

Recenzja rozprawy doktorskiej

mgra inż. Jarosława Szkoły

pt. „Uczenie federacyjne z uwzględnieniem niepewności danych:
zaawansowane metody agregacji oparte na przedziałowych zbiorach
rozmytych”

napisanej pod kierunkiem prof. dr hab. Barbary Pękałi

1. OPINIA OGÓLNA

Przedłożona do recenzji rozprawa doktorska mgra inż. Jarosława Szkoły dotyczy uczenia federacyjnego w warunkach niepewności, niekompletności i niejednorodności danych. Rozprawę stanowi cykl dziewięciu współautorskich publikacji z lat 2022–2026 wraz z omówieniem. Prace łączy problem uwzględniania jakości modeli lokalnych oraz niepewności danych przy konstruowaniu modelu federacyjnego.

Rozprawę oceniam pozytywnie. Za jej zasadniczą wartość uznaję konsekwentne rozwijanie metod agregacji modeli z wykorzystaniem przedziałowych zbiorów rozmytych, całek rozmytych i miar entropii oraz powiązanie agregacji z elastyczniejszą reprezentacją modelu lokalnego. Cykl ma czytelny kierunek rozwoju: od podstawowej architektury uczenia i wagi klientów do integracji kilku mechanizmów uwzględniających niepewność. Obejmuje również zastosowanie systemu regulowego do analizy postury człowieka.

Praca zawiera usterki w zapisie matematycznym oraz fragmenty wymagające dokładniejszego opisu eksperymentów. Najważniejsze z nich wskazuję poniżej. Przy ocenie całości cyklu uwzględniam zarówno te ograniczenia, jak i wkład koncepcyjny, konstrukcję algorytmów oraz przeprowadzone badania. Zastrzeżenia wymagają wyjaśnienia i korekty, ale nie przekreślają w mojej ocenie zasadniczego osiągnięcia rozprawy.

2. OCENA MERYTORYCZNA

2.1. Znaczenie problematyki podjętej w rozprawie

Podjęta tematyka jest istotna dla rozwoju metod uczenia maszynowego na danych rozproszonych. W uczeniu federacyjnym poszczególni uczestnicy mogą dysponować zbiorami o różnej liczebności, odmiennych rozkładach klas i zróżnicowanej jakości. Wzajemnie wyłączenie liczbą obserwacji nie wykorzystuje bezpośrednio informacji o niepewności danych ani jakości lokalnych predykcji. Poszukiwanie zasad agregacji uwzględniających te cechy jest zatem dobrze uzasadnione.

Dobór aparatu matematycznego odpowiada rozpatrywanemu problemowi. Przedziałowe zbiory rozmyte umożliwiają reprezentowanie nieprecyzyjnej informacji, natomiast miary nicaddytywne i całki rozmyte dostarczają narzędzi łączenia ocen z uwzględnieniem zależności między źródłami. Wartościowy jest również wybór modeli, których konstrukcję można interpretować: regresji logistycznej i systemów regułowych. Znaczenie rozprawy widzę przede wszystkim w opracowaniu sposobów połączenia tych narzędzi w ramach wspólnego procesu uczenia.

2.2. Cele i metodyka badawcza

Cel badań jest czytelny: rozwinąć metody uczenia federacyjnego, w których niepewność danych jest uwzględniana zarówno przy uczeniu modeli lokalnych, jak i przy ich agregacji. Kolejne publikacje rozpatrują reprezentację danych, reguły uczenia, ważenie klientów, zastosowanie całek Sugeno i Choqueta oraz włączenie miar entropii. Takie rozłożenie programu na etapy jest logiczne i umożliwia ocenę kolejnych rozszerzeń.

Badania mają charakter konstrukcyjno-eksperymentalny. Głównym zbiorem danych jest Wisconsin Diagnostic Breast Cancer, obejmujący 569 przypadków. Analizowane są m.in. braki danych i nierównowaga klas u klientów. Zaletą późniejszych prac jest rozszerzanie porównań: P7 przedstawia powtórzenia eksperymentów i odchylenia standardowe, a P9 zestawia sześć wariantów łączących uczenie klasyczne lub przedziałowe z różnymi organizacjami uczenia. W P8 wyraźnie wskazano, że test nie uczestniczy w uczeniu.

