

**RADA NAUKOWA DYSCYPLINY
INFORMATYKA TECHNICZNA I TELEKOMUNIKACJA POLITECHNIKI WARSZAWSKIEJ**

zaprasza na
OBRONĘ ROZPRAWY DOKTORSKIEJ

mgr. inż. Łukasza Tomasza Neumanna

która odbędzie się w dniu **8 stycznia 2024 roku**, o godzinie **12:00** w trybie hybrydowym

Temat rozprawy:

„Training instabilities in neural network-based sequential data modeling”

Promotor: dr hab. inż. Robert Nowak, prof. uczelni – Politechnika Warszawska

Recenzenci: dr hab. inż. Jarosław Bilski – Politechnika Częstochowska

dr hab. inż. Dariusz Mrozek – Politechnika Śląska

prof. dr hab. inż. Andrzej Obuchowicz-Uniwersytet Zielonogórski

Obrona odbędzie się zdalnie na platformie MS Teams/ w Sali nr 116. Osoby zainteresowane uczestnictwem w obronie proszone są o zgłoszenie chęci uczestnictwa w formie elektronicznej na adres sekretarza komisji: dr hab. inż. Piotr Gawrysiak, email: piotr.gawrysiak@pw.edu.pl, do dnia 5 stycznia 2024. godz. 17:00.

Z rozprawą doktorską i recenzjami można zapoznać się w Czytelni Biblioteki Głównej Politechniki Warszawskiej, Warszawa, Plac Politechniki 1.

Streszczenie rozprawy doktorskiej i recenzje są zamieszczone na stronie internetowej: <https://www.bip.pw.edu.pl/Postepowania-w-sprawie-nadania-stopnia-naukowego/Doktoraty/Wszczete-po-30-kwietnia-2019-r/Rada-Naukowa-Dyscypliny-Informatyka-Techniczna-i-Telekomunikacja/mgr-inz.-Lukasz-Neumann>

Przewodniczący Rady Naukowej Dyscypliny
Informatyka Techniczna i Telekomunikacja
Politechniki Warszawskiej
dr hab. inż. Jarosław Arabas, prof. uczelni

Training instabilities in neural network-based sequential data modeling

Abstract.

This thesis presents a series of publications regarding instabilities during neural networks' optimization process (also called training) in sequential data modeling. In our study, we investigate two approaches to solving such problems: specialized neural architectures and second-order optimizers.

In the first part of our work, we focus on custom architectures reducing problems with training instabilities. First, we show that a combination of fixed state processing and custom convolutional neural network architecture is sufficient to successfully process a series of medical images. Based on this approach, we introduce the first fully-automatic allergic skin reaction recognition system. We show that our method yields accurate results comparable between patients, as opposed to current, manual state-of-the-art. Next, we propose a minimal gated recurrent neural architecture with an arbitrary, nonlinear state transformation. We theoretically show that this architecture alleviates gradient propagation issues while using fewer parameters than its competitors. Further, we present experimental results which indicate that our method often outperforms state-of-the-art solutions, especially on problems with high state drift. Interestingly, we note that the recurrent neural network's performance degrades with the depth of the state transformation. This observation holds on several tasks for two different architectures.

In the second part, we look closer at second-order optimization algorithms as an alternative to currently often-used gradient ones. To this end, we propose a modification of an evolution strategy algorithm, tailoring it for neural network training. We show that such an approach allows for optimizing both convolutional and recurrent neural networks. Notably, the proposed method can effectively train Jordan recurrent neural networks on tasks with significant gradient propagation problems, in contrast to the state-of-the-art first-order method. Finally, we conclude that combining second-order optimization with gradient information results in the best-performing models.

Overall, our research contributes to the mitigation strategies against training instabilities of neural architectures on sequential data. We introduce two new architectural and one optimization solution to the instability problems. One of the architectural solutions serves as a basis for the novel, real-life medical data analysis system.

Keywords: Sequential Data, Recurrent Neural Networks, Neural Network Training, Optimization

Niestabilności w procesie uczenia sieci neuronowych przy modelowaniu danych sekwencyjnych

Streszczenie.

W niniejszej pracy przedstawiamy serię publikacji poświęconej zagadnieniom niestabilności w procesie uczenia (optymalizacji) sieci neuronowych przy modelowaniu danych sekwencyjnych. W ramach przedstawionych prac badamy dwa podejścia do rozwiązania problemów związanych z niestabilnościami: dedykowane architektury sieci neuronowych oraz optymalizatory drugiego rzędu.

W pierwszej części pracy skupiamy się na architekturach sieci neuronowych zaprojektowanych na potrzeby niwelacji niestabilności w procesie uczenia. Pokazujemy, że połączenie z góry narzuconego, nieelastycznego sposobu przetwarzania stanu oraz dedykowanej konwolucyjnej sieci neuronowej wystarczy aby poprawnie przetwarzać serię obrazów medycznych. Na podstawie tego podejścia wprowadzamy pierwszy w pełni automatyczny system do rozpoznawania reakcji alergicznych skóry. Zaproponowana przez nas metoda otrzymuje dobre wyniki, które mogą być porównywalne pomiędzy pacjentami, w przeciwieństwie do obecnego, manualnego podejścia. Następnie prezentujemy minimalną rekurencyjną sieć neuronową z mechanizmem bramkującym oraz arbitralnym, nieliniowym przekształceniem stanu. Pokazujemy teoretycznie, że zaproponowana architektura rozwiązuje problemy z propagacją gradientu, przy użyciu mniejszej liczby parametrów niż konkurencyjne architektury. Ponadto demonstrujemy wyniki eksperymentalne, które wskazują, że nasz model często uzyskuje lepsze wyniki w porównaniu do obecnie używanych modeli, w szczególności na problemach z dużym dryfem stanu. W wyniku eksperymentów dla dwóch różnych architektur na kilku różnych zadaniach dochodzimy do ciekawego wniosku - jakość wyników rekurencyjnej sieci neuronowej spada wraz z głębokością transformacji stanu.

