

Dr hab. inż. Andrzej Przybył, prof. uczelni
Katedra Sztucznej Inteligencji
Wydział Informatyki i Sztucznej Inteligencji
Politechnika Częstochowska

Częstochowa, dn. 31.03.2025 r.

**Recenzja rozprawy doktorskiej mgr inż. Sylwestra Czmiła pt.
„Algorytmy nadzorowanego inkrementalnego uczenia maszynowego oraz ocena jakości
ich klasyfikacji”**

Niniejsza recenzja została przygotowana w odpowiedzi na pismo RE.530/42/2024/RDITiT Pana dra hab. inż. Mariusza Oszusta, prof. uczelni, Przewodniczącego Rady Dyscypliny Informatyka Techniczna i Telekomunikacja z dnia 11 grudnia 2024r zawierające informację, że Rada Dyscypliny Informatyka Techniczna i Telekomunikacja Politechniki Rzeszowskiej w dniu 11 grudnia 2024 r. powołała mnie na recenzenta w/w rozprawy doktorskiej.

1. Zakres tematyczny i struktura rozprawy.

Tematyka recenzowanej rozprawy obejmuje zagadnienie nadzorowanego inkrementalnego uczenia maszynowego klasyfikatorów płytkich na podstawie danych uczących oraz oceny jakości ich klasyfikacji. Klasyfikacja danych jest jednym z ważniejszych zastosowań uczenia maszynowego. Jej zadaniem jest przewidywanie klasy dla rekordów danych, zwłaszcza takich, jakie nie były używane w procesie uczenia klasyfikatora. Istnieje wiele nadzorowanych algorytmów klasyfikacyjnych, a wśród nich można wyróżnić algorytmy tradycyjne (nieprzyrostowe) oraz algorytmy inkrementalne (przyrostowe). Główną zaletą algorytmów inkrementalnych jest zdolność do adaptacji modelu do nowych danych w trakcie ich napływu, podczas gdy w tradycyjnym uczeniu maszynowym model jest uczony na całym dostępnym zbiorze danych jednorazowo. Z tego powodu algorytmy inkrementalne są obiektem wielu prowadzonych obecnie badań naukowych.

Głównym celem recenzowanej pracy doktorskiej pracy było opracowanie i implementacja programowa i sprzętowa nowego inkrementalnego algorytmu klasyfikacji danych, który w klasie algorytmów płytkich okaże się nie gorszy od dotychczas stosowanych. W związku z tak postawionym głównym celem pracy zaproponowano pięć celów szczegółowych. Pierwszym z nich było opracowanie nowego, przyrostowego algorytmu klasyfikacji, który posiadałby minimalną liczbę „parametrów nastrajalnych”. Drugi cel to projekt i implementacja narzędzia (oprogramowania komputerowego w języku Python) do wszechstronnego testowania nowego algorytmu klasyfikacji oraz spełniającego

WPŁYNĘŁO

04. KWI. 2025

POLITECHNIKA RZESZOWSKA
im. Ignacego Łukasiewicza
WYDZIAŁ ELEKTRONIKI I INFORMATYKI
35-959 Rzeszów, ul. W. Pola 2
tel. 17 865 1764

istotne wymagania w zakresie badań benchmarkowych. Kolejne cele szczegółowe dotyczą przeprowadzenia wszechstronnych badań porównawczych nowego klasyfikatora z innymi algorytmami w oparciu o wyznaczanie i analizę wskaźników jakości klasyfikacji, zastosowaniu metod statystycznych w celach porównawczych oraz próba programowej i sprzętowej implementacji opracowanego algorytmu.

Tematykę rozprawy uważam za aktualną i bardzo ważną w zakresie dyscypliny Informatyka Techniczna i Telekomunikacja.

