

Poznań, 15 października 2021r.

dr hab. Agata Filipowska, prof. UEP  
Katedra Informatyki Ekonomicznej  
Uniwersytet Ekonomiczny w Poznaniu

## **Recenzja rozprawy doktorskiej mgra Pawła Piaseckiego pt. „Application of selected artificial intelligence methods to time series analysis”**

Promotor rozprawy: prof. UAM dr hab. Tomasz Górecki

Recenzja sporządzona na podstawie Ustawy z dnia 20 lipca 2018r. Prawo o szkolnictwie wyższym i nauce (Dz.U. z 2021r. poz. 478).

### **1. Tematyka rozprawy doktorskiej**

Przedstawiona rozprawa doktorska, na którą składa się cykl sześciu publikacji powiązanych wspólnym tematem, dotyczy obszaru sztucznej inteligencji, a dokładniej analizy szeregów czasowych. Trzy z przedstawionych publikacji dotyczą przeglądu dorobku w dziedzinie miar odległości w analizie szeregów czasowych i stanowią wartościowe prace przeglądowe. Jedna z publikacji dotyczy funkcjonalnej analizy danych, pozwalającej na przetwarzanie zgromadzonych obserwacji jako funkcji, co może zostać wykorzystane także przy opisie szeregów czasowych. Pozostałe z przedstawionych publikacji dość ogólnie powiązane są z tematem pracy: jedna z nich dotyczy analizy chemometrycznej, zaś druga sztucznych sieci neuronowych.

Autor w przytoczonych publikacjach wskazuje ogólnie na czym polega adresowany przez niego problem badawczy pisząc, że możliwe jest odróżnienie problemu analizy szeregów czasowych od innych zadań analitycznych ze względu na fakt uporządkowania chronologicznego atrybutów oraz możliwości wyszukania cechy charakteryzującej to uporządkowanie. Warto byłoby jednak odnieść się do istoty problemu badawczego motywującego do podjęcia pracy badawczej, a następnie postawić cele pozwalające na jego zaadresowanie.

### **2. Cele rozprawy doktorskiej**

Zgodnie z informacją zawartą we wprowadzeniu do pracy doktorskiej, jej celem głównym było zaproponowanie kilku metod sztucznej inteligencji, pozwalających na analizę szeregów czasowych. Cel ten został sformułowany bardzo ogólnie, bez zdefiniowania kryteriów sukcesu osiągnięcia celu.

We wstępie do pracy przedstawiono cele szczegółowe ściśle związane z celami definiowanymi w poszczególnych publikacjach. Do celów tych należały:

1. Porównanie miar odległości wykorzystywanych w analizie szeregów czasowych.
2. Zaproponowanie miary asocjacyjnej dla wielowymiarowych danych funkcjonalnych.
3. Zastosowanie wiedzy z zakresu analizy szeregów czasowych dla potrzeb analizy chemometrycznej.

#### 4. Zaproponowanie nowego klasyfikatora RAndom Neural Networks (RANNs).

Cele pracy zostały sformułowane poprawnie, jednak wskazanie w jaki sposób cele te przyczyniają się do zrealizowania celu głównego wskazanego przez Doktoranta podwyższałoby wartość recenzowanej pracy. Cele pracy pozwalają jednak na uzyskanie wkładu naukowego do dyscypliny badań.

#### 3. Wkład badawczy dysertacji

Wkład badawczy Doktoranta dotyczy wnikliwej analizy literatury w zakresie miar odległości stosowanych np. w analizie szeregów czasowych oraz badania porównawczego (benchmarku) przedstawionych metod. Jakość wyników badawczych uzyskanych przez Doktoranta oceniam bardzo wysoko.

Drugim wynikiem badań jest zaproponowanie nowej miary zależności dla wielowymiarowych danych funkcjonalnych – funkcjonalnego współczynnika Procrustes. W artykule wspomniane jest przeprowadzenie eksperymentu na danych rzeczywistych, który jednak nie został dokładnie opisany, a mógłby stanowić o aplikacyjności opracowanej miary. Z perspektywy recenzenta kontynuacja artykułu prezentująca zastosowanie miary podwyższyłaby wartość opracowanego artefaktu.

Kolejnym wynikiem prac było wskazanie użyteczności metod analizy danych w ocenie jakości soków jabłkowych z wykorzystaniem danych z analizy chemometrycznej. W artykule zawarto tylko niewielki fragment opisujący metodykę i wyniki prac przeprowadzonych przez Doktoranta, a sam artykuł stanowi wkład do dziedziny chemii analitycznej. Artykuł został jednak opublikowany w czasopiśmie o bardzo wysokim impact factor, a Doktorant jest jego współautorem w 30%.

