W266 Final Project Presentation

Bags _and_ Words: Mining Product Features and Customer Sentiment from eBags.com Reviews

Mike Frazzini

The Opportunity

- Unlock value in eBags.com's 3.5+ million customer reviews by extracting product features and related sentiment.

- Product features are valuable for product development and improvement, enhancing search and recommendation capabilities, and conveying product authority and expertise in a hypercompetitive Amazonian e-commerce world.

eBags®
Your Journey Starts Here

Add your own review

Product Reviews

Sort Reviews by [by Most Helpful ▼]

Ranked Most Helpful Review

"I'm an IT guy. I travel with a lot of stuff, parts, cables, connectors, a few more cables. And a tablet, and a laptop. My mobile office gets crowded and heavy. For years, I carried a backpack. A nice one, an expensive one. They sell them here and I probably have at least three between myself and my kids. And for many reasons, I like to continue to carry a backpack (makes me feel young) but sometimes want something that could pass for hand-held. It has to fit in that ridiculous area known as 'under the seat in front of me' or in between the table lamp and yak rug the passenger in front of you just squeezed into the overhead. The bag must be well organized, easy to navigate on the inside, stylish on the shoulder and easy to carry. This bag does all that. Mesh pockets on the inside make finding a cable or USB drive a snap. The AC Adaptor Garage has ample room for my laptop charger, a mouse and maybe a 10ft Ethernet cable. My expensive laptop is well protected in its own compartment. And unlike his big brother, the TLS Weekender, I can't fit an extra pair of shoes in this bag, but I could find room for what I need for an overnight. Important meeting, carry it like a briefcase after stowing the shoulder straps. Catching the train? Throw the staps back on and dash away. There are a few recommendations. One, I wish the zipper for the tablet compartment zipped the other way. All of the zippers culminate to the same end of the bag. It gets a bit crowder. But I love that tablet compartment. Two, there should be a sleeve in the second compartment for maybe a notepad or file folder. While the opportunity to organize excels in the front compartment, not so much in the middle. A minor grievance. Third, the backpack strap attachment system is weak. You guys engineered the heck out of this bag but put fifty cent clasps on the straps. Do better. I shore mine up with Nite-Ize S-Biners, size #2. Love the bag. Right size, right presentation. Great value and excellent engineering."

- How do you usually use this item?
  Business Travel
- What is your occupation?
  Computer, IT or Statistical
- What is your gender?
  Male

— 11/16/2015 by Eric from Pittsburgh, PA

Show Less

447 out of 457 people found this review helpful.

Did you find this review helpful?   Yes   No

"Mesh pockets on the inside make finding a cable or USB drive a snap"

The Data

- Filter 3.5+ million customer reviews for active products that also have customer ratings = 369,888 reviews

- Sentence Boundary Detection (Dan Gillick/Splitta) = 1,069,186 sentences

- Very noisy, much cleaning required:
  - Special character removal (™ and multile CR/LF's cause EOF in SBD)
  - Many meaningful mis-spellings ("backback" vs "backpack")
  - Relevant jargon ("gripe", "TSA")
  - Punctuation removal

Processing

Partitioning, Sentence Boundary Detection

Subjectivity/Objectivity Partitioning using Naive Bayes ML Model

Deep Learning for Sentiment Classification

| Partitioning, SBD | Cleansing | NB Extract Subjectivity | PoS Tag and Extract | Classification |

Special Characters, Misspellings, Punctuation Removal

Parts of Speech Tag Pattern Analysis and Extract

Subjectivity/Objectivity Partitioning

- Used to identify sentences that are more subjective and thus more likely to have product features sentiment (Yeh, 2006, Pang and Lee, 2004)

- Build Naive Bayes Classifier based on Subjective/Objective corpus (http://www.cs.cornell.edu/people/pabo/movie-review-data/)

- Initial results poor until augment with product feature list (before/after):

"I really like the separate ventilated compartment for shoes and wet swimsuit"
Subjectivity Prediction: 0.1257017008269576

"I really like the separate ventilated compartment for shoes and wet swimsuit"
Subjectivity Prediction: 0.9867362188433887

Parts of Speech Tag Pattern Analysis

- PoS Tag pattern analysis used to identify and further refine segments that have both product feature and sentiment attributes (Wang and Ren, 2002)

Word 1	Word 2	Word 3
JJ	NN/NNS	anything
RB/RBR/RBS	JJ	NOT NN or NNS
JJ	JJ	NOT NN or NNS
NN/NNS	JJ	NOT NN or NNS
RB/RBR/RBS	VB/VBN/VBD/VBG	anything

Deep Learning for Sentiment Classification

- Goal: build Weakly Supervised Deep Learning (WS-DL) model for sentiment classification (Guan, et. al, 2016). High accuracy with low supervision (labeling of 1%) with signal augmentation (customer ratings).

- Built 2 CNN models with NB ML baselines. Leveraged Denny Britz template for CNN (https://github.com/dennybritz)

- Successful WS-DL implementation was elusive, but significant accuracy improvements were achieved through signal augmentation by simply adding customer ratings in as word features.

How did this item meet your expectations?

Overall	★★★★☆
Style	★★★★★
Organization	★★★★☆
Price/Value	★★★★★
Durability	★★★★☆

Results/Analysis

- Huge improvement with data extraction by reducing field of 1,069,186 sentences down to 146,018 more "interesting" sentences.
  - However, post-facto random sample validation of original sentence set labeled for "interesting" sentences, indicates much tuning still needed:
    * Est. post-facto precision = 0.55 (model lets in uninteresting sentences)
    * Est. post-facto recall = 0.15 (model misses interesting sentences)

- Classification Accuracy (note labeled set quite small at 1,039):

Classification Model:	No Ratings Augmentation	Ratings Augmentation
Naive Bayes ML - BoW	0.8421	0.8421
CNN	0.8373	0.8660

Lessons Learned/Next Steps

- Yet to really implement weakly supervised CNN approach.

- Don't underestimate work involved in data cleansing and engineering.

- Performance is important with large amounts of data - performance optimization takes a lot of time (and so does unoptimized performance)!

- Improve subjectivity model for better sentence precision and recall.

- Further analyze and refine PoS Tag patterns for better precision on extracted features and sentiment (example: "...very **easy to find** on the carousel **with the purple color** the **spinner wheels are perfect**")

Lessons Learned/Next Steps (cont.)

- Binary classification has limited value as it can not differentiate truly negative comments since neutral/irrelevant comments had to be lumped in with truly negative comments.

- 80:20 unbalanced positive to negative labels (likely indicative of this type of data) - probably contributing to overstated accuracy ("Accuracy Paradox")

- "Light weight" and "heavy duty" are often both positive :-)