

Prof. dr hab. inż. Bogusław Cyganek  
Wydział Informatyki, Elektroniki i Telekomunikacji  
Instytut Elektroniki  
Akademia Górniczo-Hutnicza im. Stanisława Staszica w Krakowie  
[cyganek@agh.edu.pl](mailto:cyganek@agh.edu.pl)

**Recenzja rozprawy doktorskiej**  
**Pana magistra inżyniera Jędrzeja Biedrzyckiego**  
*Integration of decision trees in geometric space*

**Wstęp**

Recenzja rozprawy doktorskiej Pana mgra inżyniera Jędrzeja Biedrzyckiego zatytułowanej „*Integration of decision trees in geometric space*”, opracowanej w roku 2021 na Wydziale Elektroniki Politechniki Wrocławskiej. Promotorem pracy jest Pan dr hab. inż. Robert Burduk, prof. uczelni PWr. Przygotowanie recenzji zostało wykonane na zlecenie Przewodniczącego Rady Dyscypliny Naukowej Informatyka Techniczna i Telekomunikacja, Pana prof. dr hab. inż. Michała Woźniaka.

**1. Charakterystyka zagadnienia naukowego podjętego w pracy, tezy pracy oraz umiejscowienie podjętego zagadnienia w nauce światowej**

Żyjemy w czasach rewolucji informatycznej charakteryzującej się wytwarzaniem oraz przetwarzaniem coraz większych strumieni danych cyfrowych. Jedną z najczęściej wykonywanych czynności związanych z przetwarzaniem tych wolumenów danych jest ich klasyfikacja, czyli – ogólnie mówiąc – wyszukiwanie i rozpoznawanie ich charakterystycznych cech, co często sprowadza się do określenia tzw. etykiety, czy też przynależności do wcześniej zdefiniowanej klasy. Przykładów mamy mnóstwo, wspomnijmy chociaż wyszukiwanie podobnych tekstów, znajdowanie twarzy w obrazach – do niedawna jeden z trudniejszych problemów widzenia komputerowego, dziś wykonywany w ułamkach sekundy nawet przez prostsze telefony komórkowe – znajdowanie podobnych podobszarów w obrazach, np. do celów stereowizji, rozpoznawanie chorób na podstawie wyników analiz medycznych, oraz wiele innych. Powyższe osiągnięcia, które często kojarzone są z pojęciem rewolucji informacyjnej, stały się możliwe dzięki opracowaniu nowych i przełomowych architektur klasyfikatorów, w szczególności głębokich sieci neuronowych. Co ciekawsze, zostało wykazane że zastosowanie wielu prostszych klasyfikatorów ale operujących w jednym współpracującym zespole, w znakomitej większości przypadków przewyższa dokładnością odpowiedzi, często bardziej skomplikowanych, ale pojedynczych klasyfikatorów.

WPLYNĘŁO  
23.06.2021

RDN-IT/118/2021

Tematyka pracy Pana magistra inżyniera Jędrzeja Biedrzyckiego wpisuje się właśnie w dziedzinę opracowania nowych wersji zespołów współpracujących klasyfikatorów danych (ang. *ensemble learning*), należącej do ogólnie pojętej informatyki. Doktorant co prawda nie zajmuje się najnowszymi architekturami neuronowymi, lecz również bardzo wydajnymi zespołami klasyfikatorów drzewiastych (ang. *tree classifiers*). W szczególności, głównym zagadnieniem naukowym opisanym w rozprawie mgra J. Biedrzyckiego są zespoły klasyfikatorów drzewiastych oraz geometryczne metody określania ich obszarów decyzyjnych, jak również wynikające z tego sposoby określania tzw. obszarów kompetencji klasyfikatorów oraz metod fuzji ich odpowiedzi w procesie decyzyjnym. Przetłumaczona teza pracy została określona następująco<sup>1</sup>:

*Wykorzystanie granic decyzyjnych wytrenowanych drzew decyzyjnych pozwala na zbudowanie klasyfikatora łączonego o większej wartości miary jakości klasyfikacji niż klasyfikator łączony jakim jest las losowy lub głosowanie większościowe korzystające z tego samego zbioru wytrenowanych drzew decyzyjnych.*