Ostatnim z uzyskanych wyników prac było zaproponowanie nowego klasyfikatora wykorzystującego koncepcję łączenia sztucznych sieci neuronowych podobną do zastosowanej w przypadku lasów losowych. Pomysł Doktoranta jest bardzo ciekawy i mógłby stanowić punkt wyjścia do podjęcia dalszych badań. Przedstawiony artykuł w dość zwięzły sposób przedstawia metodę oraz jej ewaluację z wykorzystaniem zbiorów często używanych dla weryfikacji skuteczności działania nowych metod. Wartość pracy podwyższa porównanie wyników klasyfikacji z innymi, powszechnie stosowanymi klasyfikatorami.

Wkład badawczy Doktoranta wynika z przedstawionych publikacji, z których 4 stanowią kontynuację badań, zaś pozostałe dwa związane są z wynikami projektów lub stanowią najnowsze zainteresowania autora. Artykuły stanowiące skonsolidowany dorobek doktoranta dotyczący szeregów czasowych sensu stricto są następujące:

- Górecki T., Piasecki P. An Experimental Evaluation of Time Series Classification Using Various Distance Measures Archives of Data Science, Series A (Online First), 5(1), 2018.
- Górecki T., Piasecki P. A Comprehensive Comparison of Distance Measures for Time Series Classification. In: Steland A., Rafa-łłowicz E., Okhrin O. (eds) Stochastic Models, Statistics and Their Applications. Springer Proceedings in Mathematics & Statistics, vol 294, pages 409–428, 2019. Springer.

- Górecki T., Łuczak M., Piasecki P. An exhaustive comparison of distance measures in the classification of time series with 1NN method. Paper has been submitted to "Expert Systems with Applications".
- Górecki T., Piasecki P. Generalization of the Procrustes coefficient to functional data. Communications in Statistics - Simulation and Computation, 3(49): 808–816, 2019. Taylor & Francis.

Wkład badawczy Doktoranta wynosi w nich od 60% do 80%, Doktorant jest więc głównym autorem przedstawionych wyników prac.

Ponadto, Doktorant brał udział w trzech projektach, przy czym w jednym z nich był kierownikiem, a w pozostałych dwóch wykonawcą. Przedstawione publikacje, których Doktorant był współautorem, odnoszą się tematycznie do dwóch z projektów. Projektami tymi były:

- GRANT PRELUDIUM pt. „Classification of time series using similarity forests and deep neural networks”, Narodowe Centrum Nauki, program PRELUDIUM 16, nr 2018/31/N/ST6/01209.
- GRANT OPUS pt. „Spectral fingerprints as a tool for testing the quality of fruit juices”, Narodowe Centrum Nauki, program OPUS 12, nr 2016/23/B/NZ9/03591.

Recenzując dorobek Doktoranta warto także zauważyć, że uzyskał on także pierwszą nagrodę Young Scientists Awards w kategorii plakatu konferencyjnego podczas międzynarodowej konferencji LinStat'2018. Na konferencji Doktorant przedstawił plakat pt. „Selected association measures for multivariate functional data” związany z tematyką prezentowanej rozprawy.

#### **4. Struktura dysertacji**

Przedstawiona dysertacja składa się z dwóch części. W pierwszej z nich zawarto streszczenia pracy w języku polskim i angielskim oraz wprowadzenie wskazujące na motywację autora do podjęcia tematu, cele pracy oraz opis metodyki badawczej.

Drugą część pracy stanowi sześć artykułów, w których Doktorant jest wiodącym lub głównym badaczem. W pracy nie zawarto podsumowania odnoszącego się do celów pracy i wskazującego na stopień ich realizacji, przy czym tę rolę częściowo przejęło wprowadzenie do pracy.

Wartość wprowadzenia obniża fakt, że z 20 cytowanych publikacji, jedynie cztery stanowią pozycje najnowsze z roku 2019 lub późniejsze. W przytoczonych artykułach stanowiących dorobek Doktoranta, występuje więcej cytowań najnowszych, jednak warto byłoby zwrócić uwagę na ten aspekt w przyszłych pracach podejmowanych przez Doktoranta.

#### **5. Ocena metodyki badawczej**

W rozprawie brakuje dokładnego omówienia metodyki badawczej. We wstępie do pracy Doktorant wspomina, że podczas prowadzenia badań wykorzystano metody analizy statystycznej i sztucznej inteligencji, jednakże metodyka prowadzenia prac powinna także być opisana. Brak ujęcia opisu metodyki Doktorant motywuje faktem umieszczenia opisu metodyki w każdej z publikacji, w nich znajduje się jednak tylko kilka zdań, które nie pozwalają na rzetelną ocenę procesu badawczego prowadzonego przez Doktoranta.

Należy jednak zauważyć, że Doktorant opisuje technologię, jaką posłużył się w procesie dokonywania obliczeń (R) oraz środowisko wykorzystywane w eksperymentach (Intel Xeon E5-2697, komputer osobisty z procesorem Intel i9 (12 rdzeni) i 32 GB RAM).

