

Содержание.

1. Метрики качества.
2. RMSE.
3. Normalised RMSE.
4. MAPE.
5. MASE.
6. Обнаружение аномалий.
7. Онлайн мониторинг.
8. Оффлайн мониторинг.
9. Мониторинг на многомерных рядах.

Метрики качества.

Как оценить качество предсказания?

1. Проверить остатки на стационарность.
2. Проверить по критерию AIC.
3. По некоторой абсолютной метрике.
4. В сравнении с другими моделями.

RMSE.

$$RMSE = \sqrt{\sum_{i=1}^n \frac{(\hat{y}_i - y_i)^2}{n}}$$

В чем проблема RMSE?

- а) Нельзя сравнивать ряды разного масштаба.
- б) Неустойчива к выбросам.

MAPE

Средняя абсолютная ошибка, выраженная в процентах.

Устойчива к выбросам, относительна.

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{y_t - y'_t}{y_t} \right|$$

R2 - ошибка

Иначе - коэффициент детерминации.

Показывает процент дисперсии, объясненный моделью.

$$R^2 = 1 - \frac{\sigma^2}{\sigma_y^2}$$

MASE

Mean absolute scaled error.

Процент дисперсии, объясненный моделью, в сравнении с другой моделью.

Или, насколько больше данная модель объясняет разброс, чем модель в сравнении.

$$\text{MASE} = \frac{1}{T} \sum_{t=1}^T \left(\frac{|e_t|}{\frac{1}{T-1} \sum_{t=2}^T |Y_t - Y_{t-1}|} \right) = \frac{\sum_{t=1}^T |e_t|}{\frac{T}{T-1} \sum_{t=2}^T |Y_t - Y_{t-1}|}$$

В числителе стоит ошибка прогноза валидируемого метода предсказания, $e_t = Y_t - F_t$
в знаменателе средняя абсолютная ошибка метода наивного предсказания.

MASE

Так, если MASE равно 0.2 - это значит, что наша модель объясняет разброс данных в 5 раз лучше, чем наивное предсказания.

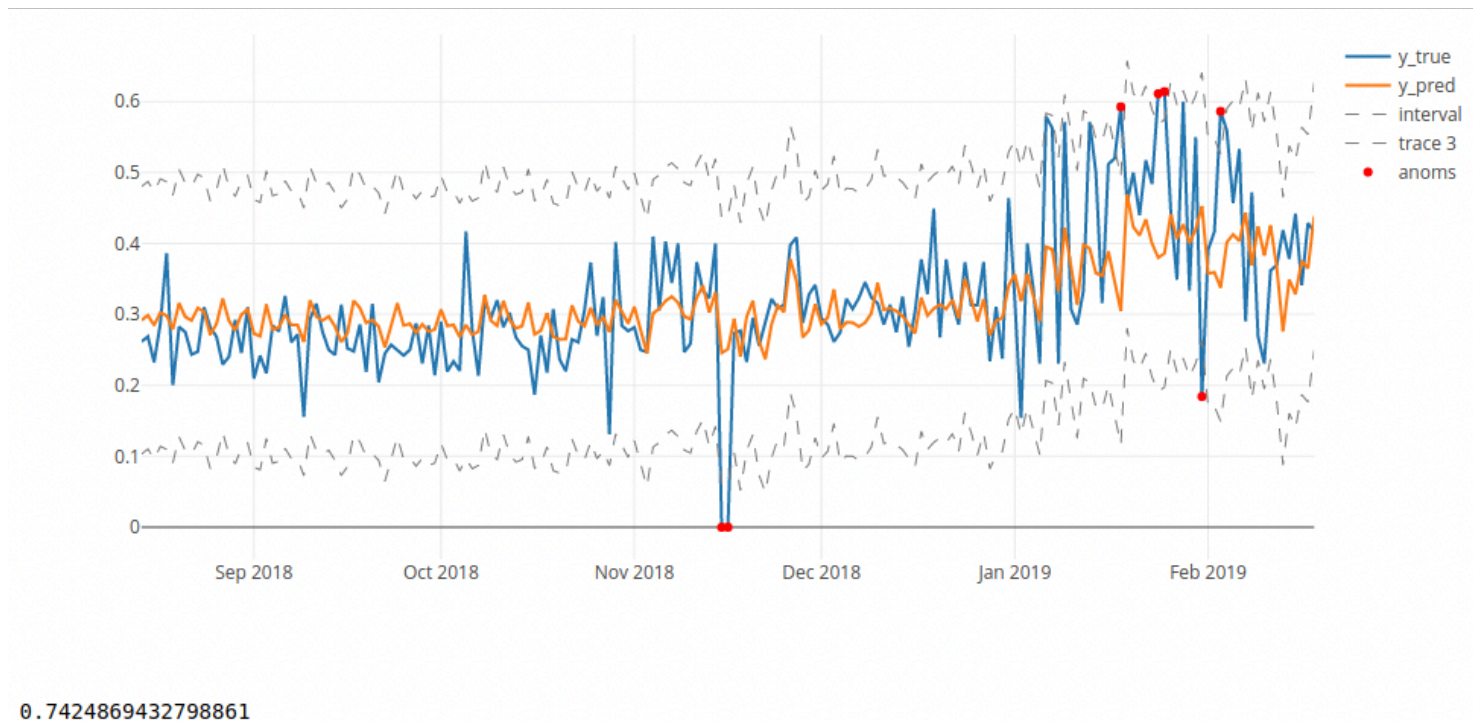
Если 1, то ничуть ни лучше.