Zagadnienie konstrukcji zespołów klasyfikatorów, w szczególności zoptymalizowanych zespołów drzew decyzyjnych, znalazło i znajduje szerokie zainteresowanie wśród badaczy z całego świata. Wynikiem są nowe metody i ciągle przyrastające publikacje z tej dziedziny. Zainteresowanie tą dziedziną wynika między innymi ze wspomnianej już rewolucji w obszarze klasyfikacji bardzo różnych rodzajów danych i, mimo mnóstwa opracowań, żadna pojedyncza metoda nie zaspakaja wszystkich potrzeb. W nurt ten wpisuje się praca badawcza, jak i opracowane metody, zaprezentowane w rozprawie doktorskiej Pana magistra inżyniera Jędrzeja Biedrzyckiego. Dlatego podjętą pracę uważam jako potrzebną i odpowiadającą bieżącym nurtom i potrzebom informatyki, a w szczególności jej gałęziom uczenia maszynowego i sztucznej inteligencji.

Rozprawa doktorska Pana magistra inżyniera Jędrzeja Biedrzyckiego liczy 132 strony i jest napisana po angielsku. Ponieważ recenzja napisana jest po polsku, część opisów metod oraz stwierdzeń zamieszczonych w pracy w przypadku ich opisu, czy też dyskusji w tej recenzji, zostało albo przytoczone w wersji oryginalnej, albo przetłumaczone na język polski. Rozprawa doktorska została podzielona na 5 rozdziałów, uzupełnionych przez obszerny spis literatury dotyczącej tematyki badań. Przedstawiona praca ma charakter teoretyczno-eksperymentalny i plasuje się w dyscyplinie Informatyka Techniczna i Telekomunikacja.

## ***2. Najważniejsze osiągnięcia naukowe przedstawione w rozprawie, ich oryginalność oraz pozycja w stosunku do stanu wiedzy***

Główne zagadnienia badawcze podjęte przez Pana magistra inżyniera Jędrzeja Biedrzyckiego oraz przedstawione w Jego rozprawie doktorskiej dotyczą geometrycznych metod konstruowania zespołów klasyfikatorów z drzewami decyzyjnymi.

W rozprawie Pan mgr inż. Jędrzej Biedrzycki przedstawił 4 autorskie algorytmy konstrukcji zespołów klasyfikatorów wykorzystujących geometrycznie wyznaczone granice decyzyjne. Do najważniejszych osiągnięć, opisanych w rozprawie doktorskiej, zaliczam:

---

<sup>1</sup> Wersja polska została tu podana na podstawie polskiego abstraktu pracy, podczas gdy wersja angielska, przedstawiona we wstępie pracy, została sformułowana następująco: *Utilization of trained decision trees' decision boundaries allows for building an ensemble of classifiers with a greater value of performance quality measure than that of the random forest or majority voting using the same set of trained decision trees.*

1. Metodę ważonego łączenia drzew decyzyjnych w zespół współpracujących klasyfikatorów w statycznej przestrzeni geometrycznej.
2. Metodę integracji drzew decyzyjnych w zespół współpracujących klasyfikatorów z wykorzystaniem dynamicznie wyznaczonych regionów kompetencji klasyfikatorów składowych.
3. Zmodyfikowana metoda dynamicznej integracja drzew decyzyjnych z wykorzystaniem regionów kompetencji.

W pracy została też zaprezentowana czwarta metoda o stosunkowo ograniczonych możliwościach:

4. Metoda integracji drzew decyzyjnych w zespół klasyfikatorów z użyciem odległości od (i) punktów centralnych oraz od (ii) granic decyzyjnych.