Recenzowana rozprawa doktorska obejmuje 113 stron i składa się z dziewięciu rozdziałów oraz bibliografii zawierającej 71 pozycji. Rozdział pierwszy wprowadza w tematykę badań, natomiast cel i zakres pracy został przedstawiony w rozdziale drugim. Rozdział trzeci obejmuje opis wybranych algorytmów uczenia nadzorowanego do klasyfikacji danych, z kolei rozdział czwarty zawiera opis wybranych metod porównywania różnych modeli klasyfikacji pod względem ich skuteczności. W rozdziale piątym opisano nowy, zaproponowany przez Doktoranta w ramach niniejszej pracy klasyfikator inkrementalny o nazwie SEVQ. Rozdział szósty to obszernie porównanie algorytmu SEVQ z innymi algorytmami. Rozdział siódmy to pobieżny opis sprzętowej implementacji algorytmu SEVQ na układzie FPGA oraz prezentacja uzyskanych wyników w porównaniu do implementacji programowej na CPU. W rozdziale ósmym zawarty został krótki opis autorskiego oprogramowania CACP. Podsumowanie i wnioski końcowe zawarte są w rozdziale dziewiątym.

Struktura rozprawy jest prawidłowa, podział treści jest czytelny.

2. Zawartość rozprawy i ocena najważniejszych wyników

Recenzowana rozprawa doktorska dotyczy bardzo ważnego i aktualnego zagadnienia jakim jest klasyfikacja danych. W zakresie klasyfikacji w ostatnich latach zyskały na popularności algorytmy głębokie, tj. modele oparte na sztucznych sieciach neuronowych, składające się z dużej liczby neuronów i wielu warstw ukrytych. Jednak, jak słusznie zauważa Doktorant, algorytmy tradycyjne (płytke) wciąż są interesujące w wielu zastosowaniach, ponieważ zwykle mają znacznie mniejsze zapotrzebowanie na zasoby obliczeniowe i pamięć niż algorytmy głębokie. A więc nadają się do praktycznego zastosowania tam, gdzie ze względu na ograniczone zasoby (pamięć, moc obliczeniowa, zużycie energii, koszt, itp.) nie mogą być użyte algorytmy głębokie. Ponadto, algorytmy płytke są często bardziej transparentne (interpretowalne) niż algorytmy głębokie, co ma duże znaczenie w takich dziedzinach, jak medycyna czy prawo. Wskazując na te zalety Doktorant w swojej rozprawie skoncentrował się na algorytmach płytkich. Naturalną konsekwencją wyboru takiego obszaru badań była próba programowej i sprzętowej implementacji zaproponowanego, nowego algorytmu

(SEVQ). Próba ta miała na celu wykazanie skuteczności działania algorytmu zaimplementowanego na platformach sprzętowych o mocno ograniczonych zasobach – w porównaniu do komputerów PC. Doktorant wykazał również przydatność opracowanego algorytmu poprzez rzetelne porównanie go z innymi, znanymi dotychczas algorytmami klasyfikacji. Pod względem analizowanych metryk jakości klasyfikacji nowy algorytm w klasie algorytmów płytkich okazał się nie gorszy od dotychczas stosowanych. Ponadto, dzięki zaproponowanym przez Doktoranta rozwiązaniom opracowany przez niego nowy algorytm SEVQ nie ma „parametrów nastrajalnych”, a klasyfikowane dane nie wymagają normalizacji. Wymienione cechy są potwierdzeniem poprawności zastosowanego podejścia i przydatności uzyskanych wyników.

Do najważniejszych osiągnięć Doktoranta zawartych w recenzowanej rozprawie należy zaliczyć:

