

Gdynia, 29 września 2023 r.

Prof. dr hab. inż. Ireneusz Czarnowski
Katedra Systemów Informatycznych
Uniwersytet Morski w Gdyni
ul. Morska 83, 81-225 Gdynia

RECENZJA

rozprawy doktorskiej mgr. Michała Bałchanowskiego

pt.: Ewolucyjna agregacja rang w systemach rekomendacyjnych.

Recenzję przygotowałem w związku z wyznaczeniem przez Radę Naukową Instytutu Informatyki Uniwersytetu Śląskiego w Katowicach w dniu 17 lipca 2023 r. mojej osoby jako recenzenta rozprawy doktorskiej mgr. Michała Bałchanowskiego.

Rozprawa doktorska została przygotowana w dyscyplinie *informatyka*. W tej samej dyscyplinie prowadzone jest postępowanie w sprawie nadania mgr. Michałowi Bałchanowskiemu stopnia doktora.

1. Problematyka naukowa rozprawy

Tematyka rozprawy dotyczy poszukiwania nowych efektywnych algorytmów rekomendacyjnych. Problem generowania rekomendacji jest aktualnym problemem i wyzwaniem z punktu widzenia algorytmicznego, ale również z uwagi na zainteresowanie sfery komercyjnej sposobami wyboru i zaproponowania potencjalnemu użytkownikowi internetowego systemu zakupowego lub grupom tych użytkowników najbardziej trafnej odpowiedzi spośród szerokiego zbioru możliwych. Konkretnie zaimplementowane rozwiązania mogą dotyczyć rekomendacji w ogólnym rozumieniu produktów oferowanych w wybranym kanale dystrybucji elektronicznej. Rekomendacje są także nieodzowną funkcjonalnością innych systemów, w tym systemów społecznościowych lub, w ogólności, serwisów internetowych, w których spośród wielu informacji celem jest wybór i prezentacja tych, które w największym możliwym stopniu będą odpowiadać preferencjom użytkownika takiego systemu lub serwisu. Wyzwaniem w poszukiwaniu dobrych jakościowo rekomendacji jest także agregowanie rekomendacji (tudzież rang) pochodzących w kilku rozwiązaniach algorytmicznych i przedstawianie tej zagregowanej rekomendacji użytkownikowi systemu jako rekomendacji docelowej. Problem agregacji rekomendacji nie jest problemem trywialnym. Trudno również wskazać jedną uniwersalną, a tym bardziej możliwie najlepszą metodę łączenia rekomendacji. Ten aspekt stanowi także zainteresowanie badawcze Doktorant, co przedstawił w recenzowanej rozprawie doktorskiej.

W szczególności Doktorant dla generowania rekomendacji zaproponował algorytm ewolucji różnicowej (ang.: *differential evolution*, DE), którego celem jest agregacja rang wygenerowanych przez bazowe algorytmy rekomendacyjne i zaprezentowanie użytkownikowi finalnego rankingu, który będzie dostosowany do jego personalnych preferencji. Sam aspekt zastosowania metaheurystyki do generowania rekomendacji stanowi także istotną unikatowość podjętych przez Doktoranta badań.

Główna teza pracy doktorskiej mgr. Michała Bałchanowskiego opiera na stwierdzeniu, że zaproponowany algorytm ewolucyjnej agregacji rang poprawia jakość generowanej agregacji, w porównaniu z wybranymi metodami zaproponowanymi w literaturze.

Podsumowując aspekt podjętej przez Doktoranta problematyki badawczej, należy stwierdzić, że wpisuje się w aktualny i ważny nurt badawczy, a sam problem został sformułowany w sposób klarowny i zrozumiały.

2. Treść rozprawy

Rozprawa doktorska została przygotowana w języku polskim i łącznie składa się z 137 stron, w tym:

- wstępu,
- ośmiu kolejnych rozdziałów, w tym rozdziału prezentującego zaproponowany algorytm oraz wyniki eksperymentu obliczeniowego,
- rozdziału podsumowującego,
- wykazu bibliografii zawierającej 193 pozycje.

Praca została także opatrzona w spis treści, wykaz oznaczeń, spis rysunków, spis tabel oraz spis algorytmów. Dodatek A rozprawy prezentuje wykresy i graficzne zestawienia uzupełniające rozdział prezentujący wyniki eksperymentu obliczeniowego.

We wstępie, po krótkim wprowadzeniu do tematyki badań, uzasadnieniu podjętej problematyki badań, Doktorant przedstawia tezę i cele rozprawy. Rozdział kończy omówienie zawartości oraz układu rozprawy.

W rozdziale drugim Doktorant przedstawia podstawowe informacje dotyczące systemów rekomendacyjnych. W rozdziale tym przedstawia On formalną definicję systemu rekomendacyjnego, taksonomię systemów rekomendacji, przegląd podstawowych metod oraz wybranych algorytmów rekomendacyjnych, oraz wskazuje wyzwania badawcze i implementacyjne związane z tej klasy systemami.

W rozdziale trzecim rozprawy Doktorant omawia wybrane miary, które wykorzystywane są do oceny jakości generowanych rekomendacji.

