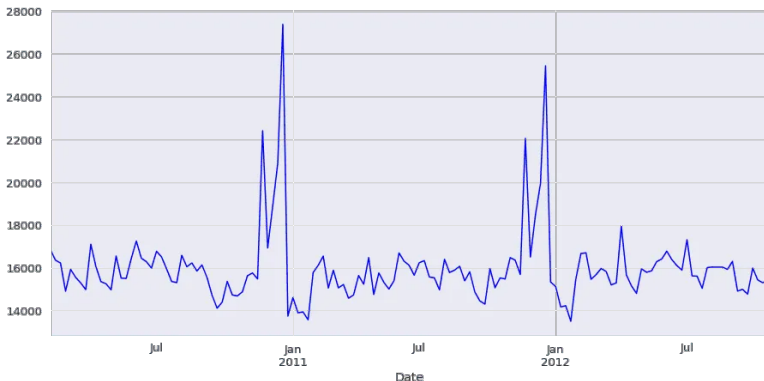


Анализ временных рядов

Временные ряды

Временной ряд — значения меняющихся во времени признаков, полученные в некоторые моменты времени.



Временные ряды. Задачи

- Прогнозирование временных рядов (в прямом и обратном времени)
- Восстановление пропущенных значений
- Фильтрация
- Классификация временных рядов или их интервалов
- Кластеризация временных рядов или их интервалов
- Использование как признака
- etc

Свойства временных рядов

- Стационарность
- Белый шум / случайное блуждание
- Иерархичность
- Один ряд или несколько
- Наличие внешних признаков

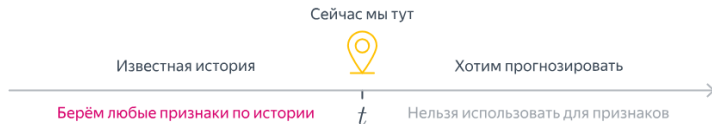
Задача предсказания временных рядов

Дан временной ряд до момента T . Построить предсказания временного ряда на моменты $T + 1, \dots, T + N$.

- классические методы (скользящее среднее, авторегрессия, ...)
- методы на основе машинного обучения

Сведение к задачи регрессии

$$y_t = f(y_{t-1}, \dots, y_{t-p})$$



Признаки

- Производные от даты (ден недели, месяц, час, выходной, ...)
- Предыдущие значения ряда
- Агрегированные предыдущие значения (скользящее окно)
- Сезонность
- Счётчики — группировка по категориальным данным
- Разность между значениями

Стратегии предсказания

Пусть требуется предсказать N значений. Как это сделать?

Стратегии:

- Рекурсивная — обучаем модель предсказывать одно следующее значение и рекурсивно вызываем N раз.
- Прямая — обучаем N моделей для каждого времени.
- Гибридная — N моделей, использующих предсказания предыдущих.
- MIMO — одна модель предсказывает N значений.

Оценка качества. Метрики

Метрики для задач регрессии:

- MSE
- MAE
- MAPE

Схемы валидации



Декомпозиция временных рядов

Декомпозиция – процедура разложения временного ряда на три временных компоненты:

- 1 Тренд T_t – плавное долгосрочное изменение временного ряда.
- 2 Сезонность S_t – циклические изменения временного ряда с постоянным периодом сезонности ss .
- 3 Ошибка R_t – непрогнозируемая случайная компонента ряда.

Обычно аддитивная или мультипликативная декомпозиция

Декомпозиция временных рядов