В знаменателе обязательно должна стоять ошибка от наивного метода, можно выбрать произвольный и сравнить.

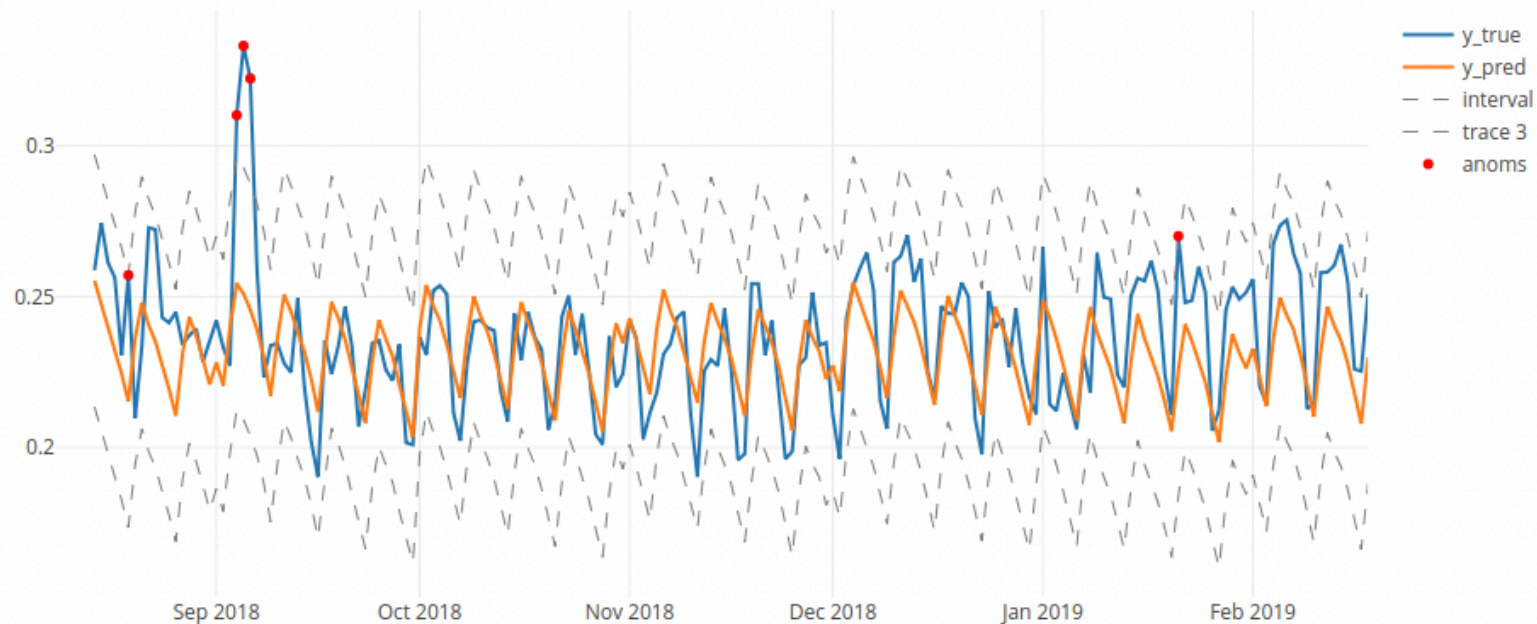
Типы аномалий.