Metody 2 oraz 3 są pochodnymi wcześniej opracowanych oraz eksperymentalnie zbadanych metody statycznej 1 i obydwie, tj. 2 oraz 3 mogą być traktowane jako jedna podgrupa. Wspólną cechą wszystkich metod jest użycie zespołów składających się z drzew decyzyjnych. Jedną z najbardziej znanych i reprezentatywnych metod tej klasy jest klasyfikator typu losowy las (ang. *random forest*), wykorzystujący losowanie podzbiorów uczących, a następnie fuzję odpowiedzi z wykorzystaniem metody głosowania większościowego (ang. *majority voting*). Są to metody, które niewątpliwie inspirują wielu naukowców, głównie ze względu na ich wszechstronność i często bardzo wysoką dokładność odpowiedzi, a w przypadku pracy badawczej p. mgra inż. Jędrzeja Biedrzyckiego są to również metody referencyjne, do których Doktorant porównuje swoje propozycje. Na podkreślenie zasługuje tutaj fakt, że metody zaproponowane przez Doktoranta w znakomitej większości przypadków przewyższają jakością wspomniane lasy losowe oraz głosowanie większościowe, jak zostało to wykazane w rozprawie między innymi na podstawie statystycznej analizy wyników klasyfikacji z wykorzystaniem ogólnie znanych zbiorów danych testowych UCI. Niewątpliwie jest to istotne osiągnięcie Doktoranta.

Główną nowością zaproponowaną w statycznej wersji ważonego łączenia drzew decyzyjnych w zespół współpracujących klasyfikatorów – metoda nr 1 – jest wykorzystanie tzw. horyzontalnego partycjonowania zbioru uczącego a następnie wykorzystanie drzew decyzyjnych jako klasyfikatorów bazowych w celu otrzymania regionów decyzyjnych. Przestrzeń cech dzielona jest tu na równe podprzestrzenie z identyczną liczbą podprzedziałów wzdłuż każdej z cech. W kolejnych krokach obliczane są punkty centralne podprzestrzeni klasyfikacyjnych, jak również obliczany jest ich rozmiar i położenie wynikowych regionów decyzyjnych. W końcu obliczane są wagi wykorzystywane w procesie określenia dominujących etykiet we właściwych regionach cech.

Rozszerzeniem metody statycznej jest metoda integracji drzew decyzyjnych w zespół współpracujących klasyfikatorów z wykorzystaniem dynamicznie wyznaczonych regionów kompetencji klasyfikatorów składowych – są to metody druga oraz trzecia na powyższej liście. Główną modyfikacją w porównaniu z wersją statyczną jest partycjonowanie przestrzeni cech wprost na podstawie reguł decyzyjnych każdego składowego klasyfikatora, będącego drzewem decyzyjnym. Po wytrenowaniu klasyfikatorów bazowych następuje obliczenie punktów centralnych każdej z nowo powstałych podprzestrzeni. Punkty centralne są następnie użyte do obliczenia wag, które są z kolei wykorzystane w procesie ważonego tym razem głosowania większościowego. Wagi te obliczane są zgodnie z zaproponowaną funkcją, bazującą między innymi na odległości euklidesowej między punktami podprzestrzeni klasyfikacyjnych.

W najbardziej zaawansowanej i najnowszej metodzie dynamicznej (nr 4 na liście) wprowadzono dodatkowe udoskonalenia polegające na wykorzystaniu do pomiaru odległości pomiędzy podprzestrzeniami średniej punktów treningowych zamiast geometrycznego środka podprzestrzeni. Kolejną modyfikacją jest wprowadzenie minimalnego progu dla podprzestrzeni, która dla każdej klasy wymaga obecności co najmniej 5% obiektów treningowych z tej klasy, w stosunku do liczby wszystkich obiektów w tej podprzestrzeni. W końcu został dodany też mechanizm selektywnego zwiększania

pierścienia przeszukiwania przestrzeni sąsiednich. Metoda ta pozwoliła na osiągnięcie najlepszych wyników wśród metod zaprezentowanych w rozprawie.

Należy wspomnieć, że najważniejsze metody przedstawione w rozprawie zostały opracowane przy współpracy z promotorem, Panem profesorem Robertem Burdukiem. Oprócz najnowszej, metody te zostały również opublikowane w znanych i dobrze cytowanych czasopismach naukowych (IEEE Access oraz MDPI Entropy).

