



POLISH-JAPANESE ACADEMY OF INFORMATION TECHNOLOGY

ポ
ー
ラ
ン
ド
日
本
情
報
工
科
大
学

Warszawa, 03.07.2023

prof. dr hab. Adam Wierzbicki
Polsko-Japońska Akademia
Technik Komputerowych

OPINIA O ROZPRAWIE DOKTORSKIEJ

MGR INŻ. MICHAŁA DANILUKA

"MULTIMODAL RECOMMENDATION SYSTEM FOR E-COMMERCE"

1. Jakie zagadnienie naukowe jest rozpatrzone w pracy i czy zostało ono dostatecznie jasno sformułowane przez autora? Jaki charakter ma rozprawa (teoretyczny, doświadczalny, inny)?

Mgr inż. Michał Daniluk sformułował 4 cele badawcze swojej rozprawy doktorskiej:

1. Opracować reprezentację danych, która uwzględnia właściwości dystrybucyjne danych behawioralnych
2. Umożliwić integrację różnych modalności danych w ramach ujednoliconej reprezentacji
3. Ustalić, czy samonadzorowane uczenie się dynamiki zachowań tworzy modele, które można transferować pomiędzy różnymi zadaniami predykcyjnymi
4. zweryfikować praktyczną przydatność modelowania zachowań opartego na własnościach dystrybucyjnych przez testy porównawcze i wdrożenie produkcyjne

Doktorant sformułował także 5 szczegółowych hipotez, które precyzują cele badawcze rozprawy. Weryfikacja hipotez 1-3 opisana jest w rozdziale 5, a hipotez 3-5 w rozdziale 6.

Rozprawa ma charakter badań stosowanych i jest wynikiem doktoratu wdrożeniowego realizowanego przez doktoranta we współpracy z firmą Synerise. Firma Synerise wdrożyła wiele wyników badań doktoranta w swojej działalności.

Cele badawcze i hipotezy rozprawy zostały sformułowane bardzo klarownie i precyzyjnie.

- 2. Czy w rozprawie przeprowadzono w sposób właściwy analizę źródeł (w tym literatury światowej, stanu wiedzy i zastosowań w przemyśle) świadczący o dostatecznej wiedzy autora? Czy wnioski z przeglądu źródeł sformułowano w sposób jasny i przekonujący?**

Rozprawa zawiera rozdziały 3, który dotyczy analizy źródeł i aktualnego stanu badań w jej dziedzinie. Rozdział ten omawia stan wiedzy w zakresie sekwencyjnych, grafowych oraz wielomodalnych algorytmów rekomendacyjnych, a także modeli danych służących do generalizacji problemu rekomendacji. Rozdział 3 kończy sekcja 3.5 „Positioning This Thesis”, która podsumowuje luki w stanie wiedzy oraz planowaną kontrybucję rozprawy – rozwój modeli danych i algorytmów rekomendacyjnych opartych na estymacji rozkładu zachowań użytkowników.

- 3. Czy autor rozwiązał postawione zagadnienia, czy użył właściwej do tego metody i czy przyjęte założenia są uzasadnione?**

Doktorant używa metodologii empirycznej, będącą podstawą uczenia maszynowego. Wszystkie algorytmy są oceniane poprzez testowanie ich na wielu publicznie dostępnych zbiorach danych, oraz poprzez porównanie z licznymi konkurencyjnymi algorytmami, które są aktualnym stanem wiedzy i techniki w rozpatrywanym zagadnieniu badawczym. Algorytmy i metody doktoranta były także testowane poprzez wzięcie udziału przez doktoranta (wraz z zespołem firmy Synerise) w wielu konkursach z dziedziny systemów rekomendacyjnych oraz „data science”.

Wątpliwości budzi możliwość replikacji wyników doktoranta, jako że algorytmy opisane w rozprawie nie są dostępne publicznie (co może być uzasadnione strategią ich komercjalizacji).

