



dr hab. Bogdan Księżopolski, prof. ALK  
Katedra Cyberbezpieczeństwa  
Akademia Leona Koźmińskiego

29.04.2026 r.

## Recenzja rozprawy doktorskiej

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

**Autor: mgr inż. Gracjan Kątek**

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

Rozprawa doktorska dotyczy zaprojektowania, implementacji oraz empirycznej weryfikacji autorskich, hybrydowych metod ekstrakcji cech tekstu w celu poprawy skuteczności detekcji dezinformacji. Podjęta tematyka jest aktualna i istotna społecznie, gdyż dezinformacja stała się jednym z kluczowych zagrożeń informacyjnych ostatnich lat, zarówno w kontekście procesów politycznych, zdrowia publicznego, jak i cyberbezpieczeństwa.

Autor jednoznacznie sformułował cel pracy: „Celem niniejszej rozprawy jest opracowanie i empiryczna weryfikacja autorskich, hybrydowych metod ekstrakcji cech tekstu, łączących podejścia semantyczne, kontekstowe i strukturalne, w celu zwiększenia skuteczności detekcji dezinformacji.”

Postawioną tezę badawczą sformułowano w sposób następujący: „Zastosowanie hybrydowych metod ekstrakcji cech, łączących reprezentacje semantyczne, logiczne, kontekstowe, przyczynowo-skutkowe oraz mechanizmy adaptacyjnego łączenia wektorów osadzeń, pozwala na istotne zwiększenie skuteczności detekcji dezinformacji w porównaniu z podejściami opartymi wyłącznie na modelach transformerowych.”

W celu weryfikacji postawionej hipotezy oraz realizacji głównego celu pracy Autor wyznaczył pięć zadań badawczych:



1. Opracowanie hybrydowej metody ekstrakcji cech uwzględniającej strukturę leksykalną oraz kontekst semantyczny.
2. Opracowanie metody łączącej informacje semantyczne, logiczne, kontekstowe oraz przyczynowo-skutkowe w celu poprawy jakości ekstrakcji cech.
3. Zaprojektowanie mechanizmu adaptacji wag, którego celem jest dostosowanie znaczenia poszczególnych typów reprezentacji.
4. Badanie skuteczności metod hybrydowych w detekcji dezinformacji w porównaniu z referencyjnym podejściem opartym wyłącznie na modelach transformerowych.
5. Przeprowadzenie eksperymentalnej ewaluacji proponowanych metod hybrydowych w zestawieniu z innymi podejściami detekcji dezinformacji opisanymi w literaturze.

Zarówno cel, hipoteza, jak i zadania badawcze zostały sformułowane poprawnie i precyzyjnie. Postawiona hipoteza jest weryfikowalna empirycznie, a jej wykazanie wymagało rozwiązania wszystkich pięciu zadań badawczych, co Autor systematycznie zrealizował w kolejnych rozdziałach pracy.

Charakter rozprawy oceniam jako projektowo-doświadczalny: część projektowa obejmuje zaprojektowanie dwóch oryginalnych architektur ekstrakcji cech (LFM i GEM), zaś część eksperymentalna obejmuje implementację referencyjną oraz ewaluację porównawczą na pięciu zbiorach danych).

**2. Czy w rozprawie przeprowadzono w sposób właściwy analizę źródeł świadcząca o dostatecznej wiedzy autora? Czy wnioski z przeglądu źródeł sformułowano w sposób jasny i przekonujący.**

Analiza źródeł została przeprowadzona w rozdziale 2 rozprawy i obejmuje 98 pozycji bibliograficznych. Pod względem ilościowym należy uznać ją za adekwatną do podjętej tematyki. Bibliografia obejmuje publikacje z renomowanych czasopism oraz konferencji międzynarodowych, w tym z baz indeksowanych w JCR (m.in. *Neurocomputing*, *Information Fusion*, *Applied Soft Computing*, *Expert Systems with Applications*, *Computers in Industry*, *IEEE Access*, *Scientific Reports*) oraz wydawnictw o ugruntowanej pozycji (Springer, Elsevier, ACM, IEEE).

Pod względem aktualności należy podkreślić, że Autor uwzględnił prace najnowsze – znaczna część pozycji bibliograficznych pochodzi z lat 2022–2025, co odzwierciedla



rzeczywiście aktualny stan badań w bardzo dynamicznie rozwijającej się dyscyplinie detekcji dezinformacji.

Wnioski z przeglądu literatury są jasno sformułowane i zostały zwięźle przedstawione w podrozdziale 2.3. Autor wskazuje trzy kluczowe luki badawcze: (1) niewystarczające wykorzystanie grafów w językach innych niż angielski, (2) niedostatek systematycznych badań porównawczych dotyczących efektywności hybrydowych metod ekstrakcji cech, (3) zależność od niewielkich, najczęściej binarnych zbiorów danych. Te trzy obserwacje stanowią logiczne uzasadnienie dla podjętych w pracy zadań badawczych.

