

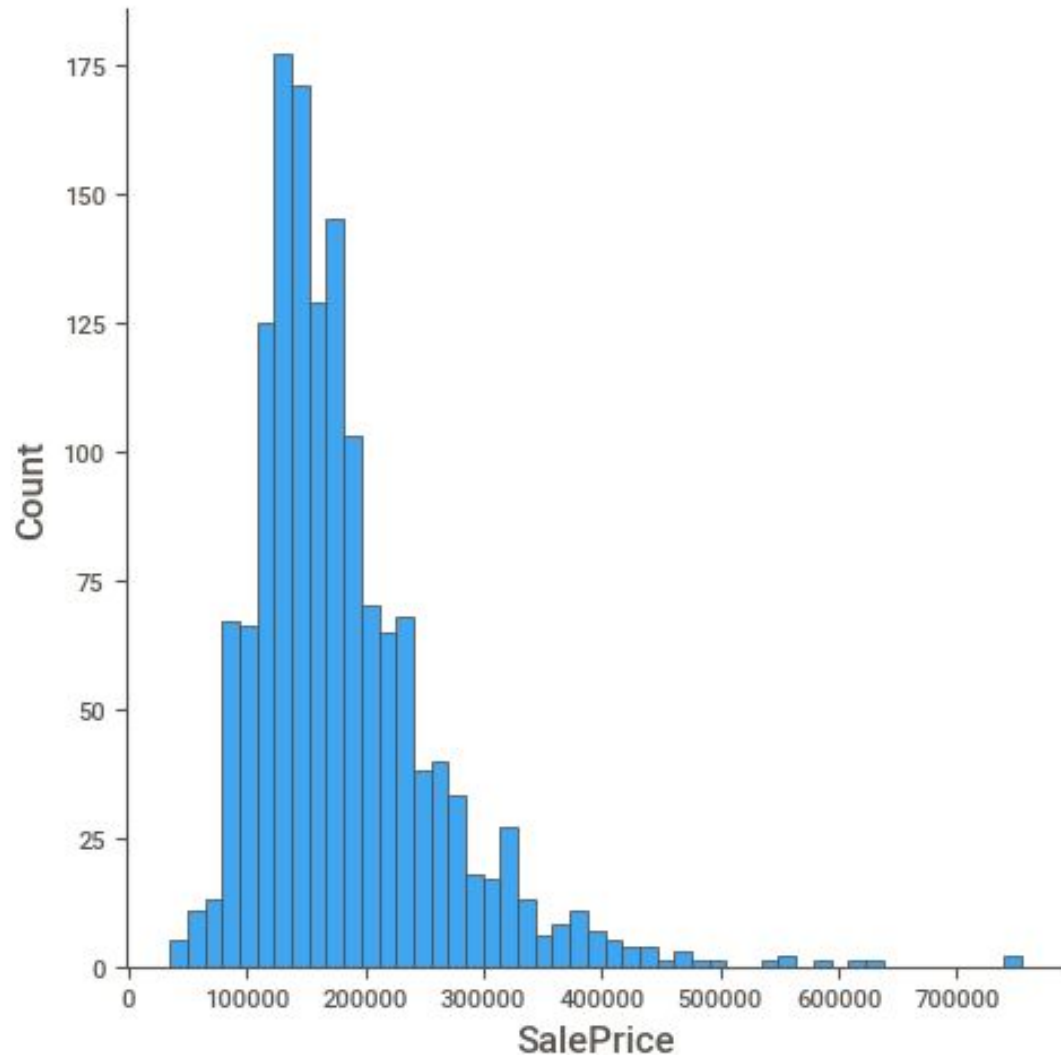
# House Prices - Advanced Regression Techniques