- Выбросы
- Множественные выбросы
- Сдвиги
- Изменение тренда
- Изменение дисперсии

Выбросы.



Множественные выбросы.

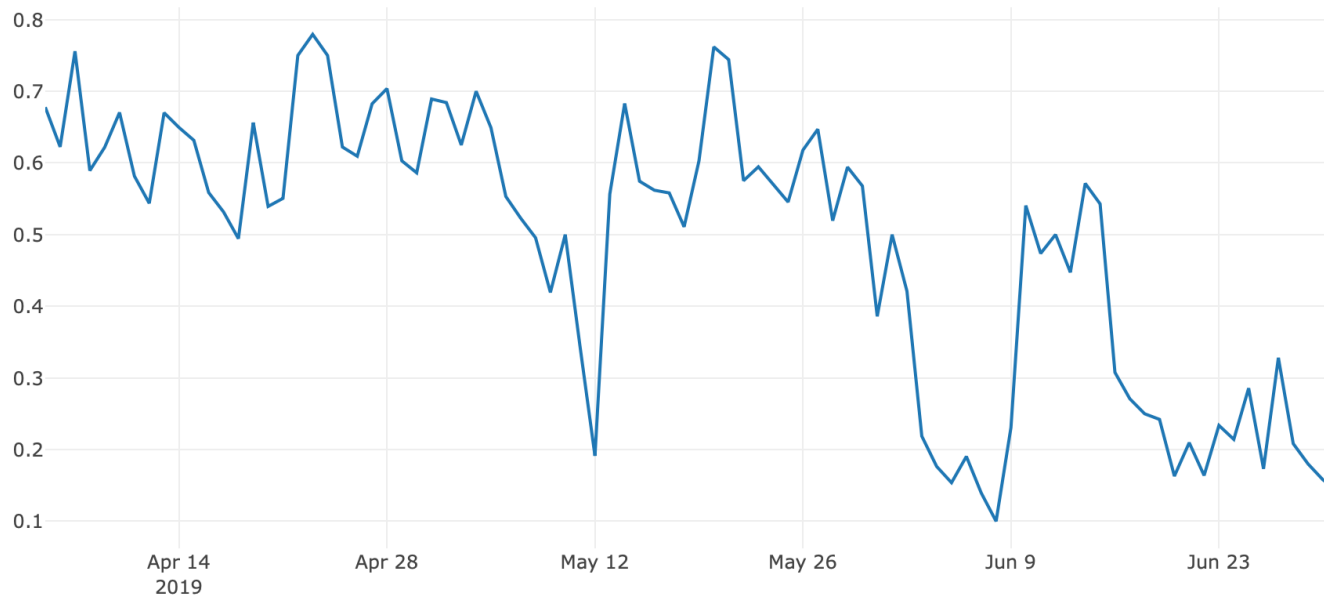


0.8742500970488297
