

Prof. dr hab. Daniel Krzysztof Wójcik

Warszawa, 24.06.2024

Pracownia Neuroinformatyki

Instytut Biologii Doświadczalnej

im. Marcelego Nenckiego PAN

02-093 Warszawa, ul. Pasteura 3

Recenzja rozprawy doktorskiej mgr inż. Mateusza Ozimka pt.:
„Zastosowanie wybranych metod entropowych w badaniach EKG”.

Choroby układu krwionośnego są jedną z głównych przyczyn zgonów na świecie i chociaż metody ich diagnostyki i leczenia systematycznie się rozwijają, cały czas potrzebujemy lepszych, tańszych, dokładniejszych, jak najmniej inwazyjnych metod diagnostycznych. Celem rozprawy mgr inż. Mateusza Ozimka jest zbadanie przepływu informacji pomiędzy rytmem serca a repolaryzacją komórek serca a przy tym licznej grupy metod analizy sygnału EKG opartych o teorię informacji pod kątem możliwości ich wykorzystania do charakterystyki i ewentualnie diagnostyki różnych chorób układu krwionośnego. Cel został osiągnięty. Autor analizuje dane pozyskane z bazy THEW (Telemetric and Holter ECG Warehouse) oraz z Instytutu Kardiologii w Aninie, z którym współpracuje. Z bazy THEW pozyskano rejestracje EKG osób zdrowych oraz chorych z syndromem wydłużonego odstępu QT (grupa LQTS) lub z chorobą wieńcową. Z Instytutu Kardiologii otrzymano zapisy analizowanych przez autora szeregów interwałów RR, QT, oraz DI osób zdrowych oraz pacjentów z kardiomiopatią przerostową.

Recenzowana rozprawa zawiera 140 stron maszynopisu w dziewięciu rozdziałach, odnosi się do około 140 publikacji i jest bogato ilustrowana ponad 100 rysunkami i wykresami. Autor zaczyna rozprawę od przedstawienia swoich motywacji i celów pracy, następnie omawia jej zakres i układ. W rozdziale drugim autor omawia badania EKG i podstawowe dane, jakie będzie analizować w swojej rozprawie. Przedstawia też trzy schorzenia układu krwionośnego, które będzie charakteryzować metodami entropijnymi oraz standardową diagnostykę tych chorób. W rozdziale trzecim autor omawia używane metody obliczeniowe, w szczególności metody liniowe używane w analizie szeregów czasowych otrzymywanych z EKG, metody rozkładu entropii, które są głównym narzędziem badawczym w rozprawie, entropie Tsallisa oraz wykorzystane metody klasyfikacji. Część metodyczną kończy rozdział 4, w którym autor omawia dane, które analizował, ich własności i pochodzenie, oraz przetwarzanie wstępne – usuwanie arytmii, niestacjonarności, wybór odprowadzenia, wyznaczanie interwałów RR, QT, DI, itp. Autor bada też tutaj aspekty metodyczne wyznaczania miar entropicznych, takie jak wpływ wyboru długości analizowanego okna sygnału, metod usuwania niestacjonarności czy algorytmów wyznaczania interwałów QT na otrzymane wyniki.



Wyniki są przedstawione w rozdziałach od 5 do 7, gdzie autor po kolei porównuje własności interwałów otrzymanych z EKG osób zdrowych i pacjentów z syndromem wydłużonego odstępu QT, z chorobą wieńcową oraz z kardiomiopatią przerostową. Najobszerniejszy jest rozdział 5, oparty na wcześniejszej publikacji doktoranta. Należy podkreślić systematyczność analizy oraz uważne omówienie istotnych aspektów klinicznych, które mogą mieć wpływ na wyniki analizy, takich jak różne przyczyny obserwacji syndromu (warianty LQTS1 i LQTS2), zbadanie wpływu leków na rejestrowaną aktywność, oraz sygnatury omdleń. W rozdziale piątym podjęta jest jedyna próba stworzenia klasyfikatora stanu zdrowia pacjenta w oparciu o otrzymane wyniki.

