# Анализ игр для приставок PS

Автор работы Бондарь Степан

#### Исходный данные

Данные взяты из датасета PlayStation Games Info, автор Evgeny Baulin

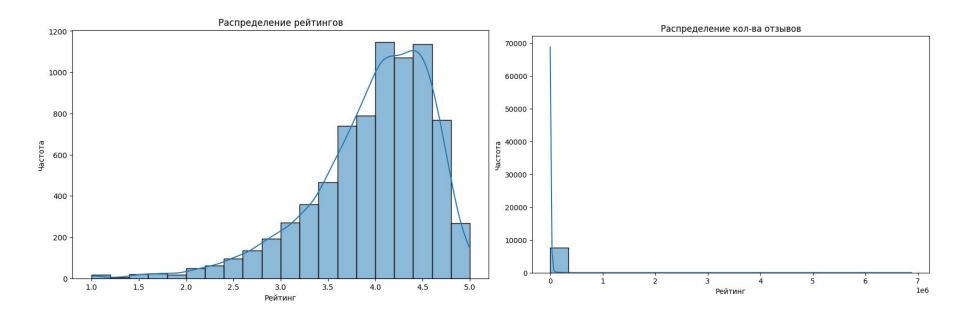
- title game title. Data in format string
- platforms platforms on which the game is available. Data in format list of strings
- rating game rating. Data in the format real number
- votes number of votes. Data in the format integer
- developer game developer. Data in the format string
- release date game release date. Data in the format date
- price price of the game in \$. Data in the format real number
- genres game genres. Data in format list of strings
- url link to the game. Data in format string

Размер датасета 8981

#### Подготовка данных

- удаляем {'url', 'title'}
- дату преобразуем в год и месяц (в синусной форме)
- с помощью MultiLabelBinarizer преобразуем { 'genre', 'platform'} в бинарные признаки
- разделяем на тестовую и тренировочную выборку
- применяем частотное кодирование к 'developper'