1. Zaprojektowanie nowego, sprawnego algorytmu klasyfikacji oferującego wyniki konkurencyjne w stosunku do innych algorytmów. Opracowany algorytm charakteryzuje się przy tym prostotą oraz wygodą stosowania, gdyż nie wymaga doboru „parametrów nastrajalnych”.
2. Opracowanie metodologii automatycznego porównywania algorytmów klasyfikacji w oparciu o istotne metryki jakości klasyfikacji. Punkt ten jest szczególnie istotny gdyż, jak słusznie Doktorant zauważył w swojej rozprawie, wiele „algorytmów posiada mocne i słabe strony, a ich skuteczność często zależy od specyfiki zbioru danych oraz charakterystyki problemu. Wybór odpowiedniego algorytmu powinien uwzględniać dostępną ilość danych, liczbę cech, zapotrzebowanie na interpretowalność modelu, a także zasoby obliczeniowe. Często dobrym rozwiązaniem jest testowanie kilku algorytmów i porównanie ich wydajności za pomocą odpowiednich metryk”.
3. Opracowanie efektywnego narzędzia w postaci oprogramowania CACP pozwalającego na zautomatyzowanie procesu porównywania nowo opracowywanych algorytmów klasyfikacji. Bardzo istotnym aspektem tego rozwiązania jest zamieszczenie jego kodu wraz z dokumentacją w ogólnodostępnym repozytorium (GitHub). Uwiarygadnia to w znaczącym stopniu zrealizowane przez Doktoranta badania oraz pozwala na ich kontynuację przez inne osoby. Należy też docenić kolejne praktyczne aspekty udostępnionego, autorskiego rozwiązania, a mianowicie automatyczne generowanie tabel zawierających wyniki badań w formatach CSV i TEX oraz gotowych wykresów z danymi w formatach graficznych PNG i EPS. Pozwala to na ich bezpośrednie i wygodne użycie w nowych publikacjach naukowych w profesjonalnym systemie składania tekstów LaTeX.

W trakcie realizacji dosyć obszernych badań trudno ustrzec się przed popełnieniem drobnych błędów lub nieścisłości. W recenzowanej rozprawie błędy te nie są znaczące i nie wpływają na poprawność ogólnych wniosków sformułowanych przez Doktoranta.

Poniżej przedstawiono szczegółową listę najważniejszych uwag merytorycznych i drobnych błędów edycyjnych wykrytych w rozprawie.

2.1. Uwagi merytoryczne

2.1.1. Na str. 34. rozprawy Doktorant zawarł następujący, a przy tym prawdopodobnie słuszny wniosek, a mianowicie „z przedstawionych eksperymentów wynika, że aby uzyskać najlepsze rezultaty, warto po kolei dostarczać punkty należące do różnych klas”. Jednak nie rozwinął bardziej tej kwestii, ani tym bardziej nie przeanalizował jakie wypływają z niej konsekwencje. Jaką korzyść daje taka wiedza, tj. czy zaniechanie tego wymogu spowoduje, że algorytm może działać nieprawidłowo, gdyż jakość jego działania może spaść do nieakceptowalnego poziomu? A jeśli tak, to w jaki sposób wymieszać takie punkty reprezentujące poszczególne klasy? Czy w praktycznych zastosowania zawsze jest możliwe dostarczanie punktów w opisanej kolejności?

2.1.2. Na str. 42. podana została informacja „Dla każdego algorytmu podano średnie wartości oraz odchylenia standardowe dla metryk wydajnościowych, takich jak ACC, PRE, SEN, F1 i AUC. Wyniki obliczono dla wszystkich zbiorów danych na każdej partii danych wydzielonej podczas 10-krotnej walidacji krzyżowej dla każdego algorytmu.” Nie podano przy tym informacji, czy zaproponowanie 10-krotnej walidacji krzyżowej wynika z braku determinizmu algorytmu wyznaczającego jakość testowanego algorytmu i chęci uśrednienia wyników? W jaki sposób dane są dzielone na zbiór uczący i zbiór testowy: czy jest to podział losowy, czy inny? Jeśli podział jest losowy, to czy stosowane są dokładnie te same zbiory dla różnych algorytmów? Informacje te są istotne z punktu widzenia możliwości oceny zastosowanej metodologii porównywania między sobą różnych algorytmów. Brak determinizmu może być istotną przeszkodą w rzetelnej ocenie porównywanych algorytmów.