Rozprawę zamykają dwa krótkie rozdziały. Rozdział ósmy podsumowuje wyniki i przedstawia wnioski ogólne autora, rozdział dziewiąty proponuje dalsze kierunki badań.

Rozprawę oceniam bardzo pozytywnie. Autor wykazał się znajomością materiału z którym pracuje, czyli chorób układu krwionośnego oraz własności danych EKG na poziomie absolutnie adekwatnym do analizy tych danych i interpretacji wyników ze zrozumieniem. Jest dobrze odczytany w powiązanej literaturze i wchodzi z nią w dialog. Dobrze rozumie stosowane metody obliczeniowe i statystyczne, a zarazem czuje ograniczenia swojego podejścia, wynikające z charakteru używanych metod czy z własności samych danych, co zwiększa zaufanie czytelnika do otrzymanych przez niego wyników. Podoba mi się, że autor wskazuje potrzebę badań prospektywnych, mam nadzieję, że będzie miał szansę w takich badaniach brać udział. Szczególnie starannie opisano przetwarzanie wstępne otrzymanych danych, co jest istotną częścią procesu analizy.

Zgodnie z obowiązującą Ustawą recenzent jest zobowiązany stwierdzić, czy recenzowana rozprawa spełnia dwa warunki:

„1. Rozprawa doktorska prezentuje ogólną wiedzę teoretyczną kandydata w dyscyplinie albo dyscyplinach oraz umiejętność samodzielnego prowadzenia pracy naukowej.

2. Przedmiotem rozprawy doktorskiej jest oryginalne rozwiązanie problemu naukowego albo oryginalne rozwiązanie w zakresie zastosowania wyników własnych badań naukowych w sferze gospodarczej lub społecznej.”

Uważam, że przedstawiony kontekst dziedzinowy jest wystarczający i oceniam go pozytywnie. Problem diagnostyki chorób serca jest ważny i w sposób twórczy i oryginalny został zbadany w niniejszej rozprawie z wykorzystaniem metod teorio-informacyjnych, co wyczerpuje drugi warunek ustawowy.

Niezależnie od mojej wysokiej oceny tej rozprawy mam kilka uwag rozmaitej natury, częściowo polemicznych, częściowo technicznych, które dla kompletności oceny tu przedstawię. Ważniejsze uwagi:

- W rozdziale siódmym autor porównuje osoby zdrowe z chorymi na kardiomiopatię przerostową budując mniejszą sieć fizjologiczną niż we wcześniejszych rozdziałach, ponieważ nie ma oryginalnych zapisów EKG. Warto byłoby porównać analizy z wcześniejszych rozdziałów tymi samymi metodami, żeby pokazać w jakim zakresie brak pełnego EKG jest ograniczające dla wykorzystania metod nieliniowych. W szczególności ciekaw jestem jak brak informacji o amplitudach sygnału wpływa na jakość klasyfikacji z rozdziału 5.10 i czy taki klasyfikator można wprowadzić dla przypadków z rozdziałów 6 i 7.



