

Prof. Michał Woźniak
Politechnika Wrocławska

Recenzja rozprawy doktorskiej

mgra Sławomira Smyl

zatytułowanej:

Methods for improving performance of neural network-based forecasting models

1. Problem badawczy i jego znaczenie

Rozprawa koncentruje się na metodach poprawy jakości modeli prognostycznych opartych na sieciach neuronowych do analizy szeregów czasowych, a jej zakres jest zgodny z dyscypliną naukową informatyka techniczna i telekomunikacja. Autor proponuje innowacyjne podejścia, w tym hybrydowe architektury łączące klasyczne wygładzanie wykładnicze z rekurencyjnymi sieciami neuronowymi oraz koncepcję modeli trenowanych na wielu szeregach jednocześnie. Praca jest *de facto* zbiorem ośmiu artykułów, poprzedzonych wstępem oraz opisem każdego z nich.

Zaproponowane metody zostały poddane ocenie pod kątem wybranych czynników, które mogą wpływać na jakość proponowanego rozwiązania. Ocena jakości zaproponowanych metod została dokonana głównie na drodze eksperymentu komputerowego. Opracowane metody mogą być podstawą do budowy efektywnych systemów predykcji szeregów czasowych i znaleźć zastosowanie np. w sektorze energetycznym.

2. Wkład autora

Zakres oraz zawartość rozprawy oceniam bardzo pozytywnie, a sformułowana teza mówiąca, że skuteczność modeli prognostycznych opartych na sieciach neuronowych, można poprawić poprzez odpowiednie przetwarzanie wstępne (tj. reprezentację danych), architektury dostosowane do danego zadania, odpowiednie procedury uczenia oraz łączenie modeli, uważam za dobrze ambitną i dobrze sformułowaną.

Rozprawa zawiera bardzo dużo ciekawych wątków i propozycji, spośród których chciałbym szczególnie wyróżnić:

- Zaproponowanie wykorzystania tzw. *cross-learning*, co umożliwia na jednoczesne uczenie pojedynczej architektury na bardzo dużej liczbie szeregów czasowych i zapobiega przetrenowaniu na małych próbach.
- Zaproponowanie normalizacji w locie (*on-the-fly preprocessing*), co pozwala na ciągłe skalowanie danych w trakcie procesu uczenia z wykorzystaniem parametrów wygładzania wykładniczego.
- Zaproponowanie "rozszerzonej" komórki rekurencyjnej dRNNCell (*dilated recurrent neural network cell*) optymalizującej analizę krótkoterminowych i długoterminowych zależności w wielokrotnej sezonowości, a także zaproponowanie na jej bazie adRNNcell (*attentive dilated recurrent neural network cell*) do analizy wielosezonowości poprzez wbudowany mechanizm uwagi służący do ważenia informacji wejściowych.
- Wprowadzenie nowatorskiej architektury wzbogaconej o tzw. globalny "kontekst" (*context track*), która pozwala na agregowanie ukrytych informacji z wybranej grupy szeregów w celu lepszego prognozowania.
- Ciekawym pomysłem było uczenie w oparciu o celowo przesunięte kwantyle treningowe, co w praktyce stanowiło swojego rodzaju formę kalibracji, uodparniając model na asymetryczne błędy i generując trafniejsze docelowe przedziały ufności.
- Zwrócenie uwagi na mechanizm *warm-up*, dzięki czemu w momencie rozpoczęcia właściwej predykcji wewnętrzny stan pamięci modelu jest już ustabilizowany i nasycony kontekstem.
- Ocena jakości zaproponowanych metod na podstawie rzetelnych i obszernych eksperymentów komputerowych.

Wyniki uzyskiwane w trakcie pracy nad rozprawą zostały zawarte w ośmiu artykułach składający się na oceniany cykl, które opublikowano w jednych z najlepszych czasopismach z zakresu uczenia maszynowego, tj. *IEEE Transactions on Neural Networks and Learning Systems*, *Neural Networks*, *Information Fusion* i *International Journal of Forecasting*. Dwie prace zostały opublikowane w materiałach konferencji indeksowanych przez CORE o rankingu A, tj. IJCNN oraz ICONIP. Warto także wspomnieć, że kod programistyczny,

umożliwiający niezależną weryfikację opracowanych metod, został udostępniony wraz z tekstami artykułów. Taki dorobek należy uznać za wybitny, szczególnie na tym etapie kariery naukowej.