- 4. Na czym polega oryginalność rozprawy, co stanowi samodzielny i oryginalny dorobek autora, jaka jest pozycja rozprawy w stosunku do stanu wiedzy czy poziomu techniki reprezentowanych przez literaturę światową?**

Doktorant wymienił w sekcji 1.9 („Statement of Contributions”) pięć oryginalnych, własnych osiągnięć. Najważniejszymi z nich są „Basemodel”, model bazowy reprezentujący rozkłady zachowań użytkowników systemów rekomendacyjnych, oraz algorytm EMDE, którego doktorant jest współautorem. Oba te osiągnięcia są obecnie wykorzystywane komercyjnie przez firmę Synerise.



POLISH-JAPANESE ACADEMY OF INFORMATION TECHNOLOGY

ポ
ー
ラ
ン
ド
日
本
情
報
工
科
大
学

Osiągnięcia opisane przez doktoranta są odzwierciedlone w publikacjach, na których opiera się rozprawa. Rozprawa nie jest pakietem publikacji, ale jej trzy główne rozdziały (5-7) są powiązane z publikacjami opisującymi omawiane metody lub algorytmy. W związku z tym, oceniając osiągnięcia doktoranta, recenzent powinien wziąć pod uwagę kolejność autorów wymienionych publikacji, lub analogiczną informację wskazującą na wkład doktoranta w opisane w rozprawie metody.

Mgr inż. Michał Daniluk jest pierwszym i głównym autorem publikacji:

1. Michał Daniluk, Tim Rocktäschel, Johannes Welbl, Sebastian Riedel: Frustratingly Short Attention Spans in Neural Language Modeling (2017). International Conference on Learning Representations (ICLR). In: Learning Representations. International Conference. 5th 2017. (ICLR 2017), Curran Associates Inc., ISBN: 9781713872719, volume 1, pp. 679-689. Poster presentation at conference track, ICLR 2017.

Według autora (rozprawa, str. 22), publikacja ta jest oceniana na 200 punktów według punktacji MNiSW. Nie jest to niestety informacja precyzyjna. Po pierwsze, praca została opublikowana przed wprowadzeniem aktualnego systemu punktacji MNiSW. Po drugie, nawet według dzisiejszych zasad, 200 punktów jest przyznawane za konferencję o rankingu A* w rankingu CORE. Konferencja ICLR w 2017 roku nie figurowała w rankingu CORE, co według dzisiejszych zasad oznacza, że publikacje na tej konferencji oceniane są na 0 punktów. Dochodzi do tego jeszcze kontrowersja dotycząca publikacji materiałów konferencyjnych (a raczej ich braku), ponieważ konferencja ICLR nie ma oficjalnego wydawcy, a materiały konferencyjne publikowane są przez system OpenReview (zostały też wydane w druku przez wydawnictwo Curran Associates). Nie ma w związku z tym numeru DOI tej publikacji, a Google Scholar nadal indeksuje ją jako publikację z serwisu arXiv. Innym zagadnieniem jest, czy artykuł – prezentowany jako poster, ale opublikowany w materiałach konferencyjnych – zasługuje na punkty. Na konferencję ICLR'2017 zgłoszono 507 prac, z czego na główną konferencję („conference track”) przyjęto 15 zgłoszeń (mniej niż 3%) jako artykuły z prezentacją ustną, a 183 zgłoszenia (ok. 36%) z prezentacją jako poster¹. Przyjęcie łącznie 39% na główną konferencję („conference track”) jest porównywalne z głównymi konferencjami w tej

¹ Acceptance rates for the top-tier AI related conferences. <https://github.com/lixin4ever/conference-acceptance-rate> Dostęp 19.08.2026

dziedzinie (ocenianymi jako CORE A lub CORE A* - na przykład, konferencja ACL przyjęta w 2023 roku ok. 40% zgłoszeń).

Powyższe uwagi nie zmieniają faktu, że praca doktoranta ma duże znaczenie w swojej dziedzinie, o czym świadczy wysoka liczba cytowań pracy: 165 cytowań w serwisie Google Scholar od 2017 roku.

Zdaniem recenzenta, to właśnie świadczy o wartości pracy, a nie omyłkowo przypisana liczba punktów MNiSW.

