



Sales prediction analysis

Jane



Objective

Predict total sales for every product and Russian shop in the next month given a 100MB historical sales dataset containing daily sales from Jan 2013 to Oct 2015.

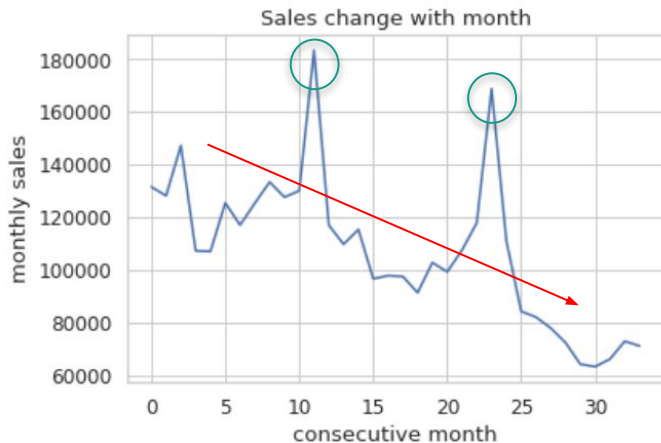
Give suggestions and recommendations based on business implications.

Approach

- Combine daily sales data to monthly data
- **Pre-processing** dataset using **seaborn**
- **Label Encoding** shop id, month and item id
- **Standard Scaler** on item price and sales
- Build **Linear Regression**, **Random Forest** and **XGBT** with **R2 score** as evaluation metric and choose best model
- Predict and export next month sale for each shop and item listed in test dataset

Deep dive into historical sales data

- Sales bloom in winter
- Decrease sales year by year



- High popularity in games and electronic devices
- Top 10 items shown below

Diablo III: Reaper of Souls (дополнение) [PC, ...
Grand Theft Auto V [PC, русские субтитры]
Grand Theft Auto V [PS3, русские субтитры]
Grand Theft Auto V [PS4, русские субтитры]
Grand Theft Auto V [Xbox 360, русские субтитры]
Sims 4 Limited Edition [PC, русская версия]
StarCraft II: Heart of the Swarm (дополнение) ...
Ведьмак 3: Дикая охота [PC, русская версия]
Ведьмак 3: Дикая охота [PS4, русская версия]
Фирменный пакет майка 1С Интерес белый (34*42)...



Results

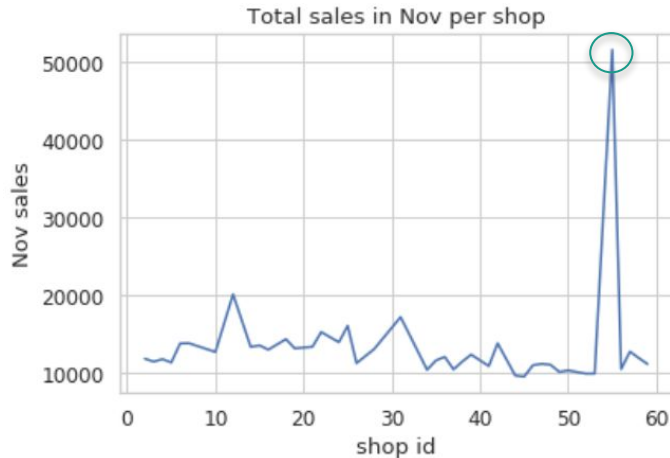
- XGBT is the best model which has a highest R2 score of 0.62, while Linear Regression works very bad.
- From feature importance, **price** is the most crucial factor (38%) on sales, then is **shops** (27%) and **month/time** (18%).

Future work

- Better tune hyperparameters in XGBT with cross-validation.
- Feature engineering. Add some features about shop, time lag.
- Try time-series model like ARIMA.

Deep dive into prediction sales data

- High demands next month
- **Top 10** predicted popular shops



- Games remain most attractive
- **Top 10** predicted popular items

ESET NOD32 Антивирус продление лицензии 3 ПК...
GUF Город дорог
GUF Дома
Pokemon XY Двойной блистер
Ratchet & Clank Trilogy [PS Vita, английская в...
БАРБОСКИНЫ ЛУЧШИЕ СЕРИИ 1 (регион)
БУКАШКИ: БУКАШКИ КАРИБСКОГО МОРЯ
ДОМ С ПАРАНОРМАЛЬНЫМИ ЯВЛЕНИЯМИ (BD)
НЕВЕРОЯТНЫЕ ПРИКЛЮЧЕНИЯ ИТАЛЬЯНЦЕВ В РОССИИ (р...
Фигурка Marvel Universe Deadpool POP 3,75"



Recommendations

- Expand quantity of production for popular items in popular shops in winter.
- Set price reduction or promotion sometime to boost sales, such as provide customers with coupons.
- Figure out why total sales decrease in the past two years. Send out public survey with tiny prize or analyze tweets and reviews.
- Add more attractive games and electronic devices.
- Why customers return some items? Analyze those products, see if we can reduce returns.
- Why customers may not go to some shops? May be due to bad service or lacking items. Enhance service experience in all shops, especially in downtown.

Thank you!