#### Анализ исходных данных



#### Корреляционная матрица

price - <mark>1.00 </mark> 0.02 0.02 0.02 0.02 0.00 0.06 0.04 0.05 0.03 0.18 0.00 0.09 0.13 0.01 0.03 0.02 0.04 0.02 0.04 0.05 0.02 0.14 0.08 0.19 0.01 0.02 0.02 0.03 0.03 0.00 0.00 0.00 0.00	1.0
year0.02 1.00 0.17 0.00 -0.25 0.60 -0.15 0.07 0.01 -0.02 -0.05 0.08 0.10 0.00 0.05 0.03 -0.07 0.02 0.02 -0.00 0.02 0.05 0.06 -0.04 -0.06 0.06 0.08 -0.01 -0.03 0.01 0	
month sin0.02 0.17 1.00 -0.06 -0.06 0.03 0.01 -0.01 -0.00 0.01 -0.00 0.01 0.00 0.00	
month cos = 0.02 0.00 -0.06 1.00 -0.03 0.02 -0.03 0.00 0.01 -0.01 0.01 0.02 0.00 0.00 0.00 0.01 0.02 0.00 0.00	
PS40.10 -0.25 -0.06 -0.03 1.00 -0.34 0.05 0.02 -0.02 0.03 0.01 0.01 0.01 0.01 0.01 0.00 0.02 -0.07 -0.03 -0.00 -0.00 -0.01 0.02 0.00 0.00 0.01 0.00 -0.00 -0.00 0.02 0	- 0.8
PS5 - 0.06 0.60 0.03 0.02 -0.34 1.00 0.08 0.08 0.04 -0.02 -0.02 0.07 0.06 -0.00 0.02 0.00 -0.05 0.02 0.03 0.01 -0.01 0.04 0.07 -0.01 0.06 0.03 0.07 -0.02 -0.04 -0.01 0.04	
Action - 0.04 -0.15 0.01 -0.03 0.05 -0.08 1.00 -0.05 0.01 0.04 -0.02 -0.08 0.13 -0.08 -0.06 -0.07 0.04 -0.02 0.01 -0.02 -0.04 -0.12 -0.06 -0.04 0.10 -0.13 -0.07 -0.10 -0.09 -0.11 0	
Adult - 0.05 0.07 -0.01 0.00 0.02 0.08 -0.05 1.00 0.06 -0.01 -0.00 0.02 -0.01 -0.02 -0.01 -0.02 -0.01 -0.02 -0.01 -0.00 -0.00 -0.01 -0.01 0.05 -0.01 -0.02 -0.03 -0.02 -0.01 -0.02 -0.01 0.	
Adventure - 0.03 0.01 -0.00 0.01 -0.02 0.04 0.01 0.06 1.00 -0.16 -0.01 -0.05 -0.03 -0.10 -0.01 -0.07 -0.02 0.06 -0.03 -0.07 0.01 -0.05 -0.02 -0.14 -0.08 -0.05 -0.11 -0.10 0.01 -0.05 -0.03	
Arcade -0.18-0.02 0.01-0.01 0.03-0.02 0.04-0.01-0.16 1.00 0.03-0.03 0.07 0.03-0.03 0.04 0.06-0.00-0.07-0.03 0.00 0.07-0.03 0.00 0.07-0.03 0.01 0.10 0.08-0.05 0.03 0.07-0.06 0	- 0.6
Board Game0.00 0.05 0.01 -0.01 0.01 -0.02 -0.02 -0.02 -0.00 -0.01 -0.01 0.01 -0.02 -0.02 -0.00 -0.01 -0.01 -0.01 -0.00 -0.00 -0.00 -0.00 -0.00 -0.00 -0.00 -0.01 -0.01 -0.01 -0.01 -0.01 -0.01 -0.01 -0.01 -0.01	2/00/2012/00/2012/00/2012
Brain Training -0.09 0.08 0.02 0.02 0.01 0.07 -0.08 0.02 0.05 -0.03 -0.00 1.00 0.04 -0.02 0.08 0.02 -0.02 0.02 -0.02 -0.00 -0.00 0.12 0.31 -0.04 -0.03 -0.02 -0.02 -0.01 -0.01 -0.01	
Casual -0.13 0.10 0.00 -0.00 0.01 0.06 -0.13 -0.01 -0.03 0.07 -0.01 0.04 0.00 0.01 -0.06 0.01 -0.06 -0.01 0.08 0.09 -0.02 -0.08 -0.09 0.08 -0.00 -0.01 0.08 0.09 -0.02 -0.08 -0.09 0.08 -0.00 -0.01 0.08 0.09 -0.02 -0.08 0.09 0.08 -0.00 0.01 0.08 0.09 0.08 -0.00 0.01 0.08 0.09 0.09	
