

Politechnika Warszawska

Wydział Elektroniki i Technik Informatycznych

DZIEKAN I RADA WYDZIAŁU ELEKTRONIKI I TECHNIK INFORMATYCZNYCH
POLITECHNIKI WARSZAWSKIEJ

zawiadamiają o

PUBLICZNEJ OBRONIE ROZPRAWY DOKTORSKIEJ

mgr. inż. Cezarego Jankowskiego

która odbędzie się w trybie zdalnym w dniu 13 listopada 2020 r. o godzinie 12.00

Tytuł rozprawy doktorskiej: „Metryki oceny jakości klasyfikatorów probabilistycznych w zagadnieniach wieloklasowych”

Promotor – prof. dr hab. inż. Andrzej Jakubiak Wydział Elektroniki i
Technik Informatycznych Politechniki Warszawskiej

recenzenci: dr hab. inż. Andrzej Majkowski Wydział Elektryczny Politechnika
Warszawska

dr hab. inż. Ewa Świercz, prof. uczelni Wydział Elektryczny Politechnika
Białostocka

Na stronie internetowej wydziału www.elka.pw.edu.pl/Wydzial/Rada-Wydzialu/Harmonogram-obron-doktorskich-streszczenia-i-recenzje znajdują się streszczenie rozprawy oraz recenzje, jak również dostęp do tekstu rozprawy umieszczonej w Bazie Wiedzy Politechniki Warszawskiej.

Sposób uczestniczenia w publicznej obronie:

https://teams.microsoft.com/l/meetup-join/19%3ameeting_M2U5MDFkZjUtZTY2Zi00ZDdhLThiZWYtZjVhYzcxZjQ5ODA3%40thread.v2/0?context=%7b%22Tid%22%3a%223b50229c-cd78-4588-9bcf-97b7629e2f0f%22%2c%22Oid%22%3a%22393ff735-7b26-4f34-bdcd-f7ccab968b42%22%7d

Dziekan



prof. dr hab. inż. Michał Malinowski

Mgr inż. Cezary Jankowski

Promotor- prof. dr hab. inż. Andrzej Jakubiak

Tytuł:” Metryki oceny jakości klasyfikatorów probabilistycznych w zagadnieniach wieloklasowych”

Streszczenie:

W rozprawie został przedstawiony autorski projekt nowej rodziny metryk oceny jakości klasyfikatorów opartej na stosunku estymaty prawdopodobieństwa klasy prawdziwej do najwyższej estymaty prawdopodobieństwa spośród pozostałych klas. Zaproponowane miary posiadają właściwości umożliwiające zastosowanie ich w zagadnieniach wieloklasowych. Pozwalają one na porównywanie wyników działania modeli, zbudowanych na podstawie danych zawierających różną liczbę klas decyzyjnych. Dostarczają dodatkowej wiedzy na temat badanych klasyfikatorów probabilistycznych. Rozprawa, oprócz opisu proponowanej rodziny metryk oceny jakości klasyfikatorów, zawiera również założenia i wyniki eksperymentów, mających na celu porównanie przedstawionych miar z dotychczas stosowanymi funkcjami oceny.

Słowa kluczowe: klasyfikacja, metryka, uczenie maszynowe, eksploracja danych

dr hab. inż. Andrzej Majkowski
Instytut Elektrotechniki Teoretycznej
i Systemów Informacyjno-Pomiarowych
Wydział Elektryczny
Politechnika Warszawska

Warszawa, 30.07.2020r.

Recenzja rozprawy doktorskiej

(recenzja opracowana na podstawie uchwały Rady Naukowej Dyscypliny Informatyka Techniczna i Telekomunikacja Politechniki Warszawskiej z dnia 2 czerwca 2020 roku)

Tytuł rozprawy: Metryki oceny jakości klasyfikatorów probabilistycznych w zagadnieniach wieloklasowych

Autor rozprawy: mgr inż. Cezary Jankowski

Promotor: prof. dr hab. inż. Andrzej Jakubiak

1. Zagadnienia naukowe rozpatrywane w pracy

Przedmiotem rozprawy jest zaprojektowanie nowych metryk oceny jakości klasyfikatorów probabilistycznych. Opracowane metryki wykorzystują stosunek estymaty prawdopodobieństwa prawdziwej klasy i najwyższej wartości spośród estymat prawdopodobieństwa pozostałych klas. Są one przystosowane do wykorzystania w zagadnieniach wieloklasowych, chociaż możliwe jest również ich zastosowanie w klasyfikacji binarnej.