0.306897496373706

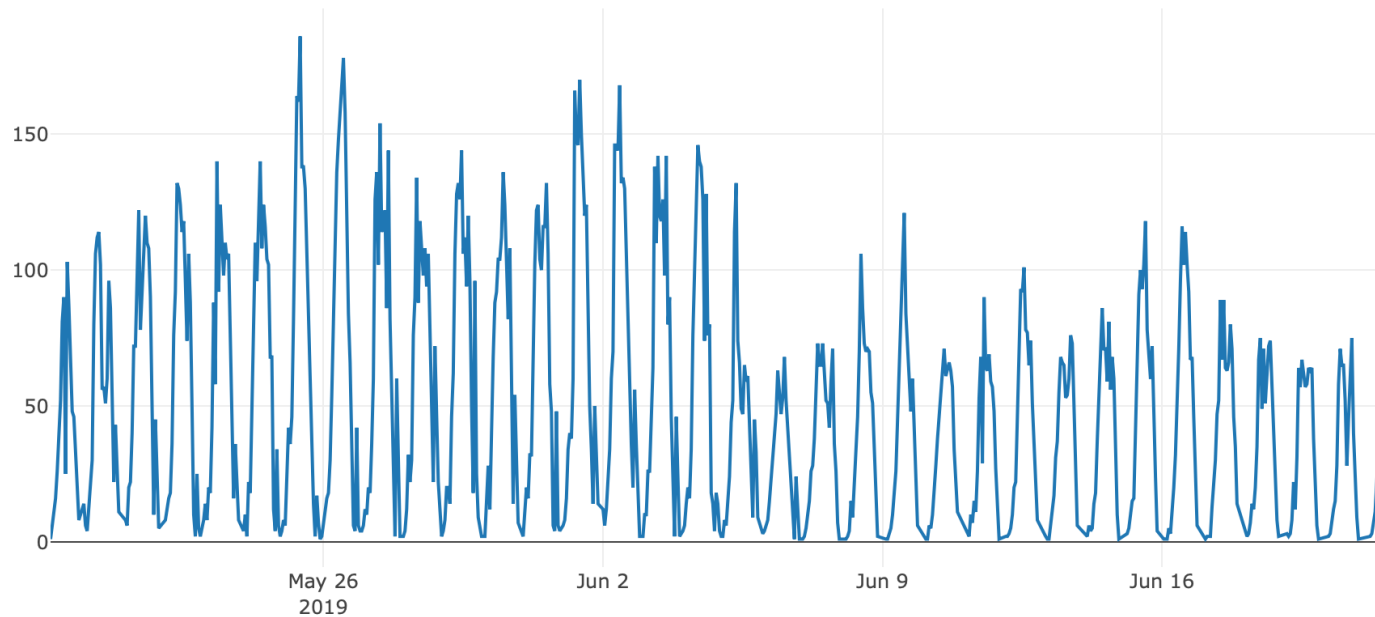
Сдвиги.



Изменение тренда.



Изменение дисперсии.

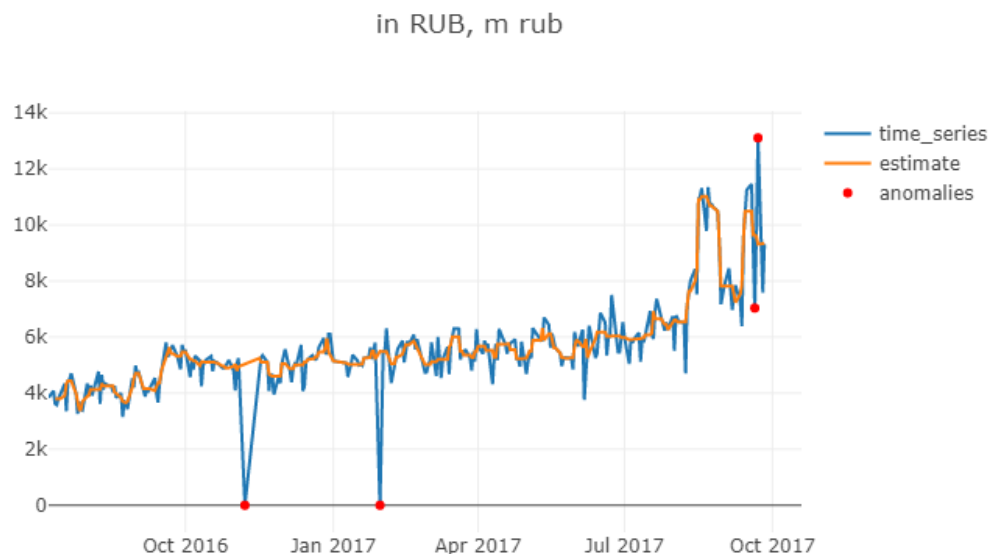


Обнаружение аномалий.

