Sentiment Analysis for Amazon Fine Foods Reviews

March, 2019

1   Team

Boran Yildirim 21401947
Deniz Evrensel 21401841
Batihan Akca 21502824
Furkan Arif Bozdag 21401748
Sekip Kaan Ekin 21401795

2   Description of the Data

URL: https://snap.stanford.edu/data/web-FineFoods.html

This dataset consists of reviews of fine foods from Amazon. The data includes all reviews from October 1999 up to October 2012. Reviews include product and user information, ratings, and a plaintext review.

Number of reviews	568,454
Number of users	256,059
Number of products	74,258
Timespan	Oct 1999 - Oct 2012

3   Description of Problem and Answer

Amazon and similar e-commerce websites are used vastly for online shopping purposes and these websites allow their users to write reviews about the products or services they received. These reviews have significant influence on the other users while deciding to buy a product or not. Therefore, it is valuable information to know the essence of a specific product's reviews. Furthermore, a classification made by the information gathered from these reviews can be applied to services such as product summary and product recommendation system. In this project, we intend to classify the usefulness of each product by studying their reviews using different learning models like Supervised Learning, Deep Learning and Unsupervised Learning. Finally, the possible outcomes of different models will be compared based on criteria which we will decide.

4   Plan to Achieve by the Milestone

We plan to implement following Supervised Learning algorithms by the Milestone.

1) Linear Regression   5) Naive Bayes
2) Logistic Regression   6) k-NN
3) Decision Tree   7) K-Means
4) SVM   8) Random Forest