dr hab. inż. Jakub Nalepa, prof. PŚ
Katedra Algorytmiki i Oprogramowania
Politechnika Śląska
Jakub.Nalepa@polsl.pl

Gliwice, 27 kwietnia 2026 r.

Recenzja rozprawy doktorskiej

Tytuł rozprawy: Zastosowanie hybrydowych metod ekstrakcji cech w problemie detekcji dezinformacji

Autor: Mgr inż. Gracjan Kątek

Promotor: Dr hab. inż. Rafał Kozik, prof. PBŚ

Promotor pomocniczy: Dr inż. Adam Marchewka, prof. PBŚ

1 Tematyka rozprawy

Recenzowana rozprawa doktorska Pana mgr. inż. Gracjana Kątka dotyczy zagadnień związanych z detekcją dezinformacji, ze szczególnym akcentem położonym na opracowanie efektywnych ekstraktorów cech z danych tekstowych. Ekstrakcja cech jest kluczowym elementem systemów uczenia maszynowego, a jakość wyekstrahowanych cech bezpośrednio wpływa na jakość działania i możliwości generalizacyjne klasyfikatorów (w przypadku niniejszej pracy, trenowanych w sposób nadzorowany). W pracy skupiono się na opracowaniu ektraktorów hybrydowych, w których wektory cech – ekstrahowane w różny sposób – mogą być ze sobą łączone, aby wypracować „holistyczny" opis surowych danych tekstowych, jednocześnie odzwierciedlając ich lokalną i globalną charakterystykę.

Uważam, że tematyka rozprawy jest ciekawa i niezwykle aktualna. Tematyka jest zgodna z obecnym nurtem badań związanych z tworzeniem systemów uczenia maszynowego w detekcji dezinformacji w danych tekstowych. Podjęta tematyka dotyczy nie tylko bardzo istotnych aspektów teoretycznych, związanych z efektywną ekstrakcją cech z danych tekstowych oraz hybrydyzacją cech wypracowanych przez pojedyncze ekstraktory, ale także jest niezwykle istotna z praktycznego punktu widzenia.

2 Ocena strony merytorycznej rozprawy doktorskiej

Rozprawa doktorska jest napisana w języku polskim i składa się z sześciu numerowanych rozdziałów oraz bibliografii, spisu tabel i rysunków. **Rozdział pierwszy** jest wprowadzeniem do tematyki rozprawy – Doktorant rozpoczął ten rozdział od uwypuklenia istotności i aktualności podjętej tematyki. Następnie przedstawiono cel pracy oraz hipotezę badawczą oraz zadania badawcze – cel i hipoteza badawcza, a także zadania badawcze zostały poprawnie zdefiniowane. Autor krótko przedstawił strukturę pracy oraz najistotniejsze wyniki uzyskane w rozprawie. Rozdział pierwszy jest zwieńczony listą publikacji, w których Doktorant był współautorem oraz projektów badawczych, w których

brał udział. Pewien niedosyt pozostawia fakt, że nie omówiono wkładu mgr. inż. Gracjana Kątka w powstanie każdej z wymienionych publikacji (w przypadku niektórych z nich ten wkład został uszczegółowiony w samym artykule). Warto jednak zauważyć, że w 3/5 pracach Doktorant był pierwszym autorem, a w 2/5 – drugim (wskazuje to na istotny wkład mgr. inż. Gracjana Kątka w powstanie tych prac). Należy również podkreślić, że trzy prace zostały opublikowane w bardzo dobrych czasopismach (dwie w *Neurocomputing*, jedna w *Logic Journal of the IGPL*).

