

Gdynia, 10 marca 2018r.

dr hab. inż. Ireneusz Czarnowski, prof. AMG
Katedra Systemów Informacyjnych
Akademia Morska w Gdyni
ul. Morska 83, 81-225 Gdynia

RECENZJA

rozprawy doktorskiej mgr Barbary Probierz

pt.: „*Automatyczna kategoryzacja wiadomości elektronicznych z zastosowaniem sieci społecznych i algorytmów mrowiskowych*”.

1. Problematyka naukowa rozprawy

Tematyka rozprawy doktorskiej dotyczy kategoryzacji wiadomości elektronicznych, a głównym celem rozprawy było opracowanie nowych algorytmów kategoryzacji wiadomości e-mail. Autorka rozprawy zaproponowała do rozwiązania problemu podejście oparte na algorytmach mrowiskowych oraz wykorzystwała wybrane własności rozwiązań sieciowych, w szczególności sieci społecznych. Zaproponowane do rozwiązania problemu metody należą do grupy metod inteligencji obliczeniowej.

Rozpatrywany przez Doktorantkę problem mieści się w obszarze zainteresowań informatyki, w szczególności jest związany z analizą i eksploracją danych tekstowych.

Hipotezą rozprawy jest przeświadczenie Doktorantki o przewadze algorytmów mrowiskowych oraz podejścia opartego na modelu sieciowym, w stosunku do innych podejść dedykowanych kategoryzacji wiadomości e-mail. Tym samym, hipotezą Doktorantki jest możliwość trafnego wskazywanie nowych folderów dla wiadomości e-mail w oparciu o zastosowanie algorytmów mrowiskowych i sieci społecznych.

Zakres merytoryczny rozprawy mieści się w obszarze aktywnym badawczo. Problem rozpatrywany w rozprawie jest często spotykany w literaturze naukowej, co potwierdza jego aktualności. Ponadto należy mieć na uwadze, że kategoryzacja wiadomości e-mail, tj. przypisywanie wiadomości do folderów ma swój wymiar praktyczny. Jest to też problem złożony obliczeniowo, co dodatkowo uzasadnia poszukiwanie nowych, bardziej wydajnych metod rozwiązujących ten problem.

2. Treść rozprawy

Recenzowana rozprawa została przygotowana w języku polskim. Obejmuje 121 stron i składa się z:

- 9 rozdziałów oraz wstępu i zakończenia (podsumowania),
- wykazu bibliografii, w której zawarto 99 pozycji literaturowych,
- spisu treści, rysunków oraz tabel.

We wstępie rozprawy znalazło się ogólne wprowadzenie do problemu. Doktorantka podkreśliła w nim ważności podjętego problemu badawczego, jego wymiar praktyczny, ale także złożoność obliczeniową. We wstępie sformułowano tezę oraz cel pracy.

Rozdział pierwszy pracy poświęcono zagadnieniu analizy danych, w szczególności analizie eksploracyjnej danych. Wprowadzono też terminologię związaną z analizą danych. Opisano wybrane modele analizy danych oraz scharakteryzowano techniki eksploracji danych. Szczególną uwagę poświęcono klasyfikacji, w tym odwołano się do wybranych algorytmów klasyfikacji, które zostały pokrótce scharakteryzowane. Rozdział pierwszy pracy zawiera także krótką charakterystykę algorytmów genetycznych.

Rozdział drugi rozprawy odwołuje się do algorytmów indukcji drzew decyzyjnych oraz do metod klasyfikacji danych opartych na zespołach klasyfikatorów. Podrozdział 2.3 dedykowano algorytmom mrowiskowym, gdzie opisano też implementację tego algorytmu związane z konstruowaniem drzew decyzyjnych oraz klasyfikatorów bazowych w algorytmie klasyfikacji opartym o tzw. lasy decyzyjne.

W kolejnym rozdziale rozprawy podano definicję sieci społecznej oraz scharakteryzowano wybrane modele sieciowe. Rozdział ten odwołuje się również do miar i wskaźników pozwalających na identyfikację właściwości sieci społecznych. Rozdział kończy krótki przegląd literatury związanej z analizą sieci społecznych.