Zakres doświadczeń uzasadnia ocenę proponowanych konstrukcji jako metod badawczych sprawdzonych w określonych scenariuszach. Dalszego potwierdzenia wymaga ich skuteczność dla większej liczby klientów i innych rodzajów niejednorodności. W omówieniu należałoby też wyraźniej oddzielić poprawę jakości klasyfikacji od odporności na braki danych, skalowalności i gwarancji prywatności. Są to różne własności, wymagające odmiennych sposobów weryfikacji.

Jedno uściślenie dotyczy konstrukcji przedziałów wejściowych. W opisach danych, m.in. P1 i P9 (s. 37 i 126 pliku rozprawy), użyto określenia „odchylenie standardowe”. Dokumentacja WDBC w repozytorium UCI podaje jednak dla odpowiedniej grupy kolumn błąd standardowy. Należy wskazać wykorzystane kolumny i ewentualne przeliczenia, ponieważ od tego zależy interpretacja szerokości przedziału. Jest to potwierdzona rozbieżność opisu, której wpływu na obliczenia nie można ustalić bez implementacji.

2.3. Struktura rozprawy i wkład Doktoranta

Przekazany dokument liczy 138 stron i zawiera streszczenia, wykaz publikacji, wprowadzenie, opis wkładu własnego, omówienie prac, pełne teksty artykułów oraz podsumowanie. Cykl obejmuje sześć publikacji konferencyjnych lub rozdziałów pokonferencyjnych i trzy artykuły czasopiśmiennicze. Struktura jest odpowiednia dla rozprawy opartej na publikacjach. Poszczególne artykuły stanowią samodzielne opracowania, a wspólne omówienie pozwala prześledzić powiązania między nimi.

Doktorant opisuje swój udział w koncepcji metod, implementacji, projektowaniu eksperymentów i przygotowaniu publikacji. Z przedstawionego opisu wynika istotne zaangażowanie

w kolejne etapy programu. Ponieważ wszystkie artykuły są współautorskie, ocenę wkładu indywidualnego odnoszę do deklaracji zawartych w rozprawie; przekazany materiał nie obejmował odrębnych oświadczeń współautorów. W samej syntezie przydatne byłoby krótkie zestawienie wskazujące dla każdej pracy nowy element, udział Doktoranta i związek z pozostałymi publikacjami.

2.4. Wykorzystana literatura

Literatura obejmuje właściwe obszary: uczenie federacyjne, teorię zbiorów rozmytych, reprezentację przedziałową, miary niepewności i agregację informacji. Artykuły posiadają własne bibliografie, dlatego oceny zakresu literatury nie należy ograniczać do osiemnastu pozycji wymienionych na końcu omówienia. Dobór piśmiennictwa świadczy o znajomości zaplecza teoretycznego badań. W opisie oryginalności należałoby precyzyjniej rozdzielić konstrukcje zaczerpnięte z literatury od ich modyfikacji i nowego sposobu zastosowania w uczeniu federacyjnym.

2.5. Ocena publikacji tworzących cykl

Stosuję oznaczenia P1–P9 zgodne z wykazem w rozprawie. Numery stron podane w recenzji oznaczają strony pliku PDF liczone od okładki, a numery wzorów, tabel i algorytmów odnoszą się do danej publikacji. Ze względu na powiązania między artykułami omawiam je w grupach odpowiadających kolejnym etapom badań.

P1–P3: architektura uczenia i ważenie klientów (s. 33–61). Prace te tworzą podstawę cyklu: wprowadzają uczenie na danych przedziałowych, wykorzystanie podobieństwa przy aktualizacji modelu oraz porównanie ważenia klientów na podstawie miar jakości. Wartością tego etapu jest powiązanie lokalnego uczenia z oceną modeli w federacji i rozpartrzenie danych niekompletnych oraz niechrównoważonych. Najważniejszego doprecyzowania wymaga zapis aktualizacji parametrów, omówiony w punkcie 2.6.