Wszystkie wymienione i zaproponowane przez Doktoranta metody świadczą o dużej wiedzy głównie w dziedzinie zaawansowanych metod klasyfikacji danych, jak również o pomysłowości oraz dojrzałości naukowej. Stanowią one istotny wkład w dziedzinę Informatyka Techniczna i Telekomunikacja, a w szczególności w dziedziny uczenia maszynowego i sztucznej inteligencji. Co więcej, zaproponowane metody mają duży potencjał aplikacyjny, jak również mogą być dalej rozwijane.

### ***3. Słabe strony rozprawy, jej główne wady oraz zagadnienia dyskusyjne***

Tak jak już to zostało wspomniane, podjęte przez Pana mgra inż. Jędrzeja Biedrzyckiego problemy naukowe są bardzo istotne we współczesnej informatyce w dziedzinie uczenia maszynowego. Na pewno będziemy świadkami dalszego i prawdopodobnie bardzo dynamicznego rozwoju tego typu metod, będącego też wynikiem wysiłkiem naukowców z całego świata. W tym kontekście należy podkreślić, że w tak technologicznie rozwiniętej dziedzinie często trudno jest się „przebić” i zaproponować nowe rozwiązania. Dzieje się tak dlatego, że przede wszystkim trzeba jednoznacznie pokazać że proponowane metody wprowadzają coś nowego i, choćby w wybranych aspektach, są nowatorskie i po prostu lepsze niż inne znane rozwiązania. Tu niewątpliwie duży sukces Doktoranta oraz Jego promotora, w postaci opracowania kilku nowoczesnych systemów zespołów klasyfikatorów, o których już była mowa, a które to rozwiązania zostały opublikowane w prestiżowych periodykach naukowych. Mam tu głównie na myśli prace: *Weighted scoring in geometric space for decision tree ensemble*, opublikowana w czasopiśmie IEEE Access (pozycja nr [16] w rozprawie) oraz *Decision tree integration using dynamic regions of competence*, w Entropy (pozycja nr [14]), obydwie opracowane w zespole J. Biedrzycki i R. Burduk i obydwie wydane w roku 2020.

W tym kontekście, i po nie najłatwiejszej lekturze tej obszernej rozprawy, należy zadać sobie pytanie jaki jest cel pisania rozpraw o takiej objętości, a jeśli już, to co powinno się w niej znaleźć i do jakiego czytelnika jest adresowana. W moim przeświadczeniu, jeżeli łatwo dostępne są artykuły prezentujące główne idee metod oraz znakomitą większość wyników, to rozprawa doktorska nie powinna zatrzymać się na ich ponownym przytoczeniu, ale wykorzystać możliwość znacznego rozszerzenia spojrzenia na dziedzinę i na pokazanie głębszych pokładów swoich rozwiązań. W takim przypadku, rozprawa tego typu może stać się rodzajem tutorialu, z walorem nowości i świeżości spojrzenia na najnowocześniejsze podejścia w danej dziedzinie, i z kolei wykorzystana np. przez innych, często młodych naukowców do pogłębienia swojej wiedzy. Tego typu doktoraty są dobrze przyjmowane i doceniane przez środowisko, co często przenosi się na częstość ich cytowań. Niestety, mimo pewnej próby podjętej przez Doktoranta w rozdziale pierwszym oraz drugim, muszę tę próbę ocenić jako nieudaną. Dla przykładu prezentacja Support vector machine (1.1.4.1) nawet nie wspomina o roli jaką funkcje kernelowe spełniają w tych klasyfikatorach, nie mówiąc już o braku jakiegokolwiek rysunku, czy też diagramu w całym tym rozdziale! Bardzo powierzchowny jest też opis Artificial neural network (1.1.4.2), w tym Algorytm 1: ANN training, który zupełnie nie wiem do czego ma służyć, tym bardziej że Autor w ogóle nie odnosi się dalej do tego typu klasyfikatorów. Aż prosiłoby się, żeby chociaż w rozdziale Image analysis (1.4.2) zamieścić choćby jeden obraz... Nie lepiej jest też z prezentacją podejścia geometrycznego do integracji

klasyfikatorów składowych w rozdziale 2, którego rola jest jeszcze bardziej istotna gdyż zagadnienie to jest w pewnym sensie fundamentalne w stosunku do głównych metod prezentowanych i analizowanych w dalszych częściach rozprawy. Niestety, w rozdziale tym nie znajdziemy żadnego rysunku – a jest to rozdział poświęcony geometrii – za to jest Algorithm 2: AdaBoost algorithm. Pytanie tylko po co, skoro w żadnym z kolejnych rozdziałów Autor ani razu do niego się nie odwołuje.

