

Are the Predictions Good?

Analyzing the trained model on Google Prediction API, we get the classification accuracy for predicting the acceptability of a car is 0.94. The accuracy rate of 0.94 is pretty high, which means that the trained model makes rather accurate predictions. I will show that the predictions from the trained model are good from different aspects.

There are 1727 data instances in the trained model, so there are enough examples when the model was trained. There are 6 features in the dataset; so I could potentially include more features that may influence the results. Of the 1727 data instances, the target variable is divided to four categories: unacceptable, acceptable, good, and very good. The table below shows the number of instances in each category.

Category	Number	Percentage of Total Number of instances
Unacceptable	1210	70.02%
Acceptable	384	22.22%
Good	69	3.99%
Very Good	65	3.76%

We can see from the table that more data instances fall into the unacceptable category than those fall into the other three categories in total. The dataset has an unbalanced distribution. That may be a reason for the model's not yet perfect accuracy rate.

The confusion matrices are presented below for the trained model in each of the four car acceptability categories.

Unacceptable:

Unacceptable	115.20	Good	0.20
Acceptable	3.60	Very Good	0.00a

Acceptable:

Unacceptable	2.80	Good	1.00
Acceptable	35.40	Very Good	0.60

Good:

Unacceptable	0	Good	6.00
--------------	---	------	------

Acceptable	1.00	Very Good	1.20
Very Good:			
Unacceptable	0	Good	0
Acceptable	0	Very good	6

Of the four categories, cars, which have very good acceptability, are predicted to have very good acceptability in the trained model. Cars, which have good acceptability, are mostly predicted to have good acceptability; however, it may be predicted to have acceptable or very good acceptability. The predictions for acceptable cars are acceptable in majority, but it also has false positives in the other three classes. Finally, the predictions for unacceptable cars are mostly accurate with the high number in unacceptable column.

To summarize, although the training dataset could be more balanced and could have more features, I would consider predictions from the trained model on Google Prediction API good. The accuracy of the classification model is 0.94 with false positives and false negatives shown in the confusion matrices. In the aspect of measuring prediction accuracy, the trained model is good with a high accuracy.