

Automatically Determining Review Helpfulness

Hyung Yul Choi

Advisor: Kristina Striegnitz

Motivation

- Too many reviews
- Automatically find the helpful reviews

75 of 86 people found the following review helpful



Not enough bang for the bucks, and steaming options

By [Brendan Moody](#) [TOP 1000 REVIEWER](#) [VINE VOICE](#) on May 24, 2013

Vine Customer Review of Free Product (What's this?)

The BP730 is LG's latest high-end Blu-ray player, and at its current \$200 than the next tier down, the 530. What does that extra \$70 get you? LG is 2D to 3D conversion and the Miracast/NFC sharing, I can't comment on both devices they require, though I will point out that both are available on the None of the other four exclusives impresses me much.

Goal



- Determining the “features” of reviews
- Learning algorithm for prediction

Research Question

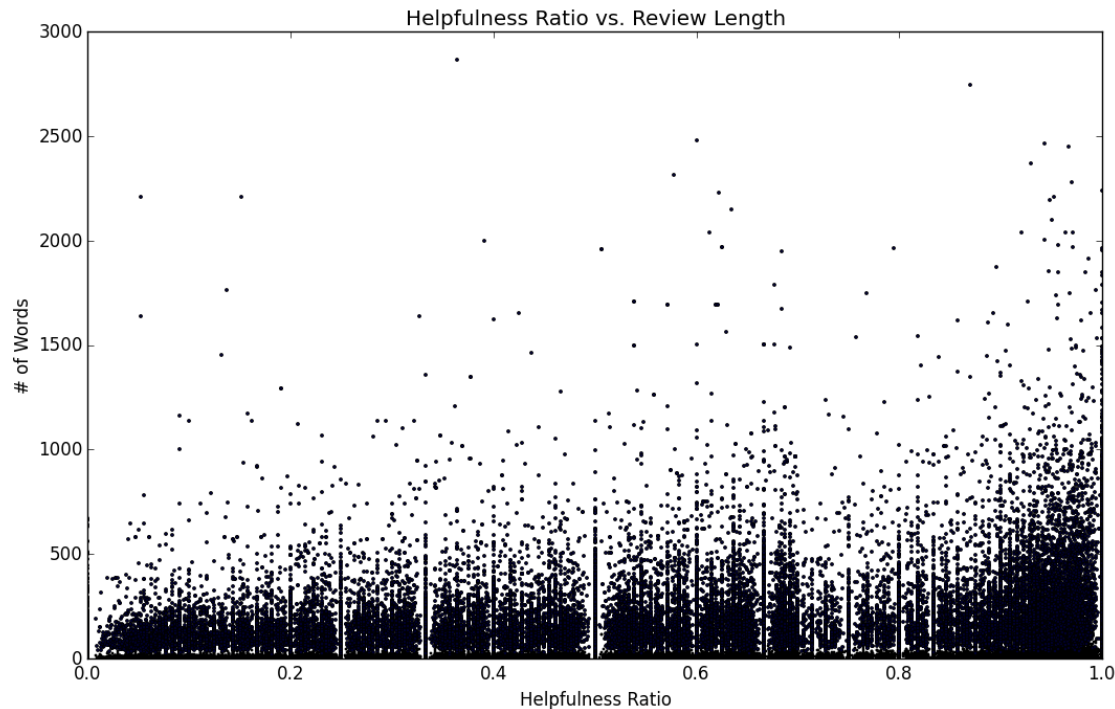
- What are the features of reviews that are indicative of their helpfulness?

Dataset

- Helpfulness Ratio =
$$\frac{\text{\# of votes found helpful}}{\text{\# of total votes}}$$
- Reviews tested for Pearson's r
 - Have at least 10 total votes **and** at least 5 sentences