Команда Optuna



# Описание данных

Распределение  
целевой  
переменной

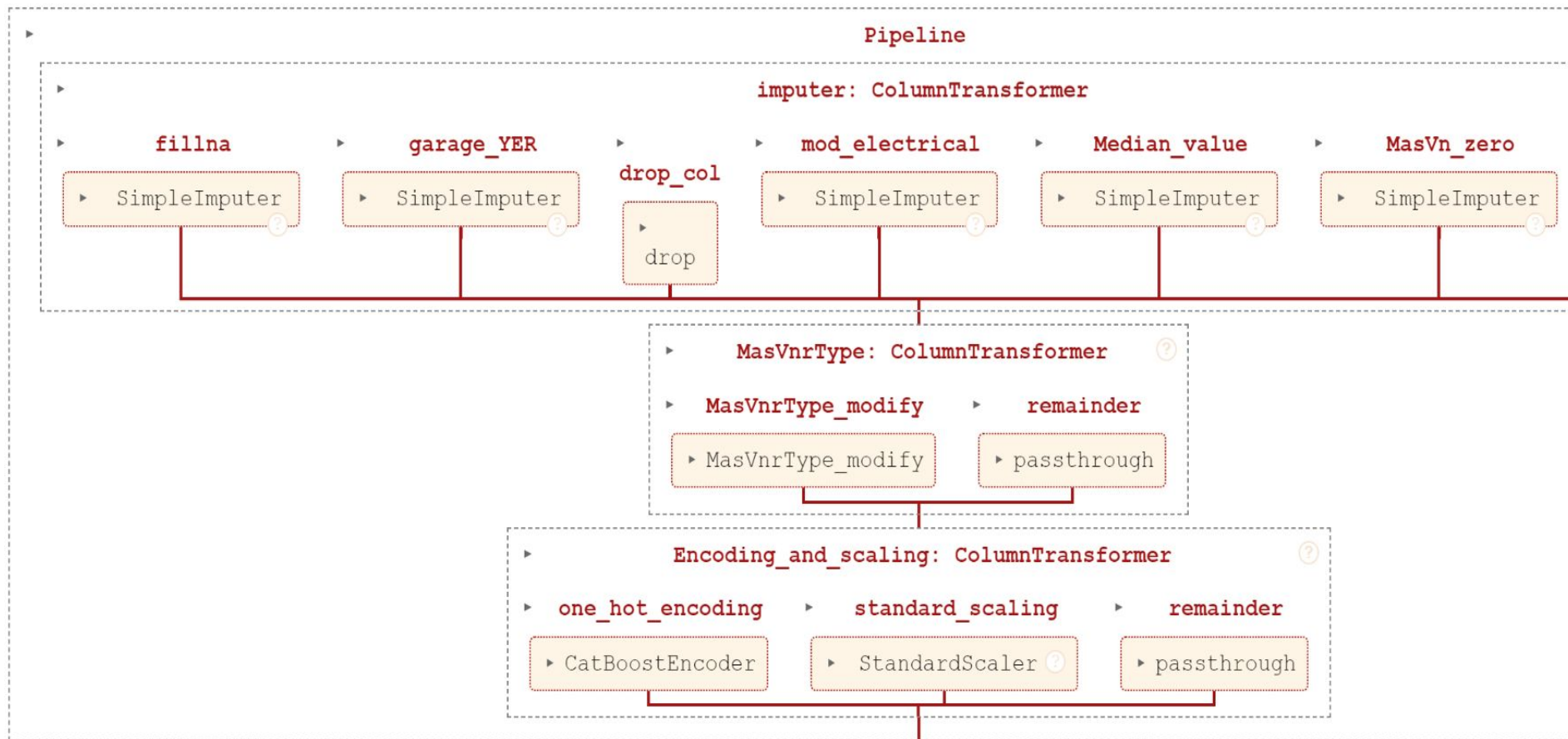


# Описание данных

Пропущенные  
значения

	NaN_count	data_type
LotFrontage	259	float64
Alley	1369	object
MasVnrType	872	object
MasVnrArea	8	float64
BsmtQual	37	object
BsmtCond	37	object
BsmtExposure	38	object
BsmtFinType1	37	object
BsmtFinType2	38	object
Electrical	1	object
FireplaceQu	690	object
GarageType	81	object
GarageYrBlt	81	float64
GarageFinish	81	object
GarageQual	81	object
GarageCond	81	object
PoolQC	1453	object
Fence	1179	object
MiscFeature	1406	object

# Предобработка



# Отбор признаков

## 1. PCA

`pca_explained_ratio = 0.93`

## 2. Feature importance

количество оставленных признаков = 20

## 3. Lasso

`alpha = 0.8`

# Модели

- **CatBoostRegressor()**
- **RandomForestRegressor()**
- **XGBoostRegressor()**
- **StackingRegressor**(estimators=[RandomForestRegressor, CatBoostRegressor])
- **VotingRegressor**(estimators=[RandomForestRegressor, CatBoostRegressor])

# Метрики

```
mean_squared_error(np.exp(y_valid), np.exp(y_pred), squared = False)
```

# Результат

## Место на Kaggle

520	Kamaev Alex		0.12477
-----	-------------	---	---------