W czwartym rozdziale Doktorantka poddała analizie zbiór danych Enron E-mail Dataset, który wykorzystwała jako zbiór referencyjny i benchmark w rozprawie. Zbiór danych został szczegółowo opisany. Następnie w podrozdziale 4.2 Autorka rozprawy odniosła się do problemu klasyfikacji dokumentów, a w szczególności do problemu klasyfikacji wiadomości e-mail i kategoryzacji ich ze względu na dedykowane dla nich foldery. W kolejnym podrozdziale zawarto przegląd literatury, na podstawie której dokonano omówienia zaproponowanych już rozwiązań kategoryzacji wiadomości e-mail. W kolejnym podrozdziale Doktorantka opisuje zastosowaną procedurę przygotowania danych do analizy, a następnie proponuje zastosowanie do opisu danych strukturę opartą na tabeli decyzyjnej.

Rozdział piąty rozprawy zawiera wyniki eksperymentów obliczeniowych klasyfikacji wiadomości e-mail ze zbioru Enron E-mail do folderów z wykorzystaniem wybranych algorytmów klasyfikacji oraz w oparciu o przyjęty sposób reprezentacji zbioru danych. Głównym celem eksperymentu była weryfikacji słuszności przyjętej struktury danych.

W szóstym rozdziale rozprawy Doktorantka zawarła opis zastosowania algorytmu mrowiskowego do kategoryzacji wiadomości e-mail. Algorytm został omówiony, a następnie przedstawiono eksperyment obliczeniowy i jego wyniki, co było podstawą oceny i porównania proponowanego podejścia z innymi algorytmami.

Rozdział siódmy rozprawy zawiera opis sieci społecznej, którą Doktorantka poddała analizie. Sieć ta została zidentyfikowana w oparciu o strukturę powiązań pomiędzy wiadomościami e-mail ze zbioru Enron E-mail. Utworzona macierz powiązań pomiędzy wiadomościami (ich nadawcami i adresatami), pozwoliła na opis tych powiązań z wykorzystaniem struktury modelu sieciowego, co docelowo umożliwiło identyfikację grup użytkowników.

Koncepcja modelu sieci została następnie wykorzystana do modyfikacji podejścia kategoryzacji wiadomości e-mail opartego na algorytmie mrowiskowym, które wcześniej zostało omówione w rozdziale szóstym pracy. Zaproponowany algorytm został poddany ocenie na drodze eksperymentów obliczeniowych. Wyniki tych eksperymentów zostały przedstawione i omówione w rozdziale ósmym rozprawy.

Rozdział dziewiąty pracy stanowi opis rozwiązania związanego z budowaniem struktury folderów dla wiadomości e-mail. Rozwiązanie to bazuje na ocenie struktury folderów innych użytkowników, co jest realizowane w oparciu o identyfikowany model tzw. „powiązań sieciowych” pomiędzy wiadomościami e-mail.

W zakończeniu pracy Doktorantka podsumowała uzyskane wyniki badań.

3. Ocena merytoryczna i najważniejsze wyniki uzyskane w pracy

Niewątpliwie podjęta w pracy problematyka związana z klasyfikacją wiadomości e-mail jest aktualna i stanowi wyzwanie zarówno dla samych użytkowników usługi poczty elektronicznej, którzy chcą dbać o porządek w swoim „inboxie”, kategoryzując wiadomości i umieszczając je w wybranych folderach; jak i badaczy poszukujących efektywnych modeli i algorytmicznych mechanizmów dla automatycznej kategoryzacji wiadomości e-mail i przypisywania tych wiadomości do dedykowanych folderów. Doktorantka podkreśla, że rozwiązanie postawionego problemu może mieć szersze zastosowanie niż tylko w kategoryzacji wiadomości e-mail. Wskazuje przy tym na problem identyfikacji spamu, jak również na zagadnienie priorytetowania wiadomości. Zawarty w pracy przegląd literatury skoncentrowano w szczególności na problemie klasyfikacji wiadomości e-mail, ich kategoryzowaniu, przypisywaniu lub przewidywaniu folderów dla tych wiadomości. Przeglądu tego dokonano sięgając do prac od roku 1989. Te najświeższe prace, do których odwołuje się Doktorantka we wspomnianym przeglądzie literatury, datowane są na rok 2015. Szkoda, że w ramach tego przeglądu literatury nie sięgnięto do prac nowszych, a w szczególności tych w których mowa o rozwiązaniach klasyfikacji wiadomości e-mail stosowanych przez jeden ze światowych koncernów.