Istnieje wiele różnych metod klasyfikacji. Z praktycznym zastosowaniem tych znanych metod wiążą się jednak istotne problemy, których rozwiązanie często okazuje się dość trudne. Nie istnieją żadne uniwersalne procedury postępowania w tym zakresie. Do każdego rozpatrywanego problemu należy zastosować podejście indywidualne. Problem optymalizacji klasyfikatorów jest szczególnie ważny w zagadnieniach wieloklasowych z różną licznością klas. Zastosowanie odpowiedniej metryki oceny jakości klasyfikacji w odniesieniu do różnych klasyfikatorów (lecz tych samych danych testowych) umożliwia ich porównanie, a co za tym idzie – pozwala na wybór najlepszego z nich. Ma to bezpośrednie przełożenie w jakości proponowanych rozwiązań. Właśnie ten problem jest poruszany w recenzowanej rozprawie doktorskiej. Główny nurt rozprawy mieści się w dyscyplinie Informatyka Techniczna i Telekomunikacja, poruszając jednocześnie pewne zagadnienia z dziedziny matematyki, a w szczególności statystyki. Praca łączy w sobie

aspekty teoretyczne (opracowanie nowych metryk) i aplikacyjne (przebadanie zaproponowanych metryk dla różnych klasyfikatorów i dla różnych baz danych).

Rozprawa doktorska mgra Cezarego Jankowskiego jest napisana w języku polskim i liczy 149 stron. Na rozprawę składa się Wstęp, Cel i zakres pracy, 3 rozdziały, Podsumowanie i Bibliografia oraz spisy rysunków i tabel. Praca zilustrowana jest 22 rysunkami, wyniki badań przedstawione są w 31 tabelach.

We wstępie autor zaprezentował problem badawczy, następnie przedstawił cel i zakres pracy oraz omówił strukturę rozprawy doktorskiej. Teza jednak nie została sformułowana w sposób jawny, a szkoda bo mogła by ona podkreślić główny cel badawczy jaki autor przed sobą postawił. Można tu też znaleźć listę skrótów i symboli ułatwiających zrozumienie pracy. Rozdział 3 rozprawy zawiera wprowadzenie teoretyczne. Autor w wyczerpujący sposób opisał problem oceny jakości klasyfikacji. Przedstawione zostały istniejące metryki pozwalające na ocenę klasyfikatorów. W rozdziale 4 zaprezentowane zostały zaproponowane przez autora metryki oceny klasyfikatorów. Omówione zostały właściwości proponowanych metryk oraz przedstawione przykłady, ukazujące potencjalną przewagę rozpatrywanych metryk nad klasycznie wykorzystywanymi. Rozdział 5 opisuje część eksperymentalną pracy. Badania dotyczyły wpływu wybranej metryki na proces klasyfikacji. Przetestowano 8 zaproponowanych, własnych metryk. Wyniki zostały porównane z 4 istniejącymi metrykami. Badania były przeprowadzane w trzech etapach: pierwszy dotyczył porównania algorytmów klasyfikacji, drugi – wyboru algorytmu klasyfikacji oraz wartości optymalnych hiperparametrów klasyfikatorów, a trzeci badań nad przebiegiem zmienności metryk w funkcji wybranych hiperparametrów klasyfikatorów. Podsumowanie oraz wnioski zostały zawarte w rozdziale 6.

2. Analiza źródeł

Analizę źródeł autor zamieścił głównie we wstępie (podstawowe definicje, wiedza historyczna) oraz w rozdziale 3 (problemy związane z klasyfikacją i metodami oceny klasyfikatorów). Bibliografia zamieszczona w pracy doktorskiej jest bardzo obszerna i zawiera 155 pozycji literaturowych. Stanowi ją bogata lista monografii oraz artykułów z czasopism i konferencji krajowych i zagranicznych. Wśród źródeł można znaleźć również te najnowsze, z ostatnich lat (3 pozycje z 2019 roku, 6 – z 2018, 6 – z 2017). Świadczy to o solidnym zapoznaniu się autora z aktualnym stanem wiedzy. Wnioski wyciągnięte z analizy źródeł zostały dobrze wykorzystane przy projektowaniu nowych metryk i porównaniu z istniejącymi. Literatura nie jest jednak uporządkowana w kolejności cytowań w pracy doktorskiej, a jest uporządkowana alfabetycznie. Utrudnia to analizę literatury podczas czytania pracy. Co więcej wiele zamieszczonych w bibliografii pozycji literatury nie jest cytowanych w ogóle w pracy.

