

Prof. Dr hab. Henryk Rybinski
Instytut Informatyki Politechniki Warszawskiej
Henryk.rybinski@pw.edu.pl

Recenzja rozprawy doktorskiej mgra Alberta Śledzianowskiego
p/t. „ The application of data mining and machine learning techniques to study eye
movement properties in the context of neurodegenerative diseases and changes in
the emotional system”

Dane ogólne

Przedstawiona do recenzji rozprawa jest napisana w języku angielskim. Składa się ze streszczenia w języku polskim i angielskim oraz dwóch rozdziałów. Rozdział pierwszy jest poświęcony wprowadzeniu w problematykę, ponadto zawiera krótki przegląd literatury oraz dyskusję na temat prac doktoranta załączonych w Rozdziale 2. Rozdział 2 stanowi zbiór dziesięciu publikacji, w których doktorant jest autorem (pozycja [3]), lub współautorem (pozostałe pozycje).

1. Podrozdział 2.1 stanowi publikacja
Śledzianowski, A., Szymański, A., Szlufik, S., Koziorowski, D. (2018), Rough Set Data Mining Algorithms and Pursuit Eye Movement Measurements Help to Predict Symptom Development in Parkinson's Disease. In: Nguyen, N., Hoang, D., Hong, TP., Pham, H., Trawiński, B. (eds) Intelligent Information and Database Systems. ACIIDS 2018. LNCS(), vol 10752. Springer
2. Podrozdział 2.2 zawiera publikację
Śledzianowski, A., Szymański, A., Drabik, A., Szlufik, S., Koziorowski, D.M., Przybyszewski, A.W. (2019). Measurements of Antisaccades Parameters Can Improve the Prediction of Parkinson's Disease Progression. In: Nguyen, N., Gaol, F., Hong, TP., Trawiński, B. (eds) Intelligent Information and Database Systems. ACIIDS 2019. LNCS(), vol 11432. Springer.
3. W podrozdziale 2.3 zawarta jest publikacja autorska
Śledzianowski, A. (2020). Predictions of Age and Mood Based on Changes in Saccades Parameters. In: Gruca, A., *et al* (eds) Man-Machine Interactions 6. ICMMI 2019. Advances in Intelligent Systems and Computing, vol 1061 . Springer.
4. Podrozdział 2.4 zawiera pracę:
Śledzianowski, A., Szymański, A., Drabik, A., Szlufik, S., Koziorowski, D.M., Przybyszewski, A.W. (2020). Combining Results of Different Oculometric Tests Improved Prediction of Parkinson's Disease Development. In: Nguyen, N., *et al* (eds) Intelligent Information and Database Systems. ACIIDS 2020. LNCS, vol 12034. Springer,

5. Podrozdział 2.5 przedstawia pracę z promotorem:
Przybyszewski A. W. & Śledzianowski A. (2020) Parkinson's disease development prediction by c-granule computing compared to different AI methods, *Journal of Information and Telecommunication*, 4:4, 425-439,
<https://doi.org/10.1080/24751839.2020.1749410>
6. W podrozdziale 2.6 przedstawiona jest praca:
Śledzianowski A., Urbanowicz K., Glac W., Slota R., Wojtowicz M., Nowak M., Przybyszewski A., Face emotional responses correlate with chaotic dynamics of eye movements, *Procedia Computer Science*, Vol. 192, 2021, pp. 2881-2892, ISSN 1877-0509,
7. Podrozdział 2.7 zawiera artykuł:
Śledzianowski, A., Nowacki, J.P., Przybyszewski, A.W., Urbanowicz, K. (2022). Detecting True and Declarative Facial Emotions by Changes in Nonlinear Dynamics of Eye Movements. In: Nguyen, N.T., *et al* (eds) Intelligent Information and Database Systems. ACIIDS 2022. LNCS, vol 13757. Springer
8. Podrozdział 2.8 przedstawia pracę:
Przybyszewski, A.W.; Śledzianowski, A.; Chudzik, A.; Szlufik, S.; Koziorowski, D. (2023). Machine Learning and Eye Movements Give Insights into Neurodegenerative Disease Mechanisms. *Sensors* 2023, 23, 2145.
9. W podrozdziale 2.9 przedstawiona jest praca
Śledzianowski A., Nowacki J.P., Sitarz K., Przybyszewski A.W., Universal machine-learning processing pattern for computing in the video-oculography, *in press*, ICCS 2023
10. W podrozdziale 2.10 przedstawiona jest praca
Śledzianowski A, Nowacki J.P., Sitarz K., Przybyszewski A.W. (2023), Novel machine learning pipeline for real-time oculometry. *in press*, ACIIDS 2023

Łączna objętość rozprawy wynosi ok. 158 stron. W ośmiu pracach doktorant jest pierwszym autorem, praca [8] jest opublikowana w czasopiśmie z listy JCR, pozostałe publikacje są rozdziałami z materiałów konferencji międzynarodowych. Do rozprawy dołączono oświadczenia współautorów o ich udziale - dotyczy to siedmiu zamieszczonych w doktoracie publikacji współautorskich [1-2, 4] oraz [6-7, 9-10]. Oświadczenia dotyczące publikacji [5,8] są zawarte w treści tych prac.

