

# Anomaly detection service

Поляков Семен

# Проект

Система сбора и анализа данных для отдела маркетинга:

Сбор и хранение данных с внешних источников

Анализ данных, с сохранением результата в БД

Визуализация и выгрузка результатов

# Обнаружение “аномалий”

В распоряжении оказались временные ряды по накопленной статистике использования продуктов JetBrains.

Поиск аномалий следующих видов:

1. Резкие отклонения
2. Перегибы тренда

# Известные решения

- 1) Смена тренда: пересечение скользящих средних
- 2) Поиск отклонений: много решений, но для других видов данных (спектры, без тренда, тренд известен, дисперсия фиксирована и т. д.)

# Модель

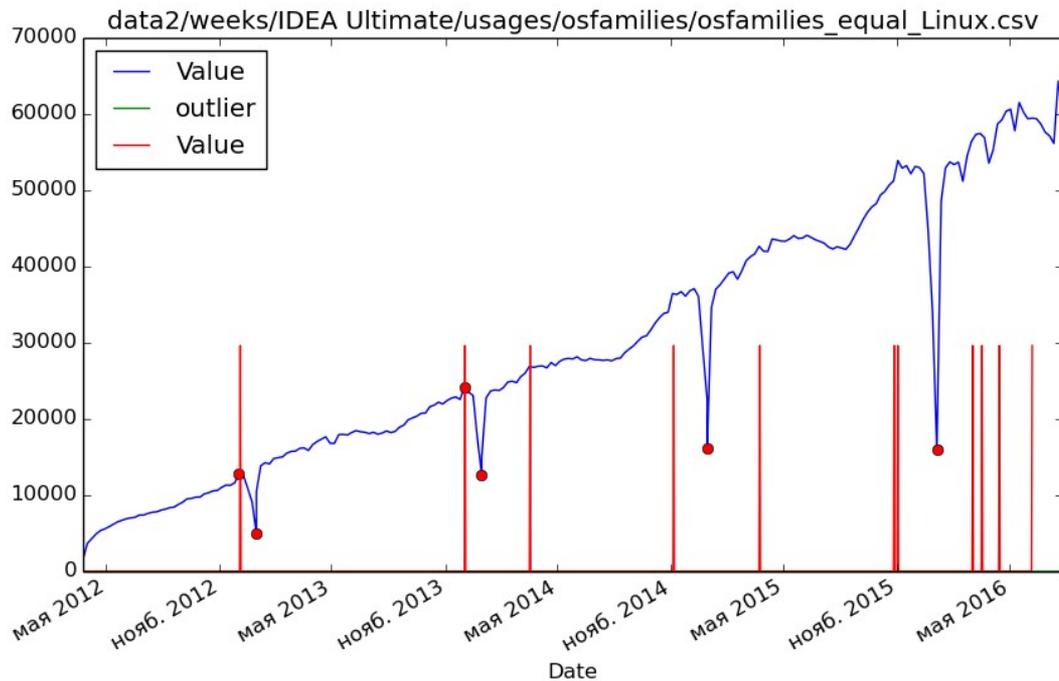
1. Аппроксимация тренда - LOWESS
2. Грубый поиск отклонений
3. Фильтрация - Standard Deviation по скользящему окну

Проблема - много параметров, для каждого ряда нужно подбирать

Решение - оптимизация параметров для максимизации “качества” найденных отклонений (эвристическая модель)

