

Edmonton, 3 października 2023

Recenzja rozprawy doktorskiej mgr. inż. Marka Gajewskiego

Niniejsza recenzja została przygotowana w odpowiedzi na prośbę zastępcy dyrektora IBS PAN Pana dr. hab. inż. Jana W. Owsieńskiego.

Ocena pracy doktorskiej:

Automatyczna klasyfikacja dokumentów tekstowych w zarządzaniu aktami spraw na użytek e-administracji.

Chciałbym rozpocząć moją ocenę pracy, wyrażając swoje uznanie dla poruszonego tematu, zastosowanych modyfikacji pojęcia zawierania się zbiorów rozmytych oraz rezultatów przedstawionej pracy. Z dużym zainteresowaniem czytałem analizę różnych podejść i rozwiązań zawierania się zbiorów z uwzględnieniem ważności elementów i ich semantycznego znaczenia, które zostały przedstawione przez mgr. inż. Marka Gajewskiego.

1. OCENA OGÓLNA:

Recenzowana rozprawa doktorska składa się z ośmiu rozdziałów. Przedstawia ona zaproponowaną przez doktoranta nową – opartą na zbiorach rozmytych – metodę klasyfikacji dokumentów. Metoda ta umożliwia klasyfikację dokumentów na kategorie i sprawy. Dodatkowo kategorie są reprezentowane jako hierarchie. Praca zawiera bardzo obszerną i dogłębną część eksperymentalną.

Oceniam tą pracę wysoko. Świadczy ona o dużym nakładzie pracy wykonanej przez doktoranta, jak również o dogłębnym zrozumieniu przedmiotu podjętych przez niego badań. Zawiera ona bardzo interesujące podejście do pojęcia zawierania się zbiorów rozmytych i jego zastosowanie do celów klasyfikacji, jak również walidację zaproponowanego rozwiązania. Treść rozprawy pokazuje, że doktorant posiada rozległą wiedzę podstawową w dziedzinach: 1) zbiorów rozmytych, a w szczególność operacji zawierania zbiorów; oraz 2) klasyfikacji dokumentów i uczenia maszynowego. Proponowane rozszerzenie operatora zawierania jest ciekawe i nowatorskie.

Wszystko to jest świadectwem głębokiego zrozumienia przez mgr. inż. Marka Gajewskiego aspektów niezbędnych do stworzenia odpowiednich operacji zawierania zbiorów rozmytych z uwzględnieniem stopni ważności elementów oraz ich analizę uwzględniającą aspekty semantyczne dotyczące interpretacji tych stopni ważności.

2. CEL PRACY:

Teza pracy doktorskiej – zastosowanie technik zbiorów rozmytych do klasyfikacji dokumentów (kategorie i sekwencje spraw) jest przykładem, i mam nadzieje zapoczątkowaniem, rozszerzenia technik uczenia maszynowego o metody oparte na zbiorach rozmytych.

Rozprawa zawiera dokładny opis podstawowych pojęć z logiki rozmytej, wyszukiwania informacji tekstowych – w szczególności zadań odkrywania i śledzenia wątku (ang. Topic Detection and Tracking), oraz zaproponowaną modyfikację operatora zawierania i jego porównywanie. Cel – pytanie badawcze – jest dobrze określony i przekonująco umotywowany. Zakres przeprowadzonych badań przedstawiony w rozprawie jest odpowiedni dla uzyskania stopnia naukowego doktora.

3. METODOLOGIA:

Rozprawa jest przykładem doskonałej pracy rozszerzającej metody zbiorów rozmytych i ich zastosowaniem do wymagającego zadania klasyfikacji dokumentów. Proponowana modyfikacja operatora zawierania jest wnikliwie analizowana i semantycznie uzasadniona.

