

Projekt badawczo - rozwojowy : Analiza porównawcza danych szeregów czasowych

Celem projektu jest stworzenie narzędzi / algorytmów do jakościowej analizy porównawczej danych szeregów czasowych. Motywacją do badań jest zapotrzebowanie biznesowe na narzędzie potrafiące dostarczać metryk do oceny zmian zachowań modeli sztucznej inteligencji / modeli statystycznych wykorzystywanych do predycji/aproksymacji danych o charakterze szeregów czasowych.

Podając konkretny przykład posiadamy model statystyczny zamieniający odczyty pomiarów urządzeń sensorycznych mierzony w Voltach na odczyty procentowe zawartości wody w medium lub przewodności elektrycznej. Wraz ze zmianą modelu i wprowadzeniem nowej wersji końcowy użytkownik potrzebuje zrozumieć trend zmian wynikających z ulepszenia modelu tak, by móc dostosować własne urządzenia / algorytmy do tych zmian. Zakładamy, że dane posiadają X serii, które mogą być zależne od siebie (np. temperatura i wilgotność). Należy wskazać jakościowe różnice między wynikami w nowej wersji np. nowy model daje średnio wyższe odczyty wilgotności przy wysokich temperaturach albo nowy model wykazuje większą korelację między parametrami A i B, itp.

W ramach pracy planowany jest przegląd istniejących metod oceny szeregów czasowych do uzyskania oczekiwanej przez nas informacji. To znaczy między innymi sprawdzenie jak dla naszego problemu sprawdzają się różne metody statystyczne, metryki odległości między seriami danych [1], klasteryzacje, SAX (Symbolic Aggregate approXimation) [3], FFT (dyskretne transformata Fouriera) [2], czy też modelowanie (exponential smoothing) oraz inne metody zaproponowane przez magistranta - autorskie lub też nie.

W projekcie zakładamy także implementację oprogramowania do oceny szeregów czasowych w oparciu o istniejące biblioteki wyliczające te metryki. Rozwiązanie to może być następnie wykorzystane komercyjnie we współpracy z partnerem wydziału.

REFERENCES

- [1] Tomasz Górecki, Maciej Łuczak, and Paweł Piasecki. "An exhaustive comparison of distance measures in the classification of time series with 1NN method". In: *Journal of Computational Science* 76 (2024), p. 102235. ISSN: 1877-7503.
- [2] Christian Herff and Dean J. Krusienski. "Extracting Features from Time Series". In: *Fundamentals of Clinical Data Science*. Ed. by Pieter Kubben, Michel Dumontier, and Andre Dekker. Cham: Springer International Publishing, 2019, pp. 85–100.
- [3] Jessica Lin et al. "A symbolic representation of time series, with implications for streaming algorithms". In: New York, NY, USA: Association for Computing Machinery, 2003. ISBN: 9781450374224. DOI: 10.1145/882082.882086. URL: <https://doi.org/10.1145/882082.882086>.