

Recenzja dorobku naukowego, organizacyjnego i dydaktycznego dr Przemysława Spurka

Niniejszą recenzję przygotowałem w związku z powierzeniem mi przez Radę Naukową Instytutu Podstaw Informatyki PAN roli recenzenta w postępowaniu habilitacyjnym dr Przemysława Spurka.

Charakterystyka przedłożonego osiągnięcia

Osiągnięciem przedłożonym przez dr Spurka jest cykl publikacji zatytułowany "Uczenie reprezentacji w metodach nienadzorowanych". Cykl składa się z 10 prac (A1-A10) i obejmuje **dwa wątki tematyczne**: zaawansowane algorytmy analizy skupień (grupowania), w klasycznym tego słowa znaczeniu, oraz uczenie reprezentacji z wykorzystaniem modeli głębokiego uczenia. Cykl klarownie wpisuje się w dyscyplinę Informatyka, a dokładniej w obszar uczenia maszynowego.

W pierwszym z wyżej wymienionych wątków Habilitant koncentrował się na wariantach metody opublikowanej w pracy A1, opartej na funkcji straty entropii krzyżowej. Metoda ta stanowi elegancką i jednocześnie prostą modyfikację algorytmu EM (*expectation maximization*), gdzie liniowa kombinacja rozkładów prawdopodobieństwa zastąpiona jest ich maksimum. Interesującą konsekwencją tej zmiany jest automatyczna penalizacja rozwiązania za obecność nadmiarowych skupień. Odpowiednie sformułowanie funkcji straty, opartej na entropii krzyżowej, pozwala na relatywnie łatwą adaptację podejścia do bazowych rozkładów prawdopodobieństwa innych niż rozkład normalny. Poza wywodem formalnym, autorzy zilustrowali działanie metody na kilku syntetycznych zbiorach danych oraz obrazach rastrowych.

Głównym przyczynkiem kolejnej pracy, A2, było uogólnienie powyższego pomysłu na grupowanie w przestrzeni różnorodności, co Habilitant osiągnął proponując ciągłe przekształcanie rozkładów normalnych poprzez podmianę argumentu funkcji gęstości rozkładu normalnego argumentem przekształconym odpowiednio dobraną funkcją (*active function* w nomenklaturze autora). Proponowane podejście umożliwia takie przekształcenie dla dowolnie wymiarowych przestrzeni, poprzez obranie jednego z wymiarów jako zmiennej niezależnej, gdzie metoda dokonuje tego wyboru autonomicznie, kierując się funkcją straty. W części empirycznej pracy porównano metodę z podejściami kontrolnymi na obrazach ręcznie pisanych znaków oraz zbiorach z repozytorium UCI.

W pracy A3 metoda ta została uogólniona na przypadek dynamicznie napływających danych (*online learning*), czyli zasadniczo strumieni danych. Habilitant zademonstrował jej skuteczność na kilku rzeczywistych zbiorach danych (ręcznie pisane znaki alfabetu chińskiego oraz opis kształtu obiektów trójwymiarowych reprezentowanych jako chmury punktów). Rozszerzenie to wymagało zaproponowania wyrafinowanego i wydajnego obliczeniowo schematu aktualizacji funkcji straty (*update* i *downdate*).

Z kolei w pracy A4 Habilitant wraz ze współautorami zademonstrował jak wcześniej wprowadzone formalizmy i metody daje się wykorzystać do budowania bardziej złożonych konfiguracji skupień. Osiąga to na dwa sposoby: 1) wykorzystując hierarchiczne podejście do grupowania, gdzie mniejsze modele Gaussowskie, opisujące części zbioru danych, łączone są w kolekcje (w oparciu o dywergencję Cauchy'ego-Schwarza, z racji wykorzystania alternatywnej definicji entropii, tj. entropii Renyi), oraz 2) przez grupowanie metodą Warda, która nie zakłada dostępności dobrze zdefiniowanych centroidów skupień. Działanie metody zilustrowane jest na zadaniach segmentacji obrazów pozyskiwanych w świetle widzialnym i obrazów mikroskopii elektronicznej.

