Learning to Classify Useful Reviews

Ertan Dogrultan   Cesar Romero   Paul Wais
{ertan,romero,pwais}@cs.ucla.edu

November 29, 2010

User Votes Inform Review Sorting and Search

837 of 855 people found the following review helpful:

⭐⭐⭐⭐⭐ **So Much **MORE** Than Just An MP3 Player!**, September 29, 2009

By **Alex Honda "onyx575"** ☑ (Los Angeles, CA USA) - See all my reviews

TOP 500 REVIEWER   REAL NAME

Amazon Verified Purchase (What's this?)

This review is from: Apple iPod touch 32 GB (3rd Generation) OLD MODEL (Electronics)

It's does so much more than just play music, which is why I wanted one. I'm not a Mac person at all but the iPod Touch...I love it.

Out of the box it found my Wi-Fi connection and had me surfing the net (the real Internet and not the cell phone web pages like the one on my Verizon phone). It was intuitive to use and I'm finally getting accustomed to using the touch screen to scroll and type away.

Elite '10   Valerie C.   🟧 Compliment
💬 144   Sex. Laugh. Pray.   💬 Send Message
🟧 113   Burbank, CA   📘 Follow This Reviewer

⭐⭐⭐⭐⭐ 2/27/2010

Best damn ice cream sandwiches in the United States of America. And they're only $1.50.

I go in Diddy Riese feeling like a million bucks sometimes. Trying to be all "creative" and "inventive", I make the top cookie in the ice cream sandwich chocolate chip, while the bottom cookie is something else, like peanut butter.
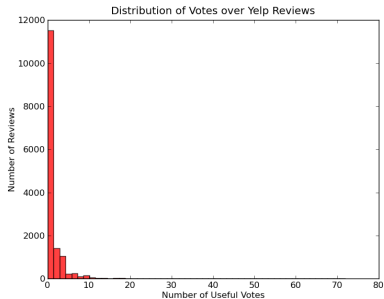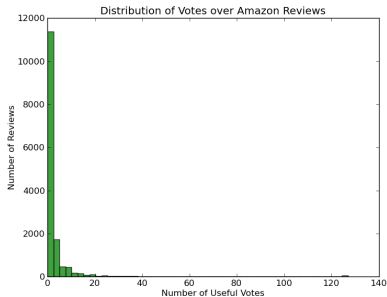
And maybe rocky road as my ice cream. Whoa! Dream BIG! =)
The combinations are endless...

Though the lines are long, you hardly notice cause they move so fast. Geez...I think I'm going again today.

Was this review ...?   Useful ● (18)   Funny ● (23)   Cool ● (21)

Unfortunately, votes are sparse ...

… and there are many useful (and not so useful) reviews without votes

Elite '10
70
379

Joie P.
eat, drink and be scary
Altadena, CA

Compliment
Send Message
Follow This Reviewer

Update - 4/21/2010   2 photos

Sorry guys! If you take the hot dogs away it's just an ice cream and cookie place. Albeit a good ice cream and cookie place but without the incredible hot dog deal to go with it my 10 yr old son said "Can you give them no stars?"

Saddest success story I've ever heard, really.

Bring the hot dogs back.

Was this review …?   Useful ⊙   Funny ⊙   Cool ⊙

★★★★★ **sjakhflbI FGNWELHUwwwwwiu eitkkhio;fkjhvndsjm,C**, November 28, 2010
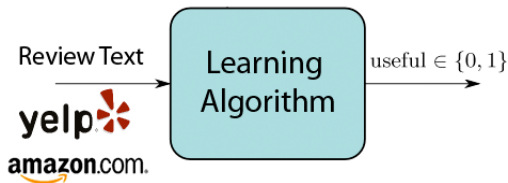By **meow meow licious** ⊡ - See all my reviews
Amazon Verified Purchase   (What's this?)
This review is from: Apple iPod touch 8 GB (2nd Generation--with iPhone OS 3.1 Software Installed) [OLD MODEL] (Electronics)
OMG THIS IS GOOD BUT IT KINDA STINKS AHHH APPLE U MAKE ME MADDDDDD

BUT I LOVE UR GUYSES MAC BOOK PRO OMG ITS TOTALLY AWESOME AND IOM WRITING THIS ON 1 RIGHT NOW

Goal: Learn to Classify Reviews as Useful (or Not)

- ► Corpus: 15,000 Yelp and Amazon reviews annotated with "useful" vote counts. [1]
- ► Test a variety of algorithms: SVM, Naive Bayes, AdaBoost, Winnow. Why does one algorithms outperform others?
- ► Also investigate domain adaptation: e.g. apply Yelp-trained model to Amazon reviews.