O ile chodzi o zaproponowane metody, to nasuwają się następujące pytania oraz uwagi:

1. W rozdziale 4.1.2 wspomniane jest użycie dokładnie pięciu, i tylko pięciu, klasyfikatorów ze zbioru klasyfikatorów (ang. *pool of classifiers*). Czemu wyłącznie ta wartość i czy nie warto byłoby zrobić choćby kilku testów ze znacznie różniącą się liczbą klasyfikatorów? Eksperymenty tego typu mogłyby pokazać ciekawe i dalsze aspekty proponowanych metod.
2. Pojęcia midpoint oraz centroid nie są precyzyjnie zdefiniowane, np. na str. 50 jest mowa o „the subspace’s centroid are used”, podczas gdy już zaraz na str. 51 jest mowa o „region spanning over the midpoint”. Podobnie jest w rozdziale 4.4.3. Te pojęcia powinny być lepiej wyjaśnione.
3. W algorytmie nr 4 (str. 54, krok 2) niejasny jest sposób wyboru cech („Choose the features that are the most informative.”). Co Doktorant ma na myśli pisząc „informative” i jaka metoda jest do tego zadania użyta i dlaczego? Jak ważny jest ten krok dla całego algorytmu? Podobne założenie jest w kroku nr 1 algorytmu nr 5.
4. Metoda kalkulacji wag, wyrażona wzorem (4.9), powinna być lepiej wyjaśniona. Zakładając, że przestrzeń jest znormalizowana do zakresu  $0 \dots 1$  i przyjmując że euklidesowska odległość  $d$  po prostu jest przycinana do wartości 1 dla punktów bardziej odległych (w końcu i tak przestrzeń jest unormowana do  $0 \dots 1$ ), można by wyeliminować czynnik delta, gdyż wtedy mamy  $1-d=0$ . Dodatkowo parametr  $s_0$ , czy nie powinien raczej być oznaczony jako  $s_i$ , skoro nie jest to jeden i jedyny punkt dla wszystkich klasyfikatorów  $\Psi_i$ ?
5. Poprzednia uwaga dotyczy również kolejnej wersji metody dynamicznej i wzoru (4.11). Wzory (4.9) oraz (4.11) są dość podobne, przynajmniej w formie, różne są jednak interpretacje poszczególnych współczynników. Dlatego lepszego wyjaśnienia wymagają różnice pomiędzy tymi dwoma wariantami dynamicznymi. Najlepiej byłoby pokazać i wyjaśnić sytuację, najlepiej w kontekście budowanych przestrzeni geometrycznych, w których najnowsza metoda przewyższa działaniem jej poprzedni wariant. Ale czy są przypadki, że jest odwrotnie, tj. że metoda z rozdziału 4.3 daje lepsze wyniki niż ta z 4.4?
6. W eksperymentach z dynamicznymi regionami kompetencji przyjęto jeszcze mniejszą liczbę klasyfikatorów bazowych  $K=3$ . Czemu tak mało i czemu tylko 3?
7. Wyjaśnienia wymaga przyjęta metodologia badawcza przedstawiona w rozdziale 4.3.2 polegająca na wyborze miary  $F$ , którą Doktorant tłumaczy dużym niezbalansowaniem wielu zbiorów uczących użytych do testów. Ale przecież w znaczącej mierze są to te same zbiory, które zostały użyte w eksperymentach z wariantem podejścia statycznego, a zaprezentowanych w rozdziale 4.1.2. W tym to przypadku Autor przekonywał nas do użycia nie miary  $F$ , ale właśnie accuracy oraz tzw. współczynnika Matthews. Czemu takie wybory?
8. Kolejne pytanie, będące pochodną poprzedniego, to dlaczego te dwie metody, tj. integracja statyczna (4.1.2) oraz dynamiczna (4.3.2), nie zostały ze sobą w żaden sposób porównane?
9. Wartościowe byłoby też zbiorcze podsumowanie wszystkich zaproponowanych metod. Czy Doktorant mógłby podać takie podsumowanie, względnie wytyczne kiedy i która metoda powinna być stosowana, np. ze względu na rodzaj danych, dokładność odpowiedzi, czy też złożoność obliczeniową.