Motywacją dla kolejnej pracy w przedłużonym cyklu, A5, była obserwacja iż rzeczywiste, skomplikowane i wielowymiarowe rozkłady prawdopodobieństwa wymagają zazwyczaj wielu komponentów Gaussowskich do ich wiernego zamodelowania, co przekłada się na znaczne wymagania obliczeniowe. Aby zaadresować ten problem, Habilitant zaproponował generalizację rozkładu normalnego (*general split Gaussian*), w którym wartości poniżej i powyżej wartości oczekiwanej modelowane są rozkładami normalnymi o z różnych wariancjach. Ten prosty zabieg przekłada się na znaczny wzrost elastyczności rozkładu w sensie możliwości dopasowywania go do danych. Habilitant zaproponował też inne uproszczenie tego schematu, w którym stopień asymetrii rozkładu zależny jest od wartości oczekiwanej, co przekłada się na dalsze zmniejszenie liczby parametrów modelu. Na szczególne uznanie zasługuje w tej pracy zilustrowanie proponowanego algorytmu na pięciu różnych zbiorach danych rzeczywistych: szeregach czasowych, danych z cytometrii przepływowej, danych tabelarycznych, danych z monitoringu turbin wiatrowych oraz segmentacji/progowania obrazów.

Nieco odrębnym tematem w pierwszym wątku tematycznym przedłożonego cyklu jest analiza składowych niezależnych (*independent component analysis*), która jest przedmiotem pracy A6. Proponowana tam metoda, podobnie jak klasyczna analiza składowych niezależnych, doszukuje się pewnej formy rozbieżności względem rozkładu normalnego w obserwowanych danych. Autor zaproponował, budując na poprzedniej pracy A5, aby skupić się na asymetrii rozkładu (w sensie skośności różnej od zera), przy jej modelowaniu z wykorzystaniem zaproponowanego w A5 rozkładu split Gaussian. Jedną z zalet tego podejścia jest lepsze niż w przypadku podstawowej metody modelowanie tak zwanych grubych ogonów rozkładów prawdopodobieństwa. Skuteczność metody została zilustrowana na danych syntetycznych ('rozplatanie' sztucznych sygnałów, tzw. *cocktail-party problem*) oraz na obrazach rastrowych.

Drugim wątkiem tematycznym przedłożonego w cyklu są metody uczenia się reprezentacji bazujące na sztucznych sieciach neuronowych, w szczególności modelach

uczenia głębokiego. Wątek ten jest reprezentowany w cyklu przez trzy tematy: modele generatywne, trójwymiarowe (3D) chmury punktów, oraz grafowe sieci splotowe.

W pracy A7 Habilitant zaproponował nowy typ autoenkodera wariacyjnego. Cechą wspólną autoenkoderów wariacyjnych jest odnoszenie rozkładu punktów generowanych w przestrzeni ukrytej (*latent space*) do pewnego rozkładu odniesienia. Oryginalnym przyczynkiem pracy A7 jest zaproponowanie i wykorzystanie nowej miary służącej do obliczania odległości pomiędzy rozkładami prawdopodobieństwa, budującej na wcześniejszych alternatywach dla podejścia wariacyjnego, tj. w szczególności na odległości Wassersteina. Proponowana metryka zasadza się twierdzeniu Cramer-Wolda, ustalającym równoważność pomiędzy równością rozkładów wielowymiarowych a równością wszystkich rzutów tych rozkładów. Nowatorskie przyczynki dotyczą 1) dekompozycji problemu poprzez pomiar komponentów odległości we wszystkich jednowymiarowych podprzestrzeniach naraz oraz 2) wykorzystanie estymacji jądrowej w powstałych w ten sposób jednowymiarowych przestrzeniach. Co istotne, metoda ta ma jawną postać analityczną i nie wymaga np. rozwiązywania problemów optymalizacyjnych, typowych dla autoenkoderów Wassersteina. Autorzy pracy wykazali także empirycznie (na znanych zbiorach rzeczywistych danych MNIST, Fashion MNIST, CIFAR10 i CELEB A) że tak zaprojektowana architektura ma bardziej pożądaną charakterystykę niż tradycyjne autoenkodery wariacyjne.