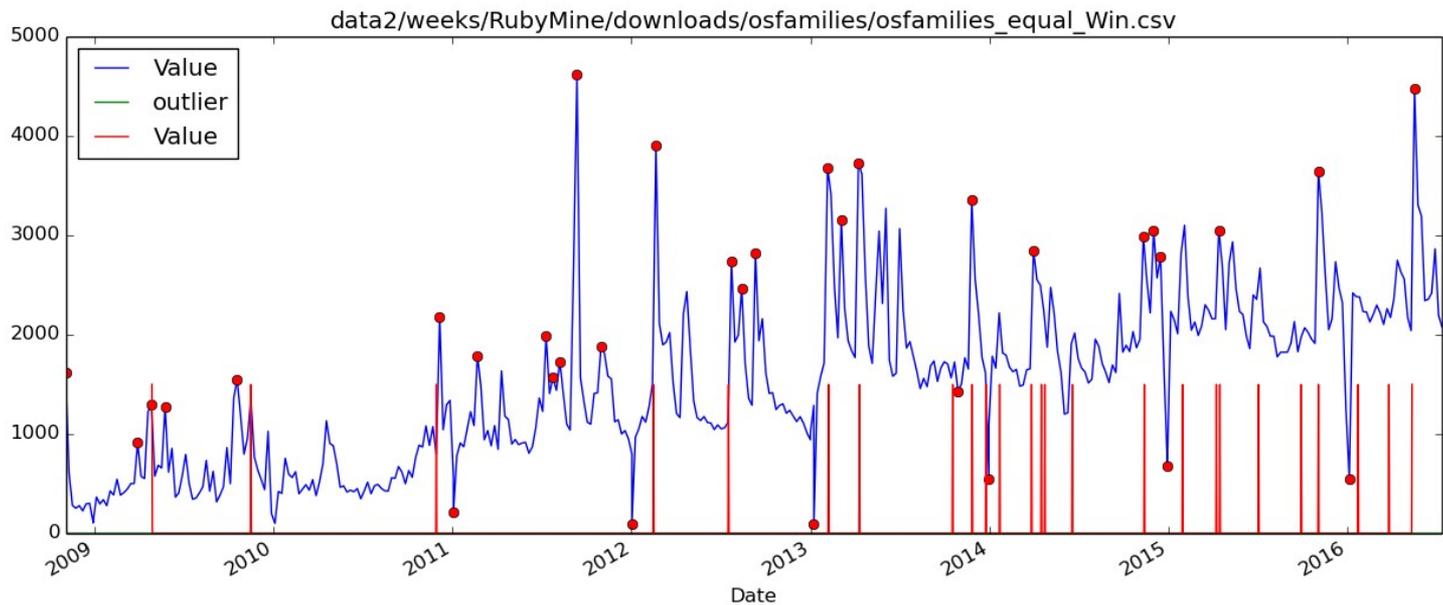
# Результаты модели

Найденные резкие отклонения



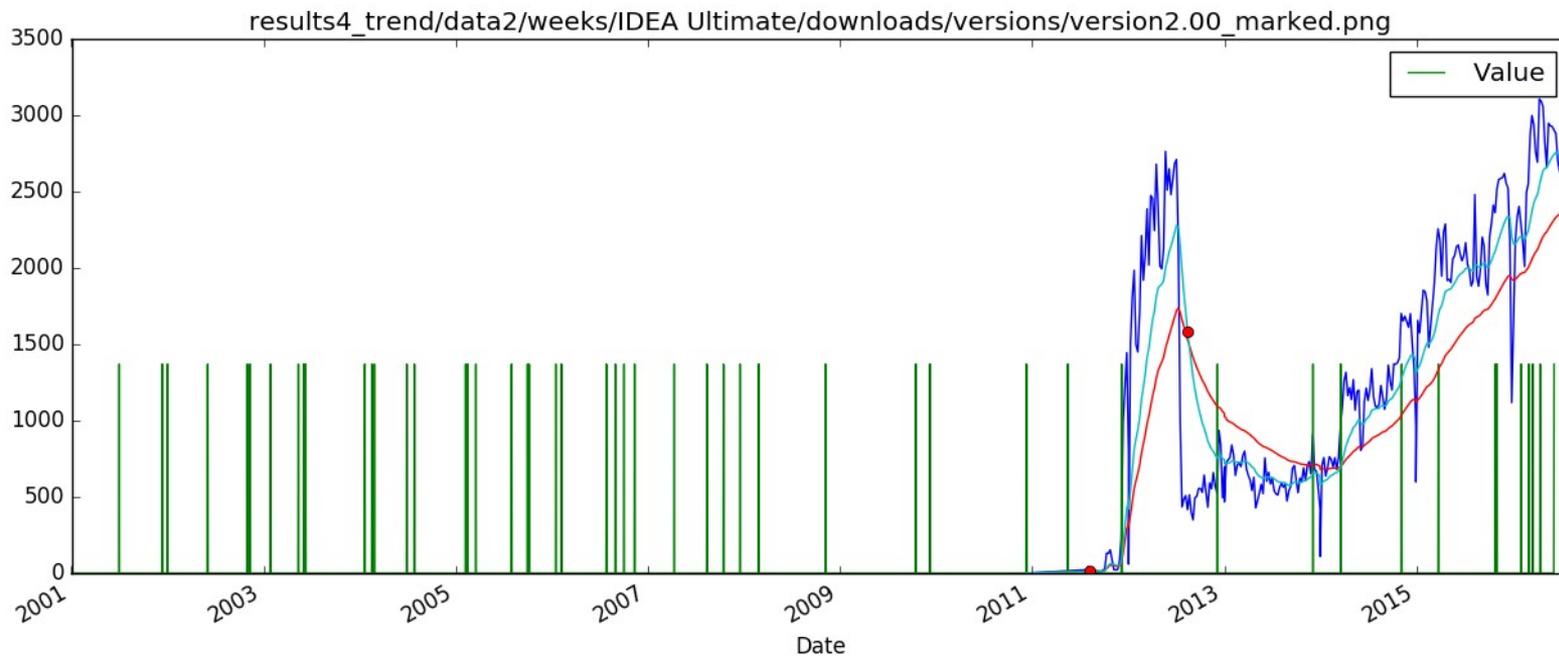
