

Wrocław, 8.07.2026

Dr hab. inż. Adrianna Kozierkiewicz  
Katedra Informatyki Stosowanej  
Wydział Informatyki i Telekomunikacji  
Politechnika Wrocławska

## **Recenzja rozprawy doktorskiej mgr inż. Michała Daniluka**

pt.

### **„Multimodal Recommendation System for E-commerce”**

Recenzja niniejsza została przygotowana na zlecenie Rady Naukowej Dyscypliny Informatyka Techniczna i Telekomunikacja Politechniki Warszawskiej na podstawie rozprawy doktorskiej przedłożonej w dyscyplinie informatyka techniczna i telekomunikacja (pismo z dnia 11.05.2026).

#### **I. Ocena tematyki rozprawy**

Przedmiotem recenzowanej rozprawy jest problem reprezentacji danych behawioralnych w systemach rekomendacyjnych dla e-commerce. Autor wychodzi od trafnej obserwacji, że dominujące w ostatnich latach podejścia sekwencyjne, inspirowane przetwarzaniem języka naturalnego, nie zawsze odpowiadają naturze danych zakupowych i interakcyjnych. W wielu scenariuszach komercyjnych kolejność kliknięć, zakupów lub ocen wynika z układu interfejsu, promocji, dostępności produktu albo wygody użytkownika, a nie z głębokiej struktury preferencji. Równie istotny jest wskazany w pracy problem reprezentacji punktowej, w której uśrednianie wektorów produktów prowadzi do utraty wielomodalnej struktury preferencji użytkownika.

Tematyka rozprawy jest aktualna, ważna naukowo i praktycznie dobrze umotywowana. Systemy rekomendacyjne stanowią dziś jeden z podstawowych komponentów platform e-commerce, mediów społecznościowych oraz systemów personalizacji treści. Jednocześnie coraz większe znaczenie mają skalowalność, możliwość wykorzystania heterogenicznych źródeł danych, odporność na problem cold-start oraz zdolność do wdrożenia rozwiązań w środowiskach przemysłowych z ograniczeniami czasu odpowiedzi. Recenzowana rozprawa lokuje się dokładnie w tym obszarze: łączy problem reprezentacji użytkownika i produktu, modelowania wielomodalnego, samonadzorowanego uczenia reprezentacji oraz walidacji w zastosowaniach produkcyjnych.

Głównym celem rozprawy jest modelowanie danych behawioralnych w systemach rekomendacyjnych za pomocą estymacji gęstości na różnorodnościach reprezentacji wektorowych produktów. Cel ten został zdekomponowany na cztery cele badawcze: opracowanie reprezentacji zachowującej własności dystrybucyjne danych behawioralnych, umożliwienie integracji heterogenicznych modalności w jednej reprezentacji, opracowanie

samonadzorowanego modelu fundacyjnego oraz walidację opracowanych rozwiązań zarówno na danych benchmarkowych jak i we wdrożeniach produkcyjnych.

Z celami pracy powiązano pięć hipotez badawczych dotyczących: przewagi reprezentacji opartych na estymacji gęstości nad agregacją punktową, możliwości skutecznej integracji wielu modalności przez konkatenację szkiców dla reprezentacji opartych na estymacji gęstości, ograniczenia potrzeby stosowania bardzo złożonych architektur przy zastosowaniu reprezentacji opartych na gęstości, transferu między typami zadań dzięki samonadzorowanemu pretrenowaniu oraz szczególnych korzyści w scenariuszach cold-start dla systemów, gdzie zastosowano reprezentację opartą na estymacji gęstości. Hipotezy zostały jasno powiązane z planem eksperymentalnym i są możliwe do weryfikacji na podstawie wyników przedstawionych w dalszych rozdziałach.

W mojej ocenie problem badawczy został trafnie wybrany i jasno sformułowany. Rozprawa dotyczy zagadnienia, które ma zarówno znaczenie poznawcze, jak i bardzo wyraźny wymiar wdrożeniowy. Przyjęta perspektywa jest oryginalna, a jednocześnie dobrze osadzona w aktualnych trendach badań nad systemami rekomendacyjnymi, multimodalnością i modelami dla danych behawioralnych. Cele pracy odpowiadają wymaganiom stawianym rozprawom doktorskim.