Doktorant jest również pierwszym autorem publikacji:

2. Daniluk, M., & Dąbrowski, J. (2023). Temporal graph models fail to capture global temporal dynamics. *Temporal Graph Learning (TGL) Workshop at NeurIPS 2023*
3. Daniluk, M., Dabrowski, J., Rychalska, B., & Gołuchowski, K. (2021). Synersize at KDD Cup 2021: Node Classification in Massive Heterogeneous Graphs. *KDD Cup OGB Challenge*.
4. Daniluk, M., Rychalska, B., Gołuchowski, K., & Dąbrowski, J. (2021). Modeling Multi-Destination Trips with Sketch-Based Model. *ACM WSDM WebTour 2021 Workshop*, March 12th, 2021 Jerusalem, Israel
5. Daniluk, M., Dabrowski, J., Rychalska, B., & Goluchowski, K. (2021). Synerise at RecSys 2021: Twitter user engagement prediction with a fast neural model. In *Proceedings of the Recommender Systems Challenge 2021* (pp. 15-21).
6. Daniluk, M., Rezaei, M., Nicolescu, R., & Klette, R. (2014, September). Eye status based on eyelid detection: A driver assistance system. In *International Conference on Computer Vision and Graphics* (pp. 171-178). Cham: Springer International Publishing.

Powyższe publikacje są związane z konkursami organizowanymi przez główne konferencje w zakresie Data Science oraz uczenia maszynowego.

Pozycja wyników uzyskanych przez doktoranta w stosunku do aktualnego stanu wiedzy i techniki jest potwierdzona licznymi sukcesami w konkursach dotyczących algorytmów rekomendacyjnych i Data Science:

- Booking.com 2021 WSDM Data Challenge
- KDD Cup 2021 MAG240M-LSC Track
- ACM RecSys Challenge 2021

We wszystkich wymienionych konkursach wykorzystywany była algorytm EMDE lub bazowy model Basemodel, które są głównym osiągnięciem doktoranta.

7. Czy autor wykazał umiejętność poprawnego i przekonującego przedstawienie uzyskanych przez siebie wyników (zwięzłość, jasność, poprawność redakcyjna rozprawy)?

Rozprawa jest napisana bardzo dobrym językiem angielskim. Ma także przejrzystą, czytelną strukturę. Wprowadzenie jest napisane bardzo dobrze, przedstawia jasno problem badawczy, jego znaczenie, oraz proponowane metody rozwiązania (w formie hipotez). Po nim następuje rozdział z przeglądem literatury. Rozdziały 5, 6 i 7 zawierają zasadnicze wyniki rozprawy. Ostatni (ósmy) rozdział podsumowuje rozprawę.

8. Jakie są słabe strony rozprawy i jej główne wady?

Rozprawa dotyczy bardzo ważnego, praktycznego problemu, jakim jest stworzenie modelu bazowego dla rekomendacji. Jest to jednocześnie cel ogólny, niezależny od proponowanego podejścia bazującego na rozkładach. Ze względu na wagę i ambicję tego zadania, należy się spodziewać, że inni badacze także podjęli ten problem. Doktorant w sekcji 3.4 („Foundation Models for Recommendation”) uzasadnia, dlaczego modele bazowe dla rekomendacji nie istnieją. Używa stwierdzenia: „Despite significant effort, no equivalent of GPT has emerged for recommendation.” (str. 49). Nie jest to jednak prawda.

Istnieje szereg prac, nie cytowanych przez doktoranta w rozprawie, które podjęły ten problem:

Wu et al. (2020) — PTUM. Podejście do pretrenowania neuronowego modelu użytkownika zbudowane na podstawie nieetykietowanych historii zachowań. Wykorzystuje ono predykcję zachowań maskowanych i predykcję zachowań przyszłych, a następnie dostraja wstępnie wytrenowany model do dalszych zadań rekomendacji.

Wu et al. (2022) — UserBERT. Samonadzorujące pretrenowanie behawioralne o dwóch celach: uczyć się kontekstu zachowania oraz dopasować sekwencje w czasie, aby uchwycić trwałe zainteresowania użytkownika.

Hou et al. (2022) — UniSRec. Międzydomenowy sekwencyjny system rekomendacji zaprojektowany w celu przezwyciężenia nieprzenoszalności identyfikatorów pozycji specyficznych dla katalogu. Wykorzystuje tekst pozycji oraz wstępny trening do budowania reprezentacji sekwencji, które można przenieść do nowych domen lub platform.

