

ПОКАЗЮК

Автоматизированная система для анализа
аномалий в декларациях

Актуальность

- Всё больше и больше данных доступны в электронном виде
- Всё больше и больше воруют
- Человеку своими силами сложно проверить согласованность деклараций

Требуется автоматизация процесса анализа деклараций

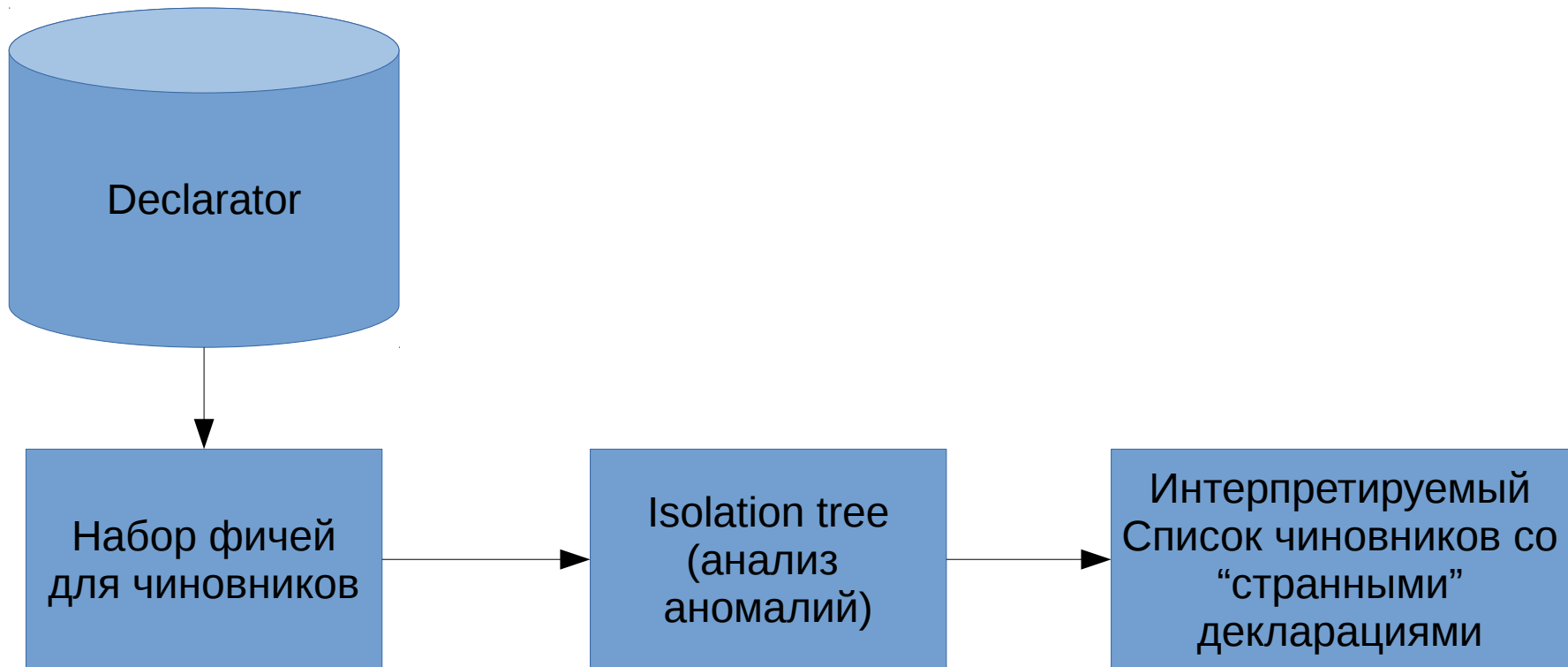
Цель и задачи

Цель: создание автоматизированной системы для анализа деклараций на предмет мошенничества

Задачи:

- Формирование списка критериев для анализа (маркеров мошенничества)
- Оценка стоимости недвижимости и транспортных средств
- Кластеризация и анализ аномалий

Архитектура



Список критериев

- **Согласованность задекларированной прибыли со стоимостью задекларированного имущества**
- **Задекларированный доход (в т.ч. у родственников)**
- **Количество машин (в т.ч. у родственников)**
- **Количество квартир (в т.ч. у родственников)**
- ...

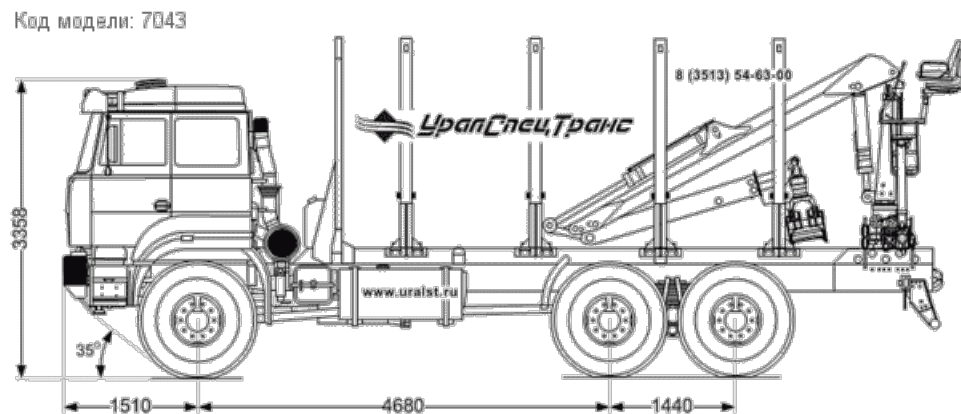
Оценка стоимости недвижимости

- Взяли статистику из <http://www.gks.ru> по годам для различных регионов:

	region	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
0	Российская Федерация	8070.0	9421.0	12051.0	14859.0	19282.0	23223.0	29742.0	42594.0	51299.0	49313.0	48035.0	42201.0
2	Центральный федеральный округ	10663.0	12210.0	14688.0	19332.0	26640.0	32294.0	40083.0	55520.0	67201.0	64316.0	63816.0	51521.0
4	Белгородская область	5111.0	5785.0	7509.0	10586.0	12735.0	15092.0	19710.0	29827.0	40886.0	42394.0	39416.0	39561.0
6	Брянская область	4658.0	4678.0	7777.0	9502.0	12890.0	14753.0	15971.0	23140.0	31619.0	32472.0	31347.0	28157.0
8	Владимирская область	4797.0	5955.0	8622.0	11127.0	14716.0	16807.0	21808.0	30422.0	39885.0	37664.0	36996.0	37452.0
10	Воронежская область	4535.0	5502.0	7054.0	9553.0	12139.0	13543.0	16515.0	25544.0	34057.0	34399.0	34858.0	34275.0
12	Ивановская область	4707.0	5311.0	7221.0	9725.0	10994.0	12419.0	15969.0	29726.0	31270.0	27560.0	32303.0	33143.0
14	Калужская область	7228.0	7356.0	10578.0	12884.0	14593.0	17159.0	24146.0	33552.0	53983.0	46242.0	45319.0	47332.0
16	Костромская область	4281.0	5286.0	8034.0	9992.0	11184.0	14971.0	18679.0	31312.0	38271.0	30638.0	29995.0	31788.0
18	Курская область	3949.0	5887.0	6442.0	8067.0	10318.0	13982.0	15974.0	24062.0	29716.0	28812.0	27316.0	27662.0

Оценка стоимости транспортных средств (проблема)

- В данных много мусора
- Данные неоднородные: декларируют от деревянных самодельных лодок и сортиментовозов (да) до яхт
- Данные часто неполны
- Нет свободного API для оценки стоимости вещи



Сортиментовоз
(7 млн руб.)

Оценка стоимости транспортных средств (решение)

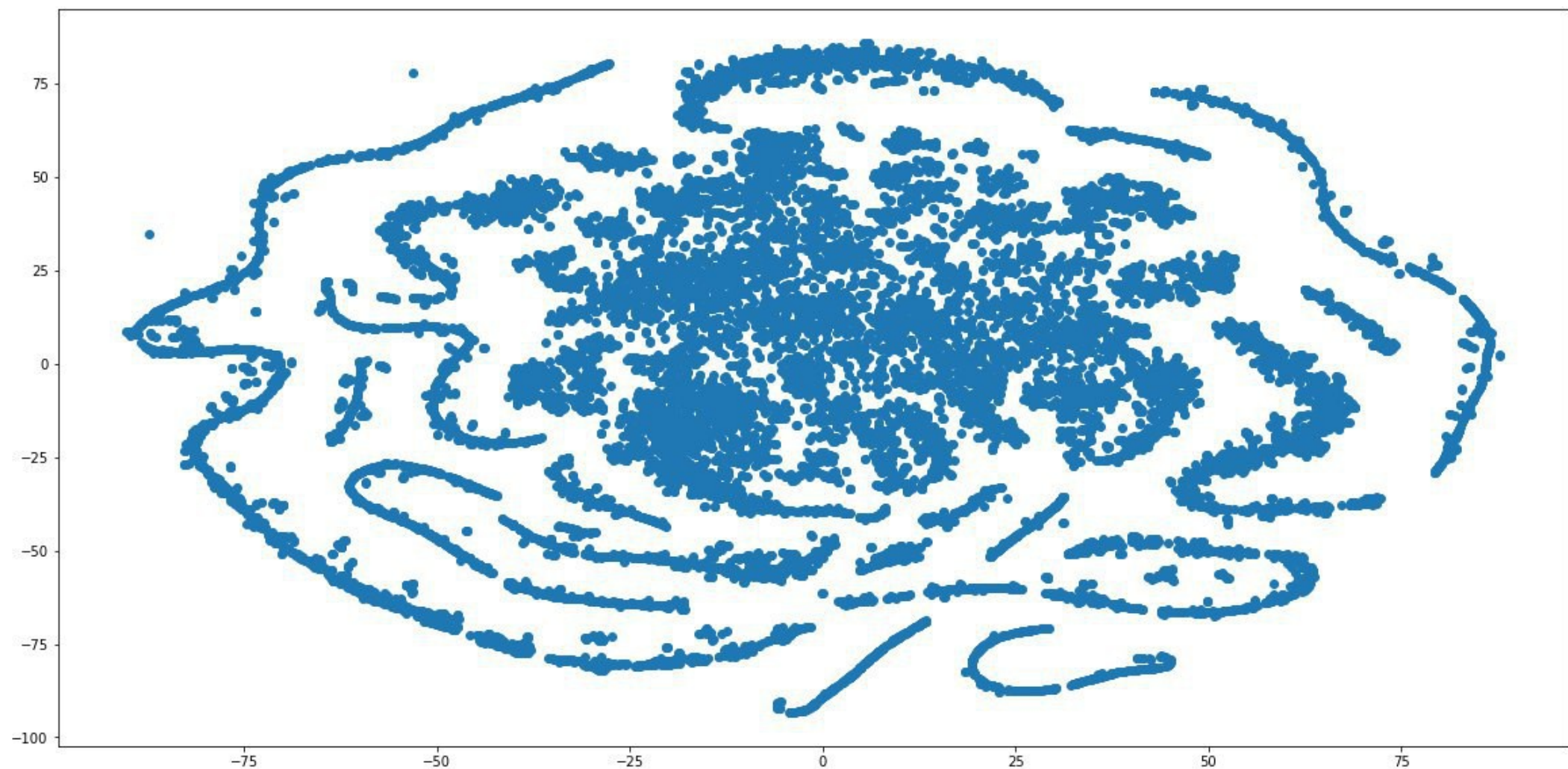
- Нет свободного API для оценки стоимости вещи, но есть Google



Кластеризация

- Посчитали признаки
- Обучили IsolationForest
- Profit

Кластеризация



Пример аномалии

Обычный представитель:
<https://declarator.org/person/4823/>

Странный представитель:
<https://declarator.org/person/14020/>

Декларация, статус, учреждение	Доход, руб.	Недвижимость, м ²	Транспорт, шт.
Антикоррупционная декларация 2014 Член Совета Федерации Федерального Собрания Российской Федерации от Брянской области - представитель в Совете Федерации Федерального Собрания Российской Федерации от исполнительного органа государственной власти Брянской области Совет Федерации Показать подробности	3 875 999	0	2 Автомобиль легковой: 2 шт.

Антикоррупционная декларация 2013 Член Совета Федерации Федерального Собрания Российской Федерации от Брянской области - представитель в Совете Федерации Федерального Собрания Российской Федерации от исполнительного органа государственной власти Брянской области Совет Федерации Показать подробности	2 476 800	0	2 Автомобиль легковой: 2 шт.
--	-----------	---	---------------------------------

Декларация, статус, учреждение	Доход, руб.	Недвижимость, м ²	Транспорт, шт.
Антикоррупционная декларация 2015 Депутат законодательного собрания Пермского края второго созыва Законодательное Собрание Пермского края Партия: ЛДПР Показать подробности	64 450 544	40 568,68 Земельные участки: 37 562 кв. м (3 шт.) Жилые дома: 1 881,8 кв. м (3 шт.) Квартиры: 7,35 кв. м (1/6 шт.) Гаражи: 159,8 кв. м (2 шт.) Иное: 957,73 кв. м (19.5 шт.)	1 Автомобиль легковой: 1 шт.

Антикоррупционная декларация 2014 Депутат законодательного собрания Пермского края второго созыва Законодательное Собрание Пермского края Партия: ЛДПР Показать подробности	449 324	3 454,07 Земельные участки: 3 400 кв. м (1 шт.) Квартиры: 22,05 кв. м (0.5 шт.) Иное: 32,02 кв. м (0.5 шт.)	1 Автомобиль легковой: 1 шт.
---	---------	---	---------------------------------

Фрагмент отчёта

ФИО	Декларатор	Аномалии
Прохоров Сергей Иванович Danger	https://declarator.org/person/14020/	<ul style="list-style-type: none">• Аномальное изменение числа объектов недвижимости у родственников• Аномальное изменение дохода• Аномальное количество недвижимости у родственников• Аномальное изменение дохода у родственников
Васильев Сергей Львович Danger	https://declarator.org/person/13374/	<ul style="list-style-type: none">• Аномальное изменение числа объектов недвижимости у родственников• Аномальное количество недвижимости у родственников• Аномальное изменение числа объектов недвижимости
Кочетков Александр Анатольевич Danger	https://declarator.org/person/15009/	<ul style="list-style-type: none">• Аномальное изменение числа объектов недвижимости у родственников• Аномальное количество недвижимости у родственников• Аномальное изменение дохода у родственников
Мамедов Хикмет Алигейдар Danger	https://declarator.org/person/12674/	<ul style="list-style-type: none">• Аномальное количество транспортных средств• Аномальное количество недвижимости у родственников• Аномальный доход у родственников
Шпаков Юрий Николаевич Danger	https://declarator.org/person/13787/	<ul style="list-style-type: none">• Аномальное изменение числа объектов недвижимости у родственников• Аномальное количество недвижимости у родственников• Аномальный доход у родственников

Результаты

- **Выявили аномалии.** Интерпретируемый результат находится в свободном доступе (на github) в сводной табличке
- **Обогатили исходные данные:**
 - получили оценку стоимости недвижимости и транспортных средств
 - Нашли мусор в транспортных данных (оценка снизу – нет результатов в поисковых запросах)

ССЫЛОЧКИ

<https://github.com/generall/declaramboom>

Перспективы развития

- Улучшить качество анализа аномалий
- Улучшить интерпретируемость (выдавать более детализированные советы)
- Собрать больше данных об имуществе (в т.ч. из других источников)
- Дальнейшее обогащение данных о транспортных средствах: классификация, восстановление данных и дедупликация
- GUI
- Досчитать стоимость транспортных средств

Конец

Спасибо за внимание!