

Recenzja pracy doktorskiej mgr. Mateusza Klimaszewskiego

„Multi-objective modularity in deep learning for natural language processing”

autor: Krzysztof Jassem

26 października 2025

## 1. Wstęp

Niniejsza recenzja dotyczy rozprawy doktorskiej mgr. Mateusza Klimaszewskiego pt. „Multi-objective Modularity in Deep Learning for Natural Language Processing”. Rozprawa, napisana w języku angielskim, ma charakter cyklu czterech publikacji naukowych. Prace te zostały zebrane w formie spójnej monografii, poprzedzonej wprowadzeniem oraz omówieniem tła teoretycznego dotyczącego zastosowania modularnego uczenia głębokiego (Modular Deep Learning, MDL) w przetwarzaniu języka naturalnego (Natural Language Processing, NLP).

## 2. Ocena merytorycznej strony rozprawy

### 2.1. Zakres badań i pytania badawcze

W rozdziale 1. Autor rozprawy formułuje cztery pytania badawcze:

*Wielozadaniowość*: Jak zaprojektować system NLP, który umożliwi użytkownikowi swobodny wybór zestawu realizowanych zadań?

*Wielodziedzinowość*: W jaki sposób, w ramach jednego zadania, skutecznie połączyć wiedzę specjalistyczną pochodzącą z różnych domen?

*Wielojęzyczność*: Czy możliwa jest efektywna integracja wiedzy z modeli generatywnych obsługujących różne języki naturalne?

*Wielomodelowość*: Czy można w sposób skuteczny scalić informacje pozyskiwane z wielu modeli generatywnych?

Zaznaczam, że powyższe sformułowania problemów badawczych stanowią parafrazę dokonaną przez Recenzenta.

Oceniam, że postawione pytania badawcze mają charakter aktualny i odnoszą się do zagadnień istotnych dla rozwoju współczesnego przetwarzania języka naturalnego. Odnoszę wrażenie, że wynikają one z realnych potrzeb praktycznych, a nie z perspektywy teoretycznej, co podkreśla ich znaczenie aplikacyjne.

### 2.2. Wkład badawczy

Podrozdział 1.2 przedstawia wkład badawczy Autora, potwierdzony publikacjami w prestiżowych konferencjach (EMNLP, LREC, ECAI, COLING). Warto podkreślić, że w większości prac udział Doktoranta jest dominujący – tak przynajmniej wynika z Jego oświadczenia. Nie znalazłem oświadczeń współautorów prac, ale w każdej pracy wchodzących w cykl Kandydat znajduje się na pierwszym miejscu.

Wkład Autora rozprawy przedstawia się następująco:

- opracowanie modularnej metody analizy morfosyntaktycznej oraz jej implementacja w ogólnodostępnym systemie *Combo*,

- zastosowanie mechanizmu *Gated Adapters* w zadaniu wielodziedzinowego tłumaczenia automatycznego,
- opracowanie koncepcji *language arithmetic*, będącej adaptacją idei *task arithmetic* do środowiska wielojęzycznego,
- wprowadzenie i eksperymentalna weryfikacja koncepcji *transferu modułów zadaniowych* pomiędzy różnymi modelami językowymi.

Przedstawione osiągnięcia mają zarówno znaczenie teoretyczne, jak i praktyczne. Autor wprowadza nowe rozwiązania w obszarze modularnego uczenia głębokiego, poszerzając możliwość adaptacji modeli językowych do wielu zadań, domen i języków.

### 2.3. Umieszczenie w literaturze

Rozdział 2 dostarcza przeglądu badań nad modularnym uczeniem głębokim oraz jego zastosowaniami w NLP. Autor wykazuje staranność w oznaczaniu referencji oraz osadzenie własnych badań w szerszym kontekście. Dzięki temu rozprawa ukazuje nie tylko oryginalny wkład Autora, ale i pełne zrozumienie miejsca, jakie ten wkład zajmuje w rozwoju dziedziny.

