Aspect Based Review Analysis

Chattoraj, Ankuran chattoraj.a@northeastern.edu

Introduction

MOTIVATION

Growing Importance of E-commerce market
- Sentiment analysis has been used for analyzing reviews for a long time
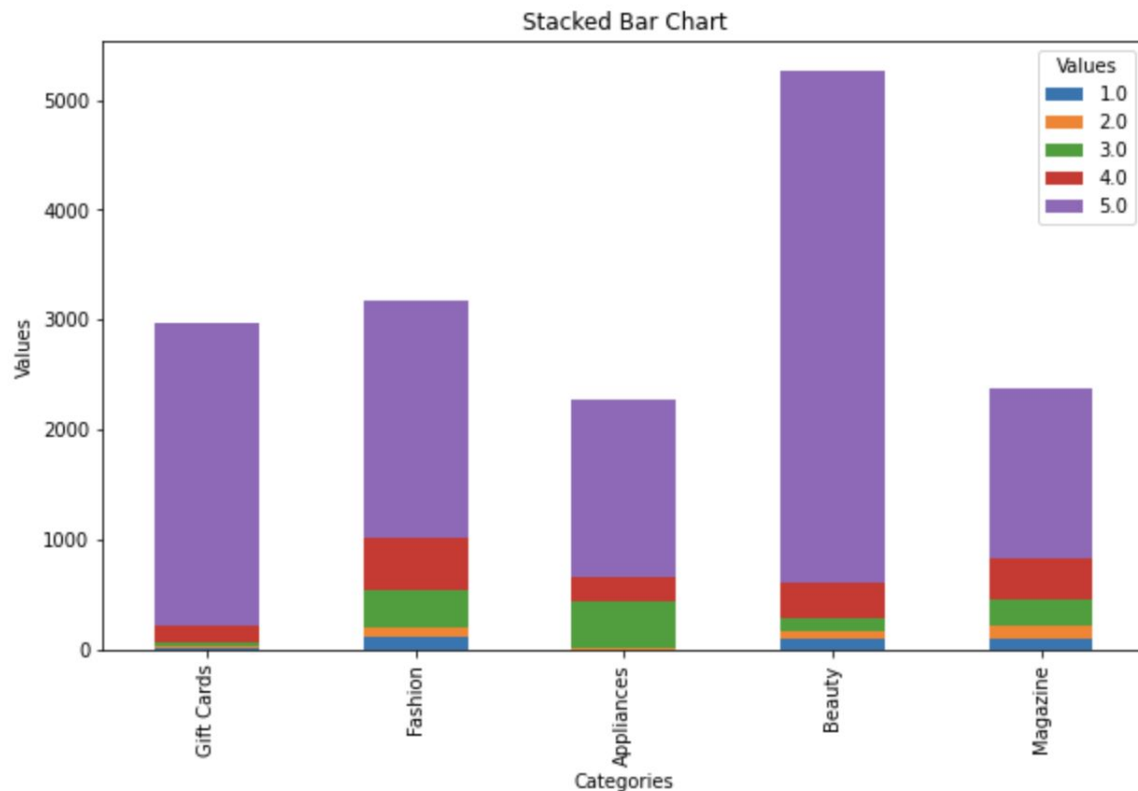
CORPUS

