

# Временные последовательности

1. Найдите набор данных с временным рядом. Например, это может быть последовательность цен акций, погоды, загруженности дорог, потребления воды и т.д. Если набор данных очень большой, сожмите его агрегируя суммой, например просуммируйте значения внутри дня, чтобы получить измерения по дням, если изначально они были по часам или минутам.
2. Проанализируйте ряд на наличие тренда, сезонности.
3. Проверьте, как на ряд влияет логарифмирование и дифференцирование. Постарайтесь сделать ряд стационарным. Запомните проделанные преобразования для применения обратных после предсказания.
4. Извлеките признаки из метки времени, например: день в году, день в недели, день в месяце, неделя в году, месяц в году и т.д. Добавьте новые признаки применив  $\sin$  и  $\cos$  к нормированным характеристикам метки времени. Сведите задачу к табличной задаче регрессии. Решите её любой моделью регрессии, можно использовать модель из прошлой лабы, либо любую другую из библиотек, если вы понимаете, как она работает.
5. Найдите в библиотеках, либо реализуйте не менее двух специальных моделей для временных рядов, например: модель Хольта-Уинтерса и SARIMA. Также используйте одну простую (наивную) модель, например, возвращающую среднее значение.
6. Обучите модели на 80% временного ряда. Предскажите оставшиеся 20% (как минимум 3 точки) ряда с помощью моделей. Если модели позволяют, постройте предсказание и на тренировочной выборке. Отобразите на графике предсказание моделей и истинные данные, а также границу начала тестовой части. Посчитайте метрику оценки для задачи регрессии. Метрику можно выбрать любую. Постройте график зависимости от времени разности между предсказаниями и реальными данными.