Rozdział czwarty pracy został poświęcony metaheurystykom, a w szczególności algorytmowi ewolucji różnicowej. W pierwszej kolejności mgr Bałchanowski przedstawia ogólne informacje dotyczące metaheurystyk, wraz z ich rysem historycznym, oraz ich zastosowań w systemach rekomendacyjnych, co czyni poprzez przegląd literatury. Podrozdział 4.3 stanowi opis algorytmu ewolucji różnicowej.

Rozdział piąty rozprawy doktorskiej mgr Bałchanowski poświęca omówieniu idei agregacji rang, podkreślając przy tym, iż jest to aspekt stosunkowo nowy i rzadko podejmowany w literaturze w kontekście systemów rekomendacyjnych. Doktorant wprowadza formalną definicję agregacji rang, a następnie odnosi się do metod agregacji informacji pochodzących z różnych źródeł i zaprezentowania jej w formie uporządkowanej listy (rankingu). W rozdziale tym Doktorant omawia także aspekt podobieństwa pomiędzy rankingami.

W rozdziale szóstym Doktorant omawia zagadnienie nazywane w literaturze jako nauką rangowania (z ang. *learning to rank*). Podkreśla on rolę algorytmów uczenia maszynowego w tworzeniu (predykcji)

listy rankingowej oraz wskazuje możliwość zastosowania tego podejścia w systemach rekomendacyjnych.

W rozdziale siódmym rozprawy Doktorant prezentuje algorytm ewolucyjnej agregacji rang (EAR). Prezentację tę poprzedza krótkim wprowadzeniem, w którym odwołuje się do motywacji podjętej pracy badawczej oraz do poprzednich rozdziałów (tj. rozdziału 5 oraz 6). Samej prezentacji algorytmu EAR dokonuje poprzez jego skrupulatne omówienie, przedstawienie pseudokodu oraz architektury systemu rekomendacyjnego opartego na idei ewolucyjnej agregacji rang. W dalszej części Doktorant omawia konstrukcję funkcji przystosowania. Podrozdział 7.4 prezentuje wersję rozszerzoną (zmodyfikowaną) algorytmu EAR, w którym do budowania rekomendacji dopasowanej do preferencji użytkownika wykorzystuje się rekomendacje ustalone dla jego najbliższych sąsiadów.

W rozdziale ósmym pracy Doktorant omówił zbiór danych *MovieLens 100k*, który został wykorzystany do przeprowadzenia eksperymentu obliczeniowego. Przedstawia także środowisko obliczeniowe oraz charakteryzuje sposób implementacji zarówno algorytmu EAR oraz innych algorytmów wykorzystanych jako referencyjne. Po omówieniu etapu parametryzacji poszczególnych wykorzystanych referencyjnych algorytmów, Doktorant omawia sposób przeprowadzenia eksperymentu obliczeniowego z wykorzystaniem EAR.

Rozdział dziewiąty rozprawy doktorskiej Doktorant poświęca wynikom eksperymentu obliczeniowego, które prezentuje w podziale na trzy grupy: wyniki badań podstawowych, wyniki badań eksperymentalnych oraz wyniki badań końcowych. W pierwszej grupie znalazły się wyniki działania algorytmów rekomendacyjnych oraz metod agregujących uzyskane na rzeczywistym zbiorze danych *MovieLens 100k*. Wyniki te zostały dość szczegółowo omówione wraz z analizą statystyczną (w tym korelacji) generowanych rekomendacji przez poszczególne porównywane algorytmy oraz dla różnych stosowanych miar oceny jakości generowanych rekomendacji. Druga grupa prezentuje wyniki badań z udziałem zaproponowanego algorytmu EAR dla różnych wariantów funkcji oceny. Wśród tych wyników są też wyniki uzyskane przez Doktoranta celem oceny różnych parametrów sterujących algorytmu EAR, w tym wpływy liczby sąsiadów w przypadku jego wersji rozszerzonej. Trzecia grupa wyników odzwierciedla te uzyskane przez porównywane algorytmy (w tym proponowany algorytm) na pełnym zbiorze danych. Doktorant w ten sposób uwypukla i potwierdza sposób dochodzenia do rozwiązań przez poszczególne algorytmy, co pozwala na zweryfikowanie stawianych pytań badawczych testami statystycznymi.

Rozdział dziesiąty pracy Doktorant poświęcił podsumowaniu uzyskanych wyników. W rozdziale tym sformułował także kierunki dalszych badań.

Spis literatury, do której odnosi się Doktorant w swojej rozprawie, nie wzbudza wątpliwości co do zasadności jej dobru i która jest adekwatna dla podjętej problematyki badań.