W drugiej części pracy skupiamy się na optymalizatorach drugiego rzędu jako alternatywie dla obecnie często używanych algorytmów bazujących na gradiencie. W ramach tych prac prezentujemy modyfikację algorytmu strategii ewolucyjnej, którą przystosowujemy do treningu sieci neuronowych. Pokazujemy, że nasz algorytm pozwala na optymalizację konwolucyjnych i rekurencyjnych sieci neuronowych. W szczególności zaproponowana metoda potrafi nauczyć rekurencyjną sieć neuronową Jordana na zadaniach z poważnymi problemami propagacji gradientu, które nie są rozwiązywane przez metody pierwszego rzędu. Pokazujemy również, że łącząc optymalizację drugiego rzędu z informacją o gradiencie otrzymujemy najlepiej działające modele.

Podsumowując, nasza praca przyczynia się do rozszerzenia strategii mitygacji niestabilności treningu sieci neuronowych przy modelowaniu danych sekwencyjnych. Wprowadzamy dwa architektoniczne i jedno optymalizacyjne rozwiązanie problemów ze stabilnością procesu optymalizacji sieci neuronowych. Jedno z architektonicznych rozwiązań stanowi podstawę nowatorskiego systemu analizy danych medycznych.

Słowa kluczowe: Dane Sekwencyjne, Rekurencyjne Sieci Neuronowe, Trening Sieci Neuronowych, Optymalizacja

Częstochowa, 10.11.2023

dr hab. inż. Jarosław Bilski, prof PCz
Politechnika Częstochowska
Katedra Inteligentnych Systemów Informatycznych
ul. Armii Krajowej 36
42 – 200 Częstochowa
e-mail: jaroslaw.bilski@pcz.pl

**Recenzja rozprawy doktorskiej
mgr inż. Łukasza Neumanna
zatytułowanej:**

**Training Instabilities in Neural Network-based Sequential
Data Modelling**

Niniejsza opinia została przygotowana na prośbę Dziekana Wydziału Elektroniki i Technik Informacyjnych Politechniki Warszawskiej prof. dra hab. inż. Michała Malinowskiego oraz Przewodniczącego Rady Naukowej Dyscypliny Informatyka Techniczna i Telekomunikacja dra hab. inż. Jarosława Arabasa prof. uczelni., z dnia 1 sierpnia 2023 roku. Promotorem pracy jest dr hab. inż. Robert Nowak prof. uczelni.

1. Problem badawczy i jego znaczenie

Problematyka przedstawiona w przesłanej do recenzji rozprawie doktorskiej mieści się w zakresie dyscypliny naukowej Informatyka techniczna i telekomunikacja. Tematyka rozprawy ma zdecydowanie charakter naukowy i współgra z innymi pracami badawczym podejmowanymi aktualnie zarówno w ośrodkach krajowych, jak i zagranicznych. Badany w pracy problem niestabilności podczas procesu uczenia (optymalizacji) sieci neuronowych w sekwencyjnym modelowaniu danych jest zagadnieniem złożonym i zależnym od wielu różnych czynników. Autor bada możliwość rozwiązania postawionego problemu na dwa sposoby: pierwszy polega na zastosowaniu opracowanych w tym celu wyspecjalizowanych architektur sieci

neuronowych, drugi na użyciu algorytmów drugiego rzędu. Autor w rozprawie zaproponował wykorzystanie różnych rodzajów sztucznych sieci neuronowych zarówno jednokierunkowych, konwolucyjnych, jak i rekurencyjnych. Przeprowadził również szereg badań dla różnych problemów syntetycznych i rzeczywistych. Na szczególną uwagę zasługuje opracowany system rozpoznawania reakcji alergicznych skóry, który pokazuje praktyczne zastosowanie prac autora. Bez wątpienia zaproponowana problematyka badawcza rozprawy ma duże znaczenie dla badania nowych architektur sieci neuronowych i algorytmów ich uczenia oraz pokazuje możliwości ich stosowania w rozwiązywaniu rzeczywistych problemów.

2. Zawartość rozprawy doktorskiej

Przedstawiona do recenzji rozprawa doktorska pod tytułem: *Training Instabilities in Neural Network-based Sequential Data Modelling (Niestabilności w procesie uczenia sieci neuronowych przy modelowaniu danych sekwencyjnych)* obejmuje w sumie 86 stron i oprócz wstępu, streszczenia w językach angielskim i polskim składa się z 4 rozdziałów podstawowych, zebranych osiągnięć naukowych autora, spisu literatury obejmującego ponad 100 pozycji i dwóch dodatków. W dodatku drugim umieszczone zostały trzy publikacje, na których bazuje praca. Rozprawa napisana jest w języku angielskim i składa się z następujących rozdziałów:

Rozdział 1. **Introduction (Wstęp)** - przedstawia pojęcie zbioru sekwencyjnego oraz podaje szereg przykładów takich zbiorów, problemy jakie występują przy ich przetwarzaniu, a także różne podejścia do modelowania danych sekwencyjnych. Następnie autor wskazuje wkład wnoszony w publikacjach będących podstawą pracy:

- wprowadzenie nowatorskiej architektury rekurencyjnej sieci neuronowej rozwiązującej problem eksplozji i zanikania gradientu przy jednoczesnym zmniejszeniu liczby parametrów w porównaniu z innymi sieciami;
- opracowanie opartej na metaheurystyce nowej metody uczenia sieci neuronowych;
- przedstawienie sposobu poprawy wydajności modelu w zadaniu klasyfikacji danych medycznych oraz zaproponowanie zautomatyzowanego podejścia do rozpoznawania reakcji alergicznych skóry.

Doktorant podaje tu również listę trzech publikacji (w czasopiśmie i na dwóch konferencjach), które są podstawą recenzowanej pracy (wszystkie zgodnie z listą ministerialną uzyskały po 140 punktów) oraz określa własny wkład w ich opracowanie (70%, 70% i 33,3%).

