Review of Large Scale Cross Category Analysis

Alex Chen

Introduction(Challenges)

1. Access to granular detail such as how many and which reviews are being read by customers
    a. Have access to data that effectively measures reviews actually read by consumers
2. **Large number of reviews from a wide range of product categories**
3. Differences in mobile purchases versus traditional method
4. What is the best order to present reviews

Overview of insights

1. Reviews have significantly less impact than expected
    a. 70% of the time, reviews aren’t read
    b. More likely to read reviews for expensive products
    c. As there are more reviews, consumers rely more on summary ratings
    d. Weigh negative reviews more than positive reviews
2. **Aesthetics and price content are the most significant dimensions**
3. **Deep learning models outperform conventional NLP models(27%)**
    a. Category-specific content does not have to be hand-coded
    b. Words to describe aesthetics for a carpet aren’t the same as for a T.V.
    c. Scalability
4. Reviews on mobile are more likely to result in conversion than on PC
5. Reordering reviews based on content can have the same effect as a 1.6% cut

Deep Learning Model

Basics

- Using Convolutional Neural Network with 4 layers
    - Layer 1: Word embeddings of product reviews
    - Layer 2: Convolutional layer
    - Layer 3: Max-over-time pooling layer
    - Layer 4: Outcome: conversion

Layer 1: Word Embeddings of Product Reviews

- Using word2vec embeddings by Google
    - Takes the text and transforms into word vectors while trying to preserve symantec distance between words
    - Stores words as a 300 dimensional vector

Layer 2: Convolution Filter

- Applying filters of length h where the filter is a one dimensional vector
- For example(“The washer is good”)
  - First applies to “The washer”, then “washer is”, etc.
- Trying sizes of 2,3,4, and 5 to try out different size word vectors

Layer 3: Pooling

- Obtaining the most important information across the window positions
- Using max to keep the maximum value in a window
- Running this process multiple times with multiple filters to create different features.(Using a total of 100 filters for each window size)

Layer 4: Output of Conversion

- Appending multiple aspects of the product purchasing together
  - Features extracted from text
  - Consumer characteristics(total product searched, and web-features used)
  - Product characteristics(price, rating, and total reviews)
- All into one model to predict the binary outcome(conversion or no conversion)

Partial Deep Learning Model

Basics

- Using deep learning to extract content features
- Passing through a classical model

$$u_{ijkt} = \overrightarrow{\theta_k}\overrightarrow{Z_{it}} + \overrightarrow{\gamma_k}\overrightarrow{X_{jt}} + \xi_j + \overrightarrow{\beta_k} * \overrightarrow{ReviewFeatures}_{jt} + \varepsilon_{jkt}$$

i = consumer
j = product
k = device used
t = time

Where θ and γ are coefficients of Z, consumer characteristics vector, and X, product characteristics vector, respectively. ξ represents other unobserved product characteristics

Quantifying effect on read reviews

- Volume
  - Number of reviews read
  - Does an increase in the number of reviews read cause a higher conversion?
- Valance
  - Effect of positive and negative reviews
  - Do negative reviews cause a lower conversion?
- **Variety**
  - Content analysis(aesthetics, price, conformance, durability, feature, perceived quality)
  - Which features have the strongest impact on conversion?

Variety

- Use of Amazon Mechanical Turk to provide labels of six dimension of information in reviews
- For example, someone might label “TV looks good but it’s too expensive” as being positive for aesthetics and negative for price
- Use of labeled data to create training and testing datasets and using conventional and deep learning models to perform sentiment analysis

Results Deep Learning model

DV: Conversion	Hit Rate	AUC
Full Deep Learning Model	98.9%	0.924
Partial Deep Learning Model	96.2%	0.873

- Full deep learning model performs better
- Coefficients of Deep Learning model aren't directly interpretable
  - Can find which ones affect conversion the most

	Mobile	PC
Log_Price	-0.787	-0.868
Total # of Reviews	0.004	0.010
Average Rating	0.570	0.840
# Products Searched	-0.135	-0.084
# Used Interactions	0.442	0.256
# Positive Reviews Read	0.039	0.060
# Negative Reviews Read	-0.088	-0.080
Obs	88118	

Results Partial Deep Learning Model

Classifier/Accuracy %	Mean	Aesthetics	Conformance	Durability	Feature	Perceived Quality	Price
SVM + BoW	0.482	0.465	0.607	0.472	0.427	0.618	0.486
NB+ BoW	0.474	0.541	0.473	0.633	0.295	0.588	0.384
Recurrent-LSTM	0.680	0.624	0.622	0.606	0.808	0.626	0.796
Recursive	0.606	0.627	0.597	0.622	0.573	0.602	0.615
Convolutional	0.846	0.854	0.813	0.826	0.868	0.840	0.872

- Deep learning models can outperform conventional machine learning methods by about 27%
- Generally, deep learning models can capture double negative sentences
  - “Without this battery, my phone is useless”

Summary

- Deep learning models perform significantly better than machine learning models
- Deep learning models require a lot less human effort
- The most impactful attributes talked about in reviews are product and aesthetics