# Результаты модели

Найденные резкие отклонения



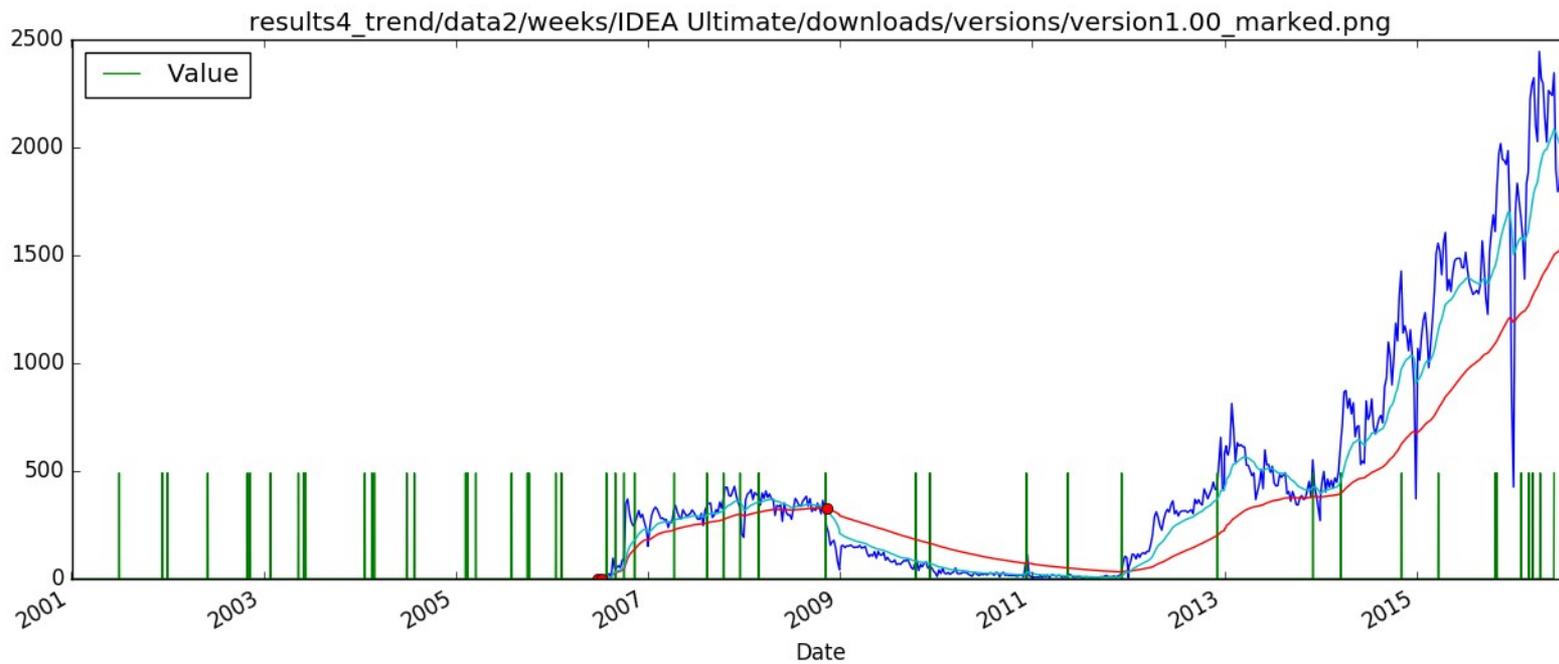
# Результаты модели

Смена тренда (вниз)