Rozdział 2. **Background (Kontekst pracy)** – składa się z trzech podrozdziałów. Pierwszy z nich opisuje metody uczenia (treningu) sieci neuronowych. Podanych jest wiele powszechnie stosowanych (w tym nowych) algorytmów pierwszego i drugiego rzędu oraz algorytmy uczenia ewolucyjnego. Podrozdział drugi zawiera zwięzły opis architektur sieci neuronowych stosowanych do modelowania danych sekwencyjnych. Do podstawowych należą: sieci rekurencyjne (RNN), sieci LSTM (Long Short-Term Memory), GRU (Gated Recurrent Unit), ich modyfikacje i inne. Podrozdział trzeci zawiera informacje nt. metod diagnostyki alergicznych reakcji skórnych, możliwości zastosowania termografii oraz automatyzacji procesu rozpoznawania reakcji alergicznych.

Rozdział 3. **Mitigating stability problems using custom neural architectures (Łagodzenie problemów ze stabilnością przy użyciu dedykowanych architektur neuronowych)** – odnosi się do publikacji oznaczonych jako P1 i P2. W pracy P1 pokazano zautomatyzowane podejście do rozpoznawania alergicznych reakcji skórnych w oparciu o sieci neuronowe. Przedstawiono procedurę postępowania oraz przeprowadzono badania kliniczne. W artykule P2 przedstawiono nową architekturę sieci rekurencyjnej opartej o komórkę DMU (Deep Memory Update), a wywodzącą się z sieci GRU. Dalej proponuje się metodę inicjalizacji otrzymanej sieci neuronowej, aby poprawić jej stabilność i skrócić czas uczenia.

Rozdział 4. **Novel Evolution Strategy-based neural network optimizer (Nowy optymalizator sieci neuronowej oparty na strategii ewolucji)** – oparty jest na pracy P3 i zawiera algorytm drugiego rzędu nazwany neural Differential Evolution Strategy (nDES). Tworzona jest populacja sieci, która podlega ewolucji. Przeprowadzono badania dla kilku problemów testowych.

Rozdział 5. **Discussion and final remarks (Dyskusja i uwagi końcowe)** – podsumowuje wnioski otrzymane w prezentowanych pracach. Przedstawia najważniejsze oryginalne osiągnięcia. Na końcu wskazuje kilka możliwości dalszego rozwijania przedstawionych rozwiązań w formie pytań.

3. Ocena pracy

Oceniana rozprawa doktorska jest w rzeczywistości połączeniem trzech opublikowanych wcześniej artykułów (podane tylko tytuły):

P1: Thermography based skin allergic reaction recognition by convolutional neural networks

P2: Least Redundant Gated Recurrent Neural Network

P3: Deep Neuroevolution: Training Neural Networks Using a Matrix-Free Evolution Strategy

W tym kontekście ocena pracy musi się do nich odnieść. Artykuł P1 przedstawia bardzo ciekawy przykład zastosowania sieci neuronowych do rozpoznawania alergicznych reakcji skórnych. Pokazano zautomatyzowane podejście do w oparciu o sieci neuronowe. Przedstawiono całościową procedurę postępowania oraz przeprowadzono badania kliniczne na rzeczywistych danych stu pacjentów. W badaniach zastosowano dwa rodzaje obrazów: w paśmie widzialnym i podczerwonym (obrazy termograficzne) przed i po iniekcji. Uzyskane obrazy podlegają segmentacji, korelacji i klasyfikacji przez konwolucyjną sieć neuronową. Do uczenia sieci zastosowano algorytm ADAM. Najważniejsze uzyskane osiągnięcia to:

- Udane zastosowanie dedykowanej architektury sieci CNN przewyższające inne rozwiązania.
- Zastosowanie różnicowych obrazów termograficznych zwiększające skuteczność rozpoznawania odczynów alergicznych skóry.
- Zbudowanie automatycznego systemu do rozpoznawania odczynów alergicznych skóry.

W artykule P2 doktorant wprowadza nową rekurencyjną architekturę neuronową o nazwie DMU (Deep Memory Update) Będącą udanym uproszczeniem architektury GRU (Gated Recurrent Units) i LSTM (Long Short-Term Memory). Jako algorytmów uczących użyto SGD i ADAM. Przeprowadzono szereg eksperymentów dla syntetycznych i rzeczywistych przykładów i porównano otrzymane wyniki z innymi architekturami uzyskując w większości lepsze rezultaty. Należy podkreślić, że:

- Zaproponowana rekurencyjna architektura DMU umożliwia dowolną transformację stanu pamięci.
- DMU jest często wydajniejsza niż inne sieci RNN.
- DMU posiada mniej parametrów wymagających uczenia niż sieci LST, GRU i inne.

Artykuł P3 zawiera rozwiązanie ewolucyjne nazwane nDES (neural Differential Evolution Strategy) i wywodzące się z algorytmu DES (opartego na połączeniu DE i CMA-ES(Covariance Matrix Adaptation Evolution Strategy)). W opisaniej metodzie użyto inicjalizacji Xavier'a i mutacji opartej o gradient. Dodatkowo opracowano warianty zaproponowanego algorytmu. Pomyślnie przeprowadzono badania kilku

problemów testowych z wykorzystaniem opracowanych metod i algorytmu ADAM. W artykule uzyskano następujące rezultaty:

- Opracowano nowy algorytm ewolucyjnego uczenia sieci neuronowych nDES w trzech wariantach.
- Uzyskano wysoką skuteczność opracowanego algorytmu również dla architektur RNN (często lepszą od algorytmu ADAM).

Bardzo pozytywnie należy również ocenić inną działalność naukową doktoranta. We współpracy z promotorem pracy i innymi autorami opublikował szereg prac związanych z sieciami neuronowymi. Publikował również w ramach ALICE Collaboration.

Dodatkowo doktorant uczestniczył w wielu projektach badawczych. Był również kierownikiem jednego z nich.