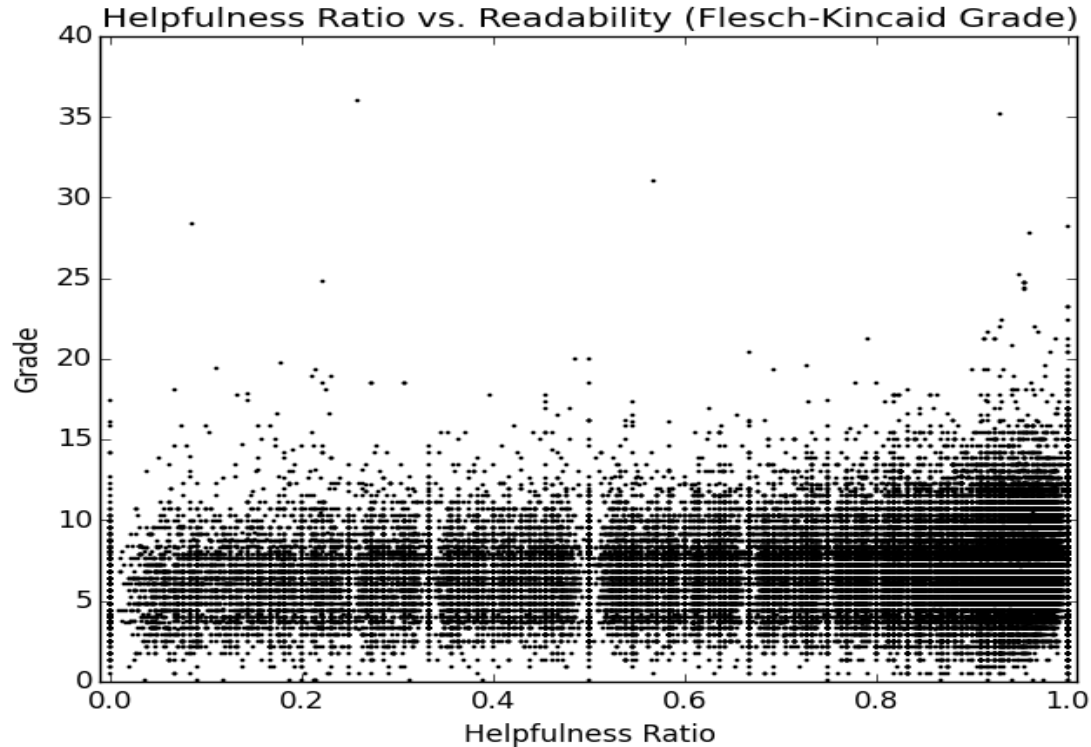
Total # of Reviews	1241778
# of Reviews, ≥ 10 Votes	167604
# of Reviews, ≥ 10 Votes and ≥ 5 sentences	116680
Average Helpfulness Ratio	0.78
Average Length of Review	108 words

Feature: length of review (# of words)



