BEST BUY PROJECT WEEK:

FORECASTING SLOW-MOVING SKUS

Team Fantastalytics

Yu-Ching Chen, Cassidy Gasteiger,

Prathyusha Pateel, Vaani Radhakrishnan

PROBLEM STATEMENT

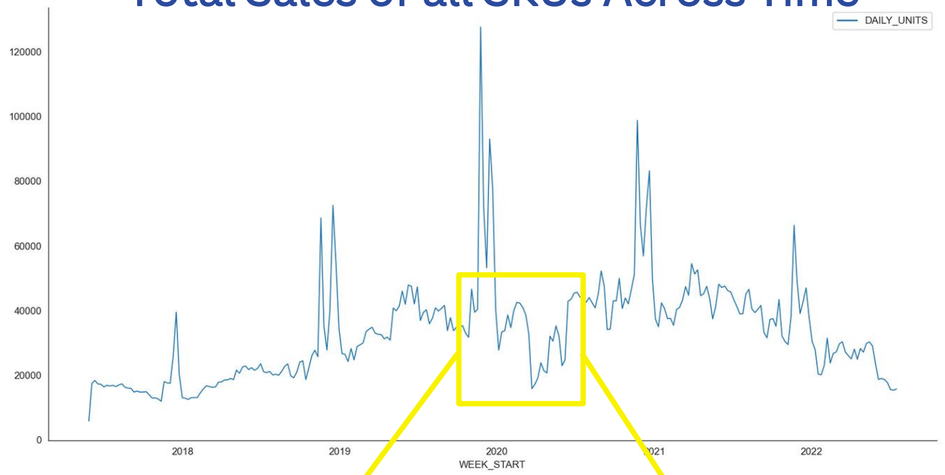
Forecast 575 unique, slow-moving, intermittent SKUs **efficiently** and **accurately**

Challenges:

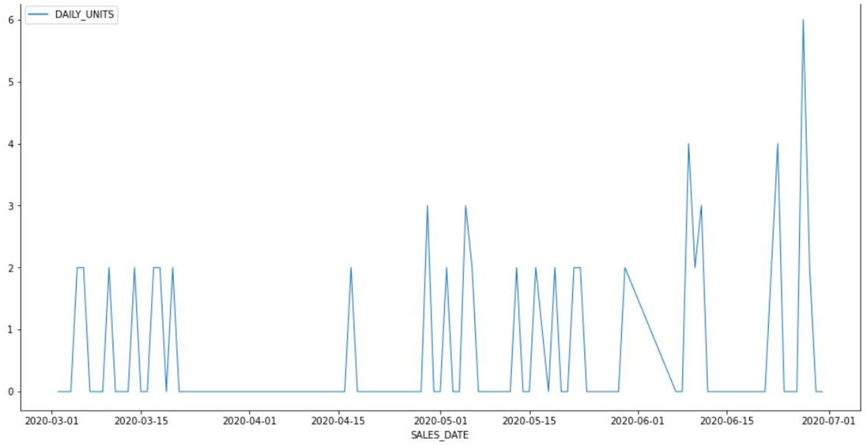
- Varied SKU selling windows
- Missing values in data features
- Model runtime

INTERMITTENT SALES DATE

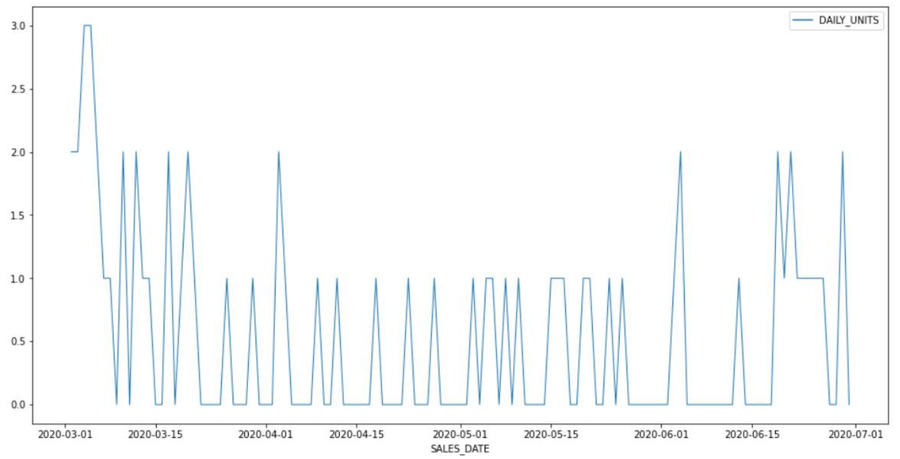
Total Sales of all SKUs Across Time



SO PAC ACCY BBHD Subclass Data



SO CAMERA ACCY Subclass Data



APPROACH

1. DATA EXPLORATION
2. MODEL BUILDING
3. MODEL SELECTION

- Researched intermittent SKU forecasting methods
- Identified and included additional datasets
- Identified key features

