



PROPOSAL FOR MASTER OF SCIENCE THESIS

academic year 2022/2023

Full name, degree held	Agnieszka Jastrzębska, PhD DSc
Place of work, telephone number, email	Zakład Strukturalnych Metod Przetwarzania Wiedzy MINI PW
Thesis topic (English)	Transfer learning for time series classification
Thesis topic (Polish)	Zastosowanie techniki transfer learning w zadaniu klasyfikacji szeregów czasowych
Area of specialisation	Data Science
Name of degree candidate (optional)	Paulina Pacyna

Aim of the work:

The goal of this thesis is to analyze the viability and evaluate the effectiveness of the use of transfer learning to time series classification. The student will review several methods and models related to this problem. The author will consider fusing transfer learning with popular deep learning architectures (plain convolutional networks, Residual Networks, Encoder architecture). Additionally, a study on the influence of transfer learning on the model's explainability (Class Activation Map method) will be conducted. The scope of work will also cover time series preprocessing methods specific to this class of problems. The study will be conducted with the use of publicly available benchmark datasets available in the repository under <https://timeseriesclassification.com>

Subject of submitted work:

The range of available approaches to the problem of time series classification has substantially increased with the progress in the domain deep learning. Initially, time series classification was solved using traditional machine learning methods combined with feature engineering algorithms designed for time series (e.g., Dynamic Time Warping with k-Nearest Neighbor classifier).

Recently, algorithms based on neural networks have started to be considered suitable for classifying time series. It was demonstrated that neural networks can identify relations between observations in ordered data (recurrent neural networks) and shapes occurring in the time series (convolutional layer). Neural-network-based approaches are nowadays in the group of the best time series classifiers. The networks will be trained on raw time series values or on characteristics obtained from a feature extraction process. Furthermore, the use of neural networks enables the possibility of employing the transfer learning technique.

Transfer learning attempts to apply patterns learned from one dataset to improve the learning process when creating a model for another dataset. In papers introducing the deep learning methods for time series classification (Fawaz et al.), the transfer learning technique has been suggested as one of interesting approaches to be analyzed in the future. In this thesis, the student will look into this topic.

The scope of work goes beyond the application of transfer learning. The student will also explore the impact of transfer learning on the explainability of the model using the Class Activation Map method (for ResNet and FCN architectures). The formerly mentioned method emphasizes parts of the time series that were essential to recognizing the class.

Reference literature:

1. Ismail Fawaz, H., Forestier, G., Weber, J. et al. *Deep learning for time series classification: a review*. Data Mining and Knowledge Discovery 33, 917–963 (2019). <https://doi.org/10.1007/s10618-019-00619-1>
2. Bagnall, A., Lines, J., Bostrom, A. et al. *The great time series classification bake off: a review and experimental evaluation of recent algorithmic advances*. Data Mining and Knowledge Discovery 31, 606–660 (2017). <https://doi.org/10.1007/s10618-016-0483-9>



3. Ruiz, A.P., Flynn, M., Large, J. et al. *The great multivariate time series classification bake off: a review and experimental evaluation of recent algorithmic advances*. Data Mining and Knowledge Discovery 35, 401–449 (2021). <https://doi.org/10.1007/s10618-020-00727-3>
4. Jiang, W. *Time series classification: nearest neighbor versus deep learning models*. SN Applied Sciences 2, 721 (2020). <https://doi.org/10.1007/s42452-020-2506-9>
5. Weiss, K., Khoshgoftaar, T.M. & Wang, D. *A survey of transfer learning*. Journal of Big Data 3, 9 (2016). <https://doi.org/10.1186/s40537-016-0043-6>

The work ~~will~~ / will not (*) be carried out in collaboration with, or on commission for external entities.
Name of external entity (if applicable):

I hereby consent / ~~do not consent~~ () to the publishing on the faculty's intranet of a full description of the proposed topic.*

.....
date and signature



ZGŁOSZENIE TEMATU PRACY DYPLOMOWEJ MAGISTERSKIEJ

rok akademicki 2022/2023

Imię, nazwisko, tytuł, stopień naukowy	dr hab. inż. Agnieszka Jastrzębska
e-mail, Instytucja, Wydział	Zakład Strukturalnych Metod Przetwarzania Wiedzy MINI PW
Tytuł zgłaszanej pracy (w jęz.polskim)	Zastosowanie techniki transfer learning w zadaniu klasyfikacji szeregów czasowych
Tytuł zgłaszanej pracy (w jęz.angielskim)	Transfer learning for time series classification
Kierunek studiów	Data Science
Imię i nazwisko dyplomanta/ki (opcjonalnie)	Paulina Pacyna

Cel pracy:

Celem pracy dyplomowej jest analiza możliwości użycia i efektywności metody transfer learning w zadaniu klasyfikacji szeregów czasowych. W pracy zostaną porównane różne techniki przygotowania danych oraz różne modele uczenia. Dyplomantka zbada użycie metody transfer learning w połączeniu z najpopularniejszymi architekturami sieci neuronowych takich jak splotowe sieci neuronowe, architektura ResNet oraz model oparty o tzw. encoder. Dodatkowo przetestowany zostanie wpływ zastosowania transfer learningu na wyjaśnialność modeli przy użyciu metody Class Activation Map. Zakres pracy obejmuje przygotowanie danych i ekstrakcję cech dla zadania klasyfikacji. W pracy dyplomowej zostaną wykorzystane publicznie dostępne dane z repozytorium pod adresem <https://timeseriesclassification.com>.

