SALES FORECAST

FOR A TOY RETAILER

TING SIT

APRIL 2020

OBJECTIVE

- Develop a one month forecast model by product and by shop, for a Toy Retailer (Clients)
- Benefits:
  - Knowing which will be sold at which shop, clients can maximize sales potential by minimizing out of stock in each shop
  - Minimize overage for products that may not have enough future demand
  - Prioritize marketing/investment effort for top sellers at fast selling shops

Model Evaluation requirements:

- Root Mean Squared Error (RMSE) in units, at monthly level

MODEL CONSIDERATION

- Regression problem: item_cnt_month (target)
- 3 models: Xgboost, Random Forest (Regressor), Rigde Linear model
- Features:
  - Item_id, item_price
  - category_id
  - Shop_id
  - Month of the year
  - Lags (1, 3, 6, 12), in units and percentage changes

RESULTS

Model	RMSE
Naïve*	3.77
Random Forest	1.40
Xgboost	2.16
Rigde	3.42

**Naïve model uses actual sales data from 12 month prior based on seasonality pattern*

- Random Forest has lowest RMSE
  - Recommend Random Forest as the winning model

CONCLUSION

1. Random Forest has the smallest RMSE and is a more scalable model as it is less likely to subject to overfitting (unlike Xgboost)
2. Client can apply the model in an ongoing basis for next month demand by product and by shop, in order to influence inventory and marketing strategy
3. RMSE of 1.4 could still add up to huge financial errors, depending on what products have higher shares of the errors. The cost of forecasting to reduce RMSE could increase substantially, and the risk of overfitting. Clients should set a level of margin of error that they would feel comfortable on making business decisions

NEXT STEPS

- Obtain Client's internal marketing effort details (e.g. markdown/promotional activities) will help creating new useful features on item/shop level predictions
- Align expectations on the RMSE level that is "good enough" for clients
- Extend model to two and three months out to observe if model demonstrate bias pattern that will help fine tune model performance