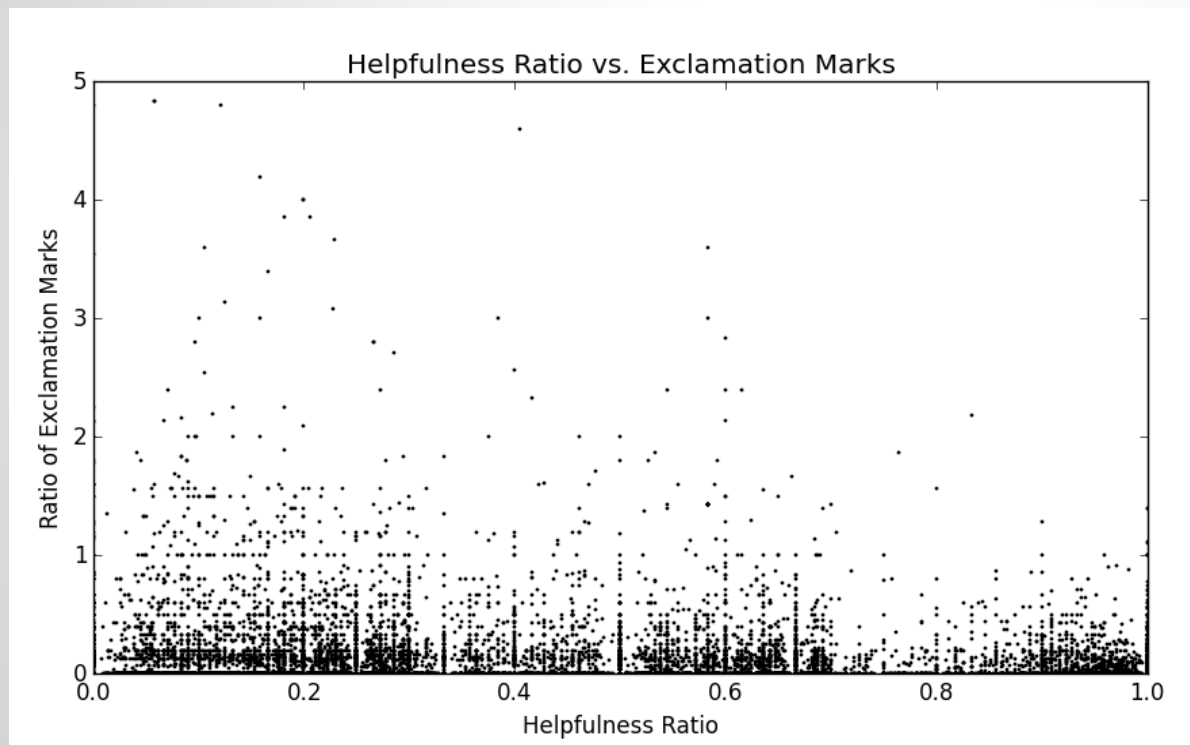
$$r = 0.26$$

Feature: Flesch-Kincaid Grade Level Test



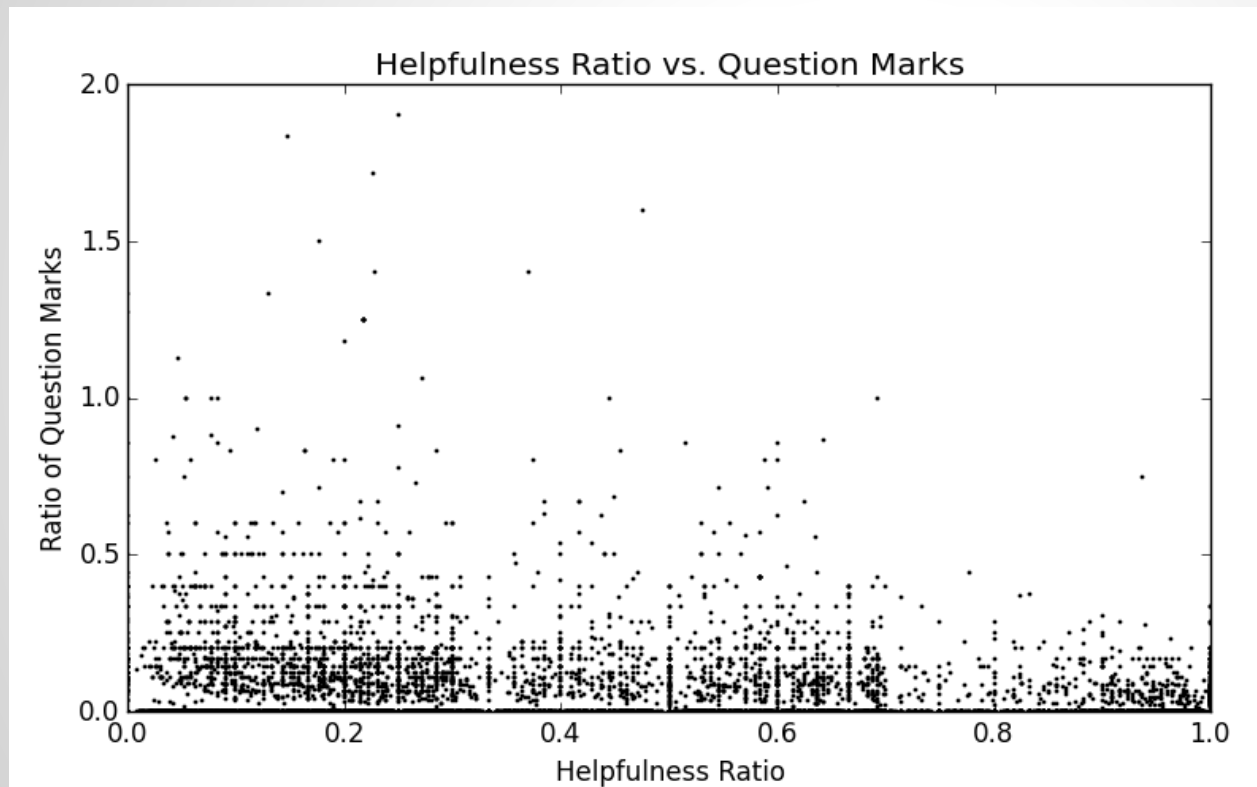
$$r = 0.17$$

Feature: punctuation, exclamation mark



$$r = -0.21$$

Feature: punctuation, question mark



$$r = -0.32$$

Other Features

Sentiment Polarity less helpful reviews use emotionally charged language	$r = -0.15$
Number of Sentences helpful reviews are longer	$r = 0.26$
Average Sentence Length sentence length has little correlation to helpfulness	$r = 0.07$
Grammatical part-of-speech Categories noun, verb, adjective use has little correlation to helpfulness	$r \approx \pm 0.05$

Results

- Prediction model
- Random baseline
accuracy: 33.3%
- Decision tree: 42.9%

Decision Tree Confusion Matrix				
		Predicted		
		Poor	Neutral	Good
Actual	Poor	308	146	133
	Neutral	217	201	169
	Good	192	187	208

Current Work

- Subsets of features
- Different # of classifications
- Different learning algorithms

Future Work

- More possible features can be explored
 - Lexical information
 - Information beyond review text

Conclusion

- Desire to collect helpful reviews
- Finding useful features
- Using features for helpfulness prediction