P4–P5: agregacja za pomocą całek rozmytych (s. 63–77). Zastosowanie całek Sugeno i Choqueta jest metodycznie uzasadnionym rozwinięciem prostego ważenia klientów. Wprowadza możliwość wykorzystania miar nieaddytywnych, dzięki czemu schemat agregacji może opisywać zależności między źródłami informacji. Etap ten jest ważny dla dalszego rozwoju cyklu. Zapisy definicji należy jednak ujednolicić z ich standardowym znaczeniem i z poprawioną postacią całki Choqueta przedstawioną w P9.

P6: federacyjny system regułowy (s. 79–92). Publikacja rozszerza badania o system wnioskowania oparty na przedziałowych regułach rozmytych i dane dotyczące postury człowieka. Jest to wartościowe uzupełnienie głównego nurtu prac, pokazujące możliwość wykorzystania idei uczenia federacyjnego w innej konstrukcji modelu. Interpretowalność reguł stanowi zaletę rozwiązania. Wyniki należy odnosić do rzeczywiście ocenianego zadania rozpoznawania postury; zakres ten jest węższy od pełnej oceny systemu wykrywającego upadki w warunkach użytkowych.

P7: agregacja wykorzystująca entropię (s. 94–106). Za istotny element pracy uznaję łączne wykorzystanie informacji o jakości modelu lokalnego i niepewności danych przy konstruowaniu agregatora. Porównanie dwóch sposobów ujmowania entropii oraz

analiza zmienności wyników pogłębiają wcześniejsze doświadczenia. Najważniejsze uwagi dotyczą zgodności zapisu całki Choqueta z przyjętym uporządkowaniem argumentów oraz precyzji wniosków o przewadze metod. Wyniki zależą od wariantu entropii i scenariusza, dlatego nie uzasadniają jednakowej przewagi wszystkich proponowanych konfiguracji.

P8–P9: reprezentacja modelu lokalnego i integracja metod (s. 108–133). Pomysł przypisania różnych reprezentantów podprzedziałom cechy zwiększa elastyczność modelu w porównaniu z pojedynczym współczynnikiem. P8 rozwija ten mechanizm w uczeniu lokalnym, natomiast P9 łączy go z federacyjną agregacją wykorzystującą entropię. Jest to logiczne domknięcie programu. Zestawienie sześciu wariantów w P9 pomaga ocenić udział poszczególnych składników, choć nie wykazuje poprawy każdej miary dla każdego klienta. Do wyjaśnienia pozostają przede wszystkim opis scalania podprzedziałów, uzgodnienie schematów ważenia i sposób obliczenia części miar jakości.

2.6. Najważniejsze uwagi wymagające wyjaśnienia

1. Znak w regule uczenia. W P1 (s. 37), a także w P9 (s. 128), zdefiniowano ujemną logarytmiczną funkcję straty, po czym zapisano dodawanie jej gradientu. Dla zwykłej regresji logistycznej, prawdopodobieństwa $p = \sigma(b)$ i etykiety y zachodzi

$$L(b) = -y \log p - (1 - y) \log(1 - p), \quad \frac{\partial L}{\partial b} = p - y.$$

Przy minimalizacji tej funkcji krok dla wyrazu wolnego ma postać $b \leftarrow b - \alpha(p - y)$, dla $\alpha > 0$. Dodanie gradientu daje przeciwny kierunek. Przykładowo dla $b = 0$, $y = 1$ i $\alpha = 0,01$ zapis z publikacji prowadzi do $b = -0,005$ i zwiększa stratę z około 0,693147 do 0,695650. Uwagi tej nie można rozstrzygnąć samą deklaracją użycia metody gradientowej: potrzebna jest poprawna definicja aktualizowanej wielkości. Kontrprzykład potwierdza błąd przy dosłownym odczytaniu wzoru, ale nie ustala, jaki znak zastosowano w kodzie.