2.1.3. W rozdziale siódmym Doktorant zamieścił opis sprzętowej implementacji algorytmu SEVQ w układzie FPGA. Wnioski z tego etapu badań, jakie zostały zamieszczone w rozprawie są generalnie słuszne, co świadczy o prawdopodobnie prawidłowym sposobie ich przeprowadzenia. Jednakże należy wskazać istotne zastrzeżenie do tej części pracy, otóż w pracy nie zamieszczono wystarczająco dużo informacji na temat sposobu sprzętowej implementacji algorytmu. W szczególności:

2.1.3.1. Nie zamieszczono kodu w języku C++ a jedynie lakoniczną informację informującą o przepisaniu „kodu algorytmu SEVQ napisanego w języku programowania Python na kod w języku C++” oraz „... przy użyciu narzędzia Vitis High-Level Synthesis (HLS), kod ten został zsyntetyzowany do kodu Verilog”. Bez znajomości tego kodu i całej metodologii postępowania nie można

rzetelnie ocenić jakości takiej implementacji. Przykładowo, jaką arytmetykę wykorzystano do obliczeń, zmiennoprzecinkową 64-bitową czy 32-bitową, a może stałoprzecinkową? Te i inne szczegółowe aspekty sprzętowej implementacji algorytmu często drastycznie wpływają na uzyskiwaną wydajność, precyzję działania, ilość konsumowanej energii elektrycznej oraz koszt sprzętowej implementacji takiego algorytmu wyrażony ilością zużywanych zasobów układu FPGA.

2.1.3.2. Wyżej wymieniona uwaga jest również istotna w kontekście zamieszczonej w tym rozdziale następującej informacji „implementacja FPGA nie pogarsza efektywności algorytmu”. Dlaczego implementacja algorytmu na FPGA miałaby pogarszać efektywność algorytmu? Dane zawarte w tabeli 7.2 wykazują różnice w metrykach pomiędzy implementacją sprzętową na FPGA oraz programową na procesorze ARM. Dlaczego takie różnice występują skoro zaimplementowany został dokładnie taki sam algorytm? Czy powodem tych różnic jest brak determinizmu algorytmu porównującego te dwie implementacje? Sposób implementacji (programowa kontra sprzętowa) nie powinien mieć żadnego wpływu na w/w metryki, dopóki nie stosuje się pewnych założeń upraszczających dla którejś z implementacji.

2.1.3.3. Porównanie czasów przetwarzania poszczególnych zbiorów danych zawarte w tabeli 7.3 wygląda wiarygodnie, jednakże podobnie jak poprzednio nie podano wystarczających szczegółów, aby to rzetelnie ocenić. W szczególności, nie podano parametrów procesora ARM – czy był to procesor ARM Cortex-A9 wbudowany w układ AMD/Xilinx ZYNQ-7? Nie opisano też sposobu dostarczania danych uczących/testujących do algorytmu zaimplementowanego na FPGA. Czy pośredniczył w tym typ CPU, czy było to zrealizowane w pełni sprzętowo i w jaki sposób? Te elementy mogą istotnie wpłynąć na uzyskiwaną wydajność algorytmu wyrażoną czasem jego przetwarzania.

2.1.3.4. Na str. 78 Doktorant zamieścił pewne istotne spostrzeżenie, a mianowicie: „W przypadku 20% zbiorów danych implementacja algorytmu SEVQ w układzie FPGA okazała się wolniejsza od implementacji na procesorze ARM. Przypadki te dotyczyły jednak bardzo małych zbiorów danych, które zawierały 80-215 rekordów opisanych przez 4-19 cech”. W pracy zabrakło jednak analizy przyczyn tego zjawiska. Wydaje się, że może być ono efektem w/w sposobu dostarczania danych uczących/testowych do algorytmu zaimplementowanego w FPGA. Trudno to jednak ocenić bez znajomości szczegółów implementacji.

Trudno się dziwić, że w pracy nie podano w/w szczegółowych informacji, skoro rozdział siódmy zawiera jedynie cztery strony tekstu, z których jedną wypełnia tabela

z wynikami. Na usprawiedliwienie można jednak dodać, że tematyka poruszana w tym rozdziale dotyczy jedynie małego wycinka badań opisywanych w ramach recenzowanej rozprawy. Z tego powodu wyżej wymienione uwagi krytyczne nie obniżają znacząco całości pracy.