## **II. Struktura rozprawy**

Rozprawa liczy 160 stron i składa się z ośmiu rozdziałów, streszczenia w języku angielskim i polskim oraz bibliografii obejmującej 90 pozycji. Układ pracy jest logiczny i prowadzi czytelnika od motywacji oraz przeglądu stanu wiedzy, przez opis proponowanych metod, aż do szerokiej walidacji eksperymentalnej i wdrożeniowej. Treść rozprawy można scharakteryzować następująco:

- Rozdział pierwszy stanowi wprowadzenie do problematyki reprezentacji danych behawioralnych w systemach rekomendacyjnych. Autor przedstawia problem reprezentacyjny, uzasadnia potrzebę odejścia od bezpośredniej analogii do modelowania języka naturalnego, formułuje cele, hipotezy oraz główne wkłady pracy. W rozdziale zawarto również opis publikacji, opis wkładu własnego oraz kontekst doktoratu wdrożeniowego.
- Rozdział drugi przedstawia podstawy teoretyczne i techniczne potrzebne do zrozumienia dalszych części pracy. Omówiono systemy rekomendacyjne, podstawowe wyzwania takie jak cold-start, rzadkość danych i skalowalność, paradygmat osadzeń, dane multimodalne, różnice między sekwencyjnym i dystrybucyjnym spojrzeniem na zachowanie użytkowników oraz metodologię ewaluacji.
- Rozdział trzeci zawiera przegląd literatury obejmujący modele sekwencyjne, grafowe sieci neuronowe dla zadania rekomendacji, rekomendację multimodalną i modele fundacyjne dla rekomendacji. Szczególnie wartościowe jest konsekwentne wskazanie wspólnego ograniczenia wielu istniejących podejść, czyli redukcji złożonej struktury preferencji do reprezentacji punktowych lub sekwencji, które nie zawsze niosą istotną informację.

- Rozdział czwarty omawia osadzenia grafowe dla danych behawioralnych, w szczególności metodę Cleora. Autor uczciwie wskazuje, że Cleora została opracowana przez innych członków zespołu Synerise, natomiast w rozprawie stanowi istotny element przygotowania reprezentacji wejściowych dla EMDE i BaseModel.
- Rozdział piąty prezentuje autorską metodę EMDE- Efficient Manifold Density Estimator czyli metodę opartą na estymacji gęstości na rozmaitościach reprezentacji wektorowych produktów. Omówiono podstawy teoretyczne, opis samej metody, jej własności. Wyniki eksperymentów demonstrują najnowocześniejszą wydajność w testach porównawczych opartych na sesjach i rekomendacjach typu k-top, a analiza ujawnia, dlaczego sieci grafów czasowych zawodzą tam, gdzie estymacja gęstości się udaje. Rozdział zawiera także podsumowanie weryfikacji hipotez H1-H3.
- Rozdział szósty przedstawia BaseModel, czyli najważniejszy indywidualny wkład Autora. Opisano w nim reprezentacje profili behawioralnych, architekturę modelu, adaptację do zadań rekomendacji, klasyfikacji i regresji, wyniki eksperymentalne na benchmarkach sekwencyjnych i RelBench, badania ablacyjne, analizę złożoności obliczeniowej oraz wdrożenia produkcyjne.
- Rozdział siódmy przedstawia wyniki walidacji EMDE w trzech konkursach uczenia maszynowego: Booking.com WSDM WebTour Challenge, KDD Cup 2021 oraz ACM RecSys Challenge 2021. Ten rozdział dobrze uzupełnia klasyczną ewaluację akademicką, ponieważ pokazuje działanie metod w porównaniu z wynikami bardzo silnych zespołów badawczych.
- Rozdział ósmy zawiera podsumowanie wkładów, ocenę hipotez, refleksję nad wynikami, ograniczenia oraz kierunki dalszych badań.

Struktura rozprawy jest spójna. Praca jest napisana jasno, a kolejne części dobrze wynikają z poprzednich. Na podkreślenie zasługuje przejrzyste rozdzielenie wkładu własnego Autora od elementów opracowanych zespołowo lub przez innych członków zespołu.

