

Подбор параметров в алгоритмах детектирования аномалий

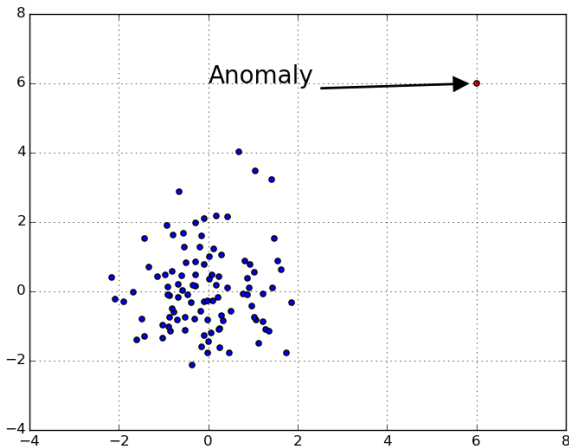
Смоляков Дмитрий

ИППИ РАН

Конференция МФТИ 2014

Определение аномалии

Аномалия - измерение, которое не соответствует некоторому ожидаемому поведению.

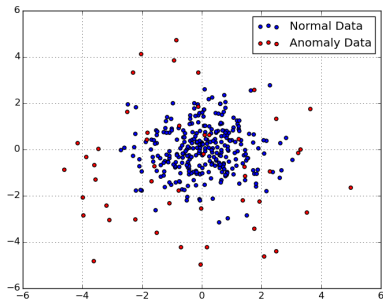


Области применения

- Поиск мошенничества на основе странного поведения
- В медицине необычные результаты обследований могут свидетельствовать о потенциальных проблемах со здоровьем
- Аномальные значения могут быть результатом ошибок в измерении
- Раннее обнаружение аномальной работы двигателя может предотвратить неожиданную поломку

Формальное определение

- Пусть Q некоторый механизм порождения данных
- $X_1 \dots X_n$ порождены механизмом Q
- $\tilde{X}_1 \dots \tilde{X}_m$ порождены каким-либо иным механизмом
- Тогда будем считать, точки $X_1 \dots X_n$ **нормальными** измерениями, а точки $\tilde{X}_1 \dots \tilde{X}_m$ **аномалиями**



Основные подходы

- Оценка распределения данных из нормального класса
- Метод k ближайших соседей
-