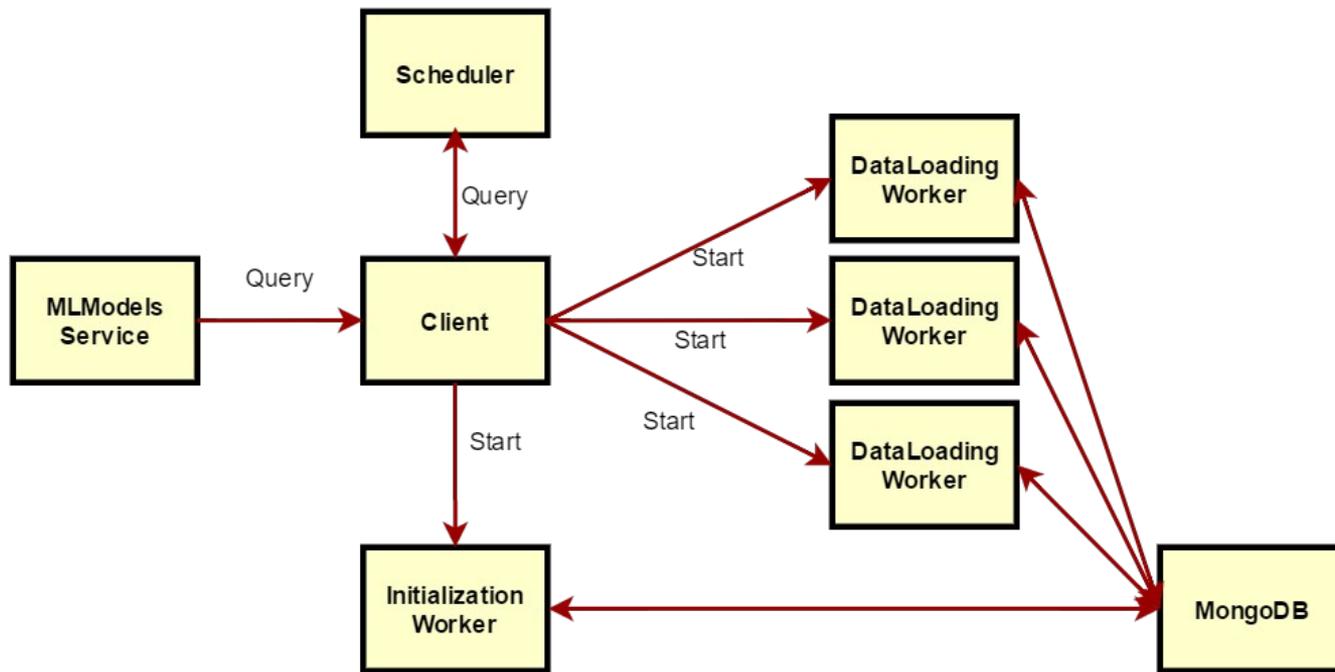
# Результаты модели

Смена тренда (вниз)

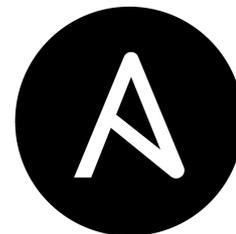
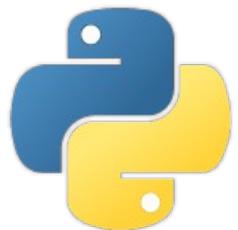


# Сервис

REST API сервис, позволяющий скачивать данные и обнаруживать в них аномалии с помощью реализованной модели.



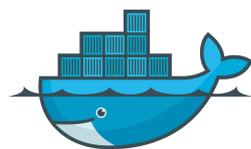
# Технологии



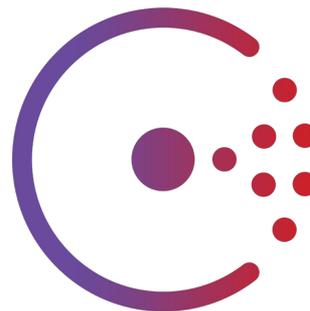
ANSIBLE



mongoDB



docker



Q&A