1. Problem badawczy i jego znaczenie

1.1 Problem rozważany w rozprawie

Przytoczone publikacje dotyczą badań mających charakter interdyscyplinarny. W sześciu przypadkach w zespole autorskim jest co najmniej jeden naukowiec w dziedzinie medycyny (poza publikacjami 2.3, 2.5, 2.9, 2.10). Moja opinia ogranicza się jedynie do

problemów dotyczących badań informatycznych w przedstawionych pracach.

Tematyka rozprawy ściśle wiąże się z metodami odkrywania wiedzy ze zbiorów danych. Dziedzina ta jest od dawna w obszarze zainteresowań statystyków, nieco krócej w obszarze zainteresowań naukowców zajmujących się metodami sztucznej inteligencji, w szczególności metodami uczenia maszynowego. Badania w tej dziedzinie mają ogromne znaczenie, zarówno praktyczne jak też teoretyczne, przede wszystkim w kontekście systemów akwizycji wiedzy, uczenia maszynowego, a także metod wnioskowania indukcyjnego przez analizę przykładów. Dlatego też od ponad 25-u lat daje się zaobserwować gwałtowny wzrost zainteresowań badaniami nad nowymi metodami analizy danych i uczenia maszynowego.

Jednym z istotnych kierunków badawczych w dziedzinie analizy danych jest problem praktycznych zastosowań metod eksploracji wiedzy w różnych obszarach badań. Dotyczy to wielu rozmaitych dziedzin. Jednym z bardziej spektakularnych zastosowań sztucznej inteligencji jest medycyna, bowiem analiza danych biomedycznych stwarza istotne narzędzia pracy zarówno w naukach medycznych, jak też w praktycznych zastosowaniach. Opiniowana praca mieści się w tym nurcie badań i dotyczy problematyki odkrywania wiedzy na potrzeby diagnozowania stopnia zaawansowania choroby Parkinsona.

Podstawowym problemem w dziedzinie medycyny jest zmieniająca się wiedza medyczna dotycząca mózgu i nowe technologie diagnozowania mózgu, w szczególności w zakresie choroby Parkinsona. Pociąga to za sobą konieczność opracowywania nowych, bardziej dokładnych metod analizy danych, wykorzystujących specyfikę nowych technologii. Istotną częścią pracy jest wprowadzenie nowych biomarkerów oraz udowodnienie ich skuteczności w automatycznej klasyfikacji symptomów choroby Parkinsona. Doktorant koncentruje się nad możliwościami zbudowania nowych modeli klasyfikacji i predykcji, których celem jest dokładniejsze diagnozowanie stanu choroby Parkinsona, w oparciu o nowatorskie biomarkery związane z ruchem gałek ocznych. W pracy autor proponuje modele oparte na technikach uczenia maszynowego, eksploracji danych oraz metodach zbiorów przybliżonych.

Tematyka rozważana przez doktoranta jest niezwykle ważna i niewątpliwie jest godna rozprawy doktorskiej.

1.2 Cel rozprawy

Zasadniczym celem pracy było opracowanie nowych metod analizy danych, prostych w zastosowaniu, pozwalających wcześniej diagnozować chorobę Parkinsona. Obszarem badań dla znalezienia takich metod są dla doktoranta metody eksploracji danych oraz uczenia maszynowego. Doktorant koncentruje się na zbudowaniu modeli predykcyjnych przez wykorzystanie zbiorów danych badawczych zawierających wyniki obserwacji biomarkerów związanych z ruchem gałek ocznych. Cel jaki postawił sobie doktorant, to udowodnienie skuteczności zaproponowanych rozwiązań w automatycznej klasyfikacji symptomów choroby Parkinsona, oraz wskazywania terapii.