10. Z kolei najnowsza i najbardziej rozwinięta druga wersja metody z dynamiczną integracją, zaprezentowaną w rozdziale 4.4, Doktorant przetestował tym razem 4 warianty z klasyfikatorami bazowymi  $K \in \{3, 5, 7, 9\}$ . Niestety, nie wiadomo dla której z tych wartości  $K$  zostały zaprezentowane wyniki działania tej metody, zaprezentowane w rozdziale 4.4.2. Tutaj już można pochwalić Doktoranta, gdyż porównał wyniki z wcześniejszym wariantem metody dynamicznej, co np. w tabeli 4.18 oznaczono jako  $\Psi_{10}$ . Tym niemniej, wyjaśnienia wymaga rozbieżność tej kolumny z tabeli 4.18 z wynikami zamieszczonymi w tabeli 4.15, gdzie właśnie są wyniki dla  $\Psi_{10}$ .
11. Jaka jest relacja złożoności obliczeniowych w stosunku do metod, do których Doktorant porównuje swoje rozwiązania?
12. Doktorant jest współautorem publikacji dotyczących integracji oraz selekcji w przestrzeniach geometrycznych liniowych klasyfikatorów typu SVM (publikacje nr [24] oraz [25]), jednakże nie ma o nich mowy w rozprawie. Czemu taki wybór i jak można integrować inne rodzaje klasyfikatorów, takich jak na przykład wspomniane SVM, w przestrzeniach geometrycznych?

Drobniejsze uwagi edycyjne:

Str. 40 – „The results in the first two papers do not cover all of the datasets” nie bardzo wiadomo, które artykuły Doktorant ma na myśli.

Str. 43 – „calculate the centroids for each label using.” chyba niedokończone zdanie.

Str. 49 – „In the presented article, a novel approach to decision tree ensembling in the geometric space was proposed” czy słowo „article” powinno być użyte w tym kontekście? Podobnie jest w pierwszym zdaniu rozdziału 4.2.5 na str. 60.

Należy jednak podkreślić, że powyższe kwestie i pytania mają raczej charakter polemiczny i dotyczą pewnych wyborów natury edycyjnej, bądź też wyjaśnienia wybranych zagadnień lub są to sugestie rozszerzenia pewnych problemów natury naukowej poruszonych w rozprawie. Nie umniejszają one jednak wartości zaproponowanych i przebadanych przez Doktoranta metod klasyfikacji danych.

#### **4. Przydatność rozprawy dla nauk technicznych**

Zbadane oraz opisane w rozprawie doktorskiej Pana magistra inżyniera Jędrzeja Biedrzyckiego metody geometrycznego konstruowania zespołów klasyfikatorów z drzewami decyzyjnymi, jak również przedstawione eksperymenty oraz statystyczna analiza wyników, wnoszą istotny wkład w dziedzinę Informatyka Techniczna oraz Telekomunikacja. Jak zostało to wykazane w rozprawie, metody te przewyższają jakościowo dobrze znane i szeroko stosowane klasyfikatory lasy losowe oraz metodę fuzji odpowiedzi z wykorzystaniem głosowania większościowego pomiędzy klasyfikatorami bazowymi. Między innymi z tego względu, jak również z wielkiego zapotrzebowania na dokładne i efektywne obliczeniowo klasyfikatory, opracowanie przez Pana magistra inżyniera Jędrzeja Biedrzyckiego oraz zaprezentowane w rozprawie algorytmy mogą znaleźć liczne zastosowania. Mogą one również być użyte jako moduły składowe bardziej złożonych systemów klasyfikacji, jak również mogą być dalej rozwijane, między innymi poprzez zastosowanie różnych wersji drzew, metod selekcji cech, jak również z wykorzystaniem innych zbiorów danych. Szczególnie ciekawa byłaby tu próba połączenia z cechami semantycznymi otrzymanymi z głębokich sieci neuronowych.

