

RDITT-PI. 14.01.2025
M. Kow

Kraków 2 stycznia 2025

Prof. dr. hab. inż. Jarosław Wąs,
Wydział Elektrotechniki, Automatyki, Informatyki i Inżynierii Biomedycznej
Akademia Górniczo Hutnicza im. Stanisław Staszica w Krakowie

Recenzja rozprawy doktorskiej

Recenzję pracy doktorskiej mgr inż. Jeremiaha Oluwagbemi Abimbola zatytułowanej „*Algorithmic methods for detecting the tempo and time signatures of music tracks*”, opracowano na prośbę prof. Andrzeja Polańskiego Przewodniczącego Rady Dyscypliny Informatyka Techniczna i Telekomunikacja Politechniki Śląskiej (na podstawie umowy UMC/4509/2024 z dnia 27 listopada 2024).

1a) Tematyka rozprawy i tytuł pracy

Tematyka rozprawy doktorskiej mgr inż. Jeremiaha Oluwagbemi Abimbola pt. „*Algorithmic methods for detecting the tempo and time signatures of music tracks*”, dotyczy problematyki prognozowania metrum z sygnałów audio z wykorzystaniem algorytmów uczenia maszynowego.

Praca obejmuje takie zagadnienia jak: stworzenie zbioru danych składającego się z kilku tysięcy próbek audio z adnotacjami (nazwanego METER2800), zbadanie modeli uczenia maszynowego i głębokiego z uwzględnieniem architektur hybrydowych pod kątem badania metrum i tempa ścieżek muzycznych. Do celów optymalizacyjnych w zakresie wykrywania metrum Doktorant zastosował algorytm genetyczny co zostało opisane podczas jednego z workshopów konferencji GECCO 2024 – lista Core A.

W mojej ocenie temat pracy poprawnie koresponduje z zawartymi treściami.

1b) Charakterystyka rozprawy i ocena układu rozprawy

Praca doktorska mgr inż. Jeremiaha Oluwagbemi Abimbola składa się łącznie z dziesięciu rozdziałów, obejmujących: wstęp, zasadniczą część pracy, podsumowanie oraz bibliografię. Całość pracy zajmuje 146 stron wraz z bibliografią i dodatkami.

Pierwszy rozdział pracy zatytułowany „Introduction” poświęcony jest określeniu celów ogólnych i szczegółowych pracy, wskazaniu problemu naukowego oraz określeniu metodyki pracy z uwzględnieniem: badań literaturowych, wypracowaniu generycznego podejścia i szeregu szczegółowych rozwiązań w modelowaniu predykcji.

W mojej ocenie układ redakcyjny pracy jest poprawny. W pierwszym rozdziale opisano tło badań z uwzględnieniem problematyki pracy, motywacji oraz metodyki. Przedstawiono podstawowe hipotezy badawcze. Następne w rozdziale 2 omówiono klasyfikację muzyki oraz w rozdziale 4 problematykę wykrywania metrum (ang. time signature). Rozdział 4 poświęcony jest prezentacji metod uczenia maszynowego. Rozdział 5 obejmuje stworzony przez Doktoranta zbiór danych o nazwie METER2800 – opublikowany przez niego w czasopiśmie „Data in Brief” Elsevier (40 pkt). Rozdział 6 propozycja współczynnika podobieństwa nazwanego MFCCSM zaproponowanego przez Doktoranta. Rozdział 7 poświęcony jest optymalizacji wykrywania metrum, zaś rozdział 8 użyciu metod uczenia maszynowego do tego celu. Rozdział 8 poświęcony jest tworzeniu metodyki porównawczej i analitycznej do detekcji metrum z wykorzystaniem złożonych metod – zostało to również częściowo opisane przez Doktoranta w czasopiśmie Eurasip Journal on Audio, Speech, and Music Processing (100 pkt). Rozdział 9 stanowi podsumowanie.

Autor posługuje się poprawną terminologią w zakresie domenowym (muzyki), a także zagadnienia uczenia maszynowego i, patrząc szerzej, stosowanych metod sztucznej inteligencji.

Bibliografia jest generalnie aktualna. Odwołania do pozycji literaturowych są prawidłowe.

Tytuły i kolejność rozdziałów jest również prawidłowa, treść rozdziałów odpowiada, w mojej ocenie, wymaganiom stawianym dla prac doktorskich.