Wśród pozycji literaturowych można znaleźć 6 pozycji autorskich (z lat 2014-2016). Jedna z nich, w której autor pracy doktorskiej jest pierwszym autorem, jest zamieszczona w znaczącym czasopiśmie. Pozycje te nie są jednak bezpośrednio związane z tematyką rozprawy doktorskiej.

3. Metody rozwiązania postawionego zagadnienia

Zagadnieniem jakie autor postawił przed sobą było opracowanie nowych metryk do oceny klasyfikatorów szczególnie w zagadnieniach wieloklasowych. Zaproponowane metryki zostały zaprojektowane tak, aby unikać wad istniejących metryk, stosowanych dotychczas w zagadnieniach wieloklasowych. Cel ten został osiągnięty, co autor udokumentował w części eksperymentalnej pracy doktorskiej. Zaproponowane metryki mogą być również wykorzystane w zadaniach klasyfikacji dotyczących zbiorów danych zawierających różną liczbę klas decyzyjnych.

Bardzo istotną częścią pracy doktorskiej było przebadanie zaproponowanych metryk i podanie w jakich warunkach je stosować. Należy w tym miejscu wspomnieć, że nie istnieje jednoznaczny sposób porównania metryk. Istotne jest zatem zrozumienie, co dokładnie określa dana miara i w jakich sytuacjach należy ją wykorzystywać. W części eksperymentalnej pracy autor próbował właśnie przedstawić te zagadnienia. Autor nie podał jednak wyraźnych wskazówek w jakich przypadkach zaproponowane metryki wnoszą dodatkową wiedzę na temat procesu klasyfikacji. Byłoby to bardzo pomocne przy wykorzystaniu zaproponowanych metryk w praktyce. Bardzo pomocny byłby też algorytm umożliwiający bezpośrednie porównanie metryk, który też nie został przedstawiony przez autora pracy.

4. Oryginalność rozprawy

Samodzielny i oryginalny dorobek autora stanowi zaprojektowanie i przebadanie nowych metryk oceny jakości klasyfikatorów probabilistycznych w zagadnieniach wieloklasowych. Opracowane metryki wykorzystują wiedzę zawartą w rozkładzie estymat prawdopodobieństwa pomiędzy klasami, bazując na stosunku estymaty prawdopodobieństwa prawdziwej klasy i najwyższej wartości spośród estymat prawdopodobieństwa pozostałych klas. Dotychczas takie miary nie zostały opracowane i zbadane. Autor przeprowadził również szereg badań, które pozwalają stwierdzić, że proponowane metryki wnoszą pewną dodatkową wiedzę na temat klasyfikatorów. Badania te są zaprezentowane w rozdziale 5. Wydaje się, że spośród proponowanych metryk, znormalizowana miara THFR oraz log THF loss są najbardziej wartościowe w analizie modeli klasyfikacji.

5. Poprawność przedstawienia uzyskanych wyników

Struktura pracy jest poprawna. Problem naukowy został poprawnie sformułowany, aczkolwiek jawne postawienie tezy według mnie byłoby wskazane. Prawdziwość przyjętych założeń została potwierdzona badaniami eksperymentalnymi, których metodologia nie budzi wątpliwości. Praca jest zredagowana dość starannie, jednak w tekście można znaleźć pewne usterki edytorskie.

Uwagi o charakterze ogólnym:

W pracy wielokrotnie jest użyty termin „model” bez wyjaśnienia, do jakiego modelu w danym miejscu autor się odnosi. Zmusza to czytelnika do zastanawiania się, co autor miał na myśli.

Do terminów anglojęzycznych podawanych w nawiasach, np. (*gradient boosting*), powinno być dołączane słowo *ang.*

W podtytułach (np. str. 43: 3.4.3. *Walidacja w procesie wyboru modelu*, str. 55: 4.2.2 *Charakterystyka*, str. 58: 4.3 *Implementacja*) należałoby dodać jakiego modelu, charakterystyki, czy implementacji czego te podrozdziały dotyczą. Ułatwia to zorientowanie się w zawartości danego podrozdziału.

