

Sosnowiec 10-08-2017

RECENZJA
rozprawy na stopień doktora nauk o zdrowiu
przygotowana na zlecenie Dziekana Wydziału Nauk o Zdrowiu
z Oddziałem Pielęgniarstwa i Instytutem Medycyny Morskiej i Tropikalnej
Gdańskiego Uniwersytetu Medycznego (znak: DNZ-255-2/2017)

Zakład Statystyki

Wydziału
Farmaceutycznego
z Oddziałem Medycyny
Laboratoryjnej
w Sosnowcu

41-200, Sosnowiec,
ul. Ostrogórska 30
www.sum.edu.pl

KIEROWNIK ZAKŁADU

dr hab. inż. Aleksander Owczarek
tel.: (+48 32) 364 13 28

aowczarek@sum.edu.pl

SEKRETARIAT

tel.: (+48 32) 364 12 90

Autor: mgr inż. Agata Joanna Czerniecka

Tytuł: Nowa metoda obliczeniowa porównywania sekwencji białek

Przedstawiona mi do recenzji rozprawa doktorska obejmuje 94 strony tekstu podzielonego na następujące rozdziały: „*Wstęp*” zawierający część teoretyczną oraz podrozdział dotyczący macierzy substytucji, „*20-wymiarowa dynamiczna reprezentacja sekwencji białkowych*” oraz „*Załączniki*” i „*Bibliografia*”. W pracy umieszczono ponadto streszczenie w języku polskim i angielskim, brak jest jednak wymaganego wykazu skrótów i oznaczeń stosowanych w tekście, spisu tabel oraz spisu rycin. Nazewnictwo poszczególnych rozdziałów zgodne jest z tematem i zawartymi treściami. Układ przedstawionej pracy odbiega od typowego sposobu prezentowania prac badawczych, brak jest np. szerszej dyskusji.

Wyniki zaprezentowano w 15 tabelach oraz na 74 rycinach. Piśmiennictwo obejmuje łącznie 72 pozycje literaturowe, zarówno polsko- jak i angielskojęzyczne, właściwie dobrane, w większości aktualne oraz prawidłowo cytowane, o istotnym znaczeniu merytorycznym.

Promotorem w przewodzie jest Pani dr hab. Dorota Bielińska-Wąż z Zakładu Informatyki Radiologicznej i Statystyki Katedry Medycyny Nuklearnej i Informatyki Radiologicznej Gdańskiego Uniwersytetu Medycznego.

Uwagi merytoryczne

Rozprawa dotyczy istotnego zagadnienia w zakresie bioinformatyki oraz biologii obliczeniowej, jakim jest ustalanie pokrewieństwa ewolucyjnego oraz analiza podobieństwa sekwencji białek. Biorąc pod uwagę rozwój medycyny i biologii, wymuszający tworzenie nowych, doskonalszych narzędzi matematycznych umożliwiających modelowanie i analizę złożonych zależności świata biologicznego, podjęcie niniejszej tematyki przez Doktorantkę uznaję za istotne oraz w pełni uzasadnione.

1. Wstęp

We wstępie, liczącym 24 stron, Doktorantka omówiła podstawowe definicje z zakresu standardowych i niestandardowych metod reprezentacji sekwencji DNA oraz sekwencji białkowych. Na stronie 8 wspomina o metodach minimalizacji zjawiska degeneracji reprezentacji sekwencji białkowych, jednak nie precyzuje nieco szczegółowiej o jakie metody chodzi (metody numeryczne, algorytmy?). W części teoretycznej Doktorantka omówiła budowę białek, drzew filogenetycznych, metod porównywania sekwencji białkowych, w tym macierzy tempa podstawień dla sekwencji białek (PAM, BLOSUM) oraz graficznych metod porównywania białek. Nie przedstawiła jednak bliżej metody UPGMA na podstawie, której miała wyznaczyć drzewo filogenetyczne do porównań z proponowaną metodą reprezentacji sekwencji białkowych. W rozdziale dotyczącym metody Virtual Genetic Code, na stronie 17 pojawiają się stwierdzenia: „...alternatywą dla tej metody jest przedstawienie dwóch sekwencji łącząc je arytmetycznie...” oraz „...inne metody wizualizacji...nazwać spiralą...”. Cennym uzupełnieniem tego fragmentu byłoby opisanie pokrótce tychże metod wizualizacji oraz prezentacja przykładowych rycin. Również na stronie 21 w podrozdziale dotyczącym metody *Magic Circle* zabrakło przykładu wykresu prezentującego różnice pomiędzy współrzędnymi odpowiadających sobie aminokwasów dwóch białek. W rozdziale dotyczącym macierzy substytucji na stronie 27 nie podano wyjaśnienia oznaczenia A we wzorach 2 – 4.

Reasumując, wstęp pracy jest napisany poprawnie, zawiera właściwe, merytoryczne treści związane z tematem badania i logicznie uzasadnia celowość podjęcia badań.