3. Najważniejsze wyniki uzyskane w pracy

Recenzowana rozprawa doktorska dotyczy agregacji rang w systemach rekomendacyjnych. W szczególności mgr Bałchanowski zaproponował nowy algorytm ewolucyjnej agregacji rang, który oparł na algorytmie ewolucji różnicowej. Należy także podkreślić, że koncepcja zaproponowanego podejścia została oparta na połączeniu nie tylko wiedzy dotyczącej systemów rekomendacyjnych i algorytmów metaheurystycznych, ale także z uwzględnieniem wiedzy dotyczącej agregacji rang oraz teorii związanej z uczeniem się rankingów. Zaproponowany algorytm został poddany ocenie na drodze eksperymentu obliczeniowego, a wyniki porównano z wynikami innych referencyjnych algorytmów i metod, co szeroko zostało omówione i skomentowane w rozprawie. Wyniki eksperymentu

obliczeniowego zostały przedstawione w sposób czytelny i rzetelny. Należy także stwierdzić, że cel pracy został osiągnięty, a teza rozprawy została potwierdzona.

Do oryginalnych osiągnięć Doktoranta zaliczam:

- zaproponowanie nowego konkurencyjnego algorytmu ewolucji agregacji rang oraz jego wersji rozszerzonej opartej na rekomendacjach najbliższych sąsiadów,
- przedstawienie algorytmu o cechach nadzorowanego algorytmu agregującego,
- zaproponowanie algorytmu agregacji rang opartego na metaheurystyce,
- przedstawianie algorytmu włączającego do procesu agregacji rankingów innych użytkowników,
- zaproponowanie alternatywnych funkcji oceny rozwiązań dla poszukiwań wektora preferencji w systemach rekomendacyjnych,
- dokonanie systematyki systemów rekomendacji i sposobów ich ewaluacji, metod agregacji rang oraz podejść wykorzystywanych w nauce rangowania.

Wybrane wyniki badań przedstawione w recenzowanej rozprawie zawarte są również w 5 artykułach naukowych, dla których mgr Bałchanowski jest współautorem oraz na które powołał się w recenzowanej rozprawie. Artykuły powstałe w latach 2020-2023 i zostały one opublikowane w materiałach konferencji naukowych oraz czasopismach naukowych.

4. Uwagi i pytania do recenzowanej pracy

Recenzowana rozprawa doktorska nasuwa również kilka pytań:

- Czy miara podobieństwa pomiędzy użytkownikami w algorytmie kNN rozszerzonej wersji EAR może mieć wpływ na jakość uzyskiwanych wyników (rekomendacji)?
- Jaka jest złożoność obliczeniowa algorytmu EAR oraz jego wersji rozszerzonej, i jak się ona ma do innych referencyjnych algorytmów?
- Czy inne (pozostałe) parametry algorytmu EAR, jak wielkość populacji czy parametry operatorów genetycznych, mogą mieć istotny wpływ na jakość uzyskiwanych wyników (rekomendacji)?

Z uwagi na prawdopodobny wpływ szeregu parametrów algorytmu EAR na jego osiągi, korzystnym byłoby przeprowadzenie wieloczynnikowej analizy wariancji, aby potwierdzić lub wykluczyć ewentualność statystycznej istotnej różnicy pomiędzy porównywanymi wersjami algorytmu EAR oraz jego wariantów wynikających z innej parametryzacji.

Ponadto, nasuwa się pytanie o możliwość zastosowania teorii konsensusu w formułowaniu sposobu agregacji rang dla systemu rekomendacyjnego. Czy takie podejście do rozwiązania problemu zdaniem Doktoranta jest możliwe?

Szkoda też, że eksperyment obliczeniowy został przeprowadzony w oparciu o jeden zbiór danych. Użycie szerszego spektrum problemów rekomendacyjnych poprzez dodatkowe zbiory danych wzmocniłoby wnioski z badań.

Ustosunkowanie się przez Doktoranta do powyższych uwag i pytań może być częścią dyskusji podczas obrony rozprawy doktorskiej.

5. Ocena redakcji i przygotowania rozprawy

Praca jest napisana w sposób przejrzysty i adekwatny dla poziomu rozprawy doktorskiej. Nie budzi wątpliwości strona językowa rozprawy. W pracy występują drobne błędy stylistyczne, interpunkcyjne i językowe, które nie ujmują jakości pracy.

6. Konkluzja

Pracę doktorską mgr. Michała Bałchanowskiego oceniam pozytywnie. Doktorant podjął się aktualnego problemu badawczego, odpowiednio go sformułował i przeprowadził badania z doбором właściwej metodologii. Przyjął też właściwą metodykę badań. Zaproponowany przez Doktoranta algorytm ewolucyjnej agregacji rang, jako oryginalny jego dorobek, jest wartościowy z punktu widzenia poznawczego, ale także przydatności praktycznej. Doktorant potwierdził, że posiada predyspozycje związane z prowadzeniem badań naukowych. Zaprezentował też ogólną wiedzę teoretyczną właściwą dla dyscypliny *informatyka* oraz odpowiednią dla osoby ubiegającej się o nadanie stopnia doktora.

Podsumowując, wyrażam opinię, że rozprawa doktorska mgr. Michała Bałchanowskiego pt. „*Ewolucyjna agregacja rang w systemach rekomendacyjnych*” spełnia wymogi stawiane przy ubieganiu się o nadanie stopnia doktora w dyscyplinie *informatyka*. Wnioskuje też o dopuszczenie rozprawy doktorskiej mgr. Michała Bałchanowskiego do obrony.