Rozdział 2. jest przeglądem stanu wiedzy w zakresie metod detekcji dezinformacji w danych tekstowych. Autor przedstawił obecnie funkcjonujące taksonomie metod detekcji dezinformacji oraz omówił wybrane metody detekcji dezinformacji. Wartościowym elementem tego rozdziału jest syntetyczne podsumowanie wybranych metod wykrywania fałszywych informacji w różnych językach (Tabela 2.1; zauważyłem, że metody są pogrupowane dla każdego języka, z jedną nieścisłością: BanFakeNews, dla języka bengalskiego, został wpleciony w grupę metod dla j. angielskiego). Warto byłoby jednak rozszerzyć to podsumowanie – cennymi informacjami, które można by zawrzeć w tej tabeli byłyby odpowiedzi na następujące pytania:

- Czy wykorzystany (w analizowanej pracy) zbiór danych jest publicznie dostępny?
- Czy wykorzystany zbiór danych jest „AI-ready", tj. czy zdefiniowano ustandaryzowany podział zbioru na podzbiory treningowe, walidacyjne i testowe?
- Jakie metryki oceny jakości technik detekcji fałszywych wiadomości wykorzystano w analizowanej pracy?
- Czy implementacje metod(y) opisywanych w analizowanej pracy są publicznie dostępne? Czy eksperymenty są w pełni reprodukowalne?

Rozdział drugi został zwieńczony syntetycznym podsumowaniem dotychczasowych podejść do detekcji dezinformacji. Pewien niedosyt pozostawia fakt, że w przeglądzie literatury niewiele jest najnowszych metod (np. z 2025 r.) – warto jednak podkreślić, że Doktorant bardzo dobrze orientuje się w literaturze związanej z tematyką rozprawy.

W **rozdziale 3.** przedstawiono dwie metody, opracowane przez Doktoranta: metodę *Learned Fusion Method* (LFM) oraz metodę *Graph Embedding Method* (GEM). W pierwszej z nich wykorzystano model DistilBERT oraz technikę *Term Frequency-Inverse Document Frequency* (TF-IDF) do ekstrakcji cech – fuzja cech wyekstrahowanych z wykorzystaniem tych podejść jest przeprowadzona z wykorzystaniem sieci typu *autoencoder*. W metodzie GEM wykorzystano różne grafy (spójności semantycznej tekstu, wynikania logicznego, przyczynowo-skutkowego oraz wiedzy) jako ekstraktory bazowe, których cechy są następnie integrowane. W kolejnym **rozdziale 4.** Autor omawia metodykę badań eksperymentalnych oraz omawia wykorzystane zbiory danych i metryki, które zostaną użyte do oceny jakości działania opracowanych detektorów. Wyniki badań eksperymentalnych zostały zaprezentowane w **rozdziale 5.**, a **rozdział 6** syntetycznie podsumowuje rozprawę.

Opracowane algorytmy oraz otrzymane wyniki eksperymentalne przedstawione w rozprawie pozwalają mi stwierdzić, że Doktorant jest dobrze przygotowany do tego, żeby prowadzić dalsze badania związane z analizą danych tekstowych, ze szczególnym uwzględnieniem metod projektowania systemów uczenia maszynowego, wykorzystujących hybrydowe ekstraktory cech. Rozprawa doktorska

mgr. inż. Gracjana Kątka, a także artykuły, których jest współautorem prezentują dojrzałość badawczą oraz wiedzę teoretyczną i praktyczną Doktoranta w dyscyplinie *Informatyka Techniczna i Telekomunikacja*. Stronę merytoryczną rozprawy oceniam pozytywnie.

2.1 Pytania i uwagi

W trakcie lektury rozprawy nasunęły mi się następujące pytania i uwagi:

1. W *Streszczeniu* (na str. 5) Autor krótko omówił hipotezę badawczą, w której czytamy, że „*... możliwe jest opracowanie hybrydowych metod ekstrakcji cech w celu podniesienia skuteczności detekcji dezinformacji w tekstach względem metod klasycznych*" – warto byłoby precyzyjniej określić, jakie „metody klasyczne" Doktorant ma na myśli.

2. W Rozdziale 3. (i w kolejnych rozdziałach) zabrakło mi bezpośredniego odniesienia do konkretnych artykułów, których współautorem był Doktorant, a które są bezpośrednio powiązane z przedstawionymi w tych rozdziałach zagadnieniami (jak wspomniano na str. 10).