1.3 Charakter rozprawy i znaczenie praktyczne badań

Praca ma przede wszystkim charakter eksperymentalny, jest jednak mocno poparta technicznym warsztatem autora, pozwalającym konstruować nowe rozwiązania techniczne. Pozytywnie oceniam też warsztat teoretyczny doktoranta. Propozycje autora są wsparte zaawansowanymi implementacjami oraz eksperymentami. Wskazują one na

1. solidny warsztat doktoranta w zakresie narzędzi uczenia maszynowego;
2. duży potencjał praktycznego wdrażania opracowanych algorytmów;

2. Wkład autora

Wkład autora obejmuje szereg ważnych elementów związanych z uczeniem maszynowym na potrzeby analizy danych otrzymywanych w procesie diagnozowania pacjentów:

1. Istotnym wkładem autora jest zidentyfikowanie i wprowadzenie nowych biomarkerów związanych z ruchem gałki ocznej. W szczególności w [1] pokazano, że jeden parametr testu wraz z informacją o długości okresu trwania choroby wystarcza do właściwego zdiagnozowania stanu choroby.
2. W pracach [1,2,4,5] zaprezentowano nowe podejście do budowania modeli służących do diagnozowania choroby Parkinsona, co pozwoliło stworzyć specjalizowane środowisko do budowania modeli w zależności od zmieniających się potrzeb, wynikających z wprowadzenia nowych biomarkerów. W

szczególności, w [5] pokazano znaczenie metod bazujących na teorii zbiorów przybliżonych, mianowicie pokazano, że zastosowanie metod RST i FRST daje większą dokładność klasyfikatorów.

3. W pracach [3,6,7] przeprowadzono dodatkowo badania nad możliwościami metod eksploracji danych uzyskanych w warunkach zmian emocjonalnych. W [3] badano efekt emocji w sesjach muzycznych, zaś w pracach [6,7] badano metody bazujące na danych uzyskiwanych przez pomiary mięśni twarzy w wyrażaniu emocji. Badania nad metodami eksploracji danych, związane z sesjami muzycznymi pozwoliły stworzyć modele o dużej precyzji. W badaniach [6,7] pokazano, że uzyskany model potwierdza autentyczność emocji zadowolenia. Zaproponowano metodę odkrywania emocji w oparciu o dane z pomiarów mięśni twarzy. W szczególności w [7] pokazano model, który rozpoznaje „udawany” uśmiech od faktycznego zadowolenia. Utworzony model okazał się bardzo przydatny w diagnozowaniu choroby Parkinsona.
4. Praca [8] jest przeglądem literatury w zakresie metod uczenia maszynowego w zastosowaniu do diagnozowania choroby Parkinsona
5. Wysoko oceniam wyniki uzyskane w [9,10]. W pracach tych zaproponowano środowisko komputerowe (aplikację), które po zainstalowaniu na komputerze pacjenta zbiera automatycznie dane z kamery, związane z ruchem gałek ocznych, pozycją twarzy, itp. Aplikacja taka może w połączeniu z opracowanymi metodami pokazanymi w [1-7] stanowić kompletne narzędzie praktyczne w diagnozowaniu choroby Parkinsona. Istotną zaletą propozycji jest prostota zastosowanych.

Zaproponowane rozwiązania mają charakter techniczny, wymagały one jednak od autora dobrego warsztatu naukowego i znajomości narzędzi sztucznej inteligencji i eksploracji danych. Wartość opracowanych metod analizy danych na potrzeby diagnostyki w chorobie Parkinsona została potwierdzona publikacjami zamieszczonymi w pracy.

3. Wiedza kandydata

W rozprawie doktorantka wykazuje się dużą wiedzą, przede wszystkim w zakresie uczenia maszynowego, w metodach analizy danych, w tym tych bazujących na zbiorach przybliżonych oraz rozmytych-przybliżonych.

4. Inne uwagi

Pracę stanowi wstęp oraz 10 publikacji. Uważam, że przedstawione artykuły stanowią dorobek, który jest dobrą podstawą do rozprawy doktorskiej.

W trakcie czytania rozprawy nasunęła mi się wątpliwość związana z brakiem odwołania się w rozprawie (np. w części 1.2 Literature Review) do doktoratu Artura Szymańskiego.

W związku z tym mam następujące pytania do doktoranta:

1. jakie są różnice pomiędzy modelami zastosowanym w rozprawie dra Szymańskiego oraz w pracy doktoranta bazującymi na RST i FRST;
2. Czy modele te są budowane w oparciu o te same dane i/lub metody;
3. Jeżeli opracowano nowe metody w opiniowanej rozprawie, czy są one konkurencyjne w stosunku do wcześniejszych, jeżeli, tak to w jakim zakresie (np. precyzja, mniejsze nakłady na stworzenie modelu, szybkość uzyskania diagnozy bez utraty dokładności, itp.).

5. Podsumowanie

Pozytywnie oceniam przedstawione w doktoracie artykuły i wkład doktoranta. Rozprawa spełnia wymagania zawarte w obowiązujących przepisach dotyczących rozpraw doktorskich, wnoszę zatem o dopuszczenie mgra Alberta Śledzianowskiego do publicznej obrony.

A. All.