Driving/Racing - 0.01 0.00 0.00 0.00 0.01 -0.00 0.08 -0.02 0.10 0.03 -0.00 -0.02 0.03 1.00 0.00 -0.02 0.03 1.00 0.01 -0.04 -0.01 0.04 -0.02 0.01 0.07 0.02 0.07 0.02 0.07 0.03 0.06 0.05 0.04 0.05 0.04 0.05 0.04 0.05 0.04 0.05 0.04 0.05 0.04 0.05 0.04 0.05 0.04 0.05 0.04 0.05 0.04 0.05 0.04 0.05 0.04 0.05 0.04 0.05 0.04 0.05 0.05	
Educational - 0.03 0.05 0.00 0.00 -0.01 0.02 -0.06 -0.01 -0.01 -0.03 -0.00 0.08 -0.00 -0.00 1.00 0.05 -0.02 -0.00 -0.02 -0.01 -0.00 0.01 0.01 -0.03 -0.03 0.04 0.01 -0.00 0.00 -0.01 -0.00 0.00 -0.01 -0.00 0.00	- 0.4
Family - 0.02 0.03 -0.00 0.01 0.00 0.00 -0.07 -0.02 -0.01 -0.04 0.02 0.02 0.21 0.01 0.05 1.00 0.03 0.01 -0.04 0.05 0.20 0.07 -0.02 -0.05 -0.06 0.03 -0.03 0.03 -0.03 0.03 -0.0	
Fighting - 0.04 -0.07 -0.01 -0.00 0.02 -0.05 0.04 -0.01 -0.07 0.06 -0.00 -0.02 -0.04 -0.04 -0.02 0.03 1.00 0.01 -0.04 -0.02 0.00 -0.05 -0.01 -0.05 -0.04 -0.05 -0.02 -0.01 -0.05 -0.04 -0.	
Fitness - 0.02 0.02 0.02 0.02 -0.00 -0.07 0.02 -0.02 -0.00 -0.02 0.00 -0.00 0.02 -0.01 -0.01 0.01 0.01 0.01 0.01 0.01 0.	
Horror - 0.04 0.02 0.00 0.01 0.03 0.03 0.01 0.00 0.06 0.07 0.000 0.06 0.04 0.02 0.04 0.04 0.01 1.00 0.00 0.02 0.03 0.02 0.03 0.02 0.03 0.02 0.03 0.02 0.05 0.05 0.03 0.04 0.04 0.03 0	
Music/Rhythm - 0.05 -0.00 0.00 0.04 -0.00 0.01 -0.02 -0.01 -0.03 -0.03 -0.03 -0.00 -0.01 -0.02 -0.01 0.05 -0.02 0.09 -0.02 1.00 0.04 -0.03 -0.01 -0.03	- 0.2
Party - 0.02 0.02 -0.01 0.01 -0.00 -0.01 -0.04 -0.01 -0.07 0.00 -0	0.06 -0.03 -0.01
Puzzle0.14 0.05 0.02 0.00 0.01 0.04 0.12 0.05 0.01 0.07 0.01 0.12 0.09 0.07 0.01 0.07 0.05 0.01 0.02 0.03 0.04 1.00 0.03 0.09 0.10 0.07 0.04 0.06 0.03 0.00 0.00	
Quiz -0.08 0.06 0.03 0.03 0.07 -0.06 -0.01 -0.05 -0.03 -0.00 0.31 -0.02 -0.02 -0.02 -0.01 -0.00 -0.03 1.00 0.03 1.00 0.03 -0.03 -0.02 -0.01 -0.00 -0.03 1.00 0.03 -0.03	
Role Playing Games - 0.19 -0.04 0.02 -0.00 0.00 -0.01 -0.04 -0.02 -0.02 -0.11 -0.01 -0.04 -0.08 -0.07 -0.03 -0.05 -0.05 -0.01 -0.06 -0.03 -0.04 -0.09 -0.03 -0.04 -0.09 -0.05 -0.04 -0.05 -0.05 -0.04 -0.05 -0.04 -0.05 -0.04 -0.05 -0.04 -0.05 -0.04 -0.05 -0.04 -0.05 -0.04 -0.05 -0.04 -0.05 -0.04 -0.05 -0.04 -0.05 -0.04 -0.05 -0.04 -0.05 -0.04 -0.05 -0.04 -0.05 -0.04 -0.05 -0.04 -0.05 -0.04 -0.05 -0.04 -0.05 -0.04 -0.05 -0.05 -0.04 -0.05 -0.05 -0.05 -0.05 -0.05 -0.05 -0.05 -0.05 -0.05 -0.05 -0.05	0.04 0.09 0.00