- Оффлайн мониторинг
- Онлайн мониторинг
- Одномерные
- Многомерные

Оффлайн мониторинг

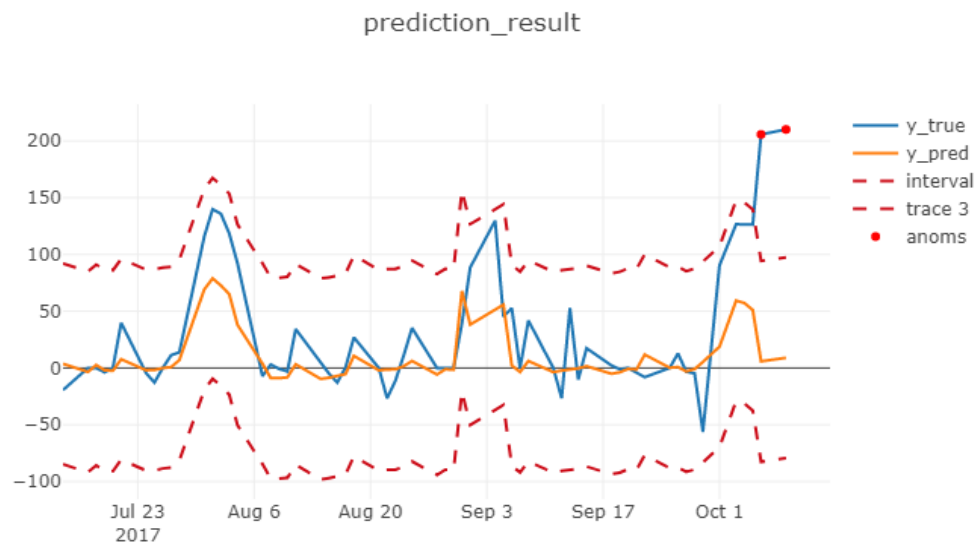
Задача: раз в n промежутков времени высылается батч из n точек, определить, какие из точек были аномальными. Хорошо использовать - медианный фильтр, stl разложение



Онлайн мониторинг

Задача: для каждой вновь пришедшей точки необходимо сделать вердикт об аномальности.

Используют: предсказание временного ряда + построение доверительных интервалов.



Онлайн мониторинг

Метод классических доверительных интервалов.

1. Получаем ошибки предсказания на train-e
2. Считаем стандартное отклонение σ
3. Откладываем от каждой точки интервалы $\pm k\sigma$

Онлайн мониторинг

Метод сезонных доверительных интервалов.

1. Получаем ошибки предсказания на train-е
2. Считаем стандартное отклонение $\sigma_i, i \in (t_1, \dots, t_s)$
3. Откладываем от каждой точки интервалы $\pm k_i \sigma_i$

Подход чистой истории ряда.

Для каждой точки.

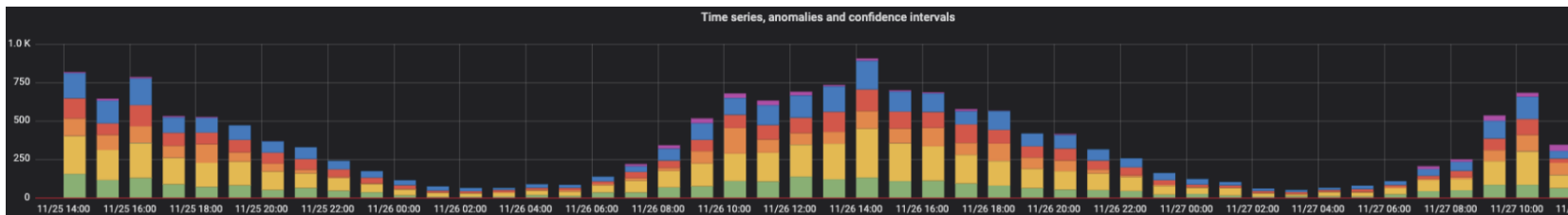
Если она аномальна, заменяем реальную точку на предсказанную, таким образом избавляемся от влияния аномалии на следующее предсказание.

Мониторинг на многомерных рядах.

Расчет stability index.

Идея: многомерный временной ряд - это по сути n-мерное распределение, меняющееся в каждый момент времени.

Population stability index - статистика, которая показывает, насколько сильно распределение в текущей точке отличается от некоторой предыдущей точки.



$$SI = \sum_{i=1}^N (p(v_t^i) - p(v_{t-k}^i)) * (\ln(p(v_t^i)) - \ln(p(v_{t-k}^i)))$$