

АНАЛІЗ ТА КЛАСТЕРІЗАЦІЯ ЗОБРАЖЕНЬ, ПРЕДСТАВЛЕНИХ У ФОРМІ ЧАСОВИХ РЯДІВ

Пригодій А.І.

Науковий керівник – доц. Кобилін О.А.

Харківський національний університет радіоелектроніки
(61166, Харків, пр. Науки, 14, каф. Інформатики, тел. (057) 702-14-19)
e-mail: anna.pryhodii@nure.ua

The development of image processing techniques leads to the emergence of all new approaches to solving clustering problems. The work is devoted to the study of the clustering of time series. The last tendencies in machine learning are considered. An overview of modern services for machine learning was conducted.

Останнім часом існує необхідність в розробці нових методів аналізу та представлення зображень. Зображення можливо представити в формі часових рядів з подальшим аналізом та кластеризацією.

Існують дві основні мети аналізу часових рядів: визначення природи ряду і прогнозування, тобто передбачення майбутніх значень часового ряду по теперішнім і минулим значенням.

Для кластеризації часових рядів можуть бути використанні різноманітні сервіси, що дозволяють проводити дослідження. Такі сервіси пропонують різні компанії, наприклад, такі як Microsoft, Amazon та Azure.

Amazon використовує рекурентні нейронні мережі (зокрема, LSTM), а також згортки і механізми уваги, обернувши все це на верхнєрівневу бібліотеку. Зокрема, GluonTS містить:

- інструменти, необхідні для побудови і навчання найбільш поширених архітектур нейронних мереж, а також компоненти для моделювання та трансформації ймовірнісних розподілів;
- механізми для завантаження і попередньої обробки даних, в тому числі автоматичної генерації ознак з часових рядів;
- кілька готових до використання state-of-the-art прогнозних моделей;
- інструменти для оцінки та порівняння різних моделей.

GluonTS дозволяє швидко отримати якісну модель часового ряду, на виході з якої, замість точкової оцінки прогнозних значень, отримуємо цілий змодельований ймовірнісний розподіл, а також надає можливість розділяти часові ряди на кластери.

Служба Azure Data Explorer (ADX) виконує безперервний збір даних телеметрії з хмарних служб або з пристроїв Інтернету речей. Ці дані можна аналізувати для отримання різних аналітичних відомостей, наприклад моніторингу працездатності служб, фізичних виробничих процесів та тенденцій використання. Аналіз виконується за часовими рядами обраних метрик для знаходження відхилень у закономірності в порівнянні з

типовою базовою закономірністю. ADX містить вбудовану підтримку створення, обробки та аналізу різних часових рядів.

API Cloud Inference дозволяє легко інтегрувати технології пошуку та аналізу Google для даних часових рядів у будь-яку програму. API Cloud Inference дозволяє:

- обробку наборів даних часових рядів;
- завантаження з JSON у внутрішні формати з ефективним запитом;
- видалити раніше відправлені набори даних з системи;
- список активних наборів даних в системі, представлених вашим проектом;
- виконати запити виведення по завантажених наборах даних.

Таким чином, можливо зробити висновок, що для аналізу та кластеризації зображень, представлених у формі часових рядів, є можливість використовувати сервіси для більш точного та ефективного кінцевого результату досліджень.

Список використаних джерел:

1. Шумейко, А. А., & Сотник, С. Л. (2012). Интеллектуальный анализ данных (введение в Data Mining). *Днепропетровск: Белая ЕА*, 212.
2. Jain, A. K., Murty, M. N., & Flynn, P. J. (1999). Data clustering: a review. *ACM computing surveys (CSUR)*, 31(3), 264-323.
3. Steinley, D. (2006). K- means clustering: a half- century synthesis. *British Journal of Mathematical and Statistical Psychology*, 59(1), 1-34.
4. Huang, Z., & Ng, M. K. (1999). A fuzzy k-modes algorithm for clustering categorical data. *IEEE Transactions on Fuzzy Systems*, 7(4), 446-452.
5. Bodyanskiy, Y., Vynokurova, O., Kobylin, I., & Kobylin, O. (2016). Adaptive fuzzy clustering of short time series with unevenly distributed observations in Data Stream Mining tasks. *Information Technology and Management Science*, 19(1), 23-28.
6. Rabotiahov, A., Kobylin, O., Dudar, Z., & Lyashenko, V. (2018, February). Bionic image segmentation of cytology samples method. In 2018 14th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET) (pp. 665-670). IEEE.
7. Oleg, K., Sergii, M., & Mykhailo, S. (2017, October). Video Clustering via Multidimensional Time-Series Analysis. In *Proceedings of the 9th International Conference on Information Management and Engineering* (pp. 60-63). ACM.