**Сравнение различных методов заполнения пропусков во временных рядах**

***Захаров И.М.* 1*, Арзангулян М.Э.* 2**

1*студент,* 2*студент*

*Московский государственный университет имени М.В.Ломоносова,*

*Физический факультет, Москва, Россия*

*E-mail:* [*z4haroff.ivan@yandex.ru*](mailto:z4haroff.ivan@yandex.ru)

Временные ряды, используемые в практических приложениях, часто содержат пропуски данных, вызванные сбоями в процессе сбора информации, техническими неполадками или человеческим фактором. Наличие пропусков значительно усложняет анализ, поскольку большинство методов работы с временными рядами требуют полных и непрерывных данных. Для решения этой проблемы разработаны различные подходы к заполнению пропусков, однако их эффективность во многом зависит от характера данных, размера и расположения пропусков, а также наличия сезонности или трендов.

В данной работе представлен сравнительный анализ нескольких популярных методов заполнения пропусков, включая линейную интерполяцию, экспоненциальное сглаживание с двусторонним подходом[1], метод k-ближайших соседей (k-NN) и инструменты открытой библиотеки Prophet[2]. В рамках исследования проведено сравнение эффективности на различных типах данных, включая синтетические и реальные временные ряды, с учетом таких факторов, как длина ряда, размер пропусков и наличие сезонности или трендов. Для оценки эффективности использовались средняя абсолютная ошибка (MAE)[3], время срабатывания алгоритма и другие метрики.

По результатам исследования можно понять, какие методы лучше всего себя проявляют в различных задачах, учитывая их эффективность и точность, что позволяет сделать выбор в пользу одного из алгоритмов в конкретных задачах. Вышепредставленные методы также интегрированы в программный комплекс для обработки временных рядов[4], вместе с указаниями их преимуществ на различных данных, что позволит пользователям выбирать наиболее подходящий метод для конкретной задачи.

******

Рис. 1 Сравнение методов заполнения пропусков

Литература:

1. Заполнение экспоненциальным сглаживанием: <https://github.com/carloalbe/fill-large-gaps-in-timeseries-using-forecasting>
2. Библиотека для работы с временными рядами Prophet: <https://github.com/facebook/prophet>
3. Rob J Hyndman and George Athanasopoulos, Forecasting: Principles and Practice: <https://otexts.com/fpp2/accuracy.html>
4. Программный комплекс для обработки временных рядов: <https://github.com/Arzangulyan/Time_series_app>
5. М.Л. Кричевский, Временные ряды в менеджменте часть 1
6. Everette S. Gardner Jr., Exponential smoothing: The state of the art—Part II