Uwagi o charakterze szczegółowym:

str. 13: ... *stosowane w literaturze jest również określenie „metrykę oceny jakości klasyfikacji”* – powinno być: metryka

str. 15: ... *wrażliwe na rozkład estymat prawdopodobieństw pomiędzy klasy.* – powinno być: klasami

str. 25: *będącego wynikiem predykcji klasyfikator probabilistyczny* – powinno być: klasyfikatora probabilistycznego

str. 28: błąd we wzorze (3.12) – powinno być: $NPV = TN / (TN + FN)$

str. 29: *Przez $P(pred = k / obs = l)$ oznaczany będzie stosunek liczby instancji, dla których predykcja była równa k i prawdziwa (to jest – obserwowana) klasa jest równa l do liczby instancji, dla których klasa decyzyjna jest równa l .* – zdanie niezgrabnie sformułowane i trudne do zrozumienia

str. 38: *Trzeba zwrócić uwagę na to, że testowania klasyfikatora za pomocą tych samych danych, które uprzednio posłużyły do jego zbudowania* – powinno być: testowanie

str. 41: *Przyjmowane jest założenie, że jakość modelu predykcyjna zbudowana na podstawie nauki...* – powinno być: modelu predykcyjnego zbudowanego

str. 49: we wzorze (4.3) brakuje logarytmowania

str. 61: *Przeprowadzono szereg eksperymentów, w których zbadano wpływ wyboru metryki na wyniki oceny jakości modelu.* – tu nie bardzo wiadomo o jaki model chodzi, należało by to uściślić. Jest to co prawda wyjaśnione później, ale należy przy pierwszym użyciu wyjaśnić o co chodzi.

Str. 65: *Porównywanie wartości różnych miar oceny jakości klasyfikacji nie w każdym przypadku może zostać uznana za konstruktywną* – powinno być: uznane za konstruktywne

[81] C. Jankowski, G. Borowik, K. Kowalski, C. Jankowski, G. Borowik i K. Kowalski, „Dyskretyzacja danych numerycznych metodami przekształceń boolowskich”, *Przegląd*

Telekomunikacyjny + Wiadomości Telekomunikacyjne, numer 10, strony 1334–1342, 2014.
– dwukrotnie są podane nazwiska autorów

[78] —, „Metryki oceny jakości klasyfikacji”, *Przegląd Telekomunikacyjny + Wiadomości Telekomunikacyjne, numer 4, strony 92–96, 2018.* – w wielu pozycjach literatury występuje kraska w miejscu autorów. Wydaje się to być niedopatrzeniem.

6. Słabe strony pracy i jej główne wady

W rozprawie doktorskiej autor nie przedstawił, szczególnie w kontekście przeglądu literatury, z czego wynika potrzeba opracowania nowych metryk oceny jakości klasyfikatorów. Od lat stosowane są z powodzeniem już istniejące metryki, również do zagadnień wieloklasowych. Również szerszego wyjaśnienia wymaga założenie wzięcia pod uwagę jednej klasy „fałszywej”. Autor nie przedstawił też w jasny, bezpośredni sposób jaką nową wiedzę można wydobyć o procesie klasyfikacji stosując zaproponowane przez siebie metryki i w jakich przypadkach uzyskujemy korzyści przez ich zastosowanie. Brak jest również porównania z wynikami innych badaczy, dotyczącymi tej dziedziny.

W części eksperymentalnej autor przeprowadził bardzo dużą liczbę eksperymentów, wymagającą dużego nakładu pracy oraz rozległej wiedzy o procesie klasyfikacji. Brak jest jednak podsumowania badań, wskazówek co do przydatności zaproponowanych metryk w konkretnych zadaniach klasyfikacji. W opisywanym powyżej kontekście wydaje się, że niezwykle przydatne byłoby opracowanie sposobu porównywania ze sobą metryk. W niniejszej rozprawie zaproponowane zostało jedynie podejście, oparte na wartościach rankingowych, wyznaczanych przez poszczególne metryki. Szczególnie przydatny byłby algorytm wyboru modelu klasyfikacji (klasyfikatora i jego hiperparametrów) na podstawie wyników więcej niż jednej metryki. Takiego algorytmu brakuje w pracy.

