International Journal of Scientific and Technical Advancements
IJSTA
ISSN: 2454-1532

Sentiment Classification of Reviews on Car Domain

Sunil Salunke[1], Melvin d'souza[2], Sulochana Devi[3]
[1, 2, 3]information Technology, Mumbai University, Mumbai, Maharashtra, India-400016
Email address: [1]sunilch79@gmail.com, [2]melvin.d21@gmail.com, [3]bishnoisulochana@gmail.com

***Abstract*—**This paper presents a working of opinion mining the online user review to generate sentiment analysis that can guide the users to take the decision. In this work, sentiment analysis on car domain is presented. In which reviews are collected from various domain related websites, then AFINN wordlist was used to get the scores for the list English words. The outcome of system will provide user with number of positive and negative sentiments. The result is presented in the form of pie charts.

***Keywords*—** Opinion mining; sentiment analysis; AFINN wordlist; data mining; sentence level; polarity.

I. INTRODUCTION

With increase in the ease of access to internet, user as lots of data and information on topics user is interested. Social media and e-commerce have created lots of buzz in recent years. Users are used to share views on particular news, issues, phones, movies, services, government policies, etc. User readings the reviews on the particular services, phones, movies, cars with huge amount of information or reviews available becomes practically impossible to go through each and every reviews and analyze pros and cons of each review. Which in turn, becomes difficult to make a decision whether to go for it or not. With ecommerce industry being a booming industry, would like to serve the people with better goods and services. For that, industry would like to analyze their customer reviews about their product and services. So that, they can make changes or take effective step to increase their businesses. With reviews generated in thousands, there is need of mining the opinion of the user's reviews, to know the positive and negative sentiments. This classification will help business and individual to make decisions.

Sentiments analysis is a type of natural language processing for tracking the mood or opinion of the public about a particular product or topic or service. Sentiment classification can be document level, feature based level and have various ways of do it by supervised approach, unsupervised approach. We are doing sentiment classification on document level with a supervised approach.

II. SYSTEM

In this section we will explain system design and text classification implemented by us. As shown in figure 1. Our approach has following steps: 1) load the review data 2) load AFINN wordlist 3) clean the review data 4) apply classification algorithm 5) results

A. System Architecture:

1) Load data

The data consists of text format, which is to be analyzed. This data is prepared by collecting the reviews from various car websites were user frequently post their reviews.

2) Load AFINN wordlist

(AFINN is a list of English words rated for valence with an integer between minus five (negative) and plus five (positive)).examples: good (+4), pathetic (-2).

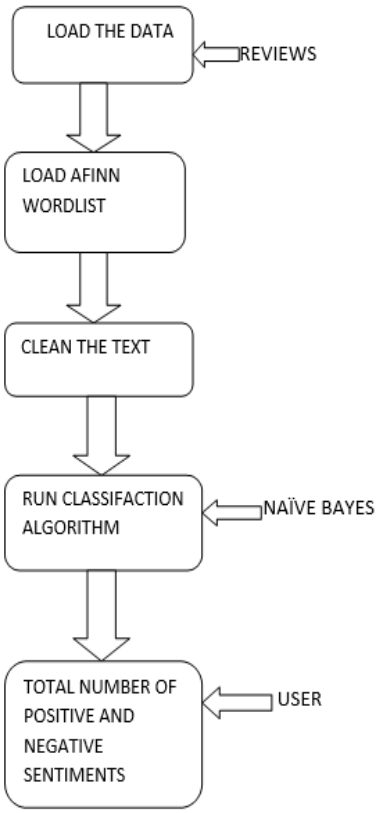

Fig. 1.

3) Clean the sentence (data)

In this process, unnecessary spaces, punctuations, characters are removed from the review data.

4) Apply naive ayes algorithm

Idea is that it looks at how the number of words in each of the four categories relates to whether the sentence is positive or negative. It then tries to guess whether a sentence is positive or negative by examining how many words it has in each category and relating this to the probabilities of those numbers appearing in positive and negative sentences.

Sunil Salunke, Melvin d'souza, and Sulochana Devi, "Sentiment classification of reviews on car domain," *International Journal of Scientific and Technical Advancements*, Volume 2, Issue 2, pp. 67-68, 2016.

5) Analyze the results

Results are analyzed by using pie charts by taking count of number of positive and negative sentiments.

III. RESULTS

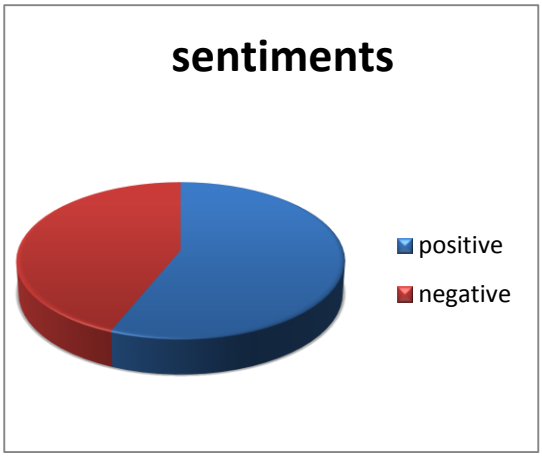

Fig. 2. Sentiments on document level.

IV. CONCLUSION

This paper presents methodology to process, analyze and summarize sentiments and opinions from sentences. This methodology was applied know the sentiments on the particular car. Our practical result, on figure 1 is an illustration of how we can visualize the results from the methodology. In particular results, particular car received more positive sentiments.

Future work

Feature based sentiment analysis can be carried out on this domain. Feature level analysis will be more effective especially considering the feature like mileage, comfort, cost. It will help companies to decide their strategy and focus.

ACKNOWLEDGMENT

We would like to express our gratitude to our project guide prof. Sulochana Devi for her guidance and support throughout the project. She has always directed us in the right direction and helped us to stick to our project's domain all the time, which in fact helped us to complete this project with much ease and in time. We are truly thankful for our Head of Department (HOD) Prof. Chhaya Narvekar for allowing us to implement our idea into its practical form wherein there were a lot of interesting and challenging things to learn. We would also like to mention and thank my friends and family for their constant support.

REFERENCES

[1] B. Liu, *Web Data Mining*, Book, 2007.
[2] B. Pang and L. Lee, *Opinion Mining and Sentiment Analysis*, Book, 2008.
[3] A. Sankhe and P. Gharpure, "Feature based sentiment analysis for online reviews in car domain," *International Journal of Current Engineering and Technology*, vol. 4, no. 2, pp. 680-684, 2014.
[4] R. M. A. Baracho, G. C. Silva, and L. G. F. Ferreira, "Sentiment analysis in social networks: A study on vehicles," vol. 938, pp. 132-143, 2012.

Sunil Salunke, Melvin d'souza, and Sulochana Devi, "Sentiment classification of reviews on car domain," *International Journal of Scientific and Technical Advancements*, Volume 2, Issue 2, pp. 67-68, 2016.