Shooter0.01 -0.06 0.03 0.03 0.00 0.06 0.10 0.03 0.14 0.10 0.01 -0.04 0.09 0.06 0.03 0.06 -0.04 0.01 -0.04 -0.01 -0.04 -0.01 -0.04 -0.10 0.03 -0.09 1.00 0.08 -0.04 0.06 -0.07 -0.06 0.07 -0.07 -0.06 0.07 -0	0.0
Simulation - 0.02 0.06 -0.03 0.02 0.01 0.03 -0.13 -0.02 -0.08 -0.08 -0.01 -0.03 0.08 0.07 0.04 0.03 -0.05 -0.01 -0.05 -0.03 -0.02 -0.07 -0.02 -0.05 -0.08 1.00 0.27 0.08 0.13 -0.01 -0.05 -0.08 0.03 -0.05 -0.08 -0.09 0.08 0.03 -0.05 -0.08 0.09 0.08 0.03 -0.05 -0.08 0.09 0.08 0.09 0.08 0.09 0.09 0.09	0.06 -0.08 -0.01
Simulator0.02 0.08 -0.02 0.01 0.00 0.07 -0.07 -0.01 -0.05 -0.05 -0.05 -0.00 -0.02 -0.00 0.03 0.01 -0.03 -0.02 -0.01 -0.03 -0.02 -0.02 -0.04 -0.01 -0.04 -0.04 0.27 1.00 -0.01 -0.02 -0.02 -0.02 -0.01	0.02 -0.05 -0.01
Sport - 0.03 -0.01 -0.03 -0.01 -0.00 -0.02 -0.10 -0.02 -0.11 0.03 -0.00 -0.02 -0.01 0.06 -0.00 0.02 -0.01 0.08 -0.04 -0.00 0.02 -0.05 -0.05 -0.05 -0.05 -0.06 0.08 -0.01	0.00 -0.10 0.01
Strategy - 0.03 -0.03 -0.03 -0.00 -0.01 -0.00 -0.04 -0.09 -0.02 -0.10 -0.07 -0.01 -0.02 -0.05 -0.00 -0.03 -0.04 -0.01 -0.04 -0.03 -0.02 -0.03 -0.02 -0.03 -0.02 -0.03 -0.02 -0.03 -0.02 -0.03 -0.02 -0.03 -0.02 -0.04 -0.03 -0.02 -0.04 -0.03 -0.02 -0.04 -0.03 -0.02 -0.04 -0.03 -0.02 -0.04 -0.03 -0.02 -0.03 -0.03 -0.02 -0.03 -0	0.08 -0.02 -0.01
Unique - 0.00 0.01 0.01 0.03 0.02 -0.01 0.11 0.01 0.01 0.01 -0.06 -0.01 -0.01 -0.01 -0.01 -0.03 -0.04 -0.01 -0.03 -0.04 0.01 -0.03 -0.01 -0.00 -0.00 -0.02 -0.04 -0.06 -0.01 -0.02 -0.04 -0.06	0.2
developer_freq0.16 0.02 0.01 -0.02 0.08 0.02 0.09 0.02 -0.08 0.19 0.01 -0.00 0.00 -0.04 -0.02 -0.04 0.08 -0.02 -0.06 -0.03 -0.06 0.00 0.02 -0.04 0.12 -0.06 -0.02 0.00 -0.04	.00 0.15 -0.01
rating - 0.16 0.10-0.00-0.01 0.06 -0.01 0.09 0.01 0.05 -0.01 -0.00 -0.01 0.05 -0.01 -0.00 -0.01 -0.07 0.08 -0.01 -0.06 0.06 0.01 -0.01 0.03 -0.03 0.00 -0.01 0.09 0.06 -0.08 -0.05 -0.10 -0.02 -0.02	0.15 1.00 0.03
	0.01 0.03 1.00
votes - 0.01 -0.06 -0.01 -0.00 -0.03 -0.00 -0.03 -0.00 -0.01 -0.00 -0.02 -0.02 -0.00 -0.01 -0.02 -0.02 -0.00 -0.01 -0.01 -0.01 -0.01 -0.00 -0.00 -0.00 -0.01 -0.01 -0.02 -0.01 -0.00 -0.01	
· , · · · · · · · <u>· · · · · · · · · · ·</u>	ba ba ud
price - year - hth_sin - tth_cos - PS4 - PS5 - Action - Adult - Arcade - I Game - I raining - Casual - Barning - Barty - Barty - Buzzle - Quiz - Quiz - Quiz - Games - shooter - ulation - nulator - Sport - Sport - Short - Inhythm - Duzzle - Quiz - Quiz - Ouiz -	er_freq rating votes
· , · · · · · · · <u>· · · · · · · · · · ·</u>	Jeveloper_freq rating votes