7. Przydatność rozprawy dla nauk technicznych

Zaproponowana rodzina metryk oceny jakości klasyfikatorów probabilistycznych może znaleźć zastosowanie w zagadnieniach, pochodzących z różnych dziedzin, np. telekomunikacji oraz teleinformatyki. Ich wykorzystanie jest szczególnie wartościowe w przypadkach wieloklasowych, takich jak na przykład automatyczne klasyfikowanie ruchu sieciowego, wykrywanie niepożądanych zachowań użytkowników w sieci czy wyznaczanie segmentów rynku w celu lepszego dostosowania oferty. Zaprojektowane metryki mogą być wykorzystane jako funkcja celu w procesie wyboru modelu predykcyjnego, a także do jego ostatecznej ewaluacji. Pozwalają one na lepsze zrozumienie wyników klasyfikacji. Umożliwiają pełniejsze odkrywanie wiedzy z danych.

8. Podsumowanie

Podsumowując stwierdzam, że tematyka rozprawy jest aktualna i ważna, autor rozwiązał postawione sobie problemy naukowe, a rozprawa świadczy o znacznej wiedzy autora i znajomości literatury z zakresu optymalizacji doboru klasyfikatorów. Uważam, że rozprawa doktorska mgra Cezarego Jankowskiego spełnia wymagania stawiane rozprawom doktorskim przez obowiązującą ustawę o stopniach naukowych i tytule naukowym oraz stopniach i tytule w zakresie sztuki z dnia 14 marca 2003 roku (z późniejszymi zmianami), jak i ustawę Prawo o szkolnictwie wyższym i nauce z dnia 3 lipca 2018 roku. Wnioskuje o dopuszczenie mgra Cezarego Jankowskiego do publicznej obrony rozprawy doktorskiej.

Andrzej Majkowski
podpis

Białystok, 5 sierpnia 2020 r.

Dr hab. inż. Ewa Świercz, prof. PB
Katedra Fotoniki, Elektroniki i Techniki Świetlnej
Wydział Elektryczny Politechniki Białostockiej
ul. Wiejska 45D, 15-351 Białystok

KWESTIONARIUSZ-RECENZJA ROZPRAWY DOKTORSKIEJ DLA RADY DYSCYPLINY
Informatyka Techniczna i Telekomunikacja

Tytuł rozprawy:

Metryki oceny jakości klasyfikatorów probabilistycznych w zagadnieniach wieloklasowych

Autor rozprawy:

Cezary Jankowski

1. **Jakie zagadnienie naukowe jest rozpatrzone w pracy/teza rozprawy/ i czy zostało ono dostatecznie jasno sformułowane przez autora? Jaki charakter ma rozprawa (teoretyczny, doświadczalny, inny)**

Jedną z głównych trudności w problemach klasyfikacyjnych jest prawidłowa ocena działania klasyfikatora. Rozprawa doktorska dotyczy problemu podejmowania decyzji przez klasyfikator (gdy dane należą do predefiniowanych klas) na podstawie oceny odpowiednio określonej miary przynależności. Celem niniejszej rozprawy jest zaprojektowanie rodziny metryk oceny jakości klasyfikatorów probabilistycznych, analiza tych metryk i porównanie rezultatów klasyfikacji z zaprojektowanymi metrykami. Zaprojektowane metryki wykorzystują stosunek estymaty prawdopodobieństwa prawdziwej klasy i najwyższej wartości spośród estymat prawdopodobieństwa pozostałych klas. Metryki budowane z wykorzystaniem koncepcji stosunku estymat prawdopodobieństw prawdziwej klasy i najwyższej wartości spośród estymat prawdopodobieństwa pozostałych klas nie były dotychczas zdefiniowane i badane. Takie metryki mogą być wykorzystane zarówno do oceny klasyfikacji binarnej, jak i klasyfikacji wieloklasowej. Zagadnienie naukowe wiąże się z udoskonaleniem metod oceny klasyfikacji i udowodnieniem, że wprowadzone metryki uwzględniające informacje o estymatach rozkładu prawdopodobieństwa zarówno w klasie prawdziwej, jak i w klasach pozostałych, zwiększają wiedzę o samym procesie klasyfikacji. Recenzent stwierdza, że cel rozprawy został przedstawiony jasno i precyzyjnie.

Doktorant zaplanował osiągnięcie przedstawionego wyżej głównego celu poprzez realizację celów cząstkowych:

- zdefiniowanie oryginalnych wzorów analitycznych opisujących proponowaną miarę,
- wybranie środowiska numerycznego Scikit-learn z bibliotekami do implementacji wybranych algorytmów klasyfikacji z zaproponowanymi miarami oceny tej klasyfikacji,

- zbadanie właściwości nowych miar klasyfikacji zastosowanych w grupie reprezentatywnych klasyfikatorów na wybranych bazach danych,
- porównanie z istniejącymi klasycznymi miarami wykorzystywanymi w zagadnieniach klasyfikacji.