Amazon Product Reviews UCSD
- 1996-2018 data
- 5 Categories

APPROACH

Multistep approach
- Supervised overall sentiment analysis and sentence-based sentiment analysis
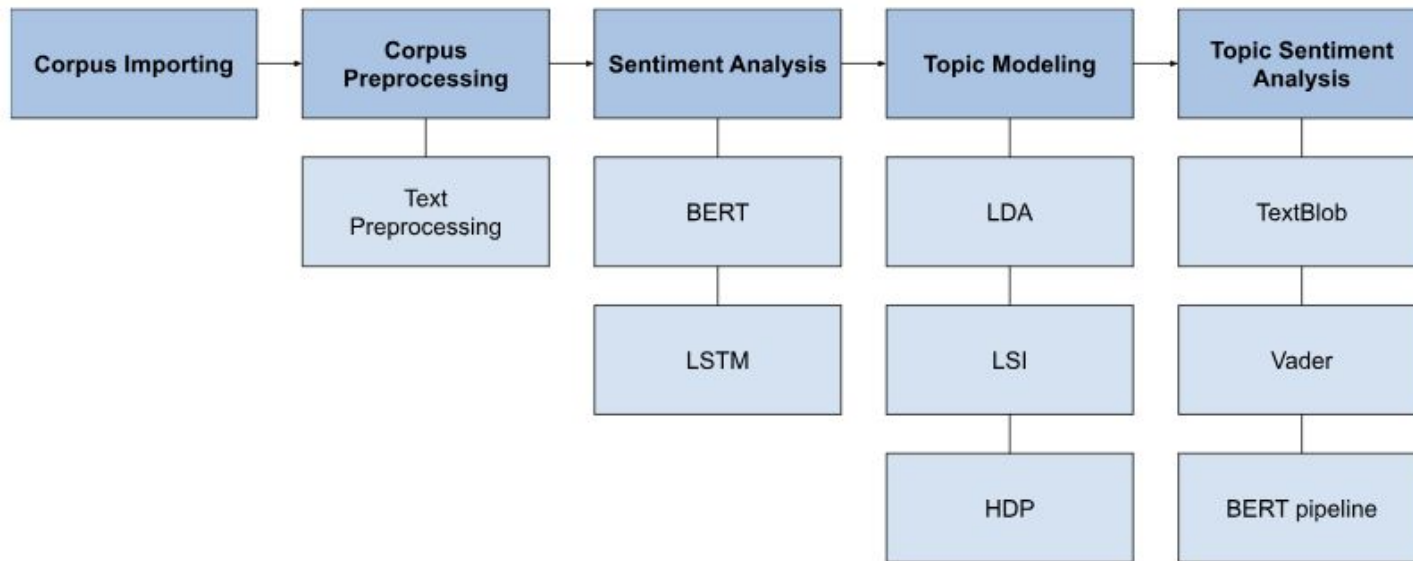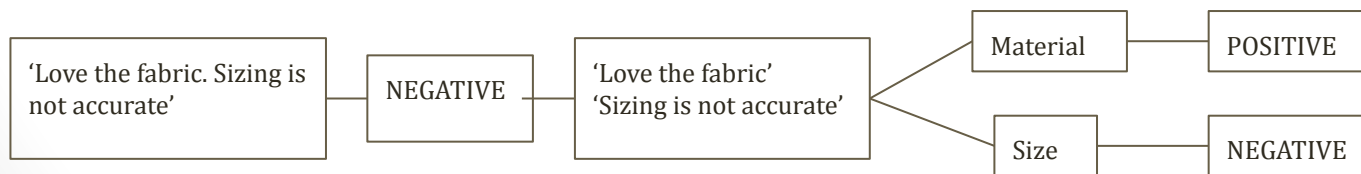
Corpus

Stacked Bar Chart

Preprocessing:

1. JSON conversion
2. Duplicate and empty review removal
3. Punctuation
4. Lowercase
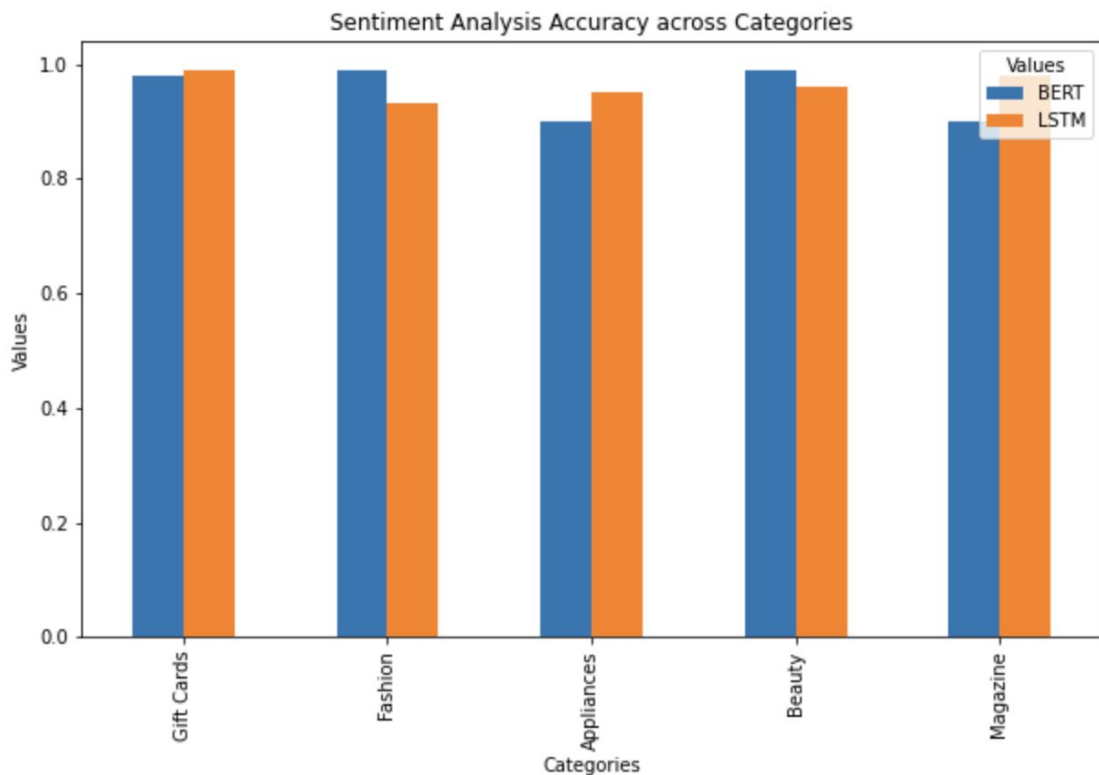5. Stopwords
6. Lemmatization
7. Sentence extraction