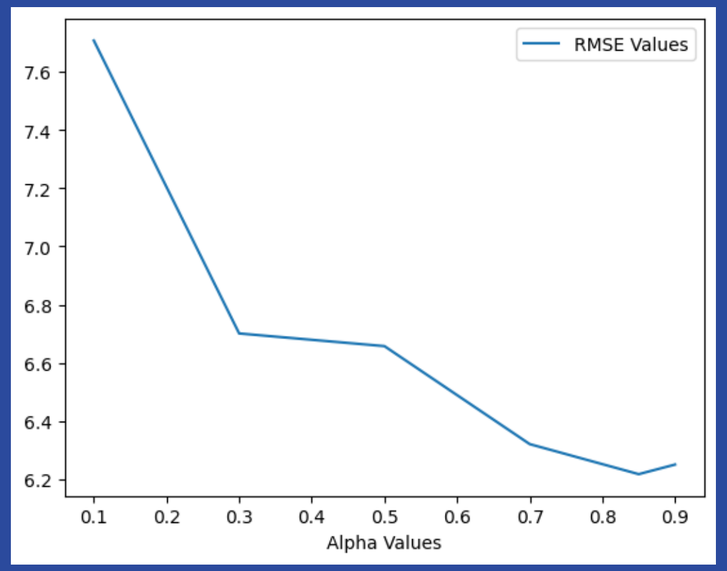
	RMSE	Runtime (s)
Neural Network (Tensorflow)	9775	513
Hierarchical time series (ARIMA)	63.78	10635.6
Simple moving average	50.08	311.2
Baseline model (mean of daily SKU sales)	22.19	N/A
XGBoost regression model	11.55	982.61
Random forest regression model	10.01	55.49
Croston time series	6.22	4.3

CROSTON TIME SERIES MODEL

MODEL

- Combines average of with average of periods of non-demand
- Good at predicting slow-moving, intermittent sales

SMOOTHING



Alpha: 0.85
Beta: 0.5

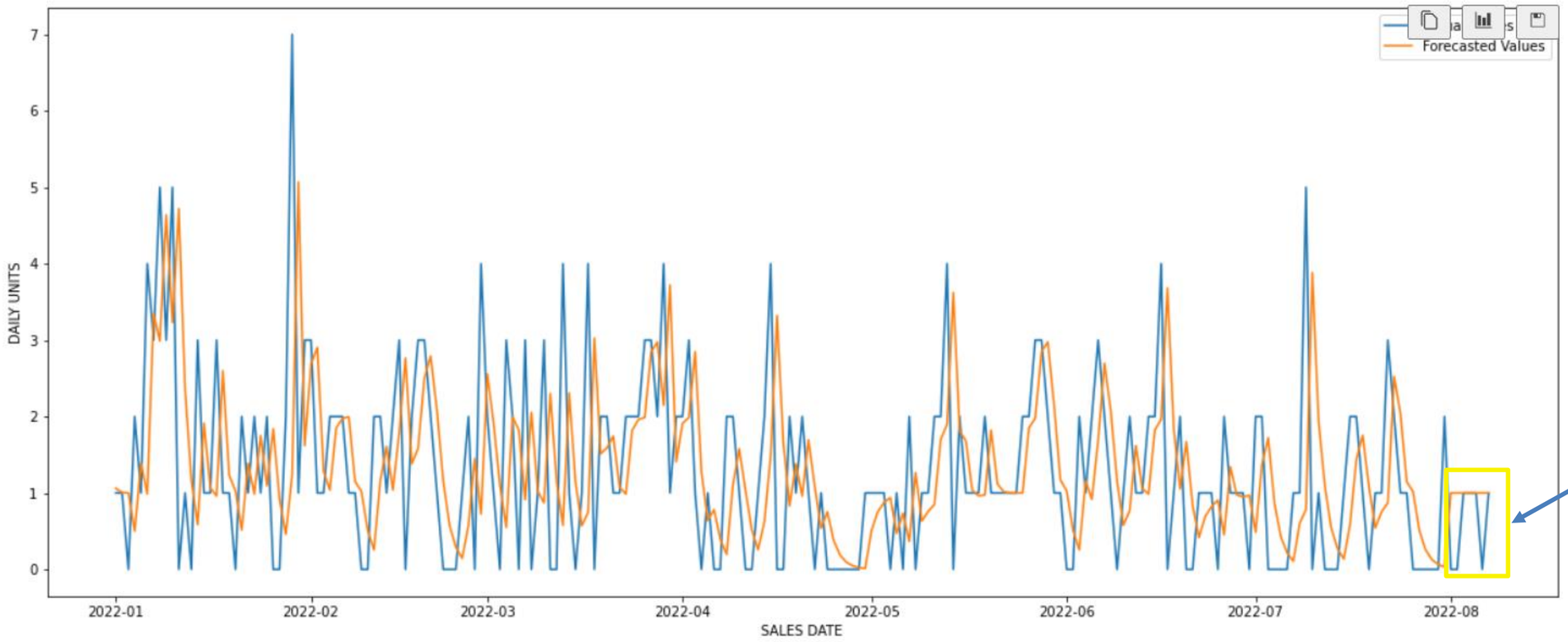
RESULTS

RMSE
4.98

RUNTIME
Predict: 1.5s
Full file: 20.1s

CROSTON MODEL PREDICTIONS

Actual vs. Predicted Daily Sale Units for SKU 7



Forecasted units

CROSTON TIME SERIES MODEL

Advantages

Fast to train and easy to scale

Easy to adjust smoothing parameters

Easily interpretable

Drawbacks

Inflexible

Does not include features

Predicts the same value for each SKU for all seven days

RECOMMENDATIONS

- If error increases in the future, try adjusting smoothing parameters
- If Best Buy is seeking greater accuracy, a combination of an XGBoost model for higher-selling SKUs in this dataset and a Croston model for the most intermittent SKUs could be even more effective