### **III. Oryginalny wkład naukowy rozprawy**

Jako najważniejsze wyniki rozprawy pragnę podkreślić następujące osiągnięcia Doktoranta.

#### **1. Reprezentacja historii behawioralnych**

Pierwszym istotnym wkładem jest rozwinięcie i walidacja podejścia opartego na estymacji gęstości na rozmaitościach reprezentacji wektorowych produktów. Metoda EMDE wykorzystuje haszowanie wrażliwe na lokalność oraz partycjonowanie zależne od gęstości danych, tworząc szkice o stałym rozmiarze. Takie szkice pozwalają zachować wielomodalną strukturę preferencji, a jednocześnie są praktyczne obliczeniowo.

Wartość tego wkładu polega nie tylko na opracowaniu nowej reprezentacji danych behawioralnych, lecz także na wykazaniu, że odpowiedni sposób reprezentowania danych może mieć większy wpływ na skuteczność systemu rekomendacyjnego niż zwiększanie złożoności architektury modelu. Wyniki przedstawione w rozdziale piątym wspierają tę tezę:

EMDE osiąga najwyższy wynik  $MRR@20$  na pięciu z sześciu zbiorów rekomendacji opartych na sesjach, a badania ablacyjne pokazują istotny spadek jakości po usunięciu struktury geometrycznej osadzeń. Co potwierdza, że proste sieci jednokierunkowe operujące na szkicach gęstościowych EMDE mogą konkurować z architekturami rekurencyjnymi, atencyjnymi i grafowymi (GRU4Rec, NARM, SR-GNN, and TAGNN).

## **2. BaseModel jako samonadzorowany model dynamiki behawioralnej**

Najważniejszym indywidualnym osiągnięciem Autora jest BaseModel, przedstawiony w rozdziale szóstym. Model uczy się przewidywać przyszłe rozkłady behawioralne na podstawie przeszłych rozkładów, dzięki czemu może zostać wykorzystany w różnych typach zadań, takich jak rekomendacja, klasyfikacja i regresja. Istotne jest to, że Autor nie ogranicza się do klasycznego zadania „next-item prediction”, lecz proponuje predykcję rozkładu jako bardziej naturalny opis ewolucji preferencji użytkowników.

Wyniki eksperymentalne zostały przeprowadzone w porównaniu z 4 różnymi metodami znanymi z literatury. W porównaniu z TIGER uzyskano poprawę rzędu 54-82% na rzadkich zbiorach Amazon, w porównaniu z HSTU na Amazon Books przedstawiono ponad dwukrotną poprawę  $HR@10$  i  $NDCG@10$  oraz przewyższono najlepszy wariant beeFormer o 6-22% w zakresie wszystkich wskaźników. Poprawa jest szczególnie wyraźna w niższych rangach: 6,2% w  $Recall@20$ , ale 18,6% w  $Recall@50$  i 21,9% w  $NDCG@100$ . Na benchmarku RelBench model osiąga lub przekracza wyniki wyspecjalizowanych metod grafowych w 10 z 12 zadań.

## **3. Prosta i skuteczna integracja multimodalna**

Kolejnym wkładem jest pokazanie, że różne modalności danych, takie jak interakcje, tekst, obrazy i metadane kategoryczne, mogą zostać połączone przez konkatencję szkiców gęstościowych. W przeciwieństwie do wielu podejść multimodalnych nie wymaga to osobnych mechanizmów fuzji. Jest to rozwiązanie praktyczne i zgodne z przemysłowymi wymaganiami prostoty oraz skalowalności.

Autor pokazuje systematyczne korzyści z dodawania modalności, między innymi poprawę o 7,7% Hit Rate na zbiorze RETAIL oraz wzrost  $Recall@1$  o 27% na zbiorze MovieLens-20M. Szczególnie istotne są wyniki w scenariuszach cold-start, gdzie metadane pozwalają rekomendować nowe produkty, dla których nie istnieje historia interakcji. Z punktu widzenia e-commerce jest to wkład o bardzo dużym znaczeniu praktycznym.

