

Prof. Dr hab. inż. Ewaryst Rafajłowicz

Członek korespondent PAN

Politechnika Wrocławska

Recenzja rozprawy doktorskiej

mgr inż. Tomasza Rutkowskiego

“Explainable Artificial Intelligence Based on Neuro-Fuzzy Approach in Application to Recommender Systems”

„Wyjaśnialna sztuczna inteligencja oparta o podejście neuronowo-rozmyte w zastosowaniu do systemów rekomendacji”

Niniejsza recenzja została napisana na zlecenie Rady Naukowej Dyscypliny Informatyka Polsko-Japońskiej Akademii Technik Komputerowych w Warszawie, w związku z przewodem doktorskim mgr inż. Tomasza Rutkowskiego.

1. Tematyka rozprawy

Recenzowana rozprawa dotyczy systemów rekomendujących, które charakteryzują się dodatkowymi cechami, w tym głównie objaśnialnością reguł wnioskowania. Podstawowym narzędziem, które wybrał Autor, do konstruowania takich systemów są systemy neuronowo-rozmyte. Trafność tego wyboru uzasadnię w dalszej części recenzji.

Doktorant sformułował cel i tezy rozprawy następująco:

„Cel: Zaproponowanie i zaimplementowanie wyjaśnialnych algorytmów w zastosowaniu do systemów rekomendacyjnych

- Teza 1: Użycie podejścia neuronowo-rozmytego w systemach rekomendacyjnych opartych o informacje o obiekcie daje możliwość generowania zrozumiałych wyjaśnień dla każdej rekomendacji.

Teza 2: Systemy rekomendacyjne oparte o podejście Neuronowo-Rozmyte mogą być jednocześnie dokładne, interpretowalne i transparentne”

Mgr T. Rutkowski przedstawił cel i tezy rozprawy w sposób jasny i weryfikowalny.

Doktorant bardzo trafnie wybrał narzędzie do konstruowania objaśnialnych systemów rekomendacyjnych w postaci sieci neuronowo-rozmytych, gdyż – z jednej strony – systemy takie są bardzo elastyczne w modelowaniu złożonych zjawisk i poddają się uczeniu, a z drugiej zaś – dają wyniki w postaci interpretowalnych reguł. Cechy te wyróżniają sieci neuronowo-rozmyte spośród innych podejść, w tym zwłaszcza od głębokich sieci neuronowych, dla których – jak dotąd – nie opracowano narzędzi, pozwalających zrozumieć dlaczego sieć „nauczyła się” rozwiązywania danego zadania w ten, a nie w inny sposób i dlaczego wybrała te, a nie inne cechy jako ważne.

Tematykę rozprawy uważam za aktualną i ważną zarówno dla teorii jak i praktyki informatyki.

2. Zawartość i kompozycja rozprawy

Rozprawa składa się ze streszczenia w języku polskim i angielskim, 5 rozdziałów, bibliografii i 6 dodatków.

Rozdział wstępny zawiera wszystkie wymagane elementy: sformułowanie i uzasadnienie celu, tezy i zakresu rozprawy oraz bardzo dobry przegląd literatury na temat:

- objaśnialnych systemów sztucznej inteligencji,
- podejść do interpretowalnych metod uczenia maszynowego,
- systemów rekomendacyjnych.

Interesująca jest koncepcja przeglądu literatury, gdyż Doktorant, niejako przy okazji, definiuje pojęcia używane w rozprawie i wskazuje źródła nie tylko stricte związane z pracą, ale daje także szeroką panoramę prawnych i etycznych aspektów sztucznej inteligencji, które prowadzą do formułowania dodatkowych wymagań, takich jak: odpowiedzialność (prawna i zawodowa), zachowanie prywatności, bezpieczeństwo i transparentność systemów oraz uczciwość wnioskowania (także w sensie biznesowym). Sposób prezentacji tych zagadnień wskazuje na szerokie czytanie mgr T. Rutkowskiego nie tylko w literaturze informatycznej, ale także w szerokiej gamie publikacji z nauk humanistycznych. Wstęp zawiera także wiele polemicznych stwierdzeń Autora w stosunku do cytowanych opinii przedstawicieli różnych nauk i wykładnię Jego własnych poglądów. Poglądy te – w sposób pośredni – rzutują na proponowane przez Doktoranta podejścia do rozwiązywania problemów.

Rozdział 2 zawiera wprowadzenie w tematykę systemów neuronowo-rozmytych, ale – jednocześnie – stanowi prezentację najważniejszej koncepcji rozprawy, a mianowicie takiego spojrzenia na te systemy, które pozwala implementować w ich języku objaśnialne rekomendacje. Autor zwraca szczególną uwagę na możliwości uczenia takich systemów (§ 2.2) oraz na

proponowane przez Niego uproszczenia znanych reguł wnioskowania rozmytego, które dają w efekcie prościej objaśnialne rekomendacje.