4. Inne uwagi

Do recenzowanej pracy można przedstawić kilka uwag i komentarzy o charakterze dyskusyjnym:

- Praca jest napisana w konwencji rozszerzony wstęp/streszczenie plus trzy publikacje. Spójny, ciągły wywód bardziej pasuje do tego typu publikacji.
- W pracy doktorant nie sformułował tezy. W jej miejscu umieścić wkład/osiągnięcia pracy.
- Rozdział drugi Background zawiera szeroki przegląd metod uczenia i architektur sieci neuronowych. Niestety opis jest skrótowy. W większości przypadków brakuje podstawowych wzorów, żeby nie wspomnieć o wyprowadzeniu choćby wybranych najważniejszych algorytmów.
- Rozdziały 3 i 4 są nadmiernym streszczeniem załączonych publikacji. W rozdziałach tych można było np. przedstawić pełne wyprowadzenie opracowanych algorytmów uczenia.
- W kontekście publikacji P1 brakuje prezentacji systemu utrzymującego ramię pacjenta w zalecanej pozycji.
- Generalnie w pracy nie ma informacji o szybkości uczenia przy użyciu badanych algorytmów (z jednym wyjątkiem) w porównaniu z innymi metodami.
- Doktorant wspomina o możliwości prostego zrównoleglenia opracowanych algorytmów, należy jednak pamiętać, że większość metod uczenia sieci neuronowych można zrównoleglić i w rezultacie będzie liczyć się stosunek wydajności tych metod.

- Występują drobne błędy językowe i niejasne sformułowania.

Wymienione powyżej uwagi i niedoskonałości nie wpływają na pozytywną ocenę merytoryczną recenzowanej pracy doktorskiej.

Należy również zauważyć, że autor w rozdziale 5. pracy wskazuje kierunki dalszych badań, co pozwala oczekiwać na kolejne jego rozwiązania i publikacje.

5. Wiedza kandydata

Autor rozprawy posiada niewątpliwie rozległą wiedzę w zakresie uczenia maszynowego, metod optymalizacji oraz badanego szczegółowego problemu rozpoznawania reakcji alergicznych skóry, którą przedstawia w kolejnych rozdziałach rozprawy. Wiedza ta jest właściwie stosowana przy wprowadzaniu nowych architektur sieci neuronowych, opracowywaniu algorytmu uczenia i przeprowadzaniu wielu badań dla różnych typów sieci neuronowych i problemów badawczych.

Podana przez autora bibliografia (ponad 100 pozycji) uwzględnia aktualny stan literatury światowej, cytowane są również nowe prace w tym zakresie.

6. Konkluzja

W konkluzji stwierdzam, że praca doktorska „*Training Instabilities in Neural Network-based Sequential Data Modelling*”, przedstawiona przez mgr inż. Łukasza Neumanna, **spełnia wymagania stosownej ustawy o stopniach naukowych i tytule naukowym**. Wnoszę o jej przyjęcie i dopuszczenie do publicznej obrony.



.....
Podpis

Gliwice, 20 listopada 2023

Dr hab. inż. Dariusz Mrozek, prof. PS
Katedra Informatyki Stosowanej
Politechnika Śląska w Gliwicach
ul. Akademicka 16
44-100 Gliwice

RECENZJA

rozprawy doktorskiej dla

Rady Naukowej Dyscypliny Informatyka Techniczna i Telekomunikacja
działającej

w Politechnice Warszawskiej

Tytuł rozprawy: Training Instabilities in Neural Network-based Sequential Data Modeling

Autor rozprawy: mgr inż. Łukasz Neumann

1. Jakie zagadnienie naukowe jest rozpatrzone w pracy (teza rozprawy) i czy zostało ono dostatecznie jasno sformułowane przez Autora? Jaki charakter ma rozprawa (teoretyczny, doświadczalny, inny)?

Przedstawiona przez Pana Łukasza Neumanna rozprawa doktorska składa się z cyklu powiązanych tematycznie publikacji wraz z towarzyszącym im opracowaniem (autoreferatem), który stanowi przewodnik po zrealizowanych pracach badawczych i szereguje wiedzę w omawianym obszarze. W ogólnym ujęciu rozprawa jest poświęcona redukcji niestabilności w procesie uczenia sztucznych sieci neuronowych wykorzystywanych do analizy danych sekwencyjnych pozyskiwanych w różnych obszarach. Główne cele rozprawy koncentrują się wokół zagadnienia poprawy stabilności procesu uczenia sieci, poprzez zastosowanie technik optymalizacji drugiego rzędu i odpowiednio zaprojektowanych architektur sztucznych sieci neuronowych. Zarówno cele pracy, jak i motywacja prowadzonych badań w tym obszarze zostały sformułowane w sposób jasny i wyczerpujący. Charakter rozprawy określiłbym jako **eksperymentalny**, ponieważ Autor:

- zaproponował szereg usprawnień dla rozwiązań architektonicznych sztucznych sieci neuronowych oraz metaheurystykę dla procesu ich trenowania,
- dla potwierdzenia słuszności przyjętych rozwiązań przeprowadził badania eksperymentalne na zgromadzonych, a także publicznie dostępnych zbiorach danych, które pozwoliły zweryfikować, iż opracowane algorytmy i rozwiązania mogą być z powodzeniem stosowane w analizie danych sekwencyjnych i pozwalają na poprawę wydajności tych procesów analitycznych.

Lektura materiału pozwala mi stwierdzić, iż założone cele rozprawy udało się osiągnąć.

2. Czy w rozprawie przeprowadzono w sposób właściwy analizę źródeł (w tym literatury światowej, stanu wiedzy i zastosowań w przemyśle) świadczący o dostatecznej wiedzy Autora. Czy wnioski z przeglądu źródeł sformułowano w sposób jasny i przekonujący?

Analiza światowej literatury i bieżącego stanu wiedzy w omawianym obszarze zostały przeprowadzone w sposób właściwy i świadczą o dostatecznej wiedzy Autora w tej dziedzinie. Analiza ta została przedstawiona w rozdziale 2 autoreferatu oraz w trzech przedstawionych publikacjach Autora, które tworzą cykl publikacyjny będący głównym osiągnięciem rozprawy. Zawartość rozdziału 2 autoreferatu, który obejmuje m.in. kategoryzację i przegląd metod uczenia sieci, a także omówienie architektur sztucznych sieci neuronowych dla modelowania danych sekwencyjnych potwierdza, iż Autor posiada szeroką wiedzę w zakresie pierwotnych i bieżących trendów odnośnie tworzenia tego typu rozwiązań, a także zna ich zalety i słabości. W rozprawie zacytowano łącznie 126 pozycji literaturowych, z których zdecydowana większość dotyczy wyżej wymienionych elementów stanu wiedzy. Rozdziały 2.1 – 2.3 oraz rysunek 1 stanowią bardzo dobry wstęp teoretyczny do całości rozprawy, a do pojęć w nich zdefiniowanych (włączając wstęp, rozdział 1) Autor nawiązuje w kolejnych podrozdziałach autoreferatu, jak również w treści poszczególnych artykułów głównego cyklu publikacyjnego. Szczególną uwagę zwraca Autor na problemy dużej liczby parametrów modeli wnioskowania opartych na sztucznych sieciach neuronowych oraz konieczności budowania i realizacji kosztownych czasowo i pamięciowo obliczeń związanych z operacjami macierzowymi (m.in. obliczenia macierzy Hessego drugich pochodnych cząstkowych). Przeprowadzony przez Autora przegląd wiedzy w tym zakresie pozwolił mu w sposób jasny i przekonujący sformułować wnioski, które zdeterminowały przyjęcie odpowiednich rozwiązań algorytmicznych i obliczeniowych.

