

dr hab. Piotr Pęzik, prof. UŁ
Wydział Filologiczny, Uniwersytet Łódzki
ul. Pomorska 171/173, 90-236, Łódź

Łódź, 1.12.2025

Recenzja rozprawy doktorskiej Pana mgr. inż. Mateusza Klimaszewskiego pt.
“*Multi-objective modularity in Natural Language Processing*” napisanej pod
kierunkiem prof. PW, dr hab. inż. Tomasza Gambina.

Przedstawiona rozprawa doktorska składa się z dwóch rozdziałów wstępnych,
części zasadniczej w postaci czterech spójnych tematycznie publikacji, których
pierwszym autorem jest Doktorant, rozdziału podsumowującego oraz bibliografii
obejmującej 204 pozycje.

We wstępie do scalonej wersji rozprawy Autor wprowadza paradygmat
modularnego uczenia głębokiego (MDL z ang. Modular Deep Learning) jako
ważny obszar badań nad modelami językowymi opartymi na głębokich sieciach
neuronowych zwłaszcza w kontekście przetwarzania języka naturalnego. Należy w
tym miejscu podkreślić wyjątkową aktualność tej tematyki. Po początkowym
zachwycie nad możliwościami ogólnego przeznaczenia generatywnych Dużych
Modeli Językowych jako uniwersalnej formy sztucznej inteligencji działającej
niezależnie od domeny czy konkretnego zadania, powszechne przekonanie o
rychłym nadejściu AGI (generycznej sztucznej inteligencji) wyraźnie osłabło.
Generatywne modele językowe operują w spektrum memoryzacji, interpolacji i
ekstrapolacji wiedzy i zdolności. Innymi słowy, wiedza i umiejętności, które nie
znalazły się w docelowej postaci w pre-treningu i post-treningu bazowych modeli

językowych (ang. foundation models) są mniej lub bardziej udanie interpolowane na podstawie danych z innych języków, pokrewnych domen czy podobnych zadań. Jak podkreśla autor adaptacja poprzez parametrycznie efektywne dostrajanie (PEFT) modeli bazowych może skutkować znaczną optymalizacją obliczeniową w konkretnych zastosowaniach. Można wykazać, że celowana adaptacja językowa, domenowa i zadaniowa modeli pre-trenowanych np. poprzez dostrajanie PEFT na dedykowanych danych może skutkować jakościową poprawą działania nawet największych modeli generatywnych. Często jest to poprawa, która decyduje o przydatności modelu w warunkach produkcyjnych, a nie jedynie demonstracyjnych.

Rozdział drugi nakreśla teoretyczne podstawy modularnego uczenia głębokiego, które Autor wywodzi z badań nad uczeniem transferowym: rozwój metod MDL wynikał z jednej strony z potrzeby strony ograniczenia negatywnego transferu zadaniowego czy językowego, a z drugiej z konieczności znacznej optymalizacji procesu dostrajania modeli głębokich do konkretnych domen czy zadań. Osobnym aspektem MDL poruszonym w tej części wstępu są metody scalania modeli (model merging), których celem jest m. in. ‘arytmetyka językowa’ (termin wprowadzony w publikacji P3), czyli wzmacnianie kompetencji językowych w różnych konfiguracjach adapterów modelu nauczyciela i studenta.

Można uznać, że w tych krótkich wstępnych rozdziałach Autor klarownie wykazuje spójność tematyczną przedstawionych w dalszej części publikacji jako prac nad modularnością głębokich modeli językowych w aspekcie wielozadaniowości, wielodziedzinowości, wielojęzyczności oraz wielomodelowości. Najzwięźlej tę spójność wyrażają jasno sformułowane na

stronie 12 cztery główne pytania badawcze, z którymi autor zmierzył się w zgłoszonych publikacjach.

