# Sales Prediction Report

The following algorithms were tried to predict **volume sales**: Random Forest, Support Vector Machine, and Gradient Boosting.

A first iteration of the Random Forest model resulted in the following variable importance:

**Figure 1**

Chart

Description automatically generated

This first iteration had the following metrics:

RMSE R-squared

1804.8486870 0.7051622

The following variables were selected for the next iteration:

* x5StarReviews
* x4StarReviews
* x3StarReviews
* x2StarReviews
* x1StarReviews
* PositiveServiceReview

**Figure 2**

Chart

Description automatically generated

The second iteration had the following metrics:

RMSE R-squared

1466.2004124 0.7710457

The second iteration of the **Random Forest** model explains **77%** of the variation in volume sales with the variables seen in **Figure 2.**

The best Support Vector Machine model had a RMSE of 2235.0525453 and R-squared of 0.3942749. The best Gradient Boosting model had a RMSE of 2167.2850344 and R-squared of 0.3530976. Therefore, the second iteration of the **Random Forest** model was **chosen** to make the predictions.

The predictions are the following:

Predictions Total Sales Volume

PC 371, 179 550

Laptop 244, 28, 8 280

Netbooks 76, 1195, 111, 12 1394

Smartphones 376, 536, 142, 204 1258

The predictions can be found in more detail in the csv document called C3\_T3\_Electronics\_Predictions.csv.