3. W algorytmach przedstawionych w pracy możemy zauważyć różne hiperparametry, np. d_{tfidf} czy współczynnik dropoutu. Autor, w niektórych miejscach, wspomina, że wartości tych hiperparametrów zostały dobrane empirycznie, np. $d_{\text{tfidf}} = 500$ – jak dokładniej wyglądał proces doboru wartości najistotniejszych hiperparametrów opracowanych metod?

4. W sekcji 3.1.2 Autor omawia architekturę sieci typu *autoencoder*, która została wykorzystana jako mechanizm „fuzji" wyekstrahowanych cech. Architektura jest stosunkowo prosta (co nie jest jej wadą) – czy Autor rozważał wykorzystanie innych architektur (lub zupełnie innych podejść, które mogłyby pozwolić na fuzję cech)? Jakie inne podejścia można byłoby wykorzystać?

5. W literaturze istnieje szereg podejść, które pozwalają radzić sobie z (ekstremalnie) niezbalansowanymi zbiorami treningowymi w treningu nadzorowanym. W rozprawie Autor (poprawnie) wykorzystał funkcję straty *Focal Loss*. Jak rozumiem, zamierzeniem Doktoranta było zweryfikowanie, które metody detekcji dezinformacji mogą być z powodzeniem wytrenowane bez dodatkowego, syntetycznego generowania danych treningowych (w procesie augmentacji danych). Jakie inne podejścia do tworzenia modeli uczenia maszynowego z wykorzystaniem niezbalansowanych próbek treningowych można byłoby wykorzystać i czy inne podejścia (oprócz wykorzystania funkcji *Focal Loss*) były rozważane/eksplorowane przez Doktoranta?

6. Opisy metod opracowanych przez Doktoranta są raczej wysokopoziomowe – bazując na tak wysokopoziomowych opisach mogłoby być (bardzo) trudno zreprodukować opisywane metody (w pracy nie znalazłem też odnośnika do repozytorium, w którym Autor zamieścił swoją implementację). Udostępnienie implementacji oraz zapewnienie, że opisy opracowywanych metod (oraz metod(y) referencyjnych, które też są raczej wysokopoziomowe) są szczegółowe są szczególnie ważne w „kryzysie reprodukowalności" badań, z którym obecnie musimy się mierzyć [1]. Czy Doktorant rozważał udostępnienie swoich implementacji (wraz ze skryptami, jednoznacznie pokazującymi jak zreprodukować badania eksperymentalne)?

7. Na str. 39 Autor wspomina, że w badaniach eksperymentalnych wykorzystano procedurę walidacji krzyżowej, ale szczegóły tego procesu nie zostały dokładnie przedstawione. Czy podzbiory

treningowe/testowe były w jakiś sposób stratyfikowane? Uważam, że Autor powinien udostępnić wyznaczone podzbiory (zwłaszcza jeśli nie są ustandaryzowane w ramach konkretnych zbiorów danych), aby zapewnić reprodukowalność eksperymentów. Czy Doktorant rozważał zweryfikowanie możliwości generalizacyjnych opracowanych algorytmów pomiędzy zbiorami (np. trening na zbiorze WELFake, test na zbiorze ISOT Fake News)?

8. W Sekcji 4.2 Autor omawia wykorzystane zbiory danych – korzystnie byłoby przedstawić dokładniejszą charakterystykę tych zbiorów (np. z rozkładami długości tekstu w poszczególnych podzbiorach w ramach walidacji krzyżowej). Zabrakło mi także przykładów tekstów (lub ich fragmentów) z każdej klasy (treści fałszywe/prawdziwe) i każdego zbioru danych. W przypadku wykorzystanego zbioru wieloetykietowego warto byłoby zaprezentować takie charakterystyki dla każdej z 13 etykiet. Czy którykolwiek z wykorzystanych zbiorów jest zbiorem „AI-ready", tj. czy istnieją dla niego ustandaryzowane podziały na podzbiory treningowe i testowe, które zostały zdefiniowane przez twórców zbioru?

