How Helpful is This Amazon Review?

Cristian Caraballo Richard Addo Hanna Kim

PROBLEM

Reviews have completely transformed the experience of online shopping. In the past, there was much risk involved for the consumer because of uncertainties in the product's quality, durability, efficiency, and etc. Now, reviews serve as a tool to minimize the distance between the product and the consumer. In theory, having a public forum on these shopping websites where previous buyers can enter in text reviews and star-based ratings should solve the problem. However, such a method does not account for the fact that undetailed textual reviews do exist and do little to inform the decision making process of a potential buyer. There have been many modifications made on shopping websites to work towards this goal of providing a potential buyer with enough information to help decide whether or not to purchase an item.

Amazon.com, for instance, poses the question, "Was this review helpful to you? Yes/No" to a potential buyer reading a review, and automatically sorts reviews based on "helpfulness". However, not everyone reading a review responds to this question, and hence inundation of unhelpful reviews still persists. Through the application of machine learning algorithms, we hope to develop a tool that can automatically measure the helpfulness of any online review to assist the consumer with the weeding out of unhelpful reviews and to promote the ones which provide substantial information.

DATA

We will use a subset of the Amazon reviews dataset from the Stanford Large Network Dataset Collection. The dataset consists of almost 35 million consumer product reviews from Amazon.com, spanning 18 years. Almost 2.5 million products are reviewed by over 6.5 million users, and over 50000 of these users have reviewed more than 50 products. Here is a snapshot of the format of the data:[1]

```
product/productId: B00006HAXW
product/title: Rock Rhythm & Doo Wop: Greatest Early Rock
product/price: unknown
review/userId: A1RSDE90N6RSZF
review/profileName: Joseph M. Kotow
review/helpfulness: 9/9
review/score: 5.0
review/time: 1042502400
review/summary: Pittsburgh - Home of the OLDIES
review/text: I have all of the doo wop DVD's and this one is as good or better than the
1st ones. Remember once these performers are gone, we'll never get to see them again.
Rhino did an excellent job and if you like or love doo wop and Rock n Roll you'll LOVE
this DVD !!
```

METHODS

$$\frac{\textit{Number of responders who chose "Yes"}}{\textit{Total number of responses to "Was this review helpful to you?"}}$$

For our project, we will define *helpfulness* as the above fraction and classify the reviews into four groups:

- Very helpful: $0.75 \leq helpfulness$
- Moderately helpful: $0.5 \leq helpfulness < 0.75$
- Helpful : $0.25 \leq helpfulness < 0.5$
- Not helpful: $helpfulness < 0.25$

We will implement three machine learning methods:

- Naïve Bayes
  The Naïve Bayes Classifier performs very well in text classification and can be trained very quickly [2]. Since we will implement this algorithm as part of the course material, we will use it as the baseline for our project.

- Random Forest Algorithm
  The Random Forest Algorithm, in addition to being fast and a good choice for large datasets like ours, works extremely well for text classification when features are well chosen.[3,4] We will utilize Random Forests as one of our two major algorithms to be evaluated against our Naïve Bayes baseline.

- Maximum Entropy
  Like Random Forests, MaxEnt has proven effective for text classification.[5] We will implement MaxEnt and evaluate its performance against our Random Forest implementation and ultimately determine which works better with the features from our dataset.

MILESTONE

We will start off the project by choosing and parsing the subsets of the dataset that we want to use for training and testing.

By the milestone date of February 17, we expect to have preliminary implementations of all three algorithms.

REFERENCES

[1] http://snap.stanford.edu/data/web-Amazon.html

[2] Huang, Jin, J. Lu, C.X. Ling. "Comparing naive Bayes, decision trees, and SVM with AUC and accuracy," *Data Mining, 2003.*

[3] https://www.stat.berkeley.edu/~breiman/RandomForests/cc_home.htm

[4] Lionel, Martin, Pearl Pu. "Prediction of Helpful Reviews Using Emotions Extraction". *Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence, 2014.*

[5] Nigam, Kamal, John Lafferty, Andrew McCallum. "Using Maximum Entropy for Text Classification". *IJCAI-99*, *Workshop on Machine Learning for Information Filtering*, 1999.