

КЛАСИФІКАЦІЯ ЧАСОВИХ РЯДІВ ІЗ ЗАСТОСУВАННЯМ РЕКУРЕНТНИХ ДІАГРАМ МЕТОДАМИ МАШИННОГО НАВЧАННЯ

Кіріченко Л.О. д.т.н., Зінченко П.П. аспірант, Кобицька Ю.О.

Харківський національний університет радіоелектроніки

Задача класифікації часових рядів є однією з найскладніших із завдань інтелектуального аналізу даних. Існує декілька підходів до класифікації часових рядів, більшість з яких засновані на розрахунку різних метрик між рядами [1]. В останні кілька років з'явився ряд досліджень, в яких для класифікації часових рядів використовується метод рекурентних діаграм. Рекурентний аналіз заснований на такій властивості процесу, як повторюваність станів [2]. Рекурентна діаграма є масивом точок, де елемент з координатами (i, j) характеризує близькість точок i та j часового ряду в фазовому просторі. В цьому випадку рекурентні властивості часового ряду представляються у вигляді геометричних структур і дозволяють візуалізувати динаміку ряду.

У роботі розглядається метод класифікації часових рядів, заснований на побудові і візуалізації рекурентних діаграм. В цьому випадку часовий ряд перетворюється на рекурентну діаграму, яка є матрицею, що складається з нулів і одиниць. Матрицю можна представити як чорно-біле зображення. Таким чином кожному часовому ряду ми ставимо у відповідність зображення. Одним з найкращих інструментів для розпізнавання і класифікації зображень є згорткові нейронні мережі. Згорткова нейронна мережа являє собою особливу архітектуру штучних нейронних мереж, націлену на ефективне розпізнавання образів. Тому для класифікації рекурентних діаграм було обрано згорткову нейронну мережу.

Метод був досліджений на модельних та реальних часових рядах. В якості модельних часових рядів в роботі було вибрано синусоїдальні часові реалізації з різними частотами коливань і різними рівнями шуму. Для проведення класифікації часові ряди були розбиті на 2 класи з різними діапазонами частот. Результати класифікації показали, що даний метод має досить високу точність. Було отримано залежності точності класифікації від рівня зашумленості синусоїд.

Для класифікації реальних часових рядів були обрані популярні дані мозкової активності [3]. Датасет містить записи мозкової активності для різних станів людини, у тому числі записи епілептичного припадку. В більшості досліджень, які застосовували ці дані, проводилася бінарна класифікація, де в один клас виділялися записи з епілепсією та порівнювалися з іншими. В представленій роботі було проведено бінарну класифікацію методом візуалізації рекурентних діаграм. Результати показали високу точність даного методу, яка в цілому не гірше, а во многих випадках краще результатів класифікації, яка була проведена другими методами.

References

1. Krisztian Buza: Time Series Classification and its Applications WIMS '18: Proceedings of the 8th International Conference on Web Intelligence, Mining and Semantics 2018 <https://doi.org/10.1145/3227609.3227690>
2. N. Marwan, M. Romano, M. Thiel, J. Kurths, “Recurrence plots for the analysis of complex system”. Physics Reports, vol. 438, no. 5-6, pp. 237-329, 2007.
3. Epileptic Seizure Recognition Data Set
<https://archive.ics.uci.edu/ml/datasets/Epileptic+Seizure+Recognition>