3. Czy Autor rozwiązał postawione zagadnienia, czy użył właściwej do tego metody i czy przyjęte założenia są uzasadnione?

Na początku realizacji rozprawy Pan Łukasz Neumann zdefiniował kilka zadań, do których realizacji konsekwentnie dążył w swoich pracach badawczych. Dotyczyły one redukcji niestabilności w procesie uczenia sztucznych sieci neuronowych podczas analizy danych sekwencyjnych. W swoich pracach Autor sięgnął do rozwiązań bazujących na odpowiednio zaprojektowanych architekturach rekurencyjnych sieci neuronowych (RNN) oraz optymalizatorach drugiego rzędu, jako alternatywie dla używanych często metod gradientowej optymalizacji statycznej. Na podstawie lektury przedłożonych prac można stwierdzić, iż postawione w rozprawie zagadnienia zostały rozwiązane w sposób właściwy. Autor osiągnął to poprzez: 1) identyfikację słabości istniejących metod optymalizacji pierwszego i drugiego rzędu oraz architektur sztucznych sieci neuronowych w kontekście ich złożoności i wydajności obliczeniowej, a także niestabilności procesu trenowania, 2) opracowanie własnych usprawnień i algorytmów, 3) badania eksperymentalne weryfikujące przydatność opracowanych metod z użyciem udostępnionych lub publicznie dostępnych zbiorów danych. Wyniki przeprowadzonych przez Autora rozprawy badań potwierdziły, iż założenia przyjęte podczas opracowania autorskich metod były słuszne i uzasadnione. W artykułach stanowiących główne osiągnięcie rozprawy przedstawiono porównanie osiągniętych

wyników z wynikami istniejących i popularnych narzędzi powszechnie używanych przez specjalistów prowadzących podobne analizy danych.

4. Na czym polega oryginalność rozprawy, co stanowi samodzielny i oryginalny dorobek Autora, jaka jest pozycja rozprawy w stosunku do stanu wiedzy czy poziomu techniki reprezentowanych przez literaturę światową?

Przedstawiona rozprawa stanowi bardzo dobre uzupełnienie bieżącego stanu wiedzy światowej w zakresie prowadzenia wydajnych analiz opartych na metodach sztucznej inteligencji. Pan Łukasz Neumann zaproponował nowatorskie rozwiązania w zakresie architektury i trenowania sztucznych sieci neuronowych, oparte na różnych podejściach, a także przeprowadził proces ich wnikliwej oceny. Na rozprawę, oprócz autoreferatu, składają się trzy publikacje – jedna w renomowanym czasopiśmie *Scientific Reports* (IF=4.996, wydawnictwa *Nature*), a także dwa artykuły opublikowane w materiałach uznanych konferencji naukowych. Wszystkie prace są ze sobą ściśle powiązane, a Pan Łukasz jest pierwszym autorem lub jego udział jest nie mniejszy niż pierwszego autora. Charakterystyka zawartości przedłożonych prac P1-P3 przedstawia się następująco:

[P1] W pracy tej przedstawiono zautomatyzowaną metodę klasyfikacji reakcji alergicznych na podstawie analizy skorelowanych obrazów przedramienia pacjenta w widmie widzialnym i termowizyjnym. W pracach badawczych procesy segmentacji i klasyfikacji były prowadzone z wykorzystaniem odpowiednio dobranych architektur sztucznych sieci neuronowych. Przedstawiono również dwa podejścia do procesu dopasowania/korelowania regionów zainteresowań. Przeprowadzone eksperymenty wykazały bardzo dobrą skuteczność opracowanego podejścia dla zebranych danych pochodzących od pacjentów Wojskowego Instytutu Medycznego w Warszawie. W pracy tej Pan Łukasz Neumann był odpowiedzialny za opracowanie i implementację metody segmentacji, klasyfikacji i dopasowania obrazów oraz prowadził badania skuteczności rozwiązania.

[P2] W pracy tej przedstawiono rekurencyjną architekturę sztucznej sieci neuronowej zwaną Deep Memory Update (DMU), w której aktualizacja poprzedniego stanu pamięci dokonywana jest poprzez głęboką transformację stanu opóźnionego i wejścia sieci. Zaproponowana architektura sieci jest w stanie nauczyć się przekształcać swój stan wewnętrzny za pomocą dowolnej funkcji nieliniowej. Zapewnia to, iż jej trenowanie jest stabilne i szybkie dzięki powiązaniu szybkości uczenia się z wielkością modułu. Badania eksperymentalne przeprowadzone z wykorzystaniem sześciu różnych tematycznie zbiorów danych potwierdziły, że zaproponowane rozwiązanie może konkurować z najnowocześniejszymi architekturami sieci, takimi jak Long Short Term Memory (LSTM), Gated Recurrent Units (GRU), a także Recurrent Highway Networks, a często je nawet przewyższa. Doktorant był współautorem opracowanej architektury oraz przeprowadził jej implementację i analizę.