1c) Ocena zastosowanego piśmiennictwa

Bibliografia składa się ze 167 pozycji związanych z tematyką pracy. Pozycje literaturowe zostały dobrane prawidłowo biorąc pod uwagę ich istotność i aktualność. Zostały one poprawnie wkomponowane w treść rozprawy.

1d) wskazanie oraz ocena celu pracy

Celem pracy było opracowanie nowych metod algorytmicznych do wykrywania tempa i metrum utworów muzycznych. W szczególności przedmiotem prac było zbadanie skuteczności różnych technik ekstrakcji cech, metod przetwarzania sygnałów i algorytmów uczenia maszynowego do wykrywania tempa i metrum. Wygenerowane w pracy metody zostały ocenione przy użyciu relatywnie dużego zbioru danych utworów muzycznych reprezentujących różne gatunki i style muzyki przygotowanego przez Doktoranta i opublikowane m.in. w czasopiśmie „Data in Brief” oraz amerykańskim zbiorze Dataverse.

W mojej ocenie cel pracy został zrealizowany w prawidłowy sposób.

1e) wskazanie oraz ocena zastosowanych metod badawczych

W prognozowaniu zapotrzebowania Autor zastosował metody klasyczne analizy metrum (ASM, BSM, MFCCSM) oraz metod uczenia maszynowego z uwzględnieniem architektur hybrydowych. W pracy zaproponowano parametryzację i poprawę szeregu metod uczenia maszynowego do problemu wykrywania metrum. Przeprowadzono szereg testów porównawczych na bazie różnych zbiorów danych, a w szczególności dedykowanej bazy danych stworzonej przez Doktoranta. Uważam, że metody badawcze zostały zastosowane w prawidłowy sposób.

1f) ocena części rozprawy doktorskiej dotyczącej omówienia wyników badań

Wyniki badań zostały opisane w praktycznej części pracy, w szczególności w rozdziałach 8, 9 oraz 10. W rozdziale 8 przeanalizowano najważniejsze parametry architektur uczenia maszynowego i metryk oceny ich skuteczności w zakresie wykrywania metrum. W rozdziale 9 uwagę poświęcono m.in. hybrydyzacji metod i tu przedstawiono najważniejsze oryginalne wyniki pracy. Z kolei rozdział 10 stanowi krótkie podsumowanie wszystkich prac.

W mojej ocenie prace merytoryczne zostały przeprowadzone i opisane prawidłowo. Autor szczegółowo opisał metody, w wielu przypadkach zostały one zilustrowane za pomocą obrazów i tabel. Podsumowując ten punkt oceny uważam, że wyniki badań zostały opisane w prawidłowy sposób.

1g) informacje dotyczące praktycznego zastosowania uzyskanych wyników badań

Badania zawarte w rozprawie mają wymiar praktyczny – zaś praca jest związana z konkretnymi zbiorami danych i analizą utworów muzycznych. Dokładne wykrywanie metrum jest przydatne dla różnych zastosowań analizy muzyki,

transkrypcji i gromadzeniu informacji o utworze. Zaproponowany przez Doktoranta model nazwany Mel-Frequency Cepstral Coefficient Similarity Matrix MFCCSM, który dla funkcji ekstrakcji z punktu widzenia dokładności wykrytego metrum. Rozwiązanie to ma, w mojej ocenie, spory potencjał rozwojowy i wdrożeniowy.

1h) informacje o ewentualnych nieprawidłowościach, które pojawiły się w ocenianej rozprawie doktorskiej

Nie zidentyfikowano poważniejszych nieprawidłowości w rozprawie.

1i) ocena czy rozprawa doktorska stanowi oryginalne rozwiązanie problemu naukowego

W pracy rozwiązano szereg zagadnień związanych z wykrywaniem metrum, które świadczą o oryginalności rozwiązania. Problemem naukowym jest skuteczne wykrywanie metrum w utworach muzycznych.

Po pierwsze zaproponowano nowatorskie podejście do wykrywania metrum poprzez wprowadzenie modelu Mel-Frequency Cepstral Coefficient Similarity Matrix. Po drugie zaproponowano hybrydową architekturę sieci np. rezydualno – konwolucyjnej do funkcji MFCC i spektrogramów do wykrywania metrum. Po trzecie zastosowano optymalizację Bayesowską i algorytmy genetyczne do poprawy parametrów uczenia się i skuteczności klasyfikacji. Po czwarte stworzono całościową metodykę łączącą metody tradycyjne z zaawansowanymi metodami uczenia maszynowego.