Wymagania odnośnie uwzględnienia hierarchii kategorii przy klasyfikacji dokumentów i ważności elementów reprezentujących te dokumenty są wyjaśnione i uzasadnione. Na szczególną uwagę zasługuje wprowadzona przez doktoranta niskowymiarowa reprezentacja dokumentów. Różne warianty modyfikacji operatora zawierania są analitycznie rozważane. Wszystko to wskazuje, że doktorant jest świadomy potrzeby dogłębnej analizy nowych rozwiązań.

Kolejnym ważnym aspektem pracy doktorskiej jest walidacja proponowanej metody klasyfikacji dokumentów. Zawiera ona badania różnych wariantów operatora/wskaźnika zawierania jak również porównanie z popularnym klasyfikatorem k-NN. Ta eksploracja dowodzi słuszności i stosowalności proponowanego operatora.

4. ORYGINALNOŚĆ I OSIĄGNIĘCIA:

Główny wkład mgr. inż. Marka Gajewskiego to: 1) propozycja metody klasyfikacji dokumentów z zastosowaniem operatora zawierania zbiorów rozmytych; 2) rozszerzenie operatora zawierania o uwzględnienie stopni ważności elementów reprezentujących dokumenty; oraz 3) obszerna i dogłębna analiza eksperymentalna.

Uważam, że praca ta jest bardzo ważna i pokazuje korzyści użycia technik zbiorów rozmytych. Przedstawione badania są bardzo oryginalne i nowatorskie.

5. INTERPRETACJA:

W rozprawie doktorskiej mgr. inż. Marek Gajewski przeprowadził walidację zaproponowanej metody/algorytmu. Przedstawione szczegóły podejścia do reprezentacji dokumentów (niskowymiarowa przestań) i różnych wskaźników zawierania stanowią integralną część pracy. Rozprawa zawiera również wyniki przeprowadzonej walidacji na dwóch dobrze znanych zbiorach dokumentów.

Recenzowana praca wyraźnie potwierdza umiejętności analityczne i interpretacyjne doktoranta.

6. PREZENTACJA:

Organizacja pracy doktorskiej jest bardzo dobra. Układ sekcji jest logiczny. Praca jest dobrze napisana i łatwa w lekturze.

7. PODSUMOWANIE:

Rozprawa doktorska potwierdza zrozumienie przez mgr. inż. Marka Gajewskiego znaczenia technik zbiorów rozmytych i ich zastosowania do uczenia maszynowego. Weryfikacja zaproponowanego podejścia zwiększa wartość pracy. Badania przeprowadzone w trakcie jej przygotowywania są dowodem na to, że mgr. inż. Marek Gajewski posiada szereg ważnych umiejętności: 1) identyfikacji celów badawczych; 2) osiągnięcia tych celów; 3) ich realizacji/implementacji; jak również 4) ich walidacji. To znakomity zestaw osiągnięć – mgr. inż. Marek Gajewski to osoba wysoko wykwalifikowana i pracowita, o doskonałych umiejętnościach analitycznych. Wszystko to, w połączeniu z intuicją badawczą wykazaną w opracowywaniu metodologii, pozwoli mgr. inż. Markowi Gajewskiemu na osiągnięcie sukcesów w przyszłości.

Podsumowując, praca ta stanowi istotny wkład w obszar zastosowań technik zbiorów rozmytych w dziedzinie uczenia maszynowego. Doktorant zaproponował oryginalną – semantycznie uzasadnioną – modyfikację operatora zawierania. Zweryfikował zaproponowane rozwiązanie na przykładzie dwóch korpusów danych. Wykazał się przy tym umiejętnością samodzielnego prowadzenia pracy naukowej. Tym samym, rozprawa doktorska mgr. inż. Marka Gajewskiego **spełnia warunki stosownych przepisów** i wnioskuje o dopuszczenie doktoranta do kolejnych etapów przewodu doktorskiego.

Podsumowując – to doskonała praca. Chciałbym pogratulować mgr. inż. Markowi Gajewskiemu efektów jego pracy.

Marek Reformat



Professor
Electrical and Computer Engineering,
University of Alberta