Recenzowaną pracę można zaliczyć do kategorii prac teoretyczno-symulacyjnych.

2. Czy w rozprawie przeprowadzono w sposób właściwy analizę źródeł/ w tym literatury światowej, stanu wiedzy i zastosowań w przemyśle/ świadczący o dostatecznej wiedzy autora. Czy wnioski sformułowano w sposób jasny i przekonujący?

Recenzowana rozprawa zawiera 155 pozycji literatury z zakresu tematyki realizowanej w pracy doktorskiej. Liczba prac z dziedziny rozprawy jest bardzo duża, a przedstawioną literaturę recenzent uważa za reprezentatywną w odniesieniu do zagadnień realizowanych w ramach doktoratu. Analiza publikacji jest właściwa. Wnioski z przeglądu literatury są formułowane w sposób przekonujący i wskazują na umiejętność korzystania z odpowiednich źródeł literaturowych.

Dorobek Autora liczy 9 publikacji, w tym jedną samodzielną. Jedna z publikacji ukazała się w czasopiśmie znajdującym się na liście A wykazu czasopism MNiSW. Przytoczone pozycje literaturowe wskazują na dużą wiedzę Autora i duże rozeznanie aktualnego stanu badań w dziedzinie, w której jest ulokowana rozprawa doktorska.

Czy autor rozwiązał postawione zagadnienia, czy użył właściwej do tego metody i czy przyjęte założenia są uzasadnione?

Celem rozprawy jest ocena jakości klasyfikatorów probabilistycznych w zagadnieniach wieloklasowych z wykorzystaniem zaproponowanych metryk. Autor wyszedł z założenia, że metryki bazujące jedynie na estymacie prawdopodobieństwa klasy prawdziwej nie w pełni pozwalają na ocenę klasyfikacji wieloklasowej. Proponowane przez Autora metryki uwzględniają dodatkowo relacje estymat prawdopodobieństwa prawdziwej klasy do najwyższej wartości estymat prawdopodobieństwa spośród pozostałych klas. Do implementacji miar wykorzystano bibliotekę uczenia maszynowego Scikit-learn. Starannie przeprowadzone eksperymenty symulacyjne potwierdzają, że zaproponowane metryki wnoszą dodatkową wiedzę, dotyczącą jakości przeprowadzonej klasyfikacji probabilistycznej. Tak postawiony problem został właściwie zbadany w rozległym scenariuszu badań symulacyjnych z analizą ograniczeń numerycznych.

4. Na czym polega oryginalność rozprawy, co stanowi samodzielny i oryginalny dorobek autora, jaka jest pozycja rozprawy w stosunku do stanu wiedzy czy poziomu techniki reprezentowanych przez literaturę światową?

Oryginalny dorobek Autora przedstawiony jest w liście zagadnień:

- Zaprojektowanie rodziny metryk klasyfikacji przy uwzględnieniu stosunku wartości estymaty prawdopodobieństwa klasy prawdziwej do najwyższej estymaty prawdopodobieństwa pozostałych klas.
- Zbadanie i krytyczna analiza reprezentatywnych modeli klasyfikacji w wybranych

bazach danych, z weryfikacją zaproponowanych miar probabilistycznych i miar klasycznych.

Zdaniem recenzenta, nie ma żadnych wątpliwości, które elementy rozprawy są oryginalnym i wyłącznym dorobkiem Doktoranta. Tematyka rozprawy doktorskiej jest ściśle skorelowana z zagadnieniami intensywnie rozwijanymi w uczeniu maszynowym czy też z zagadnieniami analizowanymi w eksploracji danych. Uczeniem maszynowym i eksploracją danych zajmuje się wiele światowych ośrodków naukowych, a dokonania Doktoranta stanowią niewątpliwie pewien wkład w rozwój tych obszarów wiedzy.

5. Czy autor wykazał umiejętność poprawnego i przekonującego przedstawienia uzyskanych przez siebie wyników /zwięzłość, jasność, poprawność redakcyjna rozprawy/

Rozprawa jest na ogół zredagowana starannie. Autor rozprawy przedstawia wyniki swoich prac w sposób poprawny i przekonujący. Zamieszczone wzory, tabele i rysunki są czytelne i ułatwiają ocenę wielowariantowych wyników badań.

