Sentiment Analysis Based on Car Reviews

Jiaqi Chen Xingtan Hu Xiaoyang Lu

Problem Statement

People's review for an automobile may be affected other than its Retail Price

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 classify the automobiles' reviews.

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

3. Using the information from USnews.com to make the classification for different Make/Model

Data

How many:
1) all the related reviews in one year
2) all related automobiles TCO prices and Retail Price.

Where: collect data from edmunds.com

Attributes: detail information of automobiles, TCO prices and Retail Price.

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

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