Bibal, Gidanz Rafterjude 2015-07805

Villamera, Christian 2015-01218

CS 180 Project

Car Evaluation Based on ???

Abstract

1. Introduction

“A good reputation for quality vehicles is essential if one is thinking of selling used cars”[7]. The decision of selling your very first or pre-loved car is definitely not an easy task. Several factors are to be considered especially the car’s condition and specifications if you want to get the most out of your car.

This project aims to help and motivate the owner/seller of the car decide whether the car can be sold at a reasonable price or if the car isn’t suited for selling at all. This project is organized under 6 main car attributes/specifications namely: Buying Price, Maintenance Price, Number of Doors, Passenger Capacity, Luggage Capacity, and the Overall Safety of the car. The following attributes plus the owner’s comments about his own car determine the car’s evaluation of whether the car to be sold is either Unacceptable, Acceptable, Good, or Very Good.

1. Short Review of Related Studies
   1. Consumer Acceptability

Consumer Acceptability as defined, results from determining the feasibility of whether a product or service will be acceptable to the consumer requires tests, surveys, pretests and even prototypes[1]. Consumer Acceptability is an important factor in the product design process. For this project, we demonstrate the application of the proposed model, car evaluation to be exact, to show that the consumer acceptability problem can be evaluated and predicted using the proposed model[2].

* 1. Machine Learning Models
     1. Hierarchical Decision Model(Decision Tree)

A Hierarchical Decision Model such as a Decision Tree is a decision support tool that uses a tree-like graph or model of decisions and their possible consequences, including chance event outcomes, resource costs, and utility. It is one way to display an algorithm that only contains conditional control statements[5]. For our project, the Decision Tree branches out but falls shortly only to the only possible outcomes: (positive, unacceptable), (positive, acceptable), (positive, good), (positive, very good), (negative, unacceptable), (negative, acceptable), (negative, good), and (negative, very good).

* + 1. Sentiment Analysis

Sentiment analysis – otherwise known as opinion mining – is a much bandied about but often misunderstood term. In essence, it is the process of determining the emotional tone behind a series of words, used to gain an understanding of the the attitudes, opinions and emotions expressed within an online mention[3].

Sentiment Analysis is the most common text classification tool that analyses an incoming message and tells whether the underlying sentiment is positive, negative, or neutral[4]. For this project, we have applied sentiment analysis to gauge the comments and opinions of the owner to his car as input and determine it either as positive or negative one.

1. Methodology and Results
   1. Methodology

For both the evaluation for the consumer acceptability on cars and the sentiment analysis on reviews,we used the datasets on the archive of UCI Machine Learning Repository.

The dataset for the consumer acceptability, *Car Evaluation Data Set,* evaluates cars according to six attributes: Buying Price, Maintenance Price, Number of Doors, Passenger Capacity, Luggage Capacity, and the Overall Safety of the car. There are four class values: unacceptable, acceptable, good, and very good. This dataset can be useful for structure discovery methods and decision analysis. It has 1728 number of instances.

The dataset for the sentiment analysis, *Sentiment Labelled Sentences Data Set,* determined if a review is positive or negative. It has 3000 reviews from IMDB, Amazon and Yelp.

For preprocessing *Car Evaluation Data Set*, we converted the values of the attributes of the cars into integer and store it in an array. Here is the conversion:

Buying Price

The project is organized under 6 main car attributes/specifications namely: Buying Price, Maintenance Price, Number of Doors, Passenger Capacity, Luggage Capacity, and the Overall Safety of the car. Each attribute accepts the following input parameters:

Buying Price: (vhigh, high, med, low)

Maintenance Price: (vhigh, high, med, low)

Number of Doors: (2, 3, 4, 5more)

Passenger Capacity: (2, 4, more)

Luggage Capacity: (small, med, big)

Overall Safety of the Car: (low, med, high)

The project also accepts a comment or an opinion of the owner to his car. The program runs through the use of Decision Trees and Naive Bayes as a machine learning technique for classification. The program then outputs a comment about the owner’s car based on the car’s attributes along with the sample training data set.

* 1. Results

Ilang percent yung ginamit mo sa data set, ilang percent yung ginamit mo as test set. Ano yung naging result nya after some sample runs

* 1. Discussion

Lagay natin dito yung accuracy nung result, limitations nya, possible improvements. GUURL PATRY IRUN NUNG CODE SA SAMPLE DATA SET DI KO MARUN

1. Conclusion

General conclusion ng buong paper

Inference:

All the variables are important for customers in assessing whether the car is in acceptable or unacceptable range.  
 Safety and Seating capacity are two main factors in rejecting the cars as unacceptable.  
 Number of doors are the least important variable in deciding the class value of the car.

1. Bibliography

Sources:

Data Set: <https://archive.ics.uci.edu/ml/machine-learning-databases/00331/>

<https://archive.ics.uci.edu/ml/datasets/car+evaluation>

<https://archive.ics.uci.edu/ml/datasets/Sentiment+Labelled+Sentences>

For RRL:

About Consumer acceptability:

[1]<https://bizfluent.com/facts-7196986-consumer-acceptability-.html>

[2]<https://books.google.com.ph/books?id=r-iMSOgr-pwC&pg=PA1548&lpg=PA1548&dq=consumer+acceptability+car+evaluation&source=bl&ots=x2kXESp_aO&sig=TQUMv7EfZgFgnfoNGqq0ghweC6U&hl=en&sa=X&ved=0ahUKEwi07Jzj7KrbAhXBxbwKHShuDpkQ6AEINzAB#v=onepage&q=consumer%20acceptability%20car%20evaluation&f=false>

Sentiment Analysis:

[3]<https://www.brandwatch.com/blog/understanding-sentiment-analysis/>

[4]<https://towardsdatascience.com/sentiment-analysis-concept-analysis-and-applications-6c94d6f58c17>

Decision Tree:

[5]<https://en.wikipedia.org/wiki/Decision_tree>

[6]<https://towardsdatascience.com/decision-trees-in-machine-learning-641b9c4e8052>

Other Sources:

[7]<https://www.sectorsdonut.co.uk/sectors/automotive/used-car-dealer/overview>