Rozprawa ma wyraźny podział na część wprowadzającą, część teoretyczną dotyczącą budowy metryk, część doświadczalną ze szczegółowym podsumowaniem kolejnych eksperymentów klasyfikacji oraz zbiorcze podsumowanie, w którym podkreślono cel pracy i posumowano wyniki. Rozprawa zawiera ponadto okazały zbiór pozycji literaturowych. Część teoretyczna ma charakter przeglądu i syntezy wiadomości pozyskanych z literatury i została przez Doktoranta opracowana na zadowalającym poziomie, co dobrze świadczy o jego znajomości aktualnego stanu wiedzy w obszarze badawczym stanowiącym zakres recenzowanej rozprawy. Dodatkowo w rozprawie zamieszczono Spis rysunków, Spis tabel oraz dwa dodatki. Dodatek A zawiera kod implementacji metryk w języku Python. Dodatek B sprawia wrażenie uzupełnienia rozdziału piątego, w którym pominięto eksperymenty w bazach wskazanych w Tabeli 5.1.

Szczególnie interesujący okazał się rozdział piąty o tytule – Eksperymenty, w którym Autor przedstawił szczegółowe wyniki wszechstronnych testów symulacyjnych, jednocześnie dokonując wielokryterialnych porównań algorytmów klasyfikacji z opracowanymi metrykami oceny.

Podsumowując, założenia badań i uzyskane wyniki tych badań Autor przedstawia w sposób przekonujący, a układ treści poza dodatkiem B, który powinien być umieszczony w rozdziale piątym, jest poprawny.

Recenzent przedstawił krótką listę swoich uwag krytycznych w następnym punkcie recenzji podkreślając jednocześnie, że nie wpływają one na dobrą merytoryczną ocenę rozprawy.

6. Jakie są słabe strony rozprawy i jej główne wady?

W wyniku krytycznej analizy tekstu rozprawy zostały dostrzeżone nieścisłości i pewna liczba błędów redakcyjnych, które prowadzą do pytań i uwag. Uwagi i pytania nie podważają waloru naukowego rozprawy.

Str. 52 - Dotyczy tekstu: *Kolejna idea jest wprowadzenie nieliniowości za pomocą logarytmu.* Jakiej zmiany oczekiwano w stosunku do metryki definiowanej wzorem 4.4 wprowadzając nieliniowość za pomocą funkcji logarytmicznej we wzorze 4.7?

Str. 53 - Jakiej korzyści oczekiwano wprowadzając miary z odwrotnym stosunkiem prawdopodobieństw tzn. $p_i(HF)/p_i(T)$, na przykład we wzorze 4.10?

Str. 54 - Nie jest jasny opis iloczynu kartezjańskiego, w wyniku którego mamy trójkę cech. Jakie cechy otrzymano i co z nich wynika, ponieważ w dalszej części rozprawy te cechy nie są analizowane i wykorzystane?

Str. 63 - Dotyczy Tabeli 5.1. Podana jest liczba atrybutów dla badanych baz, ale nie podano nazw tych atrybutów. Nie jest również wyjaśniony algorytm ich wyboru.

Str. 64 - Dotyczy Tabeli 5.2. Pomimo, że w tytule rozprawy występuje nazwa *Metryki oceny jakości klasyfikatorów probabilistycznych w zagadnieniach wieloklasowych*, Recenzent nie wszystkie klasyfikatory z Tabeli 5.2 zaliczyłby do klasyfikatorów probabilistycznych, ponieważ nie wykorzystują one bezpośrednio prawdopodobieństw. Na przykład drzewa decyzyjne wykorzystują zysk informacyjny do budowy klasyfikatora, a klasyfikator SVM wyznacza optymalny margines pomiędzy klasami. Proszę uzasadnić wybór klasyfikatorów z budową drzew i klasyfikatorów SVM do oceny metryk w klasyfikacji probabilistycznej.