Sam problem klasyfikacji wiadomości e-mail, jako szczególny przypadek problemu klasyfikacji dokumentów, jest problemem trudnym w sensie obliczeniowym. Doktorantka wspomina o tym, lecz szkoda, że nie poświęciła w swojej pracy temu zagadnieniu więcej miejsca.

W szczególności, rozwiązanie zaproponowane w pracy oparto na wykorzystaniu algorytmu mrowiskowego. Algorytm ten stanowi rozszerzenie wcześniej już zaproponowanego i prezentowanego w literaturze algorytmu. Wprowadzone rozszerzenie – pierwszy wariant algorytmu (Rozdział 6), polega na zaimplementowaniu mechanizmów przeznaczonych do badania sieci komunikacji pomiędzy poszczególnymi autorami wiadomości e-mail, gdy wiadomość ta została wysłana do grupy osób. Informacja o strukturze sieci komunikacji stanowi w tym przypadku podstawę do budowy rozwiązania, tj. pogrupowania wiadomości w foldery.

W drugim wariantcie proponowanego przez Doktorantkę algorytmu (Rozdział 8), działanie algorytmu oparto, a właściwie poprzedzono budową struktury sieciowej kontaktów pomiędzy użytkownikami. W oparciu o utworzoną strukturę sieci kontaktów Autorka rozprawy dokonuje grupowania, w konsekwencji którego uzyskuje się grupy użytkowników o podobnej strukturze folderów. W kolejnym kroku, zbiór analizowanych danych podlega przetworzeniu do struktury odpowiedniej dla poszukiwania docelowej kategoryzacji wiadomości e-mail oraz przypisania ich do właściwych folderów. To poszukiwanie zostało przeprowadzone z zastosowaniem algorytmu mrowiskowego.

Należy dodać, że dla obu wariantów algorytmu Doktorantka proponuje poszukiwanie odpowiedniej kategoryzacji wiadomości e-mail, po uprzednim przetworzeniu danych wejściowych do struktury tabeli decyzyjnej. Taka struktura jest typową strukturą wykorzystywaną w analizie eksploracyjnej, a właściwie strukturą odpowiednią dla większości algorytmów uczenia maszynowego, które są wykorzystywane w eksploracji danych. W niniejszej pracy dla przekształcenia wiadomości e-mail do struktury tabeli decyzyjnej zaproponowano dedykowane podejście, które w szczególności odnosi się do struktury wiadomości e-mail w zbiorze danych poddanych analizie, tj. Enron E-mail. Trudno tu jednak mówić o rozwiązaniu uniwersalnym, niezależnym od danych poddawanych analizie.

Wartościowym składnikiem dla pracy jest zaproponowanie rozwiązania do predykcji struktury folderów oraz sugerowania użytkownikom potencjalnych folderów w ich „inboxie”. Rozwiązanie to zaproponowano jako rozszerzenie algorytmu mrowiskowego, przeznaczonego do kategoryzacji wiadomości, o mechanizm oceny struktury folderów poszczególnych grup użytkowników. W konsekwencji, wynik tej oceny pozwala algorytmowi na sugerowanie nowych folderów użytkownikowi i takich, w których wiadomości przechowują pozostali użytkownicy grupy.

Zatem do oryginalnych i najważniejszych osiągnięć rozprawy zaliczam:

- zaproponowanie algorytmu do automatycznej kategoryzacji wiadomości e-mail opartego na algorytmie mrowiskowym,
- zaproponowanie rozwiązania do analizy struktury sieci powiązań pomiędzy wiadomościami e-mail w zbiorze takich wiadomości, a tym samym analizy struktury sieci społecznej,
- zaproponowanie nowego podejścia do predykcji folderów użytkownika na podstawie struktury folderów innych użytkowników wyznaczonych przez sieć społeczną, które to rozwiązanie ma wartość nie tylko poznawczą, ale również praktyczną.

Niemniej jednak w pracy można wskazać pewne niedopowiedzenia, których brak zapewne podwyższałby wartość merytoryczną rozprawy doktorskiej.

Szkoda, że sam problem podjęty w pracy nie został zdefiniowany w sposób formalny. W rozprawie Doktorantka podejmuje omówienia problemu dopiero w rozdziale 4, ale tylko w sposób opisowy.