Tematyka zgłaszanej pracy:

Wraz z postępem w dziedzinie uczenia głębokiego zasadniczo wzrosła liczba metod możliwych do zastosowania w zadaniu klasyfikacji szeregów czasowych. Niegdyś problem klasyfikacji szeregów czasowych był rozwiązywany przy użyciu tradycyjnych metod uczenia maszynowego w połączeniu z algorytmami inżynierii cech zaprojektowanymi dla szeregów czasowych (np. Dynamic Time Warping w połączeniu z klasyfikatorem k-Nearest Neighbor). W ostatnich latach zaczęto rozpatrywać różne modele sieci neuronowych jako narzędzie potencjalnie cenne przy rozwiązywaniu zadań klasyfikacji szeregów czasowych. Z sukcesem wykazano, że sieci neuronowe mogą identyfikować relacje pomiędzy obserwacjami w uporządkowanych danych (rekurencyjne sieci neuronowe) oraz kształty występujące w szeregu czasowym (warstwa konwolucyjna).

Podjęcia oparte na sieciach neuronowych są obecnie w grupie najlepszych algorytmów używanych dla tego problemu. W ramach pracy magisterskiej jako podejście referencyjne ustalone zostanie trenowanie sieci na surowych wartościach szeregu czasowego lub na cechach uzyskanych w procesie inżynierii cech. Główna część pracy skupi się na ocenie efektywności techniki transfer learningu w zadaniach klasyfikacji szeregów czasowych. Zostanie porównana z podejściem bazowym. Sprawdzone zostaną również techniki rozszerzania zbiorów danych w zastosowaniu do modeli uczonych przy pomocy transfer learningu.

Transfer learning polega na zastosowaniu wzorców rozpoznanych w jednym zbiorze danych do poprawy procesu uczenia przy tworzeniu modelu dla innego zbioru danych. W pracach wprowadzających metody głębokiego uczenia do klasyfikacji szeregów czasowych (Fawaz i inni), technika transfer learning została zasugerowana jako jedno z ciekawych podejść. Brak jest jednak poszerzonego studium nad tym tematem. W niniejszej pracy dyplomowej dyplomantka zbada dokładnie to zagadnienie.

Zakres pracy wykracza poza zastosowanie uczenia transferowego. Dyplomantka zbada wpływ uczenia transferowego na wyjaśnialność modelu przy użyciu metody Class Activation Map (dla architektur ResNet i FCN). Wspomniana wcześniej metoda uwypukla te części szeregu czasowego, które były istotne dla rozpoznania klasy.



Literatura pomocnicza:

1. Ismail Fawaz, H., Forestier, G., Weber, J. et al. *Deep learning for time series classification: a review*. Data Mining and Knowledge Discovery 33, 917–963 (2019). <https://doi.org/10.1007/s10618-019-00619-1>
2. Bagnall, A., Lines, J., Bostrom, A. et al. *The great time series classification bake off: a review and experimental evaluation of recent algorithmic advances*. Data Mining and Knowledge Discovery 31, 606–660 (2017). <https://doi.org/10.1007/s10618-016-0483-9>
3. Ruiz, A.P., Flynn, M., Large, J. et al. *The great multivariate time series classification bake off: a review and experimental evaluation of recent algorithmic advances*. Data Mining and Knowledge Discovery 35, 401–449 (2021). <https://doi.org/10.1007/s10618-020-00727-3>
4. Jiang, W. *Time series classification: nearest neighbor versus deep learning models*. SN Applied Sciences 2, 721 (2020). <https://doi.org/10.1007/s42452-020-2506-9>
5. Weiss, K., Khoshgoftaar, T.M. & Wang, D. *A survey of transfer learning*. Journal of Big Data 3, 9 (2016). <https://doi.org/10.1186/s40537-016-0043-6>

Praca będzie / ~~nie będzie~~ (*) realizowana przy współudziale lub na zlecenie podmiotów zewnętrznych.
Podmiot zewnętrzny (opcjonalnie):

Wyrażam zgodę / ~~nie wyrażam zgody~~ (*) na udostępnienie elektroniczne w sieci Wydziału MiNI pełnego opisu zgłaszanego tematu.

.....
data i podpis