Str. 67 - Dotyczy tekstu: *Natomiast algorytm AdaBoost uzyskał w przybliżeniu czterokrotnie gorszy wynik, jeśli za kryterium zostanie uznana miara log loss. Wskazywać to może na mniejsze skalibrowanie estymat prawdopodobieństw przynależności do klas, generowanych przez tę metodę. Jak należy rozumieć termin „skalibrowanie estymat”?*

Str. 68 - Dotyczy tekstu: *Dla metody AdaBoost regresja logistyczna co prawda osiągnęła gorszą dokładność, lecz wynik miary log loss jest ponad dwukrotnie lepszy. Myląca jest konstrukcja tego zdania.*

Str. 80 - Dotyczy Tabeli 5.14 i Tabeli 5.15. Zakres hiperparametrów można było zapisać z pominięciem niestandardowej notacji ze średnikiem, do określenia osiągniętych przez hiperparametry granic.

Recenzent proponuje, aby Autor odniósł się do zamieszczonych uwag i zauważonych nieścisłości w tekście rozprawy.

Recenzent proponuje również, aby podczas publicznej obrony rozprawy Autor przedstawił bardziej precyzyjnie algorytmy: las losowy i wzmocnienie gradientowe, ponieważ w wyniku badań Doktoranta ostatecznie te dwa algorytmy okazały się najlepsze w rankingu klasyfikatorów. Proszę podsumować rolę proponowanych metryk dla tych dwóch algorytmów.

7. Jaka jest przydatność rozprawy dla nauk technicznych

Olbrzymi wzrost ilości danych przechowywanych w bazach danych, kontenerach danych i innych skarbnicach danych spowodowało rozwój teorii i praktycznych rozwiązań pozyskania informacji i wiedzy z tych danych. Odkryta wiedza może być wykorzystana do klasyfikacji i podejmowania decyzji. Odkrywanie reguł klasyfikacji jest jedną z technik uczenia maszynowego, a ocena jakości tej klasyfikacji jest bardzo istotna. Metody uczenia maszynowego są powszechnie stosowane w inżynierii i naukach ścisłych, począwszy od systemów komputerowych po fizykę. Tematyka rozprawy wpasowuje się w obszar zagadnień uczenia maszynowego i eksploracji danych. Proponowane metryki pozwalają na bardziej wszechstronną ocenę klasyfikacji poprzez uwzględnienie zarówno estymaty klasy prawdziwej, jak i estymaty klas pozostałych. Takie podejście umożliwia pełniejsze

odkrywanie wiedzy z otrzymanych danych. Dlatego też Recenzent uważa, że podjęty przez Autora temat badań naukowych jest bardzo ważny i stanowi wkład w rozwój technik uczenia maszynowego.

8. Do której z następujących kategorii Recenzent zalicza rozprawę:

- a) niespełniająca wymagań,
- b) wymagająca wprowadzenia poprawek i ponownego recenzowania,
- c) spełniająca wymagania,
- d) spełniająca wymagania z wyraźnym nadmiarem,
- e) wybitnie dobra zasługująca na wyróżnienie.

Zadania badawcze, które postawił przed sobą Doktorant, charakteryzują się bardzo dobrym poziomem merytorycznym i zostały w sposób poprawny sformułowane i rozwiązane. Doktorant dobrał właściwe, nowoczesne i zaawansowane metody osiągnięcia zamierzonych celów rozprawy. Biorąc pod uwagę powyższą opinię proponuję zaliczenie rozprawy do kategorii

d) - spełniająca wymagania z wyraźnym nadmiarem.

Zgodnie z zaprezentowaną powyżej oceną merytoryczną rozprawy Pana mgr inż. Cezarego Jankowskiego stwierdzam, że opiniowana praca spełnia warunki i wymagania stawiane rozprawom doktorskim, określone w Artykule 13, ust. 1 Ustawy z dnia 14 marca 2003 r. o stopniach naukowych i tytule naukowym oraz o stopniach i tytule w zakresie sztuki (Dz. U. Nr 65 z 2003 r., poz. 595, z późn. zm.) oraz stosownych przepisach wykonawczych wydanych na podstawie w/w Ustawy.

W związku z Art. 179 ust. 1 Ustawy z dnia 3 lipca 2018 r. Przepisy wprowadzające ustawę – Prawo o szkolnictwie wyższym i nauce (Dz. U. z 2018 r., poz. 1669, z późn. zm.) stawiam zatem wniosek o dopuszczenie rozprawy doktorskiej mgr inż. Cezarego Jankowskiego do publicznej obrony przed Radą Dyscypliny Informatyka Techniczna i Telekomunikacja w Politechnice Warszawskiej.

Ewa Śniełwa