Kolejna praca przedłożonego cyklu, A8, zorientowana jest na zbiory danych reprezentowane jako chmury punktów (*point clouds*). Dane takie często zbiera się w robotyce i widzeniu komputerowym, ponieważ pozwalają one na charakteryzację sceny pełniejszą i bardziej niezawodną informacją o głębi (w porównaniu na przykład z konwencjonalną stereowizją). W ostatnich latach zaproponowano kilka interesujących architektur głębokich sieci neuronowych przystosowanych do przetwarzania tego typu danych. Pewnym wyzwaniem jest tu fakt iż każdy przykład, stanowiący chmurę punktów, może zawierać inną liczbę obserwacji, oraz iż punkty są nieuporządkowane, co wymaga stosowania modeli niezmienniczych ze względu na ich permutacje. Przyczynek Habilitanta w tym obszarze jest szczególnie interesujący, ponieważ zaproponowane w pracy A8 podejście zasadza się na koncepcji hipersieci (*hyper network*), czyli podejścia w którym jeden model neuronowy staje się generatorem parametrów (wag) dla innego modelu neuronowego. W kontekście pracy A8, docelowy model realizuje 'deformację' kulistego rozkładu jednorodnego tak aby odzwierciedlał on zadany kształt obiektu. Specyfika tego podejścia pozwala też łatwo przejść z reprezentacji chmur punktów na siatki triangulacyjne (*mesh*), ponieważ każdy punkt na sferze wyjściowego rozkładu jest jedynie oddalony lub przybliżony od/do środka sfery, a zatem oryginalna topologia zostaje zachowana.

W kolejnej pracy A9, która stanowi czasopismowe rozwinięcie pracy A8, Habilitant ze współautorami uogólnił podejście prezentowane w A8 na wariant w którym sieć docelowa implementuje model *continuous normalization flow*. Wymagało to zaprojektowania innego wyjściowego rozkładu prawdopodobieństwa – w miejsce rozkładu normalnego, sferycznego rozkładu log-normalnego. W efekcie generowane chmury punktów charakteryzują się większą równomiernością gęstości niż w podejściu referencyjnym. Praca obfituje w wiele wyników eksperymentalnych i wizualizacji otrzymanych chmur punktów i siatek.

Ostatnia z prac w przedłużonym cyklu, A10, proponuje architekturę grafowej sieci neuronowej przystosowaną do przetwarzania grafów, w których wierzchołki atrybutów etykietowane są współrzędnymi przestrzennymi. Motywacją jest udostępnienie modelowi bogatszej informacji o strukturze grafu, w sensie lokalizacji wierzchołków, między innymi po to aby zapewnić rozróżnialność poszczególnych sąsiadów danego węzła grafu. Innymi słowy, jest to model który reinterpretuje grafy jako obiekty zanurzone w przestrzeniach kartezjańskich. Habilitant zilustrował działanie tego modelu na przykładach grafów rozpiętych na rastrach obrazów cyfrowych, oraz grafów reprezentujących dwuwymiarowe formuły związków chemicznych (Habilitant załączył dwie wersje pracy A10, preprint arXiv oraz zaakceptowaną pracę konferencyjną). W tej drugiej domenie zadaniem była predykcja wybranych właściwości związku chemicznego (jedno zadanie klasyfikacji i dwa zadania regresji) na podstawie jego formuły reprezentowanej jako graf. Eksperyment wykazał zdecydowaną przewagę podejścia autorów nad ośmioma metodami kontrolnymi.

Ocena osiągnięcia

Prace wchodzące w skład cyklu ukazały się w większości w bardzo dobrych czasopiśmie (między innymi *Pattern Recognition*, *Journal of Machine Learning*, *IEEE Transactions on Pattern Analysis and Machine Intelligence*) i materiałach konferencyjnych renomowanych konferencji, np. ICML i IJCNN. Jedna z przedłożonych prac została przygotowana przez Habilitanta samodzielnie, a pozostałe mają w większości niewielką liczbę autorów, w tym dwie publikacje z jednym współautorem i dwie publikacje z dwójką współautorów. Dr Spurek jest pierwszym autorem siedmiu prac cyklu. Szacowany deklarowany wkład Habilitanta w większość tych prac jest wysoki, przynajmniej 50% w siedmiu pracach cyklu; załączone oświadczenia współautorów wydają się potwierdzać te deklaracje.

