Dear Danielle,

The followings illustrate the analysis and prediction we conducted on the costumer brand preferences based on the existing product attributes report to predict sales volume the in new product attribute report.

Gradient Boosting, Random Forest, and Support Vector Machine (SVM) regression model have been used for our prediction. Early screening of the data indicated that the product num and Product Price variables have no or very little correlation with other variables, therefore we decided to remove those from out predictions. Please the figure below:

Chart, scatter chart

Description automatically generated

We decided to use GBM model due to greater accuracy and efficiency-over 90%:



Based on our analysis, we projected the following sales volumes for four target product types found in the new product attributes data set:

|  |  |
| --- | --- |
| ***Product Name*** | ***Volume*** |
| ProductTypeLaptop | 467 |
| ProductTypeNetbook | 120 |
| ProductTypePC | 200 |
| ProductTypeSmartphone | 236 |

The figure below shows 5 Stare reviews are having direct and positive impact on the sale volume:

Chart

Description automatically generated with medium confidence

Armin