Najważniejsze, w moim odczuciu, są Rozdziały 3 i 4. Autor rozwija w nich koncepcję systemów rekomendujących w oparciu o wnioskowanie neuronowo-rozmyte. Rozdziały te opisują i dokumentują „kuchnię informatyczną” proponowanych systemów rekomendujących. Równoległe z prezentacją idei, Doktorant ilustruje je na przykładach dwóch klas systemów rekomendujących: wypożyczanie filmów (Rozdział 3) i gospodarowanie portfelem akcji firm (Rozdział 4). Przykłady te są oryginalne i świadomie dobrane tak, że w pierwszym z nich system rekomendujący ponosi mniejszą odpowiedzialność za nietrafne rekomendacje, w drugim zaś finansowe skutki nietrafnych rekomendacji mogą być bardzo duże. Ta istotna konstatacja doprowadziła Doktoranta do zaproponowania dwóch, istotnie różnych koncepcji tworzenia systemów rekomendujących.

W Rozdziale 3 Doktorant zaproponował koncepcje trzech objaśnialnych systemów rekomendacyjnych, oznaczanych dalej literami A, B i C. Wszystkie są wieloklasowymi systemami rekomendacyjnymi. Rozdział ten zawiera wiele oryginalnych pomysłów mgr T. Rutkowskiego, które podsumuję w dalszej części recenzji. Na tym etapie warto wskazać na dwa ciekawe sposoby transformowania danych symbolicznych (jakościowych) w dane numeryczne, potrzebne dalej do uczenia struktur neuronowo-rozmytych.

System A bazuje na wnioskowaniu typu Mamdaniego i wnioskowaniu Takagi-Sugeno-Kang's. Do optymalizacji takiego systemu Autor zastosował metodę „Grey Wolf” optymalizacji globalnej typu ewolucyjnego. Metoda ta jest relatywnie nowa i warto było zbadać jej skuteczność na przykładzie nietrywialnego zadania. Przy tej okazji Doktorant udokumentował jedną z tez rozprawy, wskazując na danych testowych konieczność szukania kompromisu między dokładnością metody oraz prostotą reguł decyzyjnych, tak by zachowały one interpretowalność.

System rekomendacyjny B poszerza A o dwa nowe elementy, a mianowicie, dopuszcza wynik w postaci rozmytej i ma wbudowany mechanizm redukcji nadmiarowości. Ten drugi aspekt zasługuje na szczególną uwagę, gdyż Doktorant zbadał wpływ kar za nadmierną rozbudowę modelu w postaciach: Akaike, FPE i reguły Schwartz'a i wskazał kryterium Akaike jako dobrze równoważące dokładność (związaną ze stopniem skomplikowania modelu) z interpretowalnością reguł decyzyjnych.

System C rozbudowany jest w kierunku dokładniejszego wnioskowania rozmytego. Jak Doktorant stwierdza, zabieg ten poprawia dokładność, ale kosztem utraty pełnej przejrzystości interpretacyjnej. Spośród systemów rekomendacyjnych A, B i C mgr T. Rutkowski rekomenduje system B.

Jak wspomniałem, system zaproponowany w Rozdziale 4 zasługuje na szczególną uwagę, gdyż w swej konstrukcji zawiera klasyfikację rekomendacji tylko do jednej klasy. Jest to ciekawa

koncepcja i bardzo dobrze dopasowana do rekomendacji o dużym znaczeniu. Jako studium przypadku Doktorant wybrał rekomendacje dotyczące zakupu akcji. Interesująca jest też koncepcja tworzenia rekomendacji na podstawie „najbliższych sąsiadów”. Oryginalnym elementem tej koncepcji jest sposób „mierzenia odległości”. Mianowicie, Doktorant zaproponował użycie stopnia aktywacji reguły jako miary podobieństwa. Ważny jest tu także sposób wyznaczania w/w stopnia aktywacji. W tym kontekście warto wspomnieć także o zaproponowanym przez Doktoranta sposobie wizualizacji danych wielowymiarowych na płaszczyźnie lub w przestrzeni 3D, którego nowym elementem jest właśnie wspomniana metoda oceny odległości.

Rozprawę kończy Rozdział 5, zawierający podsumowanie rozprawy i kilka interesujących kierunków badań, ale takich, które zdecydowanie wykraczają poza zakres rozprawy.

Warto także wspomnieć o 6 Dodatkach. Mają one ważną rolę dokumentacyjną sposobu badań i nakładu pracy. Na szczególną uwagę zasługuje Dodatek 3, w którym Doktorant przedstawia rzeczywiście interpretowalne reguły wnioskowania dla przykładu z Rozdziału 4.