Dr Spurek przystępnie prezentuje proponowane podejścia, a prace są profesjonalnie zredagowane i bogato ilustrowane. Habilitant był także autorem implementacji programistycznych niektórych z proponowanych algorytmów, co utwierdza moją pozytywną opinię o jego umiejętnościach. Doceniam także staranność w przygotowaniu autoreferatu, który opisuje przedłożony cykl w sposób bardzo przystępny i jednocześnie zwarty.

Moja ocena przyczynków merytorycznych przedłożonego cyklu jest zdecydowanie pozytywna. Habilitant zaproponował wiele nowatorskich i oryginalnych algorytmów, architektur głębokich sieci neuronowych i skojarzonych z nimi funkcji straty i innych komponentów, oraz zademonstrował użyteczność proponowanych podejść w zróżnicowanej gamie zastosowań praktycznych. Proponowane podejścia bazują na nietuzinkowych klasach modeli (np. hipersieci, sieci grafowe, autoenkodery) które reprezentują obiecujące i aktualne kierunki w uczeniu maszynowym i sztucznej inteligencji). Szczególnie doceniam bardzo dobre ugruntowanie proponowanych metod w podstawach teoretycznych, w połączeniu z demonstracją ich skuteczności na rzeczywistych problemach. Dr Spurek zaproponował metody pracujące z różnymi typami danych (obrazy, chmury punktów, grafy) oraz uczące się w różnych trybach (offline i online), co stanowi przekonującą ilustrację wszechstronności Habilitanta oraz

uniwersalności proponowanych przez niego rozwiązań. Dorobek ten jest zwarty tematycznie i stanowi przekonujące świadectwo szerokich kompetencji oraz solidnego warsztatu naukowego.

Z osobistej perspektywy szczególnie doceniam starania Habilitanta zmierzające do odkrywania i modelowania ukrytych struktur w obserwowanych danych. Uważam że jest to jedno z centralnych zagadnień we współczesnym uczeniu maszynowym. Podzielam także przekonanie Habilitanta iż studiowanie tego typu podejść w kontekście uczenia nienadzorowanego jest szczególnie uzasadnione, w odróżnieniu od uczenia nadzorowanego, gdzie jesteśmy niejako skazani na prymat konkretnej, w pewnej sensie przypadkowej (ad-hoc), zmiennej zależnej.

Zasadniczo nie mam znaczących uwag do przedłożonego cyklu publikacji. Pozwolę sobie jedynie na drobne uwagi odnośnie autoreferatu. Na str. 22, wydaje się że $n=|X|$, a zatem nie ma chyba sensu używać obu tych alternatywnych zapisów. Sformułowanie “dane które pojawiają się w sposób ciągły” (str. 13) jest dość nietypowe i potencjalnie mylące, ponieważ zdarzenia pojawiania się nowych obserwacji mają jednak charakter dyskretny (tj. nowe punkty nie ‘wyłaniają się’ stopniowo). Myślę że jednak lepiej jest mówić tu o uczeniu się ze strumieni danych (*data streams*), lub nawet używać angielskiego terminu “online”, tj. *online learning*. Dalej, wydaje się że bardziej adekwatnie byłoby zatytułować przedłożony cykl “Uczenie reprezentacji w metodach uczenia nienadzorowanego” a nie “metodach nienadzorowanych”. I ostatnia uwaga: choć nie jestem purystą językowym, to pewne sformułowania w autoreferacie nieco rażą, np. “zwraca wagi TargetNetworka” (s. 44). Te drobne uwagi nie mają wpływu na moją wysoką ocenę przedłużonego osiągnięcia.

Ocena pozostałego dorobku

Poza przedłożonym cyklem Habilitant po uzyskaniu stopnia doktora opublikował przynajmniej 30 innych prac w podobnych i powiązanych obszarach tematycznych (uczenie maszynowe, analiza skupień, analiza składowych niezależnych) oraz dotyczących zastosowań w biologii i medycynie. Także i te prace ukazały się w dobrych i bardzo dobrych czasopismach (między innymi *Knowledge-based Systems*, *PLOS One*, *IEEE Access*, *IEEE Transactions on Pattern Analysis and Machine Intelligence*) oraz w materiałach najbardziej liczących się konferencji, między innymi NeurIPS, ICONIP, i ESANN.