- Bardzo ciekawe byłoby porównanie poszczególnych grup pacjentów i osób zdrowych. Autor jest tego świadom, twierdzi, że dla uproszczenia dyskusji tego nie wykonał. Zgadza się, że rozprawa jest wystarczająco skomplikowana, a np. budowa systemu decyzyjnego wspierającego diagnostykę różnych badanych tu chorób to inny obszerny projekt. Tym niemniej zabrakło mi prób porównania otrzymanych wyników i próby bardziej obszernego zorganizowania danych we wspólnym kontekście. Jedną z relatywnie prostych rzeczy, które można byłoby wykonać, jest np. porównanie wyników otrzymanych tymi samymi metodami dla pacjentów zdrowych z obu zestawów danych. Pomogłoby to oszacować na ile można polegać na poszczególnych miarach używanych przez autora jako charakterystyk choroby a nie tylko zestawu danych. Na przykład na rysunkach 5.8 i 7.2 autor porównuje entropię jednowymiarową dla szeregów czasowych interwałów QT dla osób zdrowych i pacjentów. Nie mając danych oczywiście nie policzyłem statystyki, ale na pierwszy rzut oka wyniki dla HY i NY są spójne (nie ma różnic istotnych statystycznie – a może są?) natomiast wyniki dla SY na moje oko różnią się istotnie statystycznie. Wykonanie tego typu analiz pozwoliłoby sprawdzić na ile poszczególne miary są rzeczywiście reprezentatywne dla populacji zdrowych, co potencjalnie może być istotne w interpretacji wyników otrzymanych z wykorzystaniem tych samych miar pacjentów.
- Wyniki z rozdziałów 5 i 6 proszą się o integrację. Przecież wyniki dla pacjentów zdrowych są tożsame. Umieszczenie ich na wspólnych wykresach ułatwiłoby analizę porównawczą nie tylko pacjentów chorych ze zdrowymi ale także między chorobami. Np. Rys. 5.11 i 6.4 dają się wprost zintegrować.
- Autor stawia tezę, że używane przez niego markery entropowe dają powtarzalne rezultaty. Uważam, że nie wykorzystał wszystkich dostępnych danych, żeby to sprawdzić i uzasadnić. Patrz przykład powyższy dotyczący zdrowych pacjentów, rys. 5.8 i 7.2..
- Praca prosi się o składową modelową. Uważam, że byłaby znacznie mocniejsza, gdyby przed analizą danych doświadczalnych autor wykorzystał dane modelowe, nawet korzystając z cudzych modeli, choćby budowanych wcześniej w grupie prof. Żebrowskiego czy w innych miejscach. Analiza danych z symulacji być może pozwoliłaby powiązać otrzymane parametry przepływu informacji z konkretnymi własnościami układu krwionośnego.

Pozostałe uwagi:

- To nie jest dobra praktyka, żeby cytowania umieszczać w tytułach podrozdziałów (3.5.1, 3.5.2).
- Rozprawa jest czytelna ale z powodzeniem można było ją przedstawić w klasycznym układzie Wprowadzenia (rozdziały 1 i 2), Metod (3 i 4), Wyników (5-7) oraz Dyskusji (8-9).
- Dobrą praktyką jest rozpoczynanie kolejnych rozdziałów na stronach nieparzystych.
- Akronimy na stronie 7 (SDNN, rMSSD) nie są powszechnie znane a rozszyfrowane i zdefiniowane dopiero znaczeni później, w metodach.
- Niektóre fragmenty opisu metod są niezręczne. Np. na stronie 19 równanie (3.26) i jego opis nic nie wnoszą a tylko konfundują czytelnika. W istocie autor ma chyba na myśli nadchodzące równanie (3.28), w którym nota bene na początku powinno być chyba U_i^m .



