Abstract

Ryder Inc., is a provider of heavy transportation equipment operating more than 110K vehicles in North America alone. At the end of a vehicle’s leasing period Ryder sells these vehicles either as a retail or wholesale. Ryder would benefit from a model that will predict future selling prices (proceeds) of the vehicles based on their attributes and historical proceeds.

Methods I Design I Analysis

- Visualizing the data showed that proceeds were significantly different based on vehicle type & vehicle condition (Figure1). Hence, we decided to divide the data into 3 different groups: tractor, trailer, and truck and build a model for each.
- Correlation and association analysis were performed on, both, the continuous and categorical variables, respectively, where variables with the highest correlation/association were used in the model and the rest were removed.
  - Continuous variables: Spearman correlation analysis was performed, result in (Table2), with “Current Fair Mkt Value” variable having the highest correlation (Figure2)
  - Categorical variables: ANOVA testing was performed for the categorical variables and were one-hot-encoding for the Machine-Learning models
- After filtering the variables based on their significance, the models were trained and tested using 80% and 20% of the data, respectively, using R:
  - Models used were Linear, Random Forest, KNearestNeighbor and xgbLinear results in (Table3)

Conclusion

In conclusion, the ANOVA and Spearmen analysis helped us filter our variables down to 9, 6 and 7 variables for the models used for tractor, truck and trailer, respectively.

Using the filtered variables we were able to build our models and get R² values of 0.94, 0.93 and 0.95 for tractor, truck and trailer, respectively using the most accurate Machine-Learning Model: xgbLinear.

As a result, this model will help Ryder Inc. better understand expected vehicle proceeds, with potential pricing implications, hence improving their liquidity and reducing associated liabilities.

Acknowledgments

It has been a pleasure working with the following people who provided a great help to us, thank you:

Dr. Jaime Buitrago
Lamis Amer
George Vargas from Ryder

Numerical

- Odometer Reading
- Age in Selling Lot
- Current Fair Market Value
- Projected Mileage
- Flow Rate
- Body Height
- Body Width
- Body Length
- Number of Axles
- Initial Cost
- Basic Value
- Status

Categorical

- Vehicle ID
- Transmission Model
- Purchase Type (New/Old)
- SAM Class Code
- Engine Manufacturer
- Model Manufacturer
- Transmission Type
- Engine Horsepower
- Engine Condition
- Transmission Speed
- Rear Axle Capacity
- Engine Horsepower
- Engine Condition
- Transmission Speed
- Model Year
- Transmission Manufacturer
- SalesGr (Retail/Wholesale)
- Vehicle ID
- Transmission Model
- Purchase Type (New/Old)
- Engine Manufacturer
- Model Manufacturer
- Transmission Type
- Engine Horsepower
- Engine Condition
- Transmission Speed
- Model Year
- Transmission Manufacturer
- SalesGr (Retail/Wholesale)

Table1 (Numerical & Categorical Variables)

Figure1 (Vehicle Type vs Proceeds)

Figure2 (Fair Market Value vs Proceeds)

Table2 (Spearman Coefficient Continuous Variables vs Proceeds)

Table3 (R² value for the different models)