Dr Spurek po obronie doktoratu wygłosił 5 referatów zaproszonych i 13-krotnie prezentował wyniki swoich prac na konferencjach międzynarodowych. Jest członkiem komitetów programowych (jako recenzent) bardzo dobrych konferencji międzynarodowych, między innymi ICLR, AAAI, CVPR i NeurIPS i w komitetach programowych innych wydarzeń. Recenzował też prace zgłaszane na te konferencje.

Habilitant był kierownikiem jednego grantu typu Sonata, a obecnie kieruje projektem typu Opus. Poza tym był wykonawcą w czterech innych projektach. Dr Spurek odbył też 4 staże naukowe, z czego 2 po uzyskaniu stopnia doktora: jeden w Instytucie Podstaw Informatyki PAN, drugi w Technical University of Ostrava (Czechy). Habilitant prowadzi aktywną współpracę z innymi zespołami i ośrodkami w kraju i za granicą, czego dowodem jest publikowanie prac ze

współautorami z ośrodków zagranicznych, w tym renomowanych, na przykład University of Cambridge, University of Oxford, czy CERN.

Habilitant jest też autorem lub współautorem przynajmniej dwóch pakietów programistycznych udostępnionych w repozytorium pakietów języka programowania R.

Dr Spurek uzyskał dwie nagrody zespołowe Rektora Uniwersytetu Jagiellońskiego, a także nagrodę best paper na jednej z konferencji w której uczestniczył. Współpracował także z przemysłem w ramach dwóch projektów finansowanych przez podmioty komercyjne.

Wskaźniki bibliometryczne dr Spurka są bardzo dobre: całkowita liczba cytowań według serwisu Google Scholar na moment przygotowywania dokumentacji to 388 (razem z autocytowaniami); na moment przygotowywania niniejszej recenzji wskaźnik ten osiąga już wartość 504. Po odrzuceniu autocytowań, Habilitant może pochwalić się nadal wysoką liczbą 241 cytowań na moment składania wniosku. Indeks Hirscha Habilitanta według serwisu Google Scholar wynosi 10 gdy uwzględnić autocytowania, w przeciwnym razie 8; serwis Web of Science podaje odpowiednio wartości 6 i 5. Zatem także w kategoriach wskaźników bibliometrycznych dorobek Habilitanta prezentuje się bardzo przekonująco.

Odnosnie aktywności dydaktycznej, dr Spurek wypromował po uzyskaniu stopnia doktora 15 magistrantów i prowadził zajęcia z przynajmniej 10 przedmiotów, głównie ćwiczenia i laboratoria. Prowadził też autorskie wykłady dla studentów z kilku przedmiotów oraz seminarium dla doktorantów. Jest obecnie promotorem pomocniczym dwóch doktorantów.

Podsumowanie oceny

Całość dorobku naukowego przedłożonego przez doktora Przemysława Spurka, zarówno w części zawartej w przedłużonym cyklu, jaki w pozostałych publikacjach, oceniam bardzo wysoko. Badania prowadzone przez Habilitanta charakteryzują się nowatorskim charakterem, profesjonalizmem, wpisują się w najnowsze trendy w uczeniu maszynowym i sztucznej inteligencji, przekładają się na wartościowe implikacje teoretyczne i praktyczne, i są też zdecydowanie relewantne dla innych badaczy. Dorobek dydaktyczny i organizacyjny oraz zakres współpracy międzynarodowej Habilitanta są także znaczące i spełniają w mojej ocenie zwyczajowe i ustawowe wymagania.

Podsumowując, całokształt dorobku naukowego, organizacyjnego i dydaktycznego doktora Przemysława Spurka spełnia w mojej ocenie z nawiązką wymagania zwyczajowo i ustawowo stawiane kandydatom do uzyskania stopnia doktora habilitowanego. Moja ocena Kandydata jest zdecydowanie pozytywna, a zatem wnioskuję do Rady Naukowej Instytutu Podstaw Informatyki Polskiej Akademii Nauk o dopuszczenie dr Spurka do dalszych etapów postępowania habilitacyjnego.