9. Autor, na str. 46, podczas omawiania metryk używanych do oceny klasyfikatorów i macierzy pomyłek powinien jednoznacznie wskazać, co rozumie przez „przypadek pozytywny". Skoro celem jest detekcja dezinformacji, tę klasę można by interpretować jako klasę „pozytywną" (natomiast na str. 41, w Tabeli 4.1, etykieta 1 jest przypisana do tekstów prawdziwych).

10. W jaki sposób Autor dobrał hiperparametry wykorzystanych klasyfikatorów. Przykładowo, jaką funkcję jądrową wykorzystano w maszynach wektorów podpierających? Czy hiperparametry tej funkcji były w jakiś sposób optymalizowane? Podobne pytania można zadać w przypadku innych klasyfikatorów.

11. W tabelach, w których przedstawiono wyniki eksperymentalne, Autor powinien jednoznacznie wskazać, że przedstawiono wartości średnie, wyznaczone w ramach walidacji krzyżowej. Oprócz wartości średniej warto byłoby także przedstawić np. medianę i odchylenie standardowe dla każdej z metryk, a połączenie tabel 5.1–5.3 (i analogicznie dla pozostałych zbiorów) znacznie ułatwiłoby analizę wyników i porównanie metod.

12. W Tabeli 5.53 przedstawiono porównanie metod na zbiorze OpenFact – czy przedstawione metryki zostały wyznaczone dla dokładnie tych samych podziałów tego zbioru (na podzbiory treningowe i testowe)? Jeśli nie, to wartości przedstawione w tabeli mogą być (bardzo) mylące i zależne od tych podziałów.

13. Dlaczego, w Sekcji 5.5 (w analizie wpływu poszczególnych komponentów opracowanych metod na otrzymywane wyniki), jako miarę oceny wykorzystano metrykę dokładności (Autor sam zauważa, że zbiory danych są niezbalansowane)?

14. Na str. 85 czytamy, że usunięcie grafu przyczynowo-skutkowego powoduje polepszenie jakości działania modelu XGBoost – czy Doktorant analizował dokładniej to zjawisko? Dlaczego tak się dzieje? Taka głębsza analiza i interpretacja otrzymywanych wyników byłaby bardzo ciekawym elementem pracy.

15. Czy Autor analizował przykłady poprawnie/niepoprawnie zaklasyfikowanych tekstów? W pracy korzystnie byłoby zaprezentować i omówić takie (najciekawsze) przykłady tekstów dla każdego ze zbiorów danych.

16. Jakie najciekawsze kierunki dalszych badań można byłoby zidentyfikować na podstawie wyników otrzymanych w ramach rozprawy?

3 Ocena strony formalnej rozprawy doktorskiej

3.1 Ocena układu pracy

Układ rozprawy jest logiczny, a kolejne rozdziały wprowadzają czytelnika w najistotniejsze aspekty prac prowadzonych przez Doktoranta.

3.2 Ocena zastosowanego piśmiennictwa

Bibliografia zawiera około 98 pozycji („około", ponieważ zauważyłem duplikaty: [8] i [9]), a dobór prac jest poprawny i wkazuje na to, że Autor dobrze orientuje się w aktualnym nurcie badań związanych z tematyką rozprawy. W trakcie analizy piśmiennictwa zauważyłem drobne uchybienia dotyczące braku spójności, np. nazwy czasopism są czasami pisane w wersji skróconej (np. w [56]), a czasami pełnej (np. w [57]). W niektórych wpisach w bibliografii brakuje numerów stron, wydawnictwa lub innych informacji bibliograficznych (np. w [7], [17], [19], [41], [42], [43], [47], [54], [74], [75], [79], [87], [90], [91], [92]). Zauważyłem również niepoprawne użycie wielkich/małych liter (np. „norwegian" zamiast "Norwegian" w [78], czy „spanish" zamiast "Spanish" w [56]). Na str. 7, na której Autor krótko omawia dotychczas stosowane metody detekcji dezinformacji (i ekstrakcji cech), warto byłoby umieścić referencje do odpowiednich prac opisujących omawiane techniki. Na str. 20 (w Tabeli 2.1) możemy zauważyć niepoprawną referencję (Daria-Mihaela et al. [?]).