## **4. Szeroka walidacja eksperymentalna, konkursowa i wdrożeniowa**

Jedną z największych zalet rozprawy jest zakres walidacji. Autor nie ogranicza się do jednego zbioru danych ani jednego typu zadania. Metody są oceniane na danych benchmarkowych, w konkursach uczenia maszynowego oraz we wdrożeniach produkcyjnych. Taki zakres walidacji znacząco zwiększa wiarygodność pracy, ponieważ pokazuje różne aspekty jakości

metody: skuteczność predykcyjną, skalowalność, odporność na ograniczenia czasu odpowiedzi oraz przydatność biznesową.

Na szczególne wyróżnienie zasługują wyniki konkursowe: drugie miejsce w Booking.com WSDM Data Challenge wraz z Best Paper Award, trzecie miejsce w KDD Cup 2021 wśród około 150 zespołów oraz drugie miejsce w ACM RecSys Challenge 2021. Zespoły wyprzedzające rozwiązania Autora należały między innymi do Baidu, Google DeepMind i NVIDIA - globalnych podmiotów o znacznie większej skali działania, zapleczu badawczym oraz finansowaniu, co umożliwiało im szybszy rozwój i wdrażanie konkurencyjnych rozwiązań. Równie istotny jest wymiar wdrożeniowy. Rozprawa powstała w ramach doktoratu wdrożeniowego, a przedstawione metody zostały wykorzystane w środowiskach produkcyjnych przetwarzających, według danych zawartych w pracy, 42 miliardy zdarzeń miesięcznie dla ponad 200 organizacji na 150 rynkach. Autor podaje przykłady poprawy predykcji retencji klientów w dużym europejskim banku oraz stały wzrost zaangażowania we wdrożeniach detalicznych i telekomunikacyjnych.

Podsumowując, rozprawa zawiera oryginalny i wartościowy wkład naukowy. Jej najmocniejszą stroną jest połączenie oryginalnych metod z bardzo szeroką walidacją praktyczną. Autor wykazuje dojrzałość badawczą, dobrą znajomość literatury, umiejętność projektowania eksperymentów oraz zdolność doprowadzenia wyników naukowych do działania w systemach produkcyjnych. W świetle wyżej opisanych osiągnięć rozprawa doktorska mgr inż. Michała Daniluka stanowi oryginalne rozwiązanie określonych problemów badawczych, a uzyskane wyniki pozwalają stwierdzić, że cele pracy zostały osiągnięte.

#### **IV. Uwagi do rozprawy**

Poniższe uwagi mają w większości charakter dyskusyjny i nie obniżają pozytywnej oceny rozprawy. Wskazują raczej miejsca, w których praca mogłaby zostać doprecyzowana lub rozwinięta w dalszych badaniach.

##### **Lista uwag**

1. W przeglądzie literatury brakuje opisu metodologii przeprowadzonej analizy. Autor przedstawia wybrane rozwiązania, jednak nie wskazuje kryteriów ani przesłanek, na podstawie których zostały one wyselekcjonowane. W rozdziale brakuje również podsumowania przeprowadzonego przeglądu, w tym identyfikacji istniejących luk badawczych oraz ograniczeń obecnych metod, które uzasadniałyby podjęcie przedstawionego problemu badawczego.
2. Z przeglądem literatury związany jest również dobór wykorzystanych źródeł bibliograficznych. Autor powołuje się na prace opublikowane przed 2020 rokiem, dostępne wyłącznie w serwisie arXiv (np. [42] i [45]). Takie podejście budzi wątpliwości, ponieważ brak publikacji tych wyników w recenzowanych czasopiśmie naukowych lub materiałach uznanych konferencji, mimo upływu kilku

lat od ich udostępnienia, może wskazywać na ograniczone znaczenie naukowe tych prac lub niewystarczającą walidację przedstawionych w nich wyników.

