

Дипломная работа

По теме:

**Разработка модели для
предсказания цен на
недвижимость**

Выполнил: Досаев Савелий

Актуальность задачи

Рынок недвижимости растет и актуальность задачи только увеличивается

- ▶ Потребность в данной задаче:
- ▶ У налоговой
- ▶ У риелторов
- ▶ У банкиров

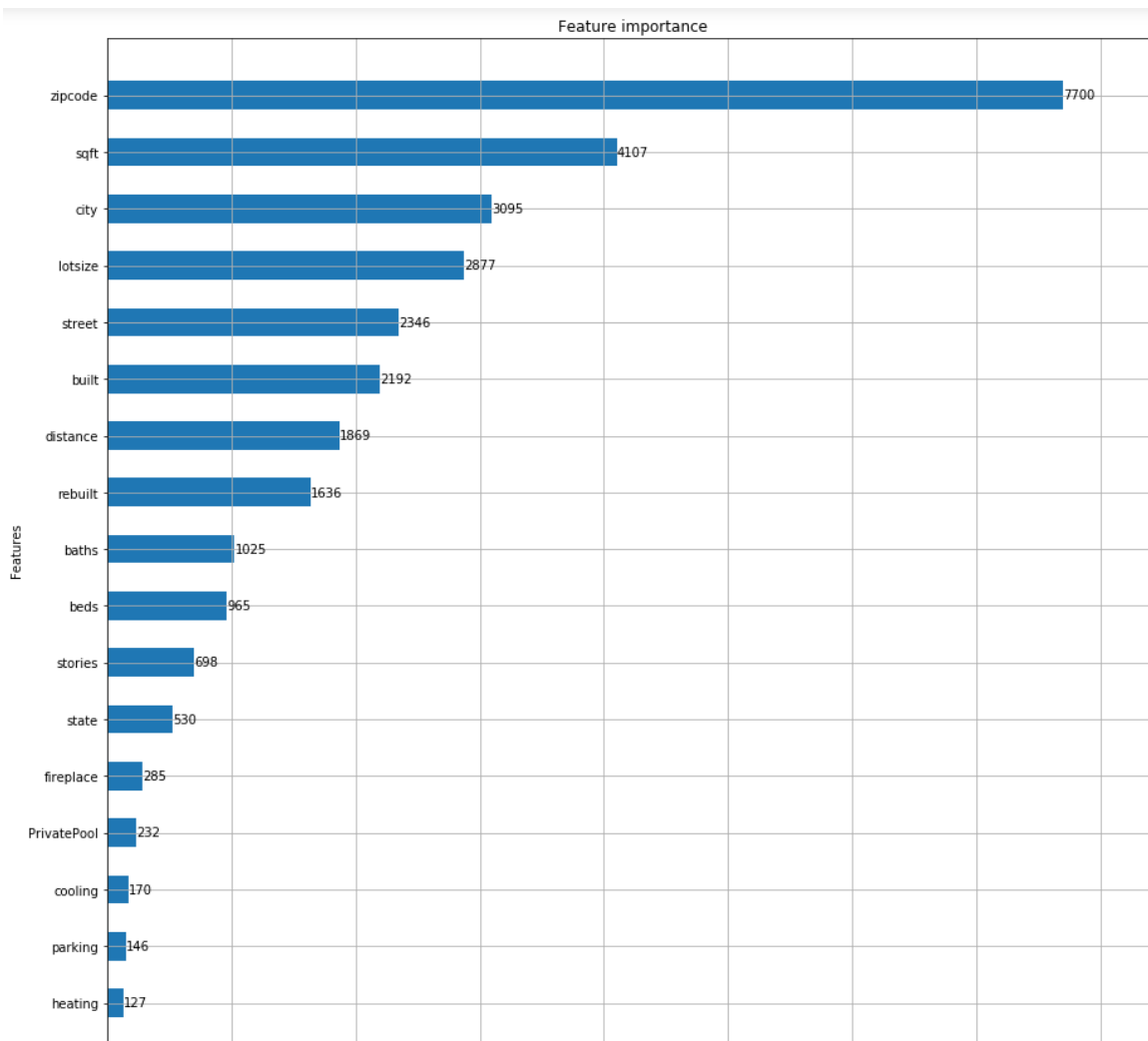
Цель работы

Разработать модель, которая точно и без ошибок предсказывает цены на недвижимость

Оценка полученного результата

- ▶ Маре(Mean absolute percentage error)
- ▶ Средний процент ошибки: принимает значения от 0 до ∞
- ▶ R2(Коэффициент детерминации)
- ▶ Доля необъяснённой дисперсии: принимает значения от $-\infty$ до 1
- ▶ Rmse(Root-mean-square deviation)
- ▶ Корень разности между предсказанным и настоящим значениями.
Принимает значения от 0 до ∞

Важность признаков в предсказании



Heating, cooling, parking, privatePool и fireplace дают очень незначительный результат

Адрес и площадь имеют наибольший вклад в предсказание

Предсказания с разным числом признаков

Точность с использованием большинства данных

	Model	r2_train	r2_test	mape_train	mape_test	d_train	d_test	rmse_train	rmse_test
0	Random Forest	62.63	60.66	48.41	48.28	35.23	36.11	28117705.45	29396602.73
1	XGB	96.90	82.31	16.34	24.49	11.07	20.78	8092559.67	19713310.75
2	LGBM	96.76	83.53	5.26	22.46	4.25	19.38	8277102.01	19021581.44

Точность с использованием адреса и площади

	Model	r2_train	r2_test	mape_train	mape_test	d_train	d_test	rmse_train	rmse_test
0	Random Forest	62.50	60.28	47.87	48.81	35.15	36.07	28656412.63	29879895.89
1	XGB	94.40	80.81	19.83	26.62	14.08	21.86	11069105.32	20765730.92
2	LGBM	96.06	82.00	10.11	24.92	7.23	20.37	9283777.05	20115727.93

Точность падает только на 2.5% по сравнению с предсказаниями по всем данным

Недвижимость для прогнозирования

- ▶ Лучшим образом для предсказания покажут себя:
 - ▶ Квартиры с площадью от 20 кв. м.
 - ▶ Любые дома до 10 этажей и с площадью до 650 кв. м.
 - ▶ Участки до 4 кв. км.
-
- ▶ Модель плохо работает на:
 - ▶ Больших участках или фермах
 - ▶ Дешевой недвижимости(<40000\$)
 - ▶ Аномально больших домах(кол-во ванн, этажей, спален>10, цена>3500000\$)

<https://github.com/Brilliance1512/dataproject>

