Sentiment Analysis on Amazon Fine Food Review Dataset

Xinyi Li, Hongzhi Liu, Sheng Wang, Siyu Zhang[1]

[1]University of Wisconsin-Madison

Abstract

some abstract here

1 Introduction

Sentiment analysis is an automated process of understanding opinions from natural language. Nowadays countless reviews are available on Internet, but almost all of them are unstructured data, making direct analysis difficult and time-consuming. Sentiment analysis can help address this issue by quantifying the information in the unstructured data, thus the polarity hidden behind the texts can be extracted for further analysis.

Recently sentiment analysis has gained much popularity recently and has been used to resolve many practical issues, such as marketing analysis, customer service, etc. In this report, we apply sentiment analysis on Amazon fine food review dataset from Stanford Network Analysis Project (SNAP) which contains 568,454 reviews from more than 200,000 users to study customer behavior and predict the ratings. There are many methods to implement sentiment analysis when our goal is to predict the level of polarity of opinions (thus can be viewed as a classification problem), such as Logistic Regression, Support Vector Machine, Neural Network.

among which we chose RNN.... since ....

2 Data Cleaning

3 Model