Biorąc pod uwagę wszystkie składowe uważam, że oceniana rozprawa zawiera oryginalne rozwiązanie problemu naukowego.

1j) ocena czy rozprawa doktorska prezentuje ogólną wiedzę teoretyczną kandydata w dyscyplinie albo dyscyplinach oraz umiejętność samodzielnego prowadzenia pracy naukowej lub artystycznej

W mojej ocenie praca wpisuje się w dyscyplinę informatyka techniczna i telekomunikacja i napisana jest poprawnie z punktu widzenia metodyki. Autor wykazał się zarówno wiedzą teoretyczną w zakresie analizy parametrów utworów muzycznych i metod sztucznej inteligencji i praktycznymi umiejętnościami w zakresie analizy danych i tworzenia algorytmiki.

Uwagi:

Poniżej przedstawiono uwagi polemiczne dotyczące strony merytorycznej pracy:

1. Sformułowanie hipotezy 1 (H1) jest mocno asekuracyjne, gdyż mówi, że wykorzystaniu MFCCSM jako narzędzia do wykrywania metrum *może potencjalnie przynieść znaczące korzyści*. Czy w świetle przeprowadzonych prac nie dałoby się tej hipotezy sformułować bardziej zdecydowanie i jednoznacznie? To pozwoli lepiej wypozycjonować i docenić wyniki badań.
2. Analogicznie aktualne sformułowanie hipotezy 2 (H2) mówi o tym, że algorytm genetyczny *może zostać użyty do optymalizacji* MCCM. Czy dałoby się sformułować tę hipotezę bardziej zdecydowanie? Brakuje mi tutaj szerszego odniesienia się do skutków i korzyści wykorzystania algorytmu genetycznego, w szczególności w porównaniu do innych metod. Owszem pewien opis jest na stronie nr 5 czy dalszych częściach pracy, ale powinien być dokładniejszy. Podobne pytanie dotyczy hipotezy 3 (H3), która dotyczy jedynie faktu, że implementacja sieci konwolucyjnej wykorzystującej schemat ResNet *posiada potencjał* do efektywnego wykrywania metrum.
3. W rozdziale 4 i 8 przydałby się dokładniejszy opis architektur hybrydowych. Ostatnio to bardzo ważny kierunek rozwoju sieci. To pozwoliłoby lepiej uzasadnić dobór architektur wybranych do pracy z uwzględnieniem różnych dostępnych opcji. Autor stosuje przykładowo architekturę hybrydową RESNET-LSTM, ale nie jest jasne dlaczego wybrał akuratnie tę możliwość.
4. W rozdziale 10 zabrakło mi podsumowania dotyczącego wkładu w dyscyplinę Informatyka Techniczna i Telekomunikacja. To znaczy, z mojego punktu widzenia, brakuje konkretnie wylistowania wprost najważniejszych, zdaniem Autora, szczegółowych osiągnięć pracy z punktu widzenia informatyki technicznej. Na stronie tytułowej pracy powinna być jasno wskazana dyscyplina pracy Informatyka techniczna i Telekomunikacja – wskazano jedynie ogólnie program Informatyka i specjalność Informatyka.
5. W rozdziale 10 powinno znaleźć się również dokładniejsze uzasadnienie realizacji celów pracy.

Wniosek końcowy

Rozprawa stanowi oryginalne rozwiązanie problemu naukowego i wskazuje na wysoki poziom wiedzy teoretycznej i praktycznej Kandydata w dyscyplinie naukowej Informatyka Techniczna i Telekomunikacja.

Przedstawiona praca doktorska spełnia w pełni warunki określone w art. 187 ustawy z dnia 20 lipca 2018 r. Prawo o szkolnictwie wyższym i nauce (j.t. Dz.U.

z 2020 r. poz. 85 z późn. zm.). W szczególności zgodnie z wymaganiami ww. ustawy oceniana rozprawa pozytywnie prezentuje wiedzę teoretyczną Kandydata oraz umiejętność samodzielnego prowadzenia pracy naukowej, przedmiotem rozprawy jest oryginalne rozwiązanie problemu naukowego. W opinii recenzenta praca może więc być dopuszczona do publicznej obrony.

Prof. dr hab. inż. Jarosław Wąs
jaroslaw.was@agh.edu.pl

