Blackwell Electronics-Volume Sales

Blackwell Electronics Sales Team (BEST) is interested in knowing what the projected sales volume will be for several products. PC’s, laptops, netbooks, and smartphones are new products for consideration, and customer or service reviews may have an impact on their sales volume. Understanding how products fare against in each other in the market will help Blackwell Electronic determine which products will be profitable.

Blackwell Electronic Data Team (BEDT) used data analytics to provide insight to which new products will be profitable on the market. BEDT analyzed historical sales data to create sale volume predictions of several new products types. Variables that were highly correlated with each other (correlations > 0.90) or not positively correlated with volume were removed. Four Star (0.72) and Positive Review (0.61) had the highest correlations with Volume.

BEDT created and tested the following machine learning models for regression analysis, Support Vector Machine (SVM), Random Forest (RF), and Gradient Boost Machines (GMB). An optimal model was then selected based on performance (Root Mean Square Error and R-Squared) and used to predict sale volume from new product list.

Results (Root Mean Square Error (RMSE) Mean Absolute Error (MAE) (See appendix for full results)

RF (mtry=10)

Trained Model

Model Performance Testset (Resample)

RMSE Rsquared MAE

218.9073 0.8854617 80.7824

GBM (int.dep.=7, n.minobsin. =1, n.trees=100

Model Performance Testset (Resample)

RMSE Rsquared MAE

179.7829 0.9297963 74.5361

SVM (cost =0.1)

Model Performance Testset (Resample)

RMSE Rsquared MAE

347.1323 0.6938 154.2098

All of the models had RMSE values that increased slightly from trained model to prediction model performance. Out of the three models, GBM performed the best. GBM had the highest R2 and lowest RMSE and MAE values. These values indicate GBM has low model prediction error and the model predictions are close to the observed values.

Predictions of the sale volume based on GBM

|  |  |
| --- | --- |
| **ProductType** | **Volume** |
| PC 1 | 485.99 |
| PC 2 | 164.72 |
| Laptop 1 | 195.93 |
| Laptop 2 | 54.91 |
| Laptop 3 | 20.32 |
| Netbook 1 | 64.93 |
| Netbook 2 | 1292.35 |
| Netbook 3 | 134.33 |
| Netbook 4 | 31.68 |
| Smartphone 1 | 404.80 |
| Smartphone 2 | 605.18 |
| Smartphone 3 | 84.810 |
| Smartphone 4 | 123.55 |

Impact of service reviews on sales volume

Data suggest the impact positive service reviews has one sales volume is gradual increase. Sales department should look at products with high positive service reviews as they will tend to have high Negative Service Reviews had several outliers. Strong negative service reviews do not equate to higher sale volumes.

Lesson learned

It is important for businesses to be aware of collinearity and variables that do not contribute to model prediction. These variables should be removed as they can affect product predictions. Dummy variable function is helpful converting factors into numeric data that can be analyzed with correlation. I also learn that model optimization and feature selection are tedious process.