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

W mojej ocenie Autor rozwiązał postawione zagadnienia w sposób poprawny i kompletny. Zaproponowana metodyka badawcza jest spójna z postawioną tezą oraz adekwatna do charakteru analizowanego problemu. Szczególnie warto zwrócić uwagę na następujące elementy.

- **Procedura eksperymentalna jest dobrze zaprojektowana** (rys. 4.1), obejmuje preprocessing, walidację krzyżową, ekstrakcję cech metodą referencyjną oraz autorską, klasyfikację z wykorzystaniem czterech alternatywnych algorytmów (XGBoost, SVM, MLP, drzewo decyzyjne) oraz ocenę wyników z użyciem zestawu metryk zarówno globalnych (accuracy), jak i wrażliwych na niezbalansowanie klas (balanced accuracy,  $F1$  dla każdej klasy).
- **Dobór miar oceny jest poprawny**, w szczególności zastosowanie zbalansowanej dokładności na zbiorach niezbalansowanych (Infotester, OpenFact, SWAROG) jest właściwym wyborem metodologicznym.
- **Analiza statystyczna istotności różnic pomiędzy rozwiązaniami** została przeprowadzona z użyciem testu Friedmana oraz testów post-hoc Wilcozona z korekcją Holma–Bonferroniego, co jest bardzo dobrą praktyką w analizie porównawczej klasyfikatorów.
- **Przyjęcie modelu DistilBERT** jako metody referencyjnej ze względu na jego dwujęzyczne wersje.

Mam jeden komentarz dotyczący metodologii. W metodzie LFM Autor ograniczył liczbę tokenów wejściowych do 256 (zamiast natywnych 512 dla DistilBERT), motywując to redukcją zużycia zasobów obliczeniowych. Choć w literaturze pojawiają się prace



wskazujące, że dla pewnych domen krótsze sekwencje mogą być wystarczające, to jednak średnia długość tekstów w zbiorach np. WELFake i ISOT znacząco przekracza okno 256 tokenów po tokenizacji. Dyskusyjne jest, czy ta redukcja nie powoduje utraty informacji istotnych dla detekcji dezinformacji, szczególnie w przypadku artykułów, w których kluczowe elementy retoryczne (wnioski, źródła) znajdują się w drugiej połowie tekstu. Pomocna byłaby analiza porównująca wyniki dla 256 i 512 wejściowych tokenów.

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

W mojej ocenie samodzielny, oryginalny dorobek Autora obejmuje cztery kluczowe elementy.

Po pierwsze, opracowanie metody **LFM (Learned Fusion Method)** stanowiącej zintegrowane rozwiązanie łączące reprezentację semantyczną z modelu DistilBERT z reprezentacją statystyczną TF-IDF, przy czym integracja realizowana jest przez autorską sieć neuronową typu encoder (FusionEncoder) z dwiema warstwami liniowymi, funkcją aktywacji ReLU, mechanizmem Dropout oraz funkcją straty Focal Loss. Choć poszczególne komponenty są znane z literatury, ich konkretna kombinacja oraz ewaluacja wzajemnego wpływu komponentów stanowią wkład poznawczy. W szczególności wykazano, że samo TF-IDF bez DistilBERT prowadzi do spadku dokładności o 6.29% (SVM) do 15.49% (drzewa decyzyjne), zaś sam DistilBERT bez TF-IDF do spadku o 1.22%–1.92%, co jednoznacznie potwierdza komplementarność obu reprezentacji.

Po drugie, opracowanie metody **GEM (Graph Embedding Method)** opartej na czterech komplementarnych grafach reprezentujących różne aspekty struktury tekstu: (1) graf spójności tekstu, (2) graf wynikania logicznego (NLI), (3) graf przyczynowo-skutkowy, (4) graf wiedzy zewnętrznej (BERTopic + RSS Google News). Końcowa reprezentacja powstaje przez konkatencję wektorów cząstkowych. Takie podejście stanowi oryginalny wkład w kontekście prac z literatury, ponieważ pojedyncze grafowe metody detekcji dezinformacji są szeroko opisane, jednak integracja czterech zróżnicowanych perspektyw analizy semantyczno-logiczno-kontekstualnej w jednej hybrydowej metodzie ekstrakcji cech nie jest w literaturze powszechnie reprezentowana.

Po trzecie, zaprojektowanie i implementacja mechanizmu **adaptacyjnego łączenia reprezentacji** (FusionEncoder w metodzie LFM), który automatycznie dostosowuje wagi pomiędzy reprezentacją semantyczną a leksykalną w procesie treningu sieci. Mechanizm



ten pozwala na uniknięcie ręcznego ważenia komponentów, co jest typowym ograniczeniem prac wcześniejszych.