2. Indeksowanie w definicji całki Choqueta. W P7, we wzorze (9) na s. 100, dla malejąco uporządkowanych argumentów użyto zbiorów końcowych $A_{(i)} = \{x_{(i)}, \dots, x_{(n)}\}$. Przy zapisanych tam różnicach wartości potrzebne są zbiory początkowe. Poprawna postać to

$$C_m(h) = \sum_{i=1}^n (h_{(i)} - h_{(i+1)}) m(\{x_{(1)}, \dots, x_{(i)}\}), \quad h_{(n+1)} = 0,$$

gdzie $h_{(1)} \geq \dots \geq h_{(n)} \geq 0$. Dla dwóch argumentów równych 0,8 i addytywnej miary z masami 1/2 zapis P7 daje 0,4, podczas gdy całka względem miary unormowanej powinna zachować stałą 0,8. W P9, we wzorze (7) na s. 128, zbiory początkowe są już podane poprawnie. Zastrzeżenie dotyczy więc określonego zapisu w P7; należy wyjaśnić jego relację z wykonanymi obliczeniami.

3. Dwa schematy ważenia w P9. We wzorze (1) na s. 125 określono wagi

$$w_i = \frac{(1 - E_i)Q_i}{\sum_{k=1}^r (1 - E_k)Q_k},$$

natomiast w punkcie 3.2.2 na s. 129 gęstości miary zdefiniowano za pomocą $|Q_i - E_i|$. Są to różne konstrukcje. Przykładowo dla $Q_1 = Q_2 = 0,8$, $E_1 = 0,2$ i $E_2 = 0,4$ pierwszy schemat

daje wagi (4/7; 3/7), a drugi gęstości (0,6; 0,4). Tekst powinien wskazać ich wzajemną rolę i wariant odpowiadający tabelom wyników. Stwierdzam tu niejednoznaczność specyfikacji; samo współlistnienie dwóch formuł nie dowodzi błędu wyników.

4. Warunek scalania podprzedziałów. W algorytmie 1 publikacji P9 (s. 127) występuje warunek $\eta_k - \eta_{k+1} \leq \psi$. Dla uporządkowanych granic $\eta_k \leq \eta_{k+1}$ i dodatniego progu jest on spełniony zawsze, zatem nie wybiera wyłącznie wąskich podprzedziałów. Jeżeli kryterium ma dotyczyć szerokości, powinno mieć postać

$$\eta_{k+1} - \eta_k \leq \psi.$$

Ponadto opis wyjścia algorytmu podaje szerokość $\leq \psi$, podczas gdy tekst obok mówi o przedziałach przekraczających próg. Te zapisy wymagają uzgodnienia. Warto również doprecyzować, jak podczas agregacji dopasowywane są reprezentanty modeli klientów, jeżeli ich lokalne podziały mają różne granice lub liczbę elementów.

5. Zgodność miar klasyfikacji. W tabeli 2 publikacji P9 (s. 130) podano $ACC = 0,983$, $SENS = 0,972$ i $SPEC = 0,950$. Dla miar wyznaczonych z jednej binarnej macierzy pomyłek, przy udziale klasy dodatniej π , obowiązuje

$$ACC = \pi SENS + (1 - \pi) SPEC, \quad 0 \leq \pi \leq 1.$$

Dokładność musi więc leżeć między czułością a swoistością. Podane wartości nie spełniają tego warunku, a różnica przekracza dokładność zaokrąglenia. Jeżeli jednak są to osobno uśredniane wyniki z prób o zmiennym udziale klas, taka nierówność między średnimi może wystąpić. Potrzebne jest zatem podanie sposobu uśredniania i macierzy pomyłek. Bez tych informacji nie przesądzam, czy źródłem rozbieżności jest tabela, czy sposób zestawienia wyników.

6. Rozdzielenie walidacji od końcowej oceny. W P9 (s. 125) lokalna ocena jakości wpływa na przyjęcie modelu globalnego i jego dalszą agregację, natomiast opis eksperymentu (s. 130) wymienia podział na uczenie i test w proporcji 90% : 10%. Należy doprecyzować, z jakich danych obliczano oceny sterujące uczeniem i czy końcowy test był od nich niezależny. Ta uwaga dotyczy kompletności opisu protokołu, a nie dowiedzionego przecieku danych. W P8 (s. 118) wyłączenie testu z uczenia opisano wprost; podobnie jednoznaczny zapis byłby przydatny dla całego procesu federacyjnego.