Niewątpliwie, pierwsze rozdziały pracy nie wnoszą istotnych informacji dla rozprawy, jedynie wprowadzają pewne wybrane definicje i odwołują się do wybranych technik eksploracji danych oraz algorytmów i metod. Dla przykładu, dość szeroko przedstawiona jest „opisowa analiza danych” – podrozdział 1.1, chociaż nie wnosi ona nic zasadniczego dla pracy, a opisane modele nie zostały wykorzystane w dalszej części pracy. Podobnie jest z rozdziałem dedykowanym modelom w analizie danych oraz opisem metod eksploracyjnych, które nie są potem wykorzystywane w dalszej pracy. Obszerny podrozdział poświęcony klasyfikacji mógłby być przez Doktorantkę przygotowany bardziej starannie pod kątem terminologicznym oraz opisu procesu budowania klasyfikatorów, w tym omówienia znaczenia danych uczących i testowych. Tego istotnego elementu, który zapewne mógłby pomóc w rozumieniu procesu budowania klasyfikatora, w pracy niestety zabrakło. W rozdziale tym niepotrzebnie poświęcono uwagę różnym algorytmom klasyfikacji, nie do końca wyczerpując opisy związane z ich działaniem, które zresztą nie wszystkie zostały użyte w dalszej pracy. Tym samym pominięto inne algorytmy, które wykorzystano na etapie eksperymentów obliczeniowych.

Ponadto należy zauważyć, że Doktorantka dość szorstko obchodzi się z niektórymi ugruntowanymi terminami. Należy tu wskazać chociażby na fakt przypisania eksploracji danych do jednej z dziedzin informatyki lub nazywania jej nauką. Chciałbym również podkreślić, że w polskim nazewnictwie powinniśmy omijać słów: „klastrowanie”, czy „klaster”, a raczej używać w zamian słów „grupowanie” i „grupa”. Podobnie, czy drzewa decyzyjne są ważnym narzędziem w uczeniu maszynowym? Raczej są jednym z narzędzi uczenia maszynowego, równie ważnym jak inne. Z drugiej strony, to nie z uwagi na swoją prostotę i popularność są „stosowane do dnia dzisiejszego”, lecz z uwagi na ich efektywność rozwiązywania problemów klasyfikacji. Podobnie Doktorantka w rozdziale 1 stwierdza, że „klasyfikację realizuje się przy użyciu drzew decyzyjnych, które buduje się na podstawie danych uczących”. W tym samym rozdziale podkreśla też, że celem klasyfikacji jest predykcja wartości danego atrybutu w oparciu o zadany zbiór treningowych. Niestety Doktorantka kilkakrotnie w pracy stosuje skróty myślowe i stosuje kolokwializm.

Szkoda, że Doktorantka badania oparła tylko na jednym wybranym zbiorze danych, tj. Enron E-mail. Znane i dostępne są przecież inne zbiory danych, które wykorzystywane są jako benchmarki

w problemie kategoryzacji wiadomości e-mail, takie jak np: Kaggle Dataset, SRI Email Dataset, czy EuAll Dataset. Zresztą ten ostatni zbiór danych był też wykorzystywany przez R. Bekkermana i innych (w spisie literatury Doktorantka umieściła tę pracę pod numerem 10).

Przeprowadzenie eksperymentów w oparciu o większą liczbę różnych zbiorów danych, pozwoliłoby na sformułowanie bardziej ogólnych wniosków oraz na potwierdzenie, że proponowane podejście jest efektywne niezależnie od danych i ich źródła.

Autorka rozprawy dość szczegółowo opisuje wykorzystane środowisko programowe RSES, jak ma to miejsce we wprowadzeniu do rozdziału 5, choć nie uzasadnia wyboru tego środowiska. W rozdziale 6.1, gdzie wspomniano o wykorzystaniu środowiska WEKA, nie wskazano jakie konkretnie algorytmy zostały użyte. Analizując wyniki z tabeli 6.4 należy zatem przyjąć, że wszystkie eksperymenty, poza algorytmem mrowiskowym, zostały przeprowadzone właśnie w oparciu o algorytmy środowiska WEKA.