Po czwarte, **eksperymentalna walidacja** metod hybrydowych na pięciu zbiorach danych o zróżnicowanej charakterystyce, czyli zbiory anglojęzyczne i polskie, zbiory binarne i wieloetykietowe, zbiory zbalansowane i silnie niezbalansowane. Wartością dodaną jest przeprowadzenie analizy porównawczej z najnowocześniejszymi metodami opisanymi w literaturze, w tym z modelami GPT-4 na zbiorze OpenFact. Warto zwrócić uwagę, że zaproponowana metoda GEM osiąga wyższą skuteczność niż szereg wariantów GPT-4.

Pozycja rozprawy w stosunku do stanu wiedzy światowej jest dobra. Autor ma istotny dorobek publikacyjny bezpośrednio powiązany z tematyką rozprawy, w tym dwie publikacje w czasopiśmie *Neurocomputing* (IF=6.5), co należy uznać za znaczące osiągnięcie na etapie doktoratu. Wyniki uzyskane na zbiorze OpenFact ( $F1 = 0.9047$  dla GEM) lokują się w czołowej grupie rezultatów dla tego zbioru i są wyraźnie konkurencyjne wobec dużych modeli językowych typu GPT-4.

Warto również podkreślić, że przeprowadzone w pracy weryfikacje empiryczne osiągają wysoki poziom gotowości technologicznej (TRL około 5), co oznacza, że proponowane komponenty zostały zweryfikowane w środowisku zbliżonym do rzeczywistego.

##### **5. Czy autor wykazał umiejętność poprawnego i przekonującego przedstawienia uzyskanych przez siebie wyników (zwięzłość, jasność, poprawność redakcyjna rozprawy)?**

Praca została napisana w języku polskim, na bardzo dobrym poziomie językowym. Nie zauważyłem istotnych problemów językowych, a terminologia techniczna stosowana jest poprawnie (z wprowadzeniem oryginałów anglojęzycznych w nawiasach przy pierwszym wystąpieniu). Praca została zredagowana w sposób bardzo dobry, zarówno warstwa wzorów matematycznych, jak i opis architektur metod LFM oraz GEM są jasne i czytelne.

Struktura pracy jest logiczna i zgodna ze standardami rozpraw doktorskich z dyscypliny informatyki technicznej i telekomunikacji: wstęp z motywacją, hipotezą i celem, przegląd literatury, opis autorskich metod, metodyka badań, prezentacja wyników, podsumowanie. Każdy z rozdziałów został zwieńczony spójnym i logicznym podsumowaniem. Zaproponowane architektury LFM oraz GEM zostały bardzo dobrze zilustrowane diagramami, a wyniki eksperymentów zaprezentowane w formie tabel są przejrzyste i pozwalają na bezpośrednią analizę porównawczą metod.





obliczeniowych (DistilBERT zamiast pełnowymiarowego LLM). To czyni je atrakcyjnymi dla praktycznych wdrożeń, w tym w środowiskach produkcyjnych z ograniczonymi zasobami.

Po drugie, **walidacja na zbiorach polskich** (Infotester, OpenFact, SWAROG) ma istotną wartość praktyczną dla rynku polskiego. Skuteczna detekcja dezinformacji w językach mniej zasobnych pozostaje wyraźnie słabiej rozwiązana w literaturze niż w przypadku języka angielskiego.

Część prezentowanych wyników została już opublikowana w wysoko punktowanych czasopismach międzynarodowych (Neurocomputing, IF=6.5), co jest dodatkowym potwierdzeniem ich wartości naukowej. Ponadto badania prowadzone były w ramach projektu SWAROG finansowanego przez NCBiR w programie INFOSTRATEG, co potwierdza ich aplikacyjny charakter oraz wskazuje na potencjał wdrożeniowy.

Myślę, że komercjalizacją zaproponowanego rozwiązania mogą być zainteresowane zarówno podmioty z sektora mediów cyfrowych, platformy społecznościowe, jak i instytucje publiczne odpowiedzialne za bezpieczeństwo informacyjne państwa.

Podsumowując, przydatność uzyskanych wyników dla nauk inżynierjno-technicznych oceniam jako istotną.

#### **8. Do której z następujących kategorii Recenzent zalicza rozprawę:**

- a) nie spełniająca wymagań stawianych rozprawom doktorskim przez obowiązujące przepisy,
- b) wymagająca wprowadzenia poprawek i ponownego recenzowania,
- c) spełniająca wymagania,
- d) spełniająca wymagania z wyraźnym nadmiarem,
- e) wybitnie dobra, zasługująca na wyróżnienie.**

Uważam, że rozprawa doktorska mgr inż. Gracjana Kątka **jest wybitnie dobra i spełnia z wyraźnym nadmiarem wymogi stawiane** rozprawom doktorskim przez obowiązujące przepisy. Należy podkreślić, że część wyników rozprawy została już opublikowana w literaturze międzynarodowej, a w szczególności w czasopiśmie *Neurocomputing* indeksowanym w JCR.

dr hab. Bogdan Książkowski