3. Najpoważniejszym mankamentem pracy jest niewystarczająca dbałość o poprawność i konsekwencję formalną. W wielu fragmentach pracy występują niespójności oraz nieścisłości dotyczące stosowanej terminologii, notacji i przedstawianych definicji, co znacząco utrudnia śledzenie toku rozumowania oraz właściwą interpretację zaprezentowanych wyników. Poniżej przykłady:
  - Na stronie 56 przez  $d$  oznaczono wymiar wektora, z kolei we wzorze 4.10 oraz 5.1 mają one kompletnie inne znaczenie
  - Na stronie 54 przez  $s$  oznaczono koszyk zakupowy, z kolei na stronie 57  $s$  oznacza funkcje podobieństwa
  - Na stronie 40 przez  $k$  oznaczono ilość rekomendacji, z kolei na stronie 60 jest to wielkość hiperkrawędzi, a na stronie 63 ilość iteracji
  - Wzór 5.7 wymagałby wyjaśnienia użytej notacji
  - Na stronie 78 funkcja  $f_\theta$  traktuje  $S_{past}$  jako parametr, z kolei we wzorze 6.3 i 6.5 zdefiniowaną dwie różne funkcje oznaczając je w ten sam sposób.
  - Na stronie 28 przez  $m$  oznaczono liczbę użytkowników, a w dalszej części pracy (np. strona 98)  $m$  oznacza modalność.
4. W pracy nie przedstawiono analizy statystycznej dla części wyników eksperymentalnych. Jej przeprowadzenie wzmocniłoby wiarygodność uzyskanych rezultatów oraz końcowych wniosków, zwłaszcza w kontekście weryfikacji postawionych hipotez badawczych.
5. Badania dotyczące wyników multimodalnych przedstawione w rozdziałach 5.6.5 i 5.6.6 warto uzupełnić o bardziej szczegółowy opis wykorzystanych modalności danych. Autor posługuje się ogólnym określeniem „other available modalities”, nie precyzując, jakie modalności były dostępne dla poszczególnych zbiorów danych. Nie jest również jasne, czy wszystkie analizowane zbiory zawierały ten sam zestaw modalności oraz w jakim stopniu liczba i rodzaj dostępnych modalności mogły wpływać na uzyskane wyniki.
6. Autor opisuje sposób konstruowania reprezentacji multimodalnej, jednak nie przedstawia jednoznacznie, na czym polega architektura modelu EMDE MM względem podstawowego EMDE. Z opisu można wnioskować, że różnica sprowadza się do zastosowania multimodalnego szkicu  $S_{multi}$  jako wejścia  $S_{input}$ , przy niezmienionej architekturze sieci, jednak warto byłoby tę zależność wyraźnie sformułować.
7. Autor przyjmuje założenie, że zmiany preferencji użytkowników mają zazwyczaj charakter stopniowy („behavioral drift”), dzięki czemu przewidywanie przyszłych rozkładów preferencji jest łatwiejsze niż przewidywanie kolejnych pojedynczych

interakcji (strona 96). Założenie to wydaje się intuicyjne i stanowi istotny element uzasadnienia proponowanej metody, jednak warto byłoby odnieść je do konkretnych wyników badań z zakresu modelowania zachowań, użytkowników, psychologii poznawczej lub zachowań konsumenckich.