Szczegółowych informacji brakuje również przy opisie eksperymentów. Nie do końca wiadomo jaki przyjęto plan eksperymentów. Czy do oceny algorytmów klasyfikacji wykorzystano metodę k-krotnej walidacji krzyżowej, czy inne podejście? Wielkość zbiorów uczących i testowych w przeprowadzonych eksperymentach nie została wskazana przez Doktorantkę. Brak takich informacji oraz innych informacji, np. związanych z doбором i ustaleniem parametrów dla algorytmu genetycznego, bądź mrowiskowego, może ograniczać powtarzalności eksperymentów obliczeniowych. W pracy posłużono się jedynie stwierdzeniem o standardowych ustawieniach dla algorytmów (jak ma to miejsce np. na stronie 75 pracy). A przecież dobór parametrów dla wspomnianych algorytmów ewolucyjnych może mieć istotny wpływ na ich zbieżność. Nie podano również wartości przyjętych parametrów dla innych algorytmów, w tym dla algorytmów indukcji klasyfikatorów bazowych oraz klasyfikatorów opartych na zespołach.

W pracy w sposób nieszablonowy podano dokładność klasyfikacji jako wartość dziesiętną, gdy zwyczajowo podaje się ją w ujęciu procentowym. Poza tym Doktorantka raz stosuje określenie dokładność klasyfikacji, w innym poprawność dokładności klasyfikacji, a jeszcze w innym miejscu trafność przypisania wiadomości do folderów. Szkoda przy tym, że nie zdefiniowano „dokładności klasyfikacji”. Szkoda również, że nie podano w oparciu o jakie dane jakość klasyfikacji została wyznaczona, czy w oparciu o dane uczące, czy w oparciu o dane testowe. Domniemać należy, że w oparciu o dane testowe. Szkoda również, że podając wyniki eksperymentów, nie podano odchylenia standardowego dla tych wyników. Pozwoliłoby to na szerszą ocenę rozkładu wyników, oczywiście w przypadku algorytmów niedeterministycznych.

Wyniki badań przedstawione w rozprawie zostały opublikowane w materiałach konferencyjnych wydanych w serii *Lecture Notes in Computer Science* wydawnictwa Springer-Verlag oraz były prezentowane na międzynarodowych konferencjach naukowych.

Recenzowana rozprawa doktorska nasuwa również następujące pytania:

- Doktorantka ocenę proponowanych algorytmów oparła jedynie na wskaźniku dokładności klasyfikacji. Jakie inne mierniki i wskaźniki można by wykorzystać w ocenie proponowanych podejść?
- Interesującym jest ocena proponowanych metod pod kątem ich złożoności obliczeniowej. Jaka jest złożoność algorytmu aACDT oraz algorytmu automatycznej kategoryzacji wiadomości elektronicznych opartego na algorytmie mrowiskowym oraz sieci społecznej i jak się ona ma do złożoności obliczeniowej algorytmu ACDT?

- Na ile należałoby zmienić podejście do rozwiązania problemu „Email Foldering Problem”, gdyby był on definiowany jako problem wieloetykietowej klasyfikacji wiadomości e-mail¹ (ang.: *multi-label classification*)?
- Na ile zaproponowane algorytmy są algorytmami w łatwy sposób dającymi się przenieść na grunt obliczeń równoległych?

Na powyższe pytania wnioskuje o odpowiedzi podczas publicznej obrony rozprawy doktorskiej, jeśli zostanie ona przyjęta do dalszego procedowania.

4. Ocena struktury i redakcji rozprawy

Można przyjąć, że organizacja pracy jest przejrzysta. Niemniej jednak w przyjętym układzie pracy w pierwszych rozdziałach Doktorantka przywołuje terminologię i zagadnienia dotyczące eksploracji danych, następnie omawia wybrane algorytmy odkrywania wiedzy z danych, po czym przechodzi do charakterystyki algorytmów indukcji drzew decyzyjnych (rozdział 2) i omówienia zagadnień związanych z sieciami społecznymi (rozdział 3). W takim układzie trudno jest rozpoczynając czytanie pracy wnioskować o jej głównych przesłankach. Oczywiście cele i teza pracy zostały zawarte we wstępie, ale bardziej czytelnym byłoby, aby także w pierwszych rozdziałach przedstawić tło problemu, następnie sformułować go, po czym omówić znane podejścia do rozwiązania stawianego problemu. W recenzowanym opracowaniu o problemie Doktorantka pisze dopiero w podrozdziale 4.2, któremu nadano tytuł „Przypisanie wiadomości do folderów”. Szkoda, że nie nazwano go po prostu „Zdefiniowanie problemu badawczego” lub „Sformułowanie problemu badawczego”.