[P3] W pracy tej przedstawiono ewolucyjną metodę uczenia głębokich sieci neuronowych, bazującą na optymalizacji wag sieci, nazwaną neuronową strategią ewolucji różnicowej (nDES), opartą na Strategii Ewolucji Różnicowej (DES) łączącej Ewolucję Różnicową (DE) ze Strategią Ewolucji Adaptacji Matrycy Kowariancji (CMA-ES). W zaproponowanym podejściu zastosowano inicjalizację populacji

opartą na współczynnikach Xaviera, przetwarzanie wsadowe i mutacje oparte na gradiencie. Badania przeprowadzone na zbiorze danych FashionMNIST dla zaproponowanej strategii nDES oraz literaturowego optymalizatora ADAM potwierdziły, iż zaproponowane rozwiązanie jest skuteczne i może zostać zastosowane w analizie problemów o dużej wymiarowości, a ponadto może zostać łatwo zrównoleglone w procesie obliczeniowym. Ponownie, Autora rozprawy przeprowadził implementację, analizę i testowanie przedstawionego rozwiązania.

Zaproponowanie odpowiednich architektur sztucznych sieci neuronowych umożliwiających m.in. zautomatyzowanie procesu klasyfikacji reakcji alergicznych na podstawie obrazów przedramienia pacjenta w widmie widzialnym i termowizyjnym, metod dopasowania wysegmentowanych regionów zainteresowań, a także alternatywnych metod optymalizacji konwolucyjnych i rekurencyjnych sieci neuronowych uważam za istotne osiągnięcie Autora i zaliczam do oryginalnych wyników przedstawionych w rozprawie. Wyniki przeprowadzonych prac badawczych zostały opublikowane w 3 artykułach bądź w liczących się w dziedzinie informatyki czasopismach (*Scientific Reports*, 140 pkt. MEiN), bądź w materiałach konferencyjnych (wszystkie za 140 pkt. MEiN). Dorobek ten uzupełnia 12 innych artykułów opublikowanych w materiałach konferencyjnych i 3 wystąpienia plakatowe. Na uwagę zasługuje udział w licznych projektach o charakterze naukowym oraz inne publikacje. Świadczy to w mojej opinii o istotności podjętego problemu oraz wyraźnym wkładzie Pana Łukasza Neumanna w rozwój tego obszaru informatyki.

5. Czy Autor wykazał umiejętność poprawnego i przekonującego przedstawiania uzyskanych przez siebie wyników /zwięzłość, jasność, poprawność redakcyjna rozprawy/?

Realizując pracę Pan Łukasz Neumann wykazał dobre opanowanie umiejętności przedstawiania uzyskanych przez siebie wyników. Same idee zostały zaprezentowane w sposób dość jasny, sformalizowany i poparty przykładami, i co niezwykle istotne, poprzedzone szeroką analizą rozwiązań dotychczas zaprezentowanych na światowym forum naukowym. Oceny skuteczności rozwiązania dokonano w oparciu o liczne i publicznie dostępne dane sekwencyjne (m.in. obrazy medyczne reakcji alergicznych skóry). Wyniki oceny skuteczności opracowanych rozwiązań danej klasy zostały przeanalizowane i skomentowane w przedstawionych artykułach P1-P3 przedłożonej rozprawy pokazując, że poszczególne rozwiązania i modyfikacje dotychczasowych metod umożliwiają poprawę jakości i wydajności osiąganych wyników w porównaniu z wybranymi i dostępnymi metodami. Od strony redakcyjnej zarówno autoreferat, jak i prace P1-P3 są w większości napisane w dobrym stylu i czyta się je z łatwością, chociaż znalazłem w samym autoreferacie również kilka zdań, w których należałoby poprawić stylistykę wypowiedzi.

6. Słabe strony rozprawy i jej główne wady?

Przedstawione prace są bardzo ciekawe i dotyczą istotnych problemów budowania i działania algorytmów analizy wykorzystujących sztuczne sieci neuronowe. Uzupełnia je autoreferat w języku angielskim, który zawiera najważniejsze konkluzje wypływające z przeprowadzonych prac badawczych. Nie znalazłem w tych dokumentach istotnych uchybień. Uwaga dotycząca poprawy stylistyki zdań w autoreferacie nie ma

charakteru znacząco krytycznego i nie umniejsza znaczeniu osiągnięć Autora rozprawy. Mam natomiast kilka uwag i pytań, na które odpowiedź chętnie bym poznał:

Q1. Dlaczego w pracy nie zdecydowano się na postawienie konkretnej tezy badawczej? Byłby to element mile widziany.

Q2. Dlaczego w temacie rozprawy i jej tekście Autor mówi o modelowaniu danych sekwencyjnych, a nie np. o ich analizie?

Q3. Co dokładnie Autor ma na myśli mówiąc o „rzędzie wielkości” w zdaniu „Using an evolution strategy on problems with order of magnitude higher than previously reported ...” – wielkość danych, liczbę parametrów czy złożoność procesu obliczeniowego?

Q4. Co oznacza symbol n we wzorach w rozdziale 2.1 i 2.1.1 autoreferatu?

U1. W odniesieniu do pojęcia *sieci neuronowych* będącego kluczowym elementem przedłożonej rozprawy, Autor powinien używać określenia *Sztuczne sieci neuronowe*. Rozumiem, że użyto skrótu myślowego.

7. Jaka jest przydatność rozprawy dla nauk technicznych?

Uważam, że przedłożona rozprawa doktorska Pana Łukasza Neumanna wpisuje się w bieżące problemy informatyki klasycznej i biomedycznej. Opracowanie rozwiązań dla stabilizowania procesu uczenia maszynowego pozwoliło Autorowi na osiągnięcie dobrych wyników jakościowych przy zachowaniu satysfakcjonującej efektywności i wydajności czasowej w stosunku do istniejących rozwiązań opublikowanych w światowej literaturze, co przekłada się bezpośrednio na tworzenie lepszych rozwiązań w tym obszarze. W ten sposób zaproponowane rozwiązania rozszerzają spektrum istniejących metod stosowanych w analizie danych z wykorzystaniem popularnych obecnie sztucznych sieci neuronowych. Potwierdzają to publikacje, których Pan Łukasz Neumann jest autorem, opublikowane przez wiodące wydawnictwa, takie jak *Nature*, *IEEE* oraz *Springer*.

