Prediction of Automobile Reviews

Jiaqi Chen Xingtan Hu Xiaoyang Lu

Problem Statement

Nowadays, people’s reviews for automobile is not only based on its price and depreciation. The ownership cost is a key factor.

Hypothesis

People’s review for an automobile may be affected by TCO[1]

[1] **True Cost to Own (TCO)** is proprietary data that helps you estimate the total five-year cost of buying and owning a vehicle — including some items you may not have taken into consideration.

Method

1. Using sentiment analysis to analysis the automobiles' reviews.

2. Using the concept from edmunds.com to achieve the TCO Price

3. Using the knowledge for Machine Learning, try to predict the review expectation(AFINN)

Data

How many:

1) 15 model for each types of automobile.
   Convertible, Coupe, Hybrid, Luxury, SUV……
2) all related automobiles TCO prices.

Where: collect data from edmunds.com

* the information quoted from usnews.rankingsandreviews.com to make the classification for different types of automobile.

Related Work

Edmunds.com's True Cost to Own (TCO)

The True Cost to Own calculations use the following set of assumptions:

- Ownership expenses are estimated for a five-year period
- You will drive 15,000 miles per year
- You are financing the vehicle using traditional financing, not lease financing
- You have an above-average credit rating for the purpose of determining your finance rate
- You are making a 10% down payment on the vehicle at purchase
- Your loan term is 60 months

Sentiment Analysis(lexicon analysis)

Mechine Learning(mlpy)

http://mlpy.sourceforge.net/docs/3.4/

Timeline

October 22 : Proposal Presentation

October 29 : Collecting Data

November 5 : Filter and Organize Data

November 12 : Method implementation

November 19 : Analysis of results

November 26 : Conclusions and reviews

December 3 : Presentation