2.7. Język i strona formalna rozprawy

Omówienie jest zasadniczo zrozumiałe, a artykuły pozwalają prześledzić rozwój badań. Tekst wymaga jednak ujednolicenia terminologii i notacji oraz ograniczenia powtarzających się, bardzo szerokich deklaracji skuteczności. W szczególności należy konsekwentnie odróżniać dane przedziałowe, przedziałową reprezentację parametrów i skalarne reprezentanty. Korekta powinna także uzgodnić wykaz publikacji z danymi zamieszczonymi w samych artykułach. Drobne usterki redakcyjne mają mniejsze znaczenie niż wskazane wyżej kwestie precyzji opisu metod.

3. PYTANIA I ZAGADNIENIA DO DYSKUSJI PODCZAS OBRONY

Za najważniejsze zagadnienia do omówienia podczas obrony uważam:

1. Jakie reguły aktualizacji parametrów i definicje agregatorów zastosowano w implementacji oraz jak odnoszą się one do wskazanych zapisów w publikacjach?
2. Jak uzgadniane są lokalne podziały przedziałowe przed agregacją i jaka jest rola dwóch schematów ważenia opisanych w P9?
3. Jak obliczano miary w tabeli 2 publikacji P9 i jak zapewniono niezależność końcowej oceny od decyzji podejmowanych podczas uczenia?
4. Który element rozwiązania Doktorant uważa za swój najważniejszy wkład i jakie porównanie najlepiej wyodrębnić korzyść z tego elementu?

4. WNIOSKI I KONKLUZJA KOŃCOWA

Rozprawa przedstawia spójny program badań nad uczeniem federacyjnym z uwzględnieniem niepewności danych. Do najważniejszych osiągnięć zaliczam rozwój sposobów ważenia i agregacji modeli lokalnych, wykorzystanie entropii w konstrukcji agregatora oraz połączenie tych rozwiązań z reprezentacją modelu opartą na podprzedziałach. System regułowy opisany w P6 poszerza zakres podjętych badań. Łączny dorobek cyklu oceniam jako wartościowy wkład w rozwój metod uczenia maszynowego i agregacji informacji.

Przedstawiony materiał wykazuje znajomość właściwego aparatu matematycznego i informatycznego, umiejętność formułowania problemów badawczych oraz projektowania metod i eksperymentów. Na podstawie treści publikacji i przedstawionego opisu udziału Doktoranta oceniam, że praca potwierdza jego umiejętność samodzielnego prowadzenia badań. Wskazane uwagi odnoszą się do precyzji części zapisów oraz zakresu uzasadnienia wniosków eksperymentalnych. Wymagają rzeczowego wyjaśnienia, lecz nie zmieniają mojej pozytywnej oceny całości rozprawy.

Konkluzja. Stwierdzam, że rozprawa doktorska mgr inż. Jarosława Szkoły pt. „Uczenie federacyjne z uwzględnieniem niepewności danych: zaawansowane metody agregacji oparte na przedziałowych zbiorach rozmytych” spełnia wymagania określone w art. 187 ustawy z dnia 20 lipca 2018 r. – Prawo o szkolnictwie wyższym i nauce (t.j. Dz.U. z 2024 r. poz. 1571 ze zm.). Przedstawia oryginalne rozwiązanie problemu naukowego dotyczącego uwzględniania niepewności danych w uczeniu i agregacji modeli federacyjnych, prezentuje ogólną wiedzę teoretyczną Kandydata oraz potwierdza umiejętność samodzielnego prowadzenia pracy naukowej. Wnoszę do właściwej Rady Dyscyplin Uniwersytetu Rzeszowskiego o przyjęcie rozprawy i dopuszczenie mgr inż. Jarosława Szkoły do publicznej obrony w postępowaniu o nadanie stopnia doktora w dziedzinie nauk inżynieryjno-technicznych, w dyscyplinie informatyka techniczna i telekomunikacja.