Wang et al. (2022) — TransRec. Przenoszalny system rekomendacji trenowany na podstawie multimodalnego sprzężenia zwrotnego – głównie tekstu i obrazów elementów – zamiast wymagać

od systemów źródłowych i docelowych współużytkowania identyfikatorów użytkowników lub produktów.

Sun et al. (2023) — UniM²Rec. Wielodomenowy, multimodalny system wstępnego trenowania, który wspólnie uczy się reprezentacji zawartości elementów i preferencji użytkownika na podstawie interakcji między domenami; koncentruje się na transferze, gdy zawartość elementów jest niekompletna lub zaszumiona.

Gong et al. (2025) — BehaveGPT. Jawna propozycja modelu bazowego zachowań użytkownika: Transformer wstępnie trenowany na dużą skalę pod kątem kolejnego zachowania, długoterminowego generowania zachowań i adaptacji międzydomenowej. W artykule przedstawiono korzyści ze stosowania modelu bazowego w porównaniu do innych rozwiązań.

Gong, J., Ding, J., Meng, F., Yang, C., Chen, H., Wang, Z., Lu, H.-R., & Li, Y. (2025). BehaveGPT: A foundation model for large-scale user behavior modeling.

<https://www.semanticscholar.org/paper/8bb77c2b5a9a731f1af352d5e6ea3baf0d6108d9>

Hou, Y., Mu, S., Zhao, W. X., Li, Y., Ding, B., & Wen, J.-R. (2022). Towards universal sequence representation learning for recommender systems. arXiv. <https://arxiv.org/pdf/2206.05941>

Sun, W., Xie, R., Bian, S., Zhao, W., & Zhou, J. (2023). Universal multi-modal multi-domain pre-trained recommendation. arXiv. <https://arxiv.org/pdf/2311.01831>

Wang, J., Yuan, F., Cheng, M., Jose, J., Yu, C., Kong, B., Wang, Z., Hu, B., & Li, Z. (2022). TransRec: Learning transferable recommendation from mixture-of-modality feedback. arXiv.

<https://arxiv.org/pdf/2206.06190>

Wu, C., Wu, F., Qi, T., Lian, J., Huang, Y., & Xie, X. (2020). PTUM: Pre-training user model from unlabeled user behaviors via self-supervision. In Findings of the Association for Computational Linguistics: EMNLP 2020. <https://www.aclweb.org/anthology/2020.findings-emnlp.174.pdf>

Wu, C., Wu, F., Qi, T., & Huang, Y. (2022). UserBERT: Pre-training user model with contrastive self-supervision.

<https://www.semanticscholar.org/paper/ecc25444efd0a40ee620dd6a60787f00c890a408>

Szczególnie praca (Gong, 2025) powinna być bezwzględnie znaleźć się w analizie literatury. Doktorant powinien być podjąć próbę porównania Basemodel w stosunku do tego rozwiązania.

9. Jaka jest przydatność rozprawy dla Informatyki?

Informatyka techniczna jest nauką stosowaną. Rola badań stosowanych jest dlatego kluczowa dla tej dyscypliny, ponieważ pozwalają one nie tylko na empiryczną walidację metod teoretycznych, ale także wskazują na istotne, nowe problemy badawcze, inspirujące kolejne badania teoretyczne. Recenzowana rozprawa jest doktoratem wdrożeniowym, ale nie ogranicza się do rozwiązywania wąskich, praktycznych problemów firmy. Zamiast tego, opisane wyniki badań rozpatrują bardzo podstawowy problem, jakim jest utworzenie modelu bazowego zachowań użytkowników, który mógłby zostać wykorzystany do uczenia z przeniesieniem (ang. *transfer learning*) w celu rozwiązywania innych problemów rekomendacji.

Co więcej, opisane przez doktoranta podejście, wykorzystujące reprezentacje rozkładów wyborów użytkowników, wnosi nową perspektywę do budowy systemów rekomendacyjnych.

WNIOSKI

Rozprawa ma wady, z których najważniejszą jest niepełny przegląd literatury. Jednak wady te nie umniejszają osiągnięć opisanych w rozprawie.

Podsumowując stwierdzam, że **rozprawa spełnia wymagania stawiające rozprawom doktorskim przez obowiązujące przepisy.**



prof. dr hab. Adam Wierzbicki