## 6. Kwestie do dyskusji

Metodyka prac:

- Cel rozprawy został zdefiniowany w dość ogólny sposób, konieczne byłoby zatem przedstawienie problemu badawczego rozwiązywanego przez Doktoranta. Jaki problem badawczy adresowany był przez Doktoranta? Czy była dokonywana redefinicja problemu po przeprowadzeniu szerokich prac przeglądowych?
- Doktorant w artykułach korzysta z różnych metodyk badawczych, konieczne byłoby ich omówienie oraz pokazanie zastosowania w różnych typach badań prowadzonych przez Doktoranta.

Zastosowania metod:

- Zastosowanie metod analizy szeregów czasowych do analizy danych ze spektrometru wymaga odpowiedniego odwzorowania tych danych – w jaki sposób przedstawiono dane? Jak przebiegała ich analiza i na czym polegało wykorzystanie metod analizy szeregów czasowych w tym scenariuszu?
- W jaki sposób zaproponowany klasyfikator RAndom Neural Networks (RANNs) dostarcza wartości dodanej w porównaniu z innymi podejściami do analizy szeregów czasowych? Jakie mogłyby być scenariusze jego wykorzystania?

## 7. Podsumowanie

Rozprawa doktorska prezentuje wiedzę teoretyczną Doktoranta w dziedzinie nauk ścisłych i przyrodniczych, w dyscyplinie informatyka oraz umiejętność samodzielnego prowadzenia pracy naukowej. Przedstawione publikacje prezentują umiejętność dokonania syntezy oraz oceny wyników badawczych uzyskanych przez innych badaczy (publikacje przeglądowe), lecz także opracowania własnych rozwiązań postawionych problemów. Jakość przedstawionych publikacji oceniam wysoko, a każda z nich stanowi oryginalne podejście autora do rozwiązywanego problemu.

Przedmiotem rozprawy doktorskiej jest oryginalne rozwiązanie problemu naukowego dotyczącego analizy szeregów czasowych. Doktorant sprawnie posługuje się aparatem badawczym, zaś wyniki prezentuje w przyjęty w dyscyplinie sposób. Uzyskane wyniki stanowią wkład naukowy do dyscypliny informatyka.

Rozprawę doktorską stanowi zbiór opublikowanych i powiązanych tematycznie artykułów naukowych, z których:

- jeden artykuł został opublikowany w zeszytach niemieckiego towarzystwa zajmującego się analizą danych (Archives of Data Science Series A, German Classification Society/Gesellschaft für Klassifikation (GfKL),
- jedna publikacja w Springer Proceedings in Mathematics & Statistics,

- jeden artykuł został opublikowany w Communications in Statistics - Simulation and Computation (Taylor and Francis), impact factor: 1.118.

Na dorobek Doktoranta bezpośrednio związany z tematyką rozprawy składa się artykuł podsumowujący pracę nad analizą metod odległości dla szeregów czasowych zgłoszony do czasopisma „Expert Systems with Applications” (IF: 6.954), przy czym w przedstawionej dokumentacji brakuje informacji nt. przyjęcia artykułu do publikacji.

Dodatkowo, w dokumentacji znajdują się dwa artykuły prezentujące wiedzę doktoranta, natomiast ich relacja do głównego celu dysertacji powinna być dokładniej wyjaśniona. Jedna z tych publikacji została uwzględniona w materiałach konferencyjnych, zaś druga pochodzi z czasopisma Microchemical Journal, impact factor: 4.821 (czasopismo publikuje prace z zakresu chemii analitycznej i analizy chemicznej).

Rozprawę doktorską przygotowano w języku angielskim wraz ze streszczeniem w języku polskim.

Podsumowując, przedstawiona rozprawa spełnia warunki określone w art. 187 Ustawy z dnia 20 lipca 2018r. Prawo o szkolnictwie wyższym i nauce (Dz.U. z 2021r. poz. 478) dla dziedziny nauk ścisłych i przyrodniczych, w dyscyplinie informatyka.

W rozprawie Doktorant, mimo zawartych w recenzji komentarzy krytycznych nie umniejszających wartości uzyskanych wyników, a jedynie wskazujących obszary do dalszej pracy, wykazał się rzetelną wiedzą i znajomością technik prowadzenia badań w dyscyplinie naukowej oraz umiejętnością rozwiązywania problemu badawczego. Cel rozprawy został zrealizowany, podobnie jak cele szczegółowe stawiane w poszczególnych publikacjach.

**W związku z powyższym występuję z wnioskiem o uznanie rozprawy doktorskiej jako spełniającej wymagania obowiązującej ustawy oraz dopuszczenie Doktoranta do jej publicznej obrony.**