8. Do której z następujących kategorii Recenzent zalicza rozprawę:

a/ nie spełniająca wymagań stawianych rozprawom doktorskim przez obowiązujące przepisy

b/ wymagająca wprowadzenia poprawek i ponownego recenzowania

c/ spełniająca wymagania

d/ spełniająca wymagania z wyraźnym nadmiarem

e/ wybitnie dobra, zasługująca na wyróżnienie

Reasumując, bardzo dobre wyniki osiągnięte przez Pana Łukasza Neumanna w trakcie realizowanych przez niego badań pozwalają potwierdzić, iż główne cele rozprawy przedstawione w rozdziale 1.1 autoreferatu. Wyniki badań pokazują, że techniki oraz metody zaproponowane przez Pana Łukasza Neumana mogą

przyczynić się do znaczącej poprawy procesów analizy danych sekwencyjnych. Wartość tych metod została dostrzeżona przez środowisko naukowe, co potwierdzają opublikowane prace, wchodzące w skład przedstawionego cyklu głównego. Uważam zatem, że **przedstawiona praca spełnia wymagania** stawiane rozprawom doktorskim określone w obowiązujących przepisach. Wnoszę zatem o przyjęcie rozprawy doktorskiej i dopuszczenie jej do publicznej obrony.



Dr hab. inż. Dariusz Mrozek, prof. PS
Katedra Informatyki Stosowanej
Politechnika Śląska w Gliwicach

prof. dr hab. inż. Andrzej Obuchowicz

Zielona Góra, 8 października 2023

Uniwersytet Zielonogórski

Instytut Sterowania i Systemów Informatycznych

RECENZJA

rozprawy doktorskiej Pana mgr inż. Łukasza Neumanna

*Training instabilities in neural network-based sequential data
modeling*

opracowana na zlecenie

Rady Dyscypliny Naukowej Informatyka Techniczna i Telekomunikacja

Politechniki Warszawskiej

1. Problem badawczy i jego znaczenie

Szybki zanik gradientu w propagowanym wstecz sygnale błędu w procesie uczenia sieci ma szczególnie negatywny wpływ na proces uczenia zarówno sieci głębokich, jak i rekurencyjnych. W tym drugim przypadku prowadzi do destabilizacji procesu uczenia i wpływa na krótkotrwałość wpływu sygnałów wcześniejszych na aktualny stan sieci, czego pragniemy uniknąć modelując dane sekwencyjne. Kierunki rozwiązania tego problemu są w zasadzie dwa: modyfikacja struktury sieci, albo oparcie procesu uczenia o metody optymalizacji wyższego rzędu, lub metody metaheurystyczne. Dotychczas modyfikacje architektury sieci wiązały się ze wzrostem liczby ich parametrów swobodnych, a zatem silnie wzrastał wymiar przestrzeni, w której realizuje się proces uczenia/optymalizacji. Poszukiwanie kolejnych rozwiązań powyższego problemu, wokół którego skoncentrowała się uwaga doktoranta w prezentowanym dorobku naukowym, nadal jest istotnym zagadnieniem badawczym wartym podjęcia.

2. Oryginalne osiągnięcia doktoranta

Zasadnicze osiągnięcia doktoranta dotyczą dwóch bardzo interesujących propozycji rozwiązania wyżej wspomnianego problemu.

Pierwszą stanowi propozycja nowej architektury sieci neuronowej w postaci minimalnej rekurencyjnej sieci neuronowej z mechanizmem bramkującym oraz arbitralnym, nieliniowym przekształceniem stanu. Spośród znanych z literatury architektur sieci neuronowych odpornych na problem zanikania/eksplozji gradientu, proponowana przez doktoranta architektura wyróżnia się znacznie mniejszą liczbą parametrów sterujących, a zatem proces uczenia/optymalizacji sieci przebiega w przestrzeni o niższej wymiarowości. Porównawcze badania eksperymentalne na wybranych zagadnieniach benchmarkowych wskazują na konkurencyjność proponowanego modelu sieci w stosunku do najbardziej popularnych modeli, a niejednokrotnie okazuje się być modelem bardziej efektywnym.

Drugą propozycją doktoranta jest zastosowanie algorytmu ewolucji różnicowej z adaptacją współczynnika skalującego inspirowanego metodą CMA-ES, w skrócie DES, opracowanego przez panów prof. Arabasa i mgr Jagodzińskiego, do uczenia głębokiej sieci neuronowej. Samo zastosowanie technik ewolucyjnych do uczenia sieci neuronowej ma już ponad trzydziestoletnią tradycję. W przypadku sieci głębokich, efektywnym w sensie uniknięcia problemu zanikania/eksplozji gradientu, przydatne są metaheurystyczne techniki drugiego rzędu, tzn. wykorzystujące informacje płynące z estymaty macierzy kowariancji. Do tych metod należy jeden z najbardziej popularnych w ostatnim czasie algorytmów: CMA-ES. Podstawowym problemem zastosowania CMA-ES w procesie uczenia sieci głębokiej jest niezwykle wysoki, często nieakceptowalny, koszt obliczeniowy, ponieważ w każdej generacji poszukiwania ewolucyjnego konieczne jest wyznaczanie i przekształcanie macierzy kowariancji dla mutacji Gaussowskiej. Algorytm DES, stawiający swoistego rodzaju fuzję metod CMA-ES i algorytmu ewolucji różnicowej, pozwala na uniknięcie kosztownych operacji macierzowych. Doktorant zaadaptował z sukcesem algorytm DES do uczenia sieci neuronowych zarówno konwolucyjnych, jak i rekurencyjnych. Wykazał, że proponowany proces uczenia sieci jest efektywny szczególnie w zadaniach z poważnymi problemami propagacji gradientu.