Wątpliwości nie budzi strona językowa rozprawy, aczkolwiek w pracy występują błędy gramatyczne i językowe, jednak nie są one krytyczne dla opracowania. Sam sposób przygotowania graficznego oraz edytorskiego rozprawy nie budzi uwag.

W pracy występują tabele, które rozciągają się na kilka stron, w tym tabela 4.1 i tabela 4.2. Podobnie jest z rysunkami 9.6-9.13. Takie wielostronicowe zestawienia tabelaryczne lub graficzne mogłyby stanowić załączniki do pracy, co dałoby większą przejrzystość pracy i prezentację wyników.

Ponadto Autorka rozprawy niejednokrotnie przywołuje różne algorytmy podając ich akronimy bez ich wprowadzenia, jak ma to miejsce np. dla PERC, OLMAN, SNA, C-SVC i kilku innych. Poza tym, w szczególności przy prezentacji wyników, Doktorantka stosuje oznaczenie algorytmów, które nie zostały wprowadzone lub zostały wprowadzone z innymi oznaczeniami. Tak jest np. w Tabeli 5.2, Tabeli 6.1 oraz na Rysunku 5.1, gdzie wskazuje się na algorytm Winnow, oraz Wide-margin Winnow oraz WMW, mając na myśli ten sam algorytm.

Pewnym mankamentem pracy jest również brak konsekwencji w stosowaniu innych oznaczeń. Dla przykładu, dość kluczowe było zdefiniowanie odznaczeń dla opisu danych, gdy w pracy nie do końca jest jasne, czym jest n , czy jest długością ciągu, czy „identyfikatorem pożądanego obiektu”; czy X_n to to samo co X_i ; czy m to liczba klas, czy długość ciągu, czy może liczbą wybranych atrybutów; czym jest k (strona 19); itd. Poza tym nie jasne i niekonsekwentne jest stosowanie zapisów „wektor cech”, „wektor atrybutów”, „wektor wartości atrybutów”. Analogicznie Autorka rozprawy raz pisze o zbiorze A , jako zbiorze reprezentującym wyniki eksperymentu, w innym miejscu jako o zbiorze wartości

¹ Problem wieloetykietowej kategoryzacji wiadomości e-mail był między innymi tłem eksperymentów obliczeniowych opisanych w pracy:

Li Li, Houfeng Wang, Xu Sun, Baobao Chang, Shi Zhao, Lei Sha, Multi-label Text Categorization with Joint Learning Predictions-as-Features Method. Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, pages 835–839, Lisbon, Portugal, 17-21 September 2015.

atrybutów i jeszcze w innym miejscu jako o zbiorze atrybutów. Niewątpliwie problem niejednoznaczności i niekonsekwentnego stosowania oznaczeń mógłby być wykluczony, gdyby Doktorantka swoją pracę wyposażyla w spis oznaczeń i skrótów. Takie rozwiązanie pomogłoby czytelnikowi, jak i samej Autorce opracowania.

Doktorantka nie do końca trzymała się także jednolitego formatu zapisu bibliografii, gdyż niektóre z pozycji literaturowych w spisie bibliografii są niekompletne.

Poza wspomnianymi powyżej uwagami, redakcję rozprawy oceniam pozytywnie.

5. Konkluzja

Przedstawione powyżej uwagi mają charakter dyskusyjny. Na pewno zabrakło bardziej precyzyjnego podejścia przy pracowaniu rozprawy. Niemniej jednak w rozprawie można doszukiwać się wartościowych wyników badań. Zawartość przedstawionej rozprawy stanowi nowy wkład merytoryczny do dziedziny wiedzy sztucznej inteligencji, w szczególności dla obszaru zastosowań narzędzi uczenia maszynowego, tj. eksploracji danych.

Doktorantka wykazała się umiejętnością przedstawiania i rozwiązywania problemów naukowych.

Podsumowując, uważam, że rozprawa doktorska pt. *„Automatyczna kategoryzacja wiadomości elektronicznych z zastosowaniem sieci społecznych i algorytmów mrowiskowych”* autorstwa mgr Barbary Probierz, spełnia warunki stawiane rozprawom doktorskim w myśl ustawy o tytule naukowych stopniach naukowych. W konsekwencji wnioskuję o dopuszczenie rozprawy do publicznej obrony.