Podsumowując, opracowane metody i algorytmy geometrycznego konstruowania zespołów klasyfikatorów z drzewami decyzyjnymi oraz otrzymane przez Pana magistra inżyniera Jędrzeja

Biedrzyckiego wyniki eksperymentalne, mają istotne znaczenie dla nauk technicznych, zarówno w aspekcie osiągnięć teoretycznych, jak i w aspekcie możliwości aplikacyjnych.

## 5. Charakterystyka działalności naukowej, w tym publikacyjnej, Doktoranta

Oprócz przedstawionych w poprzednich rozdziałach analiz dotyczących metod zaprezentowanych w rozprawie, jak również ich ewaluacji eksperymentalnej, istotnym aspektem jest działalność naukowa oraz publikacyjna Doktoranta. Z nadesłanych informacji, jak również z analizy materiałów dostępnych w Internecie, wynika, że p. mgr inż. Jędrzej Biedrzycki jest autorem lub współautorem co najmniej sześciu publikacji naukowych. Na szczególne podkreślenie zasługują tu publikacje w istotnych czasopiśmie naukowych z listy JCR opracowane wraz z promotorem, Panem profesorem Robertem Burdukiem.

1. J. Biedrzycki and R. Burduk. *Decision tree integration using dynamic regions of competence*. Entropy, 22(10):1129, 2020 (IF=2.419).
2. J. Biedrzycki and R. Burduk. *Weighted scoring in geometric space for decision tree ensemble*. IEEE Access, 8:82100 – 82107, 2020 (IF=3.745).
3. J. Biedrzycki and R. Burduk. *Integration of decision trees using distance to centroid and to decision boundary*. Journal of Universal Computer Science, 26:720 – 733, 2020 (IF=1.066).

Należy podkreślić, że w każdej z wymienionych publikacji Doktorant jest pierwszym z autorów. Pozostałe trzy publikacje zostały opublikowane w materiałach konferencyjnych.

W przypadku doktoratu jest to dobry dorobek publikacyjny świadczący o istotnym zaangażowaniu naukowym, jak również o aktywności badawczej oraz rozpoznawalności prac Pana magistra inżyniera Jędrzeja Biedrzyckiego w światowym środowisku naukowym.

## Podsumowanie

Podsumowując stwierdzam, że **postawione w rozprawie zagadnienia badawcze zostały prawidłowo rozwiązane, a postawiona w rozprawie teza naukowa została wykazana**. Uzyskane rezultaty **stanowią oryginalny własny wkład Pana magistra inżyniera Jędrzeja Biedrzyckiego** w rozwój dyscypliny naukowej **Informatyka Techniczna i Telekomunikacja**, a w szczególności dotyczą nowoczesnych i oryginalnych metod geometrycznego podejścia do konstrukcji zespołów klasyfikatorów z drzewami decyzyjnymi. Pan magister inżynier Jędrzej Biedrzycki wykazał się przy tym dobrą znajomością matematycznych metod klasyfikacji danych, jak również ich efektywnej implementacji w środowiskach rozproszonych. Wykazał się on również opanowaniem naukowego warsztatu badawczego, zarówno od strony teoretycznej, jak i eksperymentalnej. Wszystko to świadczy o dojrzałości naukowej Pana magistra inżyniera Jędrzeja Biedrzyckiego.

Recenzowaną pracę oceniam jako **spełniającą** wymagania stawiane rozprawom doktorskim. Wnioskuje o jej przyjęcie oraz o dopuszczenie Pana magistra inżyniera Jędrzeja Biedrzyckiego do publicznej obrony.

