

**Анализ временных рядов с сезонностью
на примере данных объёма трафика "Яндекс.Радара"**

А. С. Армаев, А.А.Балагура

Ключевые слова: временной ряд, Яндекс.Радар, статистические методы для временного ряда, машинные методы для временного ряда.

Временные ряды могут иметь сложную структуру [1, с.15] включающую трендовую составляющую, множественную сезонность и влияние внешних факторов. Примером таких рядов являются ряды, отражающие пользовательскую активность в интернете. В данной работе исследуется временной ряд, отражающий динамику поисковых визитов в системе Яндекс.

Цель исследования - анализ временных рядов с несколькими сезонными периодами методами машинного обучения и статистическими методами на примере данных о Яндекс-визитах, а также построение и выбор наиболее точной прогнозной модели.

Информационной базой исследования послужил датасет, содержащий ежедневные данные об объеме поисковых визитов в Яндексе за период с 09.07.2023 по 11.02.2025, полученный из источника, связанного с сервисом Яндекс.Радар. Горизонт прогнозирования составляет 30 дней.

Для данного временного ряда выполнен предварительный анализ, который показал наличие следующих паттернов: сложного восходящего тренда, ярко выраженной 7-дневной сезонности, годовой сезонности, влияние внешних факторов – выходных и праздничных дней и пр. Проведена подготовка данных, включая дифференцирование ряда для достижения стационарности и конструирование набора внешних признаков: циклических (день года/месяца/недели, месяц), бинарных (выходной, праздник), категориальных (день недели с one-hot кодированием) и порядковых (год).

Построены и оценены различные модели: SARIMA [5, с298] линейная регрессия [3, с.129], градиентный бустинг (CatBoost) [4, с.42], а также их ансамбли. Лучшие результаты получены на ансамбле моделей линейной регрессии и CatBoost (MAPE = 3.2%, RMSE = 7.4 млн.виз.)

Полученные результаты подтверждают известную эффективность комбинированного подхода [2, с.225], объединяющего статистические методы и методы машинного обучения.

1. Андерсон Т. В. Статистический анализ временных рядов = The statistical analysis of time series / пер. с англ. И. Г. Журбенко и В. П. Носко ; под ред. Ю. К. Беляева. — Москва : Мир, 1976. — 755 с.
2. Брюс П. Практическая статистика для специалистов Data Science: Пер. с англ. / П. Брюс. — Санкт-Петербург : БХВ-Петербург, 2018. - 304 с.
3. Введение в статистическое обучение с примерами на языке Python / Г. Джеймс, Д. Уиттен, Т. Хасты, Р. Тибширани, Дж. Тейлор ; пер. с англ. А. Ю. Гинько. — Москва : ДМК Пресс, 2024. — 846 с.
4. Кугаевских А. В. Классические методы машинного обучения / А. В. Кугаевских, Д. И. Муромцев, О. В. Кирсанова. – СПб : Университет ИТМО, 2022. – 53 с.
5. Магнус Я. Р. Эконометрика. Начальный курс / Я. Р. Магнус, П. К. Катышев, А. А. Пересецкий. – 6-е изд., перераб. и доп. – Москва : Дело, 2004. – 576 с.