Oprócz potwierdzenia jakości zaproponowanych metod przez ich publikacje w prestiżowych czasopismach i konferencjach, warto podkreślić, że zaproponowane metody były poddane ocenie ramach konkursów, na których doktorant zajmowały czołowe miejsca:

- Zwycięstwo (w 2018 r.) w prestiżowym konkursie *M4 Forecasting Competition* (zarówno w kategorii prognoz punktowych, jak i przedziałowych), gdzie zaproponowano połączenie standardowego modelu wygładzania wykładniczego z zaawansowanymi sieciami typu *Long Short-Term Memory* (LSTM) w ramach jednej wspólnej struktury. Jakość zaproponowanych modeli była o ok. 10% lepsza niż model referencyjny.
- Zajęcie trzeciego i czwartego miejsca w konkursie *Global Energy Forecasting Competition* 2017 (GEFCom2017). Zaproponowane algorytmy wykorzystywały drzewa decyzyjne (*Gradient Boosting* i *Quantile Random Forest*) oraz sieci neuronowe do budowy zespołu modeli. Kwantyle prognoz zespołu uzyskiwano poprzez uśrednienie kwantyli prognoz poszczególnych modeli wchodzących w jego skład.
- Pierwsze w historii pokonanie eksperckiego zespołu statystycznych modeli przez model uczenia maszynowego (zespół sieci RNN i modelu bayesowskiego) w międzynarodowym konkursie CIF2015.

3. Poprawność

Struktura pracy nie budzi zastrzeżeń. Układ jest jasny, a wprowadzenie w tematykę oraz syntetyczny opis poszczególnych prac jest bardzo pomocny w lepszym zrozumieniu dokonań doktoranta.

Lektura rozprawy prowadzi do sformułowania poniższych uwag.

- W pracach w pomijane są zmiennych egzogenicznych (pogody, wydarzenia sportowe i dni świąteczne), które są istotne np. przy predykcji zapotrzebowania na energię elektryczną. Z jednej strony jest to zrozumiałe i uzasadnione brakiem dostępu do wiarygodnych danych historycznych. Czy doktorant dostrzega inne kierunki w tym zakresie, tj. jak uwzględnić te dane, ew. ich predykcję, poza faktycznym dostępem do wiarygodnych repozytoriów?

- W pracy zaproponowano modele zespołowe bazując na wybranych fragmentach danych, które były arbitralnie, bądź losowo dobierane (np. w konkursie M4 nie było dostępu do metadanych). Czy zamiast osądu eksperckiego, bądź losowości można zwiększyć różnorodność puli modeli np. poprzez uwzględnienie metacech danego szeregu (np. wskaźnik zmienności lub sezonowości)?
- Czy rozważano wykorzystanie metod, które inspirowane byłyby dynamiczną selekcją zespołów klasyfikatorów (*Dynamic Ensemble Selection*), ew. czy wg doktoranta takie podejście miałyby sens w zastosowaniu do predykcji szeregów czasowych?
- Ważnym elementem jest wykrycie zmiany parametrów modelu w czasie, zwane w szeregach czasowych *change point detection*. W prezentowanych pracach nie korzystano z tego mechanizmu. Jakie były tego przyczyny, mimo iż szeroko dyskutowano wpływ przestarzałych danych na model, stosując np. wygładzane wykładnicze, które pozwala na uwzględnienie łagodnych dryftów w rozkładach danych? Reasumując, czy zamiast modyfikowania mechanizmu uwagi nie warto zastosować detektora *change points*, bądź opracować model hybrydowy?
- Doktorant dość krytycznie odniósł się do wykorzystania SSM (*State Space Models*) do predykcji szeregów czasowych, choć kilka prac pokazuje, że modele (np. MAMBA) dość dobrze radzą sobie z tym zadaniem (np. *Deep State Space Models for Time Series Forecasting* z NeurIPS 2018). Co więcej, tego typu modele cechują się niską złożonością obliczeniową. Prosiłbym o rozwinięcie tego wątku.