Osiągnięciem godnym szczególnego wyróżnienia jest propozycja, zawarta w artykule

opublikowanym w Scientific Reports, w pełni automatycznego systemu do rozpoznawania reakcji alergicznych skóry. Doktorant, jako członek pięciosobowego zespołu, odpowiadał za najistotniejsze elementy związane z budową środowiska rozpoznawania i klasyfikacji obrazów wizyjnych i termowizyjnych. W ramach projektowania tego środowiska musiał uwzględnić ewentualne problemy niestabilności procesu uczenia. Problemom tym zapobiegł poprzez narzucenie z góry nieelastycznego sposobu przetwarzania stanu oraz zaadaptowanej do rozwiązywanego problemu konwolucyjnej sieci neuronowej. Nie mniej, w mojej opinii, to utworzenie pełnego automatycznego narzędzia, wymagającego rozwiązania wielu istotnych problemów związanych z poszczególnymi etapami procesu rozpoznawania obrazów, a nie związanych z głównym tematem rozprawy, stanowi osiągnięcie doktoranta zasługujące na szczególne podkreślenie i wyróżnienie. Każde z wyżej wymienionych opracowań związane było z koniecznością przeprowadzenia, niemal na każdym etapie konstrukcji, szerokiego zakresu eksperymentalnych badań porównawczych związanych w doborze najkorzystniejszych rozwiązań szczegółowych spośród co najmniej kilku alternatywnych możliwych realizacji.

3. Poprawność rozprawy

Przedstawioną do recenzji rozprawę doktorską stanowi zbiór powiązanych tematycznie trzech publikacji naukowych:

- P1 Ł. Neumann, R. Nowak, J. Stępień, E. Chmielewska, P. Pankiewicz, R. Solan i K. Jahnz-Różyk: Thermographic based skin allergic reaction recognition by convolutional neural networks – *Scientific Reports* 12.1 (2022).
- P2 Ł. Neumann, Ł. Lepak i P. Wawrzyński: Least redundant gated recurrent neural network, *IEEE Int. Joint Conf. on Neural Networks (IJCNN)* 2023.
- P3 D. Jagodziński, Ł. Neumann i P. Zawistowski: Deep Neuroevolution: Training Neural Networks Using a Matrix-Free Evolution Strategy – *28th Int. Conf. on Neural Information Processing (ICONIP)* 2021.

Powyższy zbiór uzupełniony został autoreferatem, w którym przedstawiono motywacje i cele rozprawy, oraz, w świetle określonych celów, pokrótce omówiono wartość poszczególnych publikacji zbioru. Publikacja artykułu w renomowanym

czasopiśmie (IF bliski 5) i dwóch referatów na konferencjach o ustalonej znaczącej randze w środowisku badawczym, świadczą o wartości uzyskanych wyników przez doktoranta i współautorów, oraz o ich uznaniu w oczach światowych ekspertów—recenzentów tegoż czasopisma i obu konferencji.

Powyższa forma rozprawy doktorskiej jest dopuszczona przez obowiązujące prawo (ustawa Prawo o szkolnictwie wyższym i nauce Dz.U. 2023 poz. 742, art. 187 ust. 3), nie mniej przygotowanie recenzji na tej podstawie stanowi poważne wyzwanie. Specyfiką publikacji konferencyjnych i w czasopiśmie naukowych jest fakt, że głównym ich celem jest prezentacja osiągniętego wyniku badawczego, konkretnego procesu jego osiągnięcia, oraz jego dyskusja. Całość objęta jest często silnym ograniczeniem objętościowym. Jakkolwiek zbiór publikacji jest powiązany tematycznie, to każda z nich jest odrębnym raportem badawczym o ściśle określonym głównym celu, często dość luźno powiązany, bądź nie powiązany z celem rozprawy. W związku z tym akcenty i skupienie autorów położone są na innych elementach publikacji niż te, których oczekiwałby recenzent rozprawy doktorskiej. Na przykład w artykule P1 autorzy skupiają się na problemie diagnozy odpowiedzi alergicznej i technikach jej wsparcia. Nie ma analizy wariantów technik możliwych do zastosowania na każdym etapie rozpoznawania obrazów i uzasadnienia tego wyboru. Jest jedynie podane konkretne rozwiązanie i to w stopniu wysokiej ogólności. Tego typu forma rozprawy doktorskiej ułatwia recenzentowi wskazanie oryginalnego rozwiązania naukowego i odniesienie się do jego wartości (patrz art. 187 ust. 2 ustawy), ale nie daje bezpośredniej możliwości oceny ogólnej wiedzy teoretycznej kandydata w dyscyplinie i umiejętności samodzielnego prowadzenia pracy naukowej (patrz art. 187 ust. 1 ustawy) w szczególności, gdy publikacje zbioru są wieloautorskie. Ocena w tym obszarze jest oparta na wierze, że za wynikami badawczymi wysokiej jakości stoi ugruntowana i szeroka wiedza doktoranta, oraz umiejętne jej stosowanie w samodzielnej pracy badawczej. Tego typu rozterek recenzenta unikamy, jeżeli rozprawę stanowi „klasyczna” praca doktorska.

4. Wiedza kandydata

Podjęty przez doktoranta w rozprawie problem naukowy wymaga wiedzy i umiejętności z wielu obszarów badawczych, w szczególności z zakresu inteligencji ob-

liczeniowej, w tym, m. in., sztucznych sieci neuronowych, uczenia maszynowego i algorytmów metaheurystycznych, ponadto z zakresu technik przetwarzania i rozpoznawania obrazów. Zawartość treści wstępnych rozdziałów zawartych w przedstawionym zbiorze publikacji naukowych wskazuje na szeroką wiedzę autora rozprawy w powyższych obszarach i umiejętność z niej korzystania. Ponadto doktorant wykazał się poprawnością terminologiczną i znajomością literatury przedmiotu.

5. Podsumowanie

Biorąc pod uwagę opinie zaprezentowane w poprzednich punktach i wymagania zdefiniowane przez artykuł 187 ust. 1 i 2 Ustawy z dnia 20 lipca 2018 r. Prawo o szkolnictwie wyższym i nauce (Dz. U. 2023 r., poz. 742 z późn. zm.)) **stwierdzam, że:** rozprawa doktorska pana mgr inż. Łukasza Neumanna zawiera oryginalne rozwiązanie problemu naukowego, kandydat posiada ogólną wiedzę teoretyczną w dyscyplinie informatyka techniczna i telekomunikacja oraz posiada umiejętność samodzielnego prowadzenia pracy naukowej.

W związku z powyższym **wnoszę** o przyjęcie rozprawy i dopuszczenie jej do publicznej obrony.