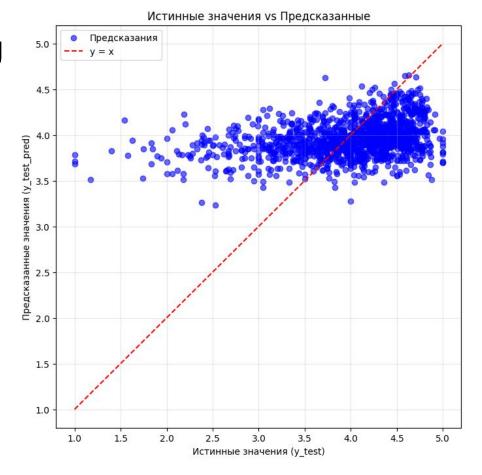
# LinearRegression() — rating

Train y (rating) MAPE: 0.1372

Train y (rating) RMSE: 0.6057

Test y (rating) MAPE: 0.1394

Test y (rating) RMSE: 0.6128



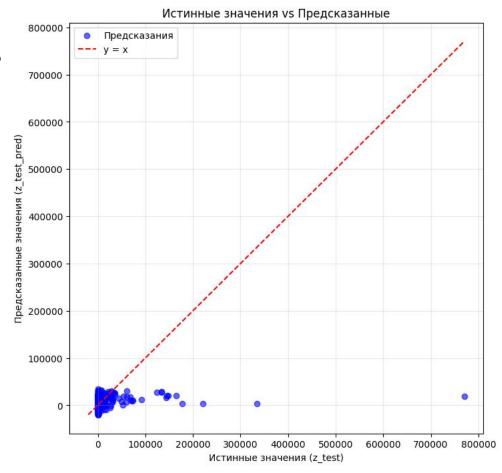
# LinearRegression() — votes

Train z (votes) MAPE: 276.1577

Train z (votes) RMSE: 85556.6565

Test z (votes) MAPE: 266.8050

Test z (votes) RMSE: 27398.9408



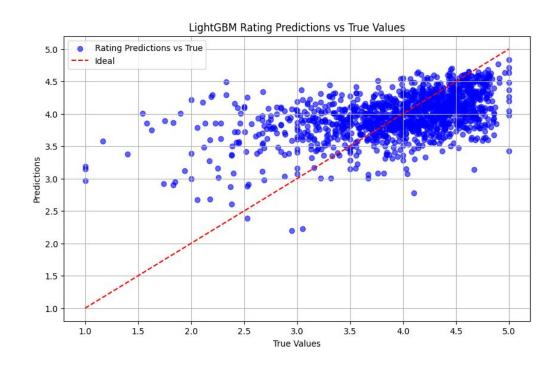
# LGBMRegressor() — rating

Train y (rating) MAPE: 0.0964

Train y (rating) RMSE: 0.4388

Train y (rating) MAPE: 0.1251

Train y (rating) RMSE: 0.5619



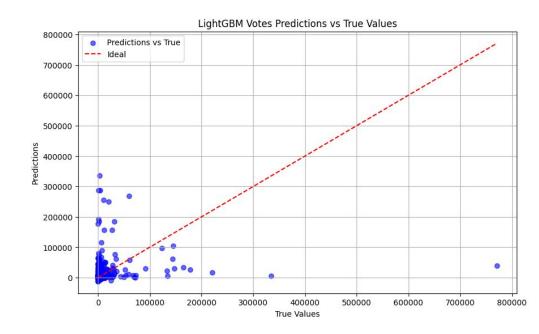
#### LGBMRegressor() — votes

Train z (votes) MAPE: 56.3689

Train z (votes) RMSE: 78343.3302

Test z (votes) MAPE: 46.4196

Test z (votes) RMSE: 33467.3918



#### Подбор гиперпараметров

Лучший результат:

Train y (rating) MAPE: 0.1060

Train y (rating) RMSE: 0.4815

Test y (rating) MAPE: 0.1249

Test y (rating) RMSE: 0.5575