Kompozycja rozprawy jest jasna, logiczna i dostosowana do jej tematyki. Zawartość rozprawy stanowi całościowy wywód, który w pełni dokumentuje Tezy 1 i 2 oraz wypełnienie jej celu.

3. Ocena najważniejszych rezultatów rozprawy.

Głównym celem rozprawy było „Zaproponowanie i zaimplementowanie wyjaśnialnych algorytmów w zastosowaniu do systemów rekomendacyjnych.„ Cel ten został osiągnięty, a proponowane podejścia zostały zweryfikowane empirycznie. Jak już wspomniałem, za najważniejszy rezultat mgr T. Rutkowskiego uważam samą koncepcję tworzenia objaśnialnych klas systemów rekomendacyjnych na bazie systemów-neuronowo-rozmytych.

Realizacja tej koncepcji wymagała od Autora dużej inwencji, także na poziomie rozwiązań bardziej szczegółowych. Zaliczam do nich:

- Opracowanie nowego sposobu oceny podobieństwa wielowymiarowych wektorów na podstawie stopnia aktywacji i użycia go do wizualizacji 2D i 3D.
- Wprowadzenie nowego sposobu objaśniania reguł działania systemu rekomendującego.
- Wkład w teorię systemów rozmytych, polegający na zaproponowaniu wariantów implikacji rozmytych, które dostosowane są do systemów rekomendacyjnych. W tym także propozycja budowania funkcji przynależności z użyciem histogramów.

Sugestie dalszych badań

Rozprawa napisana jest bardzo poprawnie po angielsku. Doktorant posiada bogate słownictwo oraz umiejętność jasnego i zwięzłego przedstawiania wyników. Nie

mam uwag szczegółowych do treści rozprawy. Zamiast nich, przedstawię swoje uwagi do dyskusji i sugestie dalszych badań.

- Warto opracować wizualizację zestawu reguł objaśniających działanie systemu rekomendującego.
- Można wyobrazić sobie zestaw („słabych”) systemów rekomendujących dla tego samego, ważnego problemu i głosowanie nad proponowanymi przez nie rekomendacjami. Podobna idea jest szeroko wykorzystywana w systemach rozpoznawania, więc ma szansę sprawdzić się także tutaj.
- Celowe może być wprowadzenie dodatkowych zmiennych, które wpływają na wynik rekomendacji, ale same nie są cechami rekomendowanych obiektów. W przypadku rekomendacji giełdowych takimi dodatkowymi zmiennymi mogą być zewnętrzne czynniki, takie jak – na przykład – pojawienie się korona-wirusa, które wywołują panikę na giełdzie. Podobne podejście sprawdziło się w problemach rozpoznawania i znane jest pod nazwą „kontekstu zewnętrznego”.
- W wielu miejscach rozprawy rozważania bazują, jawnie lub niejawnie, na współczynniku korelacji. Warto byłoby zbadać jak w tych lub podobnych przypadkach sprawdziłby się rangowy współczynnik korelacji Spearman’a.

Podkreślić należy, że powyższe sugestie daleko wykraczają poza ramy rozprawy.

KONKLUZJA

Podsumowując całokształt rozprawy doktorskiej mgr inż. Tomasza Rutkowskiego, stwierdzam, że rozprawa ta jest bardzo wartościowa, wnosi istotny wkład do aktualnego nurtu badań nad interpretowalnymi systemami rekomendacyjnymi i zawiera oryginalną, wielowariantową propozycję rozwiązania postawionego problemu naukowego.

W związku z tym stwierdzam, że rozprawa ta spełnia z nadmiarem wszystkie wymagania stawiane zwyczajowo i ustawowo rozprawom doktorskim i wnoszę o dopuszczenie jej do publicznej obrony.

Wysoką jakość rozprawy dokumentuje także 5 publikacji Doktoranta. Publikacje te są obiegu międzynarodowym, w tym 3 prace opublikowane zostały w wydawnictwach „wycenianych”

przez MNiSzW na 140 pkt. (dwie) i na 100 pkt. (jedna). W załączniku do recenzji przedstawiam wydruki w/w publikacji. Zgodnie z dostarczonym mi Regulaminem, z pełnym przekonaniem wnioskuję także o wyróżnienie tej rozprawy.

Sugeruję także Autorowi pilne (ze względu na aktualność tematyki) opublikowanie rozprawy jako książki w wydawnictwie o szerokim międzynarodowym obiegu.



Prof. Dr hab. inż. Ewaryst Rafajłowicz

Wrocław 8 marca 2020 roku

Załączniki: wydruk 3 publikacji Doktoranta