8. Opis wdrożeń produkcyjnych stanowi jedną z najmocniejszych stron rozprawy, jednak z metodologicznego punktu widzenia warto byłoby uzupełnić go o więcej informacji dotyczących sposobu prowadzenia eksperymentów produkcyjnych, w szczególności protokołów testów A/B, czasu trwania badań, liczebności grup, charakterystyki modeli bazowych oraz sposobu kontroli czynników zakłócających (np. sezonowości). Ograniczenia wynikające z poufności danych przemysłowych są zrozumiałe, niemniej przedstawienie takich informacji (choćby w ujęciu zagregowanym) dodatkowo zwiększyłoby wiarygodność zaprezentowanych rezultatów.
9. W opisie konfiguracji eksperymentów warto szerzej przedstawić procedurę doboru hiperparametrów zarówno dla proponowanych metod, jak i modeli bazowych. W szczególności przydatne byłoby wskazanie sposobu strojenia oraz kryterium wyboru najlepszej konfiguracji. Ułatwiłoby to ocenę porównywalności eksperymentów oraz zwiększyłoby możliwość odtworzenia przedstawionych wyników.
10. W sekcjach 5.6.2 oraz 6.6.3 przedstawiono zestawy modeli bazowych, jednak zabrakło wyraźnego uzasadnienia kryteriów ich doboru. Autor skoncentrował się na porównaniach z nowoczesnymi metodami, nie wyjaśniając, dlaczego wybrano właśnie te modele oraz czy rozważano uwzględnienie klasycznych algorytmów rekomendacyjnych jako dodatkowego punktu odniesienia. Krótkie przedstawienie przesłanek doboru modeli bazowych zwiększyłoby przejrzystość metodologii eksperymentów oraz ułatwiłoby ocenę kompletności przeprowadzonych porównań.
11. Na stronach 77 i 100 własności proponowanych metod przedstawiono głównie w sposób opisowy i intuicyjny. Warto byłoby rozważyć ich bardziej formalne sformułowanie oraz (w przypadku własności wynikających z przyjętych założeń, a nie bezpośrednio z konstrukcji algorytmu) uzupełnienie o odpowiednie uzasadnienie/dowód.
12. W tabeli 5.2 przedstawiono wyniki porównania różnych metod na kilku zbiorach danych. Interesującym przypadkiem jest zbiór DIGI, dla którego metoda VS-KNN uzyskała wyższą wartość wskaźnika  $MRR@20$  niż EMDE. Warto byłoby uzupełnić dyskusję wyników o próbę wyjaśnienia przyczyn tej obserwacji oraz wskazanie cech zbioru danych, które mogły wpłynąć na relatywnie słabszą skuteczność proponowanej metody.

Wskazane uwagi nie podważają głównych wyników rozprawy. Mają one przede wszystkim charakter doprecyzowujący i dotyczą formalizacji oraz prezentacji wyników. Autor w wielu miejscach sam rozpoznaje ograniczenia swojej metody, co oceniam pozytywnie.

## V. Konkluzje

Problem naukowy został trafnie wybrany i jasno sformułowany, a tematyka rozprawy jest aktualna, ważna i dobrze uzasadniona. Rozprawa pokazuje bardzo dobrą znajomość stanu wiedzy w obszarze systemów rekomendacyjnych oraz reprezentacji danych behawioralnych. Autor rozwiązał postawiony problem badawczy w sposób oryginalny, a przedstawione cele pracy zostały osiągnięte.

Na szczególnie wysoką ocenę zasługuje połączenie wkładu koncepcyjnego z walidacją na wielu poziomach: badania na benchmarkowych zbiorach danych, udział w konkursach międzynarodowych oraz we weryfikacji w rzeczywistym wdrożeniu produkcyjnym. Dodatkowo, Doktorant opublikował osiągnięte wyniki w kilku publikacjach, m.in. w materiałach konferencyjnych: *International Conference on Learning Representations (ICLR)*, (kat. A\* według listy CORE) oraz *International Conference on Neural Information Processing (ICONIP)* (kat. A według listy CORE 2020) oraz w czasopiśmie *Electronics* (z listy ministerialnej- 100pkt). Zastrzeżenia może jednak budzić ostatnia z wymienionych publikacji, opublikowana w czasopiśmie wydawnictwa MDPI, którego model wydawniczy i praktyki redakcyjne są przedmiotem dyskusji w środowisku naukowym.

Doktorant wykazał się umiejętnością samodzielnego prowadzenia badań naukowych, projektowania metod, krytycznej analizy istniejących rozwiązań, prowadzenia eksperymentów oraz interpretacji wyników. Uwagi krytyczne wskazane w recenzji dotyczą głównie zakresu formalizacji lub doprecyzowania niektórych założeń, ale nie zmieniają jednoznacznie pozytywnej oceny rozprawy.

Podsumowując, stwierdzam, że rozprawa doktorska pana mgr inż. Michała Daniluka pt. „Multimodal recommendation system for e-commerce” stanowi oryginalne rozwiązanie istotnego problemu naukowego oraz spełnia wymagania określone w Ustawie z dnia 20 lipca 2018 r. - Prawo o szkolnictwie wyższym i nauce. Wobec powyższego wnioskuję o dopuszczenie rozprawy do publicznej obrony.

Adrianna Kozienkiewicz