3.3 Uwagi redakcyjne

Rozprawa jest napisana starannie, a tekst jest poprawny pod względem językowym, stylistycznym i interpunkcyjnym. Dopatrzyłem się kilku drobnych uchybień:

- Doktorant nie ustrzegł się literówek, np. „*In this thesis, focuses...*"→„*This thesis focuses...* (str. 6), „*Generwanie cech*"→„*Generowanie cech*" (Rys. 3.1, str. 26), „*Droput*"→„*Dropout*" (str. 28), „*graf przyczynowo-skutkowego*"→„*graf przyczynowo-skutkowy*" (str. 31), „*Ekstrkcja cech*"→„*Ekstrakcja cech*" (Rys. 4.1, str. 38), „*a jego szczegółowa struktura zbioru danych została...*"→„*a jego szczegółowa struktura została...*" (str. 40), „*dla mary dokładności*"→„*dla miary dokładności*" (str. 54), „*...na uzyskanie ogólnej średnie skuteczności wyższej...*"→„*...na uzyskanie ogólnej średniej skuteczności wyższej...*" (str. 87).
- W pracy zauważyłem kilka „sierot" (pojedynczych liter na końcu wiersza, np. na str. 7).
- Autor (czasami) niespójnie używa myślników i dywizów.
- Każdy skrót/akronim (nawet „oczywisty") powinien być zdefiniowany przy pierwszym użyciu (np. SVM, KNN na str. 18).
- Jako separatora dziesiętnego, Autor powinien używać przecinka (który jest naprzemiennie używany z kropką, np. na str. 22).
- W niektórych miejscach w pracy (zwłaszcza w pobliżu odniesień do literatury) brakuje spacji (np. "*Abdelkader[69]*"→"*Abdelkader [69]*" na str. 22).

- W tekście nie zauważyłem odniesienia do wzoru 3.9.
- W pracy zauważyłem niespójne wykorzystanie wielkich/małych liter (np. svm, mlp na str. 39 – w innych miejscach pracy Autor stosuje poprawny zapis SVM). Podobnie, na str. 42–43, można zauważyć zapis Infostrateg i INFOSTRATEG.
- Zamiast „*ilość słów*" powinniśmy raczej powiedzieć „*liczba słów*" (np. na str. 41).
- Na str. 46 lepiej byłoby zapisać 2×2 zamiast 2x2.
- Najlepsze wyniki (dla każdej z analizowanych metryk) powinny być pogrubione we wszystkich tabelach – znacznie ułatwiłoby to analizę takich tabel.
- Czy, od strony praktycznej, ma sens przedstawiać 4 cyfry po separatorze dziesiętnym dla analizowanych metryk?
- Jakość niektórych rysunków, np. Rys. 5.1, mogłaby być lepsza (w przypadku tego rysunku tekst na nim zawarty jest bardzo mały, a jedna z osi jest nieopisana).

4 Konkluzja

Z pełnym przekonaniem stwierdzam, że przedmiotem rozprawy doktorskiej jest oryginalne rozwiązanie poprawnie zdefiniowanego problemu naukowego, a recenzowana dysertacja Pana mgr. inż. Gracjana Kątka **spełnia** wymagania stawiane rozprawom doktorskim przez Ustawę *Prawo o szkolnictwie wyższym i nauce, Dz. U. 2023, poz. 742 z późn. zm.* – praca prezentuje ogólną wiedzę teoretyczną i praktyczną Doktoranta w dyscyplinie *Informatyka Techniczna i Telekomunikacja* oraz umiejętność samodzielnego prowadzenia pracy naukowej.

W związku z powyższym, **wnioskuję o przyjęcie rozprawy doktorskiej oraz o dopuszczenie mgr. inż. Gracjana Kątka do publicznej obrony**.

Jakub Nalepa

Literatura

[1] S. Kapoor, A. Narayanan, Leakage and the reproducibility crisis in machine-learning-based science, Patterns (Aug. 2023). doi:10.1016/j.patter.2023.100804.
URL https://www.cell.com/patterns/abstract/S2666-3899(23)00159-9