4. Wiedza kandydata

Na podstawie lektury uważam, że doktorant posiada ugruntowaną wiedzę w dyscyplinie informatyka techniczna i telekomunikacja, w szczególności w zakresie metod uczenia maszynowego, sieci neuronowych oraz analizy szeregów czasowych. Doktorant potrafi poprawnie zaplanować i przeprowadzić eksperyment komputerowy w celu oceny jakości zaproponowanych metod. Przegląd literaturowy dotyczący zagadnień przedstawionych w rozprawie (wstęp do pracy oraz sekcje *related work* w poszczególnych artykułach) pozwala stwierdzić, że doktorant posiada aktualną wiedzę z zakresu tematyki rozprawy, a także potrafi dokonać krytycznego przeglądu źródeł w celu wskazania ciekawych kierunków badań, a także dokonać krytycznej analizy niektórych trendów, co znalazło swoje miejsce m.in. w

podsumowaniu pracy. Zawarty w dysertacji spis źródeł literaturowych jest reprezentatywny i aktualny, na czas publikacji poszczególnych prac.

5. Podsumowanie

Doktorant wykazał się w recenzowanej rozprawie właściwie stosowanym podejściem eksperymentalnym oraz bardzo dobrą znajomością aktualnej problematyki związanej z projektowaniem systemów prognozowania szeregów czasowych oraz uczeniem maszynowym. Zostało to poparte dobrymi studiami literaturowymi, obejmującymi aktualne piśmiennictwo związane z problematyką rozprawy, co świadczy o bardzo dobrej wiedzy doktoranta z tego zakresu. Dla poruszanych problemów doktorant sformułował szereg rozwiązań, które mogą być z powodzeniem stosowane do projektowania rzeczywistych systemów analizy danych, co zostało potwierdzone m.in. licznymi zwycięstwami w konkursach programistycznych związanych z analiza szeregów czasowych. Doktorant przeanalizował wyniki przeprowadzonych badań eksperymentalnych, ocenił jakość zaproponowanych metod na tle algorytmów znanych z literatury. Na podkreślenie zasługuje umiejętność krytycznego spojrzenia na otrzymane rezultaty oraz wnikliwa analiza wyników, a także zapewnienie dostępów do repozytoriów kodu i w konsekwencji możliwość zreplicowania otrzymanych przez niego rezultatów.

Recenzowana dysertacja przedstawia rozwiązanie oryginalnego problemu, wzbogacając naszą wiedzę dotyczącą predykcji szeregów czasowych. Zawarte w niej wyniki badań eksperymentalnych wskazują również na możliwość wykorzystania otrzymanych metod w praktyce.

Przedstawione w punkcie 3 recenzji uwagi mają jedynie charakter dyskusyjny, w żaden sposób nie deprecjonują osiągniętych przez doktoranta rezultatów oraz nie wpływają na wyjątkowo pozytywne wrażenie o przedłożonej rozprawie. Mogą być one przydatne w przypadku dalszego publikowania rezultatów zawartych w rozprawie. Jestem przekonany, że doktorant zdaje sobie sprawę z możliwości kontynuowania rozpoczętej w pracy tematyki, czemu dał wyraz w sformułowaniu ciekawych kierunków dalszych badań, m.in. w zakresie wykorzystanie modeli LLM, które nie są powszechnie stosowane w obszarze problemowym, rozprawy.

Reasumując, biorąc pod uwagę opinie zaprezentowane w poprzednich punktach i wymagania zdefiniowane przez art. 187 Ustawy z dnia 20 lipca 2018 r. Prawo o szkolnictwie wyższym i nauce (z późniejszymi zmianami) stwierdzam, że rozprawa mgra Sławomira Smyła pt. *Methods for improving performance of neural network-based forecasting models* spełnia wymagania stawiane pracom doktorskim, w szczególności:

- Rozprawa zawiera oryginalne rozwiązanie problemu naukowego.
- Kandydat posiada ugruntowaną wiedzę w dyscyplinie Informatyka techniczna i telekomunikacja.
- Doktorant posiada umiejętność samodzielnego prowadzenia pracy naukowej.

Wnoszę o jej przyjęcie i dopuszczenie mgra Sławomira Smyła do publicznej obrony oraz biorąc pod uwagę oryginalność zaproponowanych rozwiązań, dorobek publikacyjny i dojrzałość naukową Kandydata wnioskuję o wyróżnienie rozprawy.

Podpis