Approach

Sentiment Analysis: BERT & LSTM

Model Specifications:

	Epoch	Batch
BERT	2	16
LSTM	5	32

Sentiment Analysis Accuracy across Categories

Topic Modeling

- In-depth examination of customer reviews
- Unsupervised techniques: Latent Dirichlet Allocation, Latent Semantic Indexing, Hierarchical Dirichlet Process

Preprocessing Techniques:

- Bigrams
- Tfidf low value words removal.
- Manually identifying and removing words which don't give any topic information.

Evaluation Metrics:

- Coherence scores measure topic interpretability, crucial for assessing algorithm effectiveness.

Visualization: https://sites.google.com/view/ldavisualization/home

Topic Modeling Results
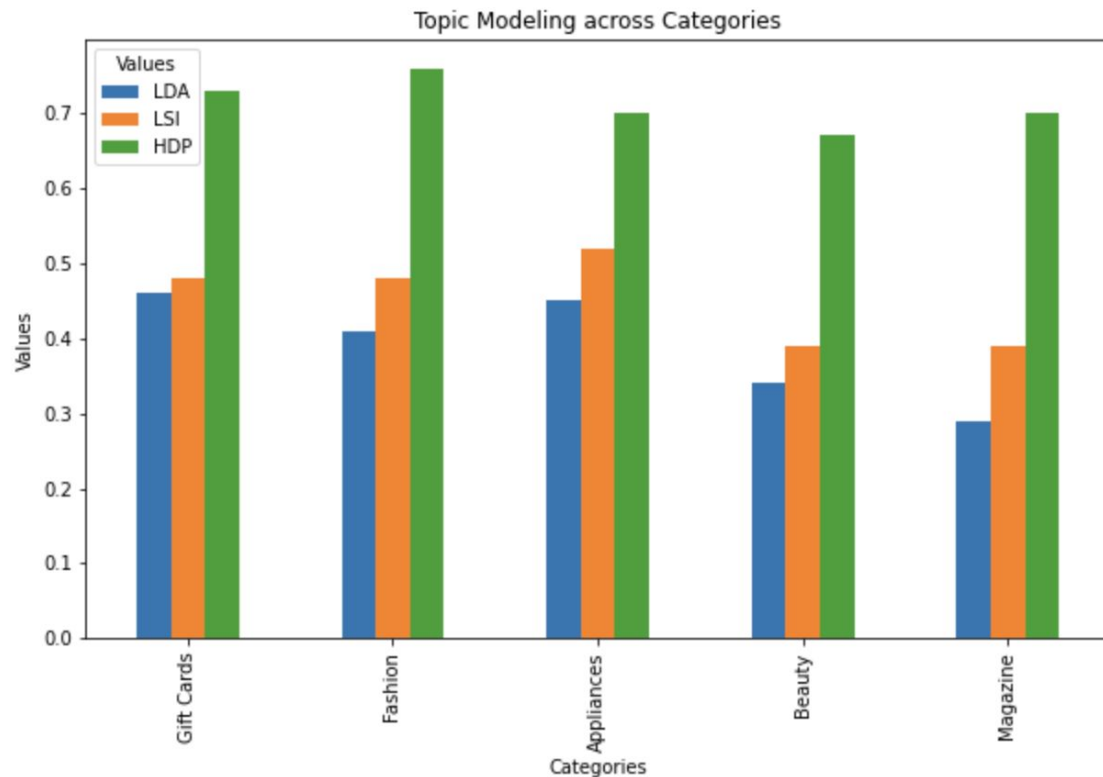
Example of generated topics:

Gift