorytm pozwoliło na zbudowanie struktury skupień lepiej odpowiadającej wykrywaniu nowych zależności - wzrost wykrytych zależności od 23% do nawet 150%."* (s. 18)

Uważam, że poniższy wynik prac można wskazać jako jeden z bardziej wartościowych: *"opracowanie architektury sytemu komputerowego do pozyskiwania meta danych z obrazów medycznych i łączenia tych danych z innymi danymi medycznymi w celu rozszerzenia procesu diagnostyki"* (s. 18).

Pozytywną uwagę zwraca też stwierdzenie: *"opracowanie adaptacji nowej metody klasyfikacji wieloetykietowej Labels Chain jako podejścia wielo perspektywicznego w celu jej zastosowania w procesie diagnostyki chorób współistniejących"* (s. 18). Publikacja na ten temat pt. *Multi-label classification methods for improving comorbidities identification* ukazała się w 2018 roku w renomowanym czasopiśmie *Computers in Biology and Medicine*. [H2.]

Podsumowując swoją recenzję deklaruję, że moje wątpliwości budzi mały stopień realizacji celu naukowego wskazanego przez habilitantkę w następującym sformułowaniu: *"Celem naukowym zrealizowanych badań było opracowanie efektywnych algorytmów, bazujących na metodach z obszaru statystycznej i eksploracyjnej analizy danych medycznych, do rozwiązywania problemów z zakresu medycyny."*

Opisane powyżej osiągnięcia naukowe habilitantki nie spełniają jednak, moim zdaniem, wymagania artykułu 16 cytowanej wyżej Ustawy. Stwierdzenie to nie daje podstawy do sformułowania na końcu tej recenzji pozytywnej konkluzji, w świetle opisanych w recenzji niejasności i wątpliwości co do znaczenia zadeklarowanego przez habilitantkę osiągnięcia naukowego.

*Leon Bobrowski*