2. 20-wymiarowa dynamiczna reprezentacja sekwencji białkowych

W rozdziale tym Doktorantka przedstawiła matematyczne podstawy proponowanej nowej metody reprezentacji sekwencji białkowych. Opis metody jest zgodny z opisem zaprezentowanym w opublikowanym przed Doktorantką artykule. W metodzie przyjęto jako nowe deskryptory 20-wymiarowego wykresu, znormalizowane główne momenty bezwładności, będące pierwiastkami z wartości własnych równania charakterystycznego podzielone przez długość sekwencji. Nasuwa się zatem pytanie, czy Doktorantka rozważała próbę rozszerzenia metody i dokonania oceny ewentualnych innych deskryptorów i ich wpływu na uzyskane drzewa filogenetyczne? W rozprawie, tak jak i w publikacji, przyjęto miarę podobieństwa par sekwencji w postaci wartości bezwzględnej różnicy sumy deskryptorów podzielonej przez sumę sumy deskryptorów odpowiedniej sekwencji białkowej. Prowadzi to do wyniku, w którym wartości 0 odpowiada pełna zgodność, a wartości 1 całkowity brak zgodności. Czy nie lepiej byłoby dokonać normalizacji w kierunku przeciwnym? Czy Doktorantka rozważała zastosowanie innej miary podobieństwa i metody normalizacji oraz próby oceny jak zmieni to uzyskane drzewa filogenetyczne?

W rozdziale dotyczącym metody Jacobiego we wzorze 21 podano błędną numerację kolejnych transformacji ortogonalnych ($\{n\}$).

3. Wyniki

Wyniki badań zostały przez Doktorantkę przedstawione w uporządkowany i przejrzysty sposób w oparciu o analizę sekwencji białkowych dostępnych w bazie PDB. Nie mniej jednak, zabrakło w rozprawie uzasadnienia, dlaczego do analizy przyjęto akurat takie reprezentacje białek, jak dehydrogenaza NADH podjednostki 4 oraz 6, czy bakulowirusów. Doktorantka przedstawiła również drzewa filogenetyczne stworzone z wykorzystaniem metody Dynamic Time Warping, jednakże trudno jest porównać wyniki obu metod ze względu na różną liczbę gatunków przyjętych do analizy. Brak jest również wyników uzyskanych nową proponowaną w rozprawie metodą dla porównania z wykorzystaniem wirusa grypy typu A. Analizę wyników utrudnia również znaczna liczba rycin i brak ich spisu.

4. *Dyskusja i wnioski*

Istotnym mankamentem pracy jest brak szerszej dyskusji, w której Doktorantka odniosłaby się do innych metod umożliwiających tworzenie drzew filogenetycznych (w tym do metody Dynamic Time Warping), w porównaniu do wyników uzyskanych z wykorzystaniem proponowanej, nowej metody dynamicznej reprezentacji sekwencji białkowych. W rozprawie zabrakło również wniosków z przedstawionych wcześniej wyników, uzyskanych dla różnych sekwencji białkowych i porównań z metodą Dynamic Time Warping.

5. *Załączniki.*

W załącznikach przedstawiono wydruki programów napisanych przez Doktorantkę w środowisku Matlab. Zabrakło jednak schematów blokowych lub np. diagramów UML prezentujących specyfikację oprogramowania. Doktorantka nie przeprowadziła także analizy złożoności obliczeniowej proponowanej metody oraz metody Jacobiego, co pozwoliłoby ocenić konieczne wymagania sprzętowe. Czy Doktorantka rozważała opracowanie aplikacji implementującej zaproponowaną przez nią metodę, która byłaby niezależna od środowiska Matlab?

Uwagi redakcyjne

Doktorantka nie ustrzegła się drobnych błędów stylistycznych oraz interpunkcyjnych: na stronie 8 jest „...porównać ze sobą białka...”, powinno być „...porównanie ze sobą białek...”, na stronie 9 jest „... można wywnioskować aktywność biologiczną”, powinno być „...można wnioskować o aktywności biologicznej...”, na stronie 19 jest „...zostało zauważony”, powinno być „...stwierdzono”

Na rysunku 1 błędnie w opisie przedstawiono grupę aminową (jest NH₃ zamiast NH₂). Wprowadzając skróty w tekście Doktorantka nie zawsze rozwijała je prezentując oryginalne nazewnictwo oraz odpowiadające polskie tłumaczenie, nie uniknęła również w kilku miejscach powtórzeń wyrazów.

Podsumowanie

Przedstawiona mi do oceny rozprawa dotyczy istotnego zagadnienia w dziedzinie bioinformatyki, jakim jest ustalanie pokrewieństwa ewolucyjnego oraz analiza podobieństwa sekwencji białek. Zaprezentowane w rozprawie wyniki odpowiadają postawionemu celowi pracy. Doktorantka przedstawiła drzewa filogenetyczne stworzone z wykorzystaniem opracowanej przez nią metody porównywania sekwencji białek.

Reasumując, mimo wcześniejszych uwag krytycznych **oceniana rozprawa doktorska Pani mgr inż. Agaty Joanny Czernieckiej spełnia warunki określone w art. 13 ust. 1 ustawy z dnia 14 marca 2003 r. o stopniach i tytule naukowym oraz o stopniach i tytule w zakresie sztuki (Dz. U. Nr 65, poz. 595, z późn. zm.).**

Na podstawie dokonanej powyższej, pozytywnej oceny, zwracam się do Wysokiej Rady Wydziału Nauk o Zdrowiu z Oddziałem Pielęgniarstwa i Instytutem Medycyny Morskiej i Tropikalnej Gdańskiego Uniwersytetu Medycznego o dopuszczenie Doktorantki do dalszych etapów postępowania w przewodzie doktorskim.

Aleksander
dr hab. n. o zdr. inż. Aleksander Owczarek
specjalista w dziedzinie epidemiologii
Owczarek