Note: Usefulness is Independent of Sentiment

65 of 74 people found the following review helpful:

⭐☆☆☆☆ **Warranty problems...**, March 24, 2010

By <u>Terry Crock</u> (Massillon, Ohio USA) - <u>See all my reviews</u>

REAL NAME

This review is from: **Apple iPod touch 8 GB (2nd Generation--with iPhone OS 3.1 Software Installed) [OLD MODEL] (Electronics)**

Same problem other people had--Apple's refusal to honor warranty. Battery was not able to hold an acceptable charge after only one month of use.

Elite '10
101
1150
Kymberli C.
Bay Area

Compliment
Send Message
Follow This Reviewer

⭐⭐☆☆☆ 11/11/2009

Came here for one of their infamous ice cream sandwiches... the line was out the door, which I bet is normal and the wait
Was this review ...? Useful • (8) Funny • (4) Cool • (7)

▶ Pang et al. [2] test Maximum Entropy, Naive Bayes, and SVM models for classifying sentiment and achieve about 80% accuracy.

---

[2] Pang, Bo and Lillian Lee and Shivakumar Vaithyanathan. "Thumbs up? Sentiment Classification using Machine Learning Techniques." Proceedings of the 2002 Conference on Empirical Methods in Natural Language Processing (EMNLP). 2002

Features for Classifying Usefulness

- Frequencies of ANEW[3] words, GRE/SAT words, logical signposting phrases, . . .
  - For example: "love" has valence 8.72; "rude" has valence of 2.5
- Frequencies of grammatical errors, typppppos, capitalization IrReGuLARIties, . . .
- URLs, IDF Scores of n-grams, product or business price, . . .

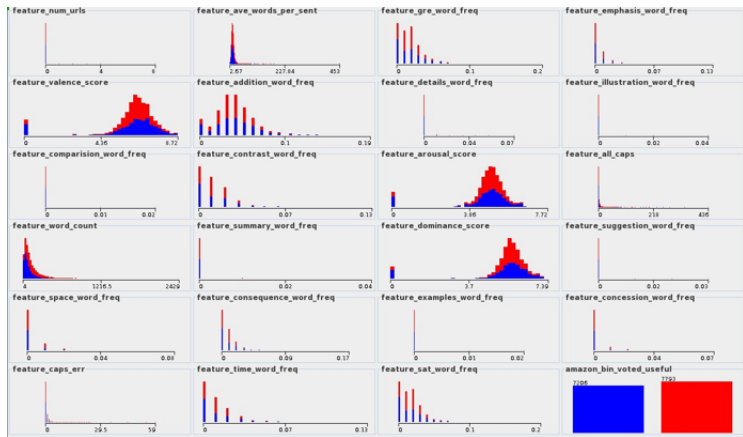[3] Sheridan Dodds, Peter and Christopher M. Danforth. "Measuring the Happiness of Large-Scale Written Expression: Songs, Blogs, and Presidents." Journal of Happiness Studies, Volume 11, Number 4, 441-456, DOI: 10.1007/s10902-009-9150-9.

Feature Distributions

Figure: The distributions of the features

Cross-validation accuracy of the algorithms

Algorithm	Accuracy
SVM (RBF)	70.13%
SVM (Linear)	75%
Adaboost	80%
Naive Bayes	68.5%
Winnow	56%

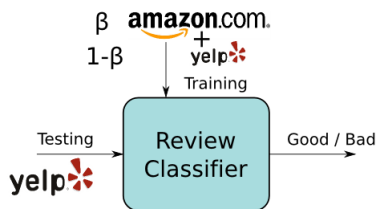Table: The 10 fold cross-validation accuracy of the algorithms

Complications

Why do different algorithms perform differently on this task?

- SVM
- Adaboost
- Naive Bayes

Future Work

- Domain adaptation experiments

- Connections with the theoretical bounds