- 2.1.4. W rozprawie Doktorant stosuje określenie „metryki wydajnościowe” do określenia wskaźników jakości takich jak dokładność, precyzja oraz czułość algorytmu klasyfikacji. Moim zdaniem jest to niefortunne określenie, gdyż może wprowadzić w błąd. W szczególności w zestawieniu z wynikami zawartymi w rozdziale siódmym dotyczącymi czasu przetwarzania poszczególnych algorytmów na różnych platformach sprzętowych. Wydajność w takim kontekście kojarzy się raczej z szybkością wyrażoną czasem przetwarzania algorytmu lub ilością przetworzonych rekordów danych na sekundę, a nie z jego dokładnością.

2.2. Uwagi edycyjne

- 2.2.1. Na str. 44 znajduje się prawdopodobny błąd redakcyjny w sformułowaniu „Algorytm XGB osiągnął najwięcej pierwszych miejsc (7 razy), drugich miejsc (7 razy) i trzecich miejsc (4 razy)”. Z analizy wymienionego rys. 6.1 wynika bowiem, że dwie pierwsze liczby to osiem i osiem, a nie siedem i siedem.
- 2.2.2. W rozdziale siódmym zawarte są niezbyt zrozumiałe sformułowania, tj. „Implementacja algorytmu SEVQ na tej płycie FPGA charakteryzuje się wykorzystaniem 10,1% dostępnych komórek LUT ... co wskazuje na efektywne wykorzystanie elementów logicznych” oraz „zużycie przerzutnika wynosi 6,27%, co sugeruje ich skuteczne wykorzystanie”. Nie jest jasne co oznaczają zwroty „efektywne” oraz „skuteczne” w kontekście określenia ilości zasobów sprzętowych zużywanych przez zaimplementowany algorytm.

3. Wniosek końcowy

Przeprowadzona przez Doktoranta w sposób właściwy krytyczna analiza rozwiązań opisywanych w literaturze świadczy o tym, że **prezentuje on wymaganą ogólną wiedzę teoretyczną osoby ubiegającej się o nadanie stopnia doktora**. Z kolei przeprowadzenie w sposób prawidłowy wszechstronnych, a przy tym dosyć szczegółowych badań dotyczących tematyki pracy wskazuje, że **Doktorant wykazuje się umiejętnością samodzielnego prowadzenia pracy naukowej, co jest wymaganiem stawianym osobie ubiegającej się o nadanie stopnia doktora**. Rozwiązania zawarte w rozprawie, w tym w szczególności zaproponowany, nowy algorytm SEVQ oraz opracowanie metody

automatycznego przeprowadzania wszechstronnych badań porównawczych przedstawiają **oryginalne rozwiązanie problemu naukowego.**

Oceniając rozprawę doktorską pt. „Algorytmy nadzorowanego inkrementalnego uczenia maszynowego oraz ocena jakości ich klasyfikacji” stwierdzam, że spełnia ona wymagania określone w ustawie Prawo o szkolnictwie wyższym i nauce z dnia 20 lipca 2018 roku z późniejszymi zmianami. Rozprawa spełnia wszystkie wymagania ustawowo przyjęte w polskim środowisku naukowym stawiane rozprawom doktorskim i wnoszę o jej przyjęcie oraz dopuszczenie do dalszych etapów postępowania o uzyskanie stopnia naukowego doktora w dziedzinie nauk technicznych, w dyscyplinie Informatyka Techniczna i Telekomunikacja.

Jednocześnie doceniam wysoki poziom merytoryczny zaproponowanych rozwiązań, oryginalność oraz wszechstronność przeprowadzonych badań, ważność i aktualność tematyki w poruszanej w rozprawie, a także wysoki, jak na Doktoranta, dotychczasowy dorobek naukowy wyrażony przez siedem publikacji naukowych oraz 42 cytowania wg Google Scholar. Sześć z wymienionych publikacji jest znaczących (za 200 lub 100 pkt), a w trzech spośród nich Doktorant jest pierwszym autorem. W oparciu o ten aspekt oceny, oczekując jednak na przebieg obrony pracy doktorskiej, dyskusję w jej trakcie oraz wyjaśnienia i odpowiedzi udzielone przez Doktoranta składam wstępnie wniosek o wyróżnienie tej rozprawy.