- Zastanawia mnie dlaczego porównując entropię próby (3.29) z entropią danych odniesienia (zastępczych) autor patrzy na ich średnią ale nie liczy chociaż SEM, podobnie dla miar S_{diff_q} . Skąd wiemy na ile wiarygodne są oszacowania z równań (3.30) i (3.31) itp. dla S_{diff_q} ?
- W równaniu (3.49) pojawiają się niezdefiniowane współczynniki ξ .
- W rozdziale 4.3 zaniepokoiło mnie dlaczego najpierw autor pisze o usuwaniu trendu wielomianowego wielomianem stopnia 3, a na rysunku 4.6 pokazuje wyniki dla stopni 2. Dopiero w rozdziale 4.8 okazuje się, że zbadał kilka sposobów usuwania niestacjonarności. Jakoś należałoby to przeorganizować.
- Podobają mi się porównania histogramów, takie jak na rysunkach 4.7-10 czy 5.3. Wydaje mi się, że takich rysunków mogłoby być więcej, są dla mnie czytelniejsze niż wybrane przez autora wykresy pudełkowe, chociaż oczywiście przyjęta konwencja jest bardziej zwarta.
- Nie rozumiem ostatniego akapitu na stronie 46 i pierwszego na stronie 47. Co to znaczy w szczególności zdanie: „Różnice pomiędzy wersjami algorytmu wyznaczania końca załamka T występują na 10 parametrów jedynie w przypadku nowej informacji szeregów czasowych interwałów RR oraz entropii własnej interwałów QT.” Czy chodzi o to, że liczone miary entropii nie zależą od algorytmu wyznaczania końca załamka T?
- Nie wiem co oznacza kolor zielony a co czarny w tabelach 4.7 i 4.8, nie opisano tego.
- Opisy wykresów nie są spójne. Na przykład na rys. 5.8 A autor pisze $\bar{S}_r$ QT HY, a na rys. 7.2 pisze $\bar{S}_r$ HY QT, itd.
- W rozdziale 5.5 autor pisze, że „otrzymane wyniki dla markerów jednowymiarowych i dwuwymiarowych wskazują na możliwość ich równoczesnego wykorzystania w celu zwiększenia jakości diagnostyki różnicującej typu LQTS”. Jak konkretnie auto wyobraża sobie taki biomarker?
- W rozdziale 5.6 autor zwraca uwagę, że „Histogramy [] wartości [entropii trójwymiarowych] nakładają się dla osób zdrowych oraz pacjentów z LQTS w znacznie mniejszym stopniu niż dla QTc .” Warto byłoby to pokazać na rysunku w stylu rysunku 5.3.
- Przykładem niespójności nazewnictwa jest entropia próby, raz nazywana SampRatio, innym razem jest to SampEnRatio. Na przykład na początku strony 77 pojawia się SampRatio, podobnie w opisie wykresu 5.41.C, ale już w podpisie rysunku 5.41 jest SampEnRatio, itd.
- Przyjęta konwencja kolorystyczna jest spójna w ramach rozdziałów. Autor używa jednak innego odcienia niebieskiego na oznaczenie pacjentów zdrowych w rozdziale 5 a w rozdziałach 6 i 7 takiego samego. Sugerowałbym albo użycie tego samego koloru wszędzie, albo takiego samego w 5 i 6 – to w końcu te same wyniki – a innego w rozdziale 7.
- Na stronie 110: Rys 9.1 $\rightarrow$ 8.1
- problemy z cytowaniami – inicjały zamiast nazwisk, np. [25], [29]; brak kompletnych cytowań, np. [52], [73], [84]. Inne dziwne, np. [60].



Pomimo przedstawionych powyżej uwag krytycznych moja ogólna ocena pracy mgr. inż. Mateusza Ozimka jest zdecydowanie pozytywna. Autor wykazał się znajomością zagadnień z teorii transferu entropii, metod numerycznych, fizyki, elektrofizjologii, kardiologii. Praca jest silnie interdyscyplinarna. Uważam, że stanowi ona istotny przyczynek do rozwoju kardiologii obliczeniowej i fizyki biomedycznej. **Uważam, że rozprawa doktorska Pana mgr inż Mateusza Ozimka spełnia zwyczajowe i formalne warunki stawiane rozprawom doktorskim i dlatego w oparciu o przepisy Ustawy Prawo o szkolnictwie wyższym i nauce wnoszę do Wysokiej Rady Naukowej Dyscypliny Nauki Fizyczne Politechniki Warszawskiej o przyjęcie rozprawy i dopuszczenie autora do dalszych etapów przewodu doktorskiego.** Jednocześnie biorąc pod uwagę szeroki zakres przeprowadzonych badań, fakt opublikowania dużej ich części w czterech publikacjach, w tym jednej pierwszoautorskiej w dobrym czasopiśmie, oraz potencjalne znaczenie translacyjne uzyskanych wyników w kardiologii klinicznej wnoszę o wyróżnienie rozprawy mgr. inż. Mateusza Ozimka.

Prof. dr hab. Daniel K. Wójcik