### 2.4. Wartość merytoryczna poszczególnych publikacji

*COMBO: State-of-the-Art Morphosyntactic Analysis (EMNLP 2021, System Demonstrations)*

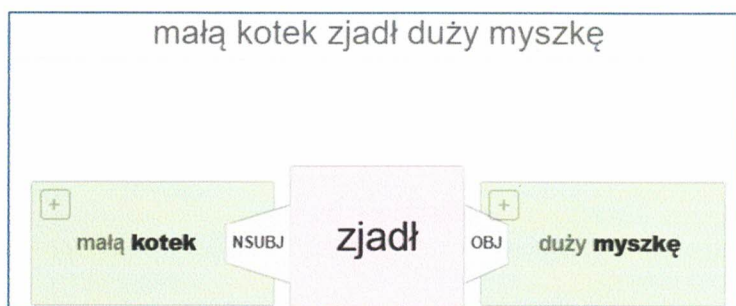
Publikacja przedstawia system *COMBO*, który jest modułową architekturą do analizy morfosyntaktycznej języka naturalnego. System wspiera następujące zadania:

- tokenizacja,
- analiza morfologiczna,
- lematyzacja,
- parsing zależnościowy.

Według Autora architektura systemu pozwala użytkownikowi końcowemu (ang. *end-user*) na elastyczne definiowanie modułów wejściowych (np. wykorzystanie cech morfologicznych lub znaczników POS) oraz wyjściowych (np. ograniczenie się tylko do parsera).

Dla mnie nieco mylące jest użycie w pracy terminu *end-user* w odniesieniu do programisty, który ma możliwość wytrenowania nowego modelu na swoich danych i zastosowania go w systemie *COMBO*. Takiej możliwości (zgodnie z moimi eksperymentami) nie ma bowiem użytkownik końcowy, który korzysta z systemu dostępnego na stronie <https://combo-demo.nlp.ipipan.waw.pl/combo-eng>.

Odporność narzędzi do analizy zależnościowej języka polskiego ciekawie bada się na przykładach krótkich zdań o strukturze nieciągłej (z którymi dobrze radzą sobie parsery regułowe). Autorowi rozprawy pozostawiam analizę rezultatów systemu *COMBO* dla zdań „Małą kotek zjadł duży myszkę” (rysunek poniżej) oraz „Która jest godzina?” (w porównaniu z wynikiem parsingu dla pytania „Która godzina jest?”).



Podsumowując, *COMBO* jest narzędziem inżynierskim, które obrazuje praktyczne zalety modułowości (jeszcze przed pojawieniem się pojęcia *Modular Deep Learning*). System zyskał uznanie międzynarodowe – znalazł się w czołówce konkursu IWPT 2021.

#### *Gated Adapters for Multi-Domain Neural Machine Translation (ECAI 2023)*

Artykuł dotyczy zastosowania metody wykorzystującej dane pochodzące z różnych domen w zadaniu tłumaczenia automatycznego. W tym celu opracowano architekturę *Gated Adapters*, w której komponent pełniący rolę routera decyduje o sposobie łączenia wyjść pochodzących z różnych adapterów domenowych.

Praca adresuje realny problem: w sytuacji, gdy tekst wejściowy nie może zostać jednoznacznie przypisany do jednej domeny, zamiast ograniczać się do pojedynczego źródła danych treningowych, możliwe jest efektywne wykorzystanie danych pochodzących z wielu domen równocześnie.

Na szczególną uwagę zasługuje *aplikacyjny charakter* metody, gdyż zaproponowane rozwiązanie może być stosowane w rzeczywistych systemach tłumaczenia automatycznego.

#### *No Train but Gain: Language Arithmetic for training-free Language Adapters enhancement*

W tej publikacji wprowadza się koncepcję *language arithmetic*, dla której inspiracją jest idea *task arithmetic*, przeniesiona tutaj na poziom reprezentacji językowych. Badania eksperymentalne pokazują, że możliwe jest połączenie adapterów wytrenowanych dla różnych języków (np. angielskiego i francuskiego) w celu uzyskania nowego adaptera dla języka trzeciego (np. hiszpańskiego), co pozwala osiągnąć poprawę jakości w scenariuszu *zero-shot* (czyli bez dodatkowych przykładów uczących).

Praca ukazuje zdolność autorów do twórczej adaptacji istniejących paradygmatów badawczych nowych obszarów zastosowań.

Wyniki mogą znaleźć praktyczne znaczenie szczególnie dla języków o mniejszych zasobach (*low-resource languages*).

#### *Is Modularity Transferable? A Case Study through the Lens of Knowledge Distillation*

Destylacja wiedzy ma na celu zwiększenie efektywności modeli językowych poprzez zmniejszenie ich rozmiarów. W omawianej pracy weryfikuje się, czy w procesie destylacji możliwe jest skuteczne przeniesienie modułów zadaniowych z modelu większego (nauczyciela) do modelu mniejszego (ucznia). Przeprowadzone eksperymenty wykazują, że jest to możliwe w scenariuszu, w którym model ucznia powstał jako destylowana wersja modelu nauczyciela, natomiast skuteczność tego podejścia istotnie maleje w przypadku modeli trenowanych niezależnie.

Badania zrealizowano na danych wielojęzycznych, obejmujących korpusy w 20 językach, oraz dla trzech reprezentatywnych zadań NLP: rozpoznawania nazw własnych (NER), identyfikacji parafrazy oraz wnioskowania w języku naturalnym (NLI). Projekt eksperymentalny potwierdza zatem uniwersalność prowadzonych analiz.

Wartość merytoryczna pracy polega na krytycznej ocenie możliwości oraz ograniczeń metod modularnego uczenia głębokiego.

### 2.5. Replikowalność wyników

W ocenie prac badawczych z dziedziny informatyki istotnym aspektem jest replikowalność uzyskanych wyników. W tym kontekście należy podkreślić, że wszystkie wyniki badań Autora zostały udostępnione w postaci pakietów oprogramowania o otwartym dostępie, co umożliwi ich niezależne odtworzenie i ponowne wykorzystanie. Takie podejście zwiększa transparentność badań oraz znacząco ułatwia dalszy rozwój prac w dziedzinie rozprawy.

## 3. Ocena formalnej strony rozprawy

### 3.1. Język

Praca napisana jest w języku angielskim na wysokim poziomie. Styl jest precyzyjny i pozbawiony zbędnych komplikacji. Autor unika zdań złożonych!

### 3.2. Układ pracy

Struktura rozprawy jest logiczna i przejrzysta. Rozdziały 1 i 2 dostarczają zarówno wprowadzenia teoretycznego, jak i kontekstu dla własnych badań. Odwołania do literatury są poprawne i spójne.

### 3.3. Estetyka

Rozprawa jest dobrze zredagowana i czytelna. Schematy rysunkowe i tabele są klarowne i ułatwiają odbiór treści.

## 4. Wpływ pozostałych osiągnięć doktoranta

Rozprawa doktorska obejmuje cykl publikacji, z których trzy zostały przygotowane we współpracy z zespołami krajowymi, natomiast jedna powstała w ramach współpracy międzynarodowej. Na uwagę zasługuje pozostały dorobek publikacyjny Doktoranta, realizowany we współpracy z badaczami zagranicznymi. Świadczy to o ugruntowanej pozycji międzynarodowej Autora.

Za jedno z najbardziej znaczących osiągnięć Autora uważam jego udział w konsorcjum międzynarodowym, które opracowało wielojęzyczny model językowy *EuroLLM*. Model ten stanowi istotny krok naprzód w zakresie przetwarzania języka naturalnego dla języków europejskich. Co istotne, jego wysoka jakość potwierdza się w zastosowaniach praktycznych, o czym miałem możliwość przekonać się osobiście, stosując go w komercyjnym systemie tłumaczenia automatycznego.

## 5. Podsumowanie recenzji

Stwierdzam, że rozprawa doktorska magistra Mateusza Klimaszewskiego spełnia wymagania stawiane pracom doktorskim i rekomenduję dopuszczenie Doktoranta do dalszych etapów przewodu doktorskiego.

Ponadto wnioskuję o nadanie rozprawie wyróżnienia, co uzasadniam następującymi, obiektywnie potwierdzonymi osiągnięciami Doktoranta:

- wszystkie publikacje składające się na rozprawę są ocenione na liście MNiSW na poziomie 140 punktów,
- Autor rozprawy jest pierwszym autorem każdej z tych publikacji,
- Kandydat jest autorem kodów źródłowych udostępnionych w formie sześciu otwartych pakietów programistycznych,
- Doktorant prowadził cztery własne projekty badawcze oraz uczestniczył w realizacji trzech kolejnych, w tym dwóch o charakterze międzynarodowym,
- Kandydat jest jedynym przedstawicielem Polski w międzynarodowym konsorcjum opracowującym wielojęzyczny model językowy EuroLLM.

Krzysztof Janem