Rozdział 3 to pierwsza publikacja z cyklu zatytułowana *COMBO: State-of-the-Art Morphosyntactic Analysis* przedstawiony na konferencji EMNLP w 2021 r. Udział doktoranta w tej publikacji określono na 60%. Artykuł przedstawia system przetwarzania języka naturalnego COMBO, którego najważniejszą, choć nie jedyną funkcją jest analiza składniowo-zależnościowa tekstów w ponad 40 językach. Wyróżnikiem tego rozwiązania jest wysoka skuteczność w zakresie predykcji rozkładów zależnościowych (SOTA w 2021 r.), optymalizacja procesu trenowania, dostępność reprezentacji wektorowych dla poszczególnych warstw strukturalnych oraz modularność użytkowa. Osobne głowice predykcyjne COMBO obsługują różne zadania z zakresu analizy morfologicznej, morfosyntaktycznej i składniowej-zależnościowej. W niezależnej kontynuacji tej pracy powstał również moduł ekstrakcji jednostek nazewniczych. Praca ta zasługuje z pewnością na bardzo wysoką ocenę a raportowane w niej wyniki można uznać za poziom światowy NLP w 2021 r. (być może z wyjątkiem podzadania lematyzacji). Po lekturze publikacji nasuwa się pytanie o aktualność tego typu rozwiązań w 2025 r. Sama przydatność badawcza parserów składniowych choćby w językoznawstwie komputerowym, czy w humanistyce cyfrowej nie budzi wątpliwości recenzenta. W części wstępnej brakuje natomiast w moim odczuciu kilkudzaniowego odniesienia do skuteczności dużo większych modeli klasy LLM w tym obszarze. Czy oparty na modelach reprezentacyjnych parser COMBO nadal jest nie tylko wielokrotnie wydajniejszy, ale też dokładniejszy od LLM-ów na tych zadaniach? Czy może adaptacja LLM-ów np. poprzez pełne lub adaptacyjne dostrojenie modeli bazowych na tym zadaniu daje porównywalne wyniki?

Rozdział 4 stanowi druga publikacja z cyklu, zatytułowana *Gated Adapters for Multi-Domain Neural Machine Translation*, przedstawioną na konferencji European Conference on Artificial Intelligence (ECAI 2023). Udział Doktoranta w tej pracy określono na poziomie 65%. Zaproponowano w niej system tłumaczenia maszynowego złożonego z tzw. mieszaniny ekspertów (ang. MoE, mixture of experts), a dokładniej modułów modelu językowego pre-trenowanych na jednej z góry założonych domen, m.in. prawo, medycyna, informatyka. W odróżnieniu od standardowych adapterów domenowych, rozwiązanie to nie wymaga odrębnego klasyfikatora dziedzinowego, który może spowalniać działanie całego systemu, a dodatkowo nie dokonuje sztywnego przypisania próbki tekstu do jednej domeny. Zaproponowane podejście polega na użyciu modelu routera destylowanego z zewnętrznego modelu RoBERTa, który pozwala na ‘wieloetykietową’ klasyfikację domenową tłumaczonego tekstu. Ewaluacja na zadaniu tłumaczenia maszynowego wykazuje przewagę tego rozwiązania nie tylko nad systemami z zewnętrznym klasyfikatorem domenowym, ale też (w przypadkach niejednorodnych tekstów) nawet nad systemami z tzw. ‘wyrocznią dziedzinową’, tzn. takimi, które znają prawdziwą domenę tłumaczonego tekstu. Praca ta stanowi doskonały przykład zalet modularności domenowej głęboki modeli językowych. Można też stwierdzić, że stanowi udaną próbę wzmocnienia pozytywnego transferu między domenami poprzez równoległą aktywację adapterów. Potencjalną wadą tego rozwiązania może być konieczność zapewnienia stosunkowo dużych ilości danych domenowych dla tłumaczonych języków.

Rozdział 5 stanowi publikacja pt. *No Train But Gain: Language Arithmetic for training-free Language Adapters Enhancement*, przedstawiona na konferencji COLING 2025 z 70% udziałem Doktoranta. Tematem pracy jest scalanie pretrenowanych adapterów językowych jako beztreningowa metoda dopasowania

językowego modeli. Podejście to rozpatrywane jest w dwóch głównych aspektach: a) scenariusz ‘zero-shot’, czyli adaptacyjne aproksymacje języków, na których nie dokonywano treningu bazowym oraz b) wzmacnianie kompetencji języków podobnie lub słabiej reprezentowanych w pre-treningu. Ewaluację przeprowadzono na docelowych zadaniach wykrywania jednostek nazewniczych, inferencji i odpowiadania na pytania. Wyniki eksperymentów są szczególnie przekonujące dla scenariusza zero-shot (braku adaptera dla danego języka) oraz dla języków słabo reprezentowanych (ang. low-resource languages). Publikacja ta stanowi przykład doskonałej znajomości tematyki adaptacji językowej głębokich modeli. Na marginesie tej pracy nasuwa się pewna wątpliwość terminologiczna, dotycząca trafności rozróżnienia między PLM-ami (Pre-trained Language Models) a LLM-ami (Large Language Models). Desygnaty tych terminów są w kontekście publikacji jasne: PLM-y to zdaje się przede wszystkim transformery reprezentacyjne, a LLM-y to nowszej generacji transformery typu text-to-text o znacznie większej liczbie parametrów. Jednakże każdy LLM jest w dosłownym znaczeniu tego terminu również PLM-em, w związku z czym to rozróżnienie (obecne również w podsumowaniu rozprawy) może być mylące bez dokładnej definicji.

Czwarta publikacja z cyklu zawarta w **Rozdziale 6** pt. *Is Modularity Transferable? A Case Study through the Lens of Knowledge Distillation* zaprezentowana została podczas konferencji LREC-COLING 2024 z 80% udziałem Doktoranta. Praca prezentuje badanie przenośności modułów wydajnego dostrajania (PEFT) między modelami nauczyciela i studenta w dwóch scenariuszach, tj. ‘łatwym’ i ‘trudnym’. Łatwość scenariusza pierwszego polega na tym, że do modelu studenta przenoszone są adaptory o kompatybilnej architekturze (zgodność między modelami *mBERT* i *distilbert-base-multilingual-cased*). W trudniejszym

scenariuszu architektury modeli są niekompatybilne. W trakcie ewaluacji udaje się potwierdzić, że przeszczepienie warstw nauczyciela skutkuje lepszym wykonywaniem szeregu standardowych zadań NLP. Co ciekawe, w trakcie eksperymentów autorzy potwierdzają hipotezę o specjalizacji warstwowej modeli: podejście SKIP polegające na transferze co n-tej warstwy modelu nauczyciela daje lepsze wyniki niż uśrednianie sąsiadujących warstw. Praca ta niewątpliwie stanowi ważne badanie modularności modelowej głębokiego uczenia.

Wnioski i konkluzje z recenzji

Przedstawiona do recenzji rozprawa doktorska zasługuje na jednoznacznie pozytywną ocenę i jednocześnie *spełnia wymogi stawiane w ustawie Prawo o szkolnictwie wyższym i nauce , Dz. U. 2023, poz. 742 z późn. zm.* Dodatkowo w ocenie Recenzenta *zasługuje na wyróżnienie*. Autor proponuje i pomyślnie waliduje oryginalne rozwiązania problemu naukowego jakim jest modularność głębokiego uczenia w obszarze przetwarzania języka naturalnego. Przedstawiony cykl publikacji, których pierwszym i głównym autorem jest Doktorant potwierdza szeroką wiedzę w dyscyplinie informatyki technicznej i telekomunikacji oraz bezsprzeczne umiejętności samodzielnego prowadzenia pracy naukowej. Zaproponowane w rozprawie podejścia w sposób kompleksowy i wieloaspektowy traktują modularność głębokiego uczenia. Doktorant projektuje i przeprowadza eksperymenty na zróżnicowanych zbiorach danych osiągając np. w zadaniu analizy składniowej języka naturalnego wyniki na poziomie światowym. Wywód przeprowadza w klarowny i konsekwentny sposób z zachowaniem poprawności metodologicznej właściwej dla dyscypliny i obszaru badań. Wobec powyższego *wniosuję o dopuszczenie Doktoranta do dalszych etapów przewodu doktorskiego.*

Prof. Jerzy