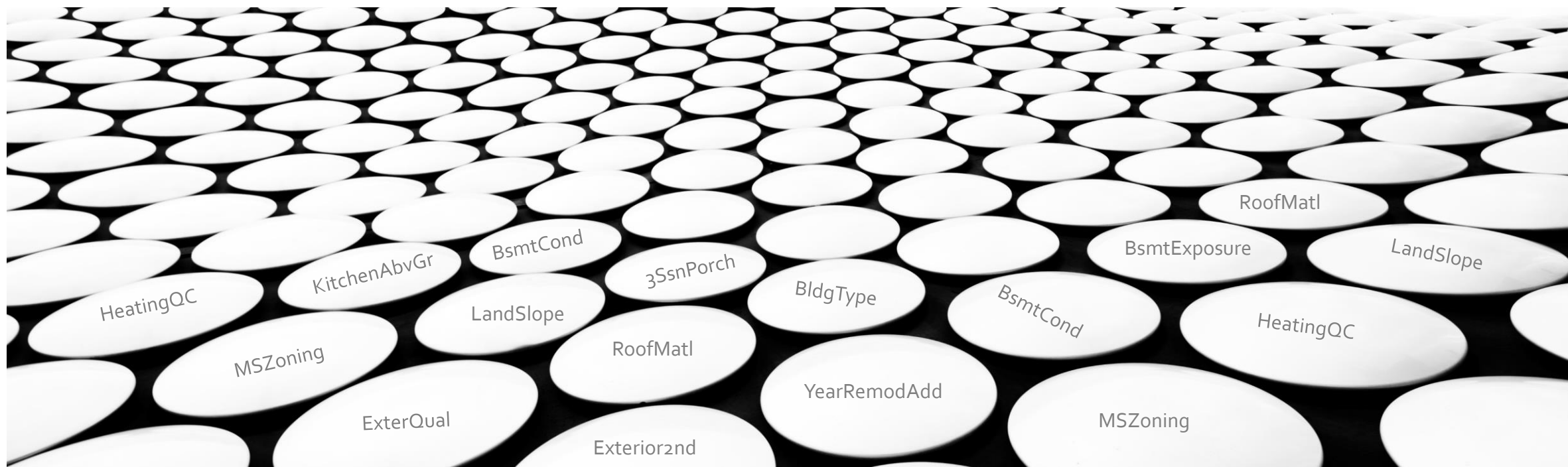


ПРОЕКТ ПРЕДСКАЗАНИЕ ЦЕН НА НЕДВИЖИМОСТЬ

ГАЛИНА ГОРЯИНОВА & АНДРЕЙ АБРАМОВ

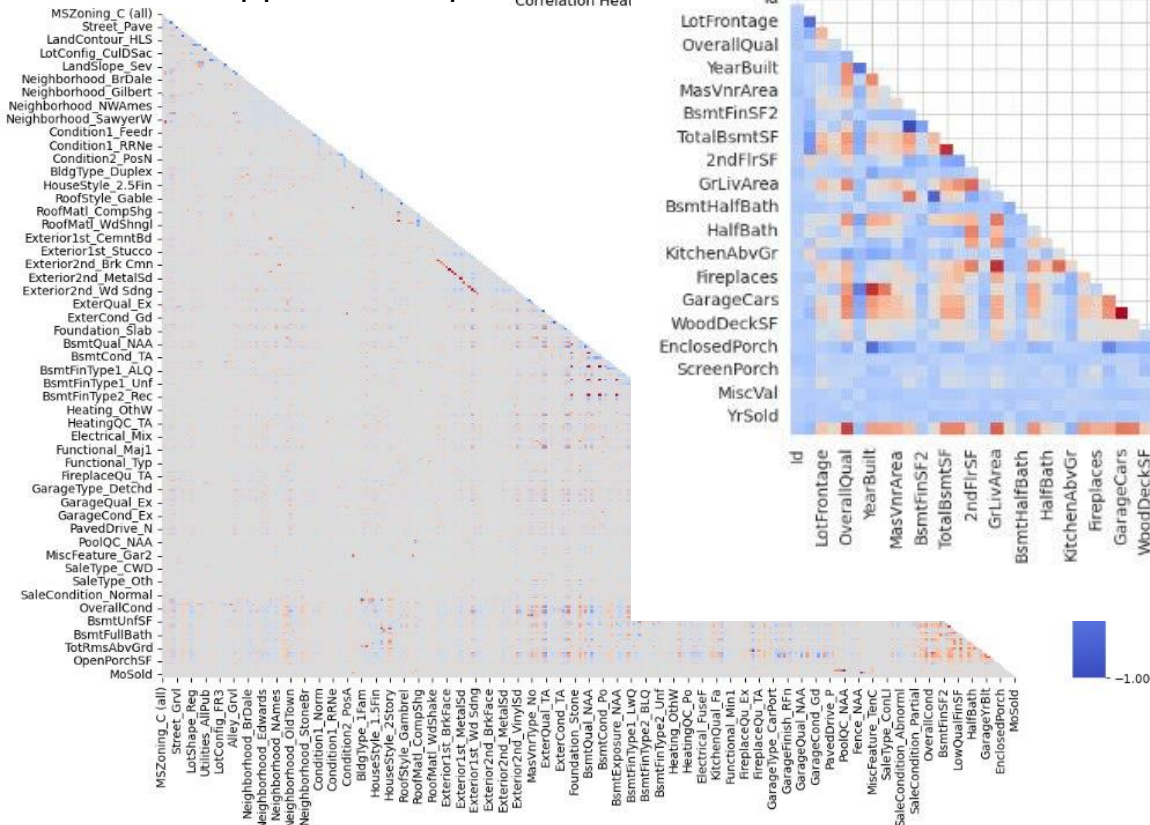


ОСНОВНЫЕ ШАГИ И ВЫРАБОТКА СТРАТЕГИИ

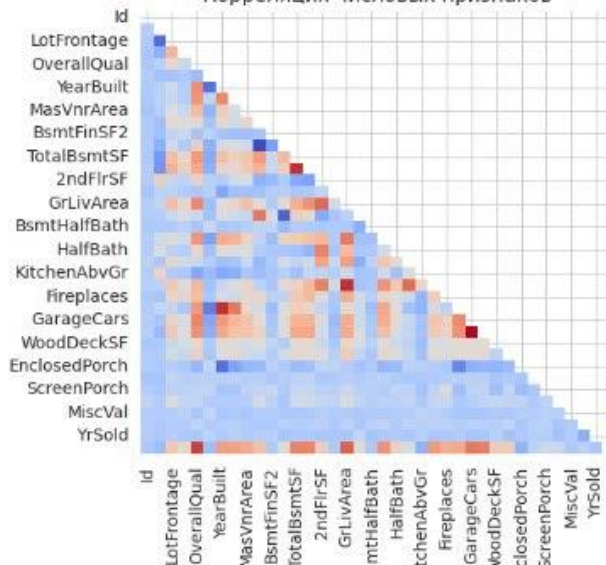


ВАЖНОСТЬ ПРИЗНАКОВ

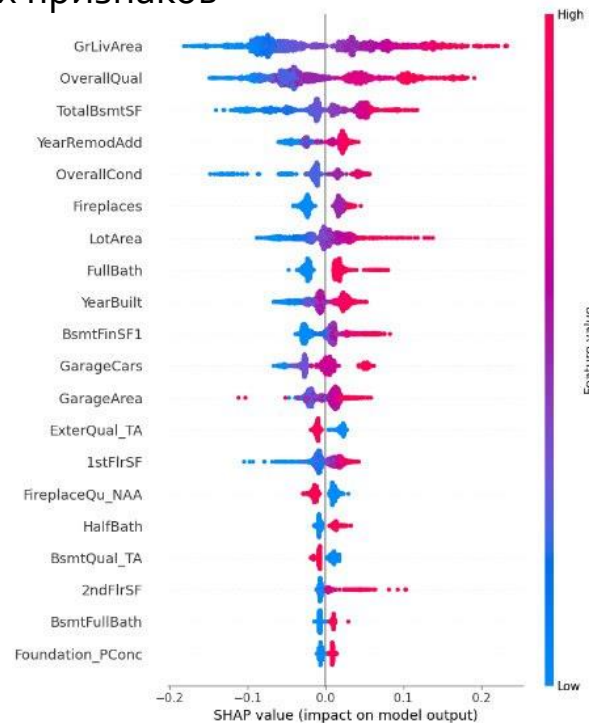
Корреляция признаков



Корреляция числовых признаков



Влияние наиболее важных числовых признаков



Влияние признаков на текущей позиции



ПРОМЕЖУТОЧНЫЕ ПАРАМЕТРЫ МОДЕЛЕЙ

Этап 2 Обучения:
Применение группы
моделей к результатам
работы моделей
предыдущего этапа

Этап 2 Тестирования:
Применение группы
моделей к результатам
работы моделей
предыдущего этапа

```
> middle_metrics
#           name           mse
1      LinearRegression 0.008769
2           Ridge      0.009481
3           Lasso      0.048565
4           Lars       0.045683
5      ElasticNet      0.033684
6      LassoLars       0.048564
7      BayesianRidge   0.012629
8      ARDRegression   0.009761
9      TweedieRegressor 0.014967
10     CatBoostRegressor 0.001688
11     XGBRegressor     0.000052
21     STEP-2: LinearRegression 0.000045
22     STEP-2: Ridge          0.000097
23     STEP-2: Lasso          0.063309
24     STEP-2: Lars           0.000052
25     STEP-2: ElasticNet     0.023572
26     STEP-2: LassoLars      0.063309
27     STEP-2: BayesianRidge  0.000045
28     STEP-2: ARDRegression  0.000045
29     STEP-2: TweedieRegressor 0.019317
30     STEP-2: CatBoostRegressor 0.000575
31     STEP-2: XGBRegressor   0.000021
40     STEP-1v: LinearRegression valid 0.016319
41     STEP-1v: Ridge valid 0.014796
42     STEP-1v: Lasso valid 0.049560
43     STEP-1v: Lars valid 0.046427
44     STEP-1v: ElasticNet valid 0.032423
45     STEP-1v: LassoLars valid 0.049559
46     STEP-1v: BayesianRidge valid 0.015413
47     STEP-1v: ARDRegression valid 0.013026
48     STEP-1v: TweedieRegressor valid 0.016108
49     STEP-1v: CatBoostRegressor valid 0.030039
50     STEP-1v: XGBRegressor valid 0.019299
60     STEP-2v: LinearRegression valid 0.019762
61     STEP-2v: Ridge valid 0.018252
62     STEP-2v: Lasso valid 0.077456
63     STEP-2v: Lars valid 0.019297
64     STEP-2v: ElasticNet valid 0.034200
65     STEP-2v: LassoLars valid 0.077456
66     STEP-2v: BayesianRidge valid 0.019762
67     STEP-2v: ARDRegression valid 0.019712
68     STEP-2v: TweedieRegressor valid 0.026051
69     STEP-2v: CatBoostRegressor valid 0.017672
70     STEP-2v: XGBRegressor valid 0.017672
> middle_answer2 = pd.DataFrame(middle_answer)
```

Этап 1 Обучения:
Применение
группы моделей к
исходным данным

Этап 1 Тестирования:
Применение группы
моделей первого
этапа обучения

Решение какую
модель лучше давать в
качестве финальной
при подаче



ВЫПОЛНЕННЫЕ ПОДАЧИ

#	Ансамбль	Основная модель	Kaggle Score	Kaggle position
1	Stacking	XGBoost Regr. (Step-2)	0.14716	2337
2	Stacking	Liner Regr. (Step-2)	0.20205	-
3	-	CatBoost Regr. (Step-1)	0.14415	2031-2
4	Stacking	CatBoost Regr. (Step-2)	0.17397	-
5	Stacking	ARDRegression (Step-2)	0.20205	-
6	Stacking	Lasso (Step-2) (с изменением одного из параметров очистки)	0.30209	-
7	Stacking	BayesianRidge (Step-2)	0.20205	-

2031

AAndrey

0.14415

8

36m

2032

Galina Goriainova

0.14415

2

3h

Your Best Entry!

Your most recent submission scored 0.14415, which is an improvement of your previous score of 0.14716. Great job!

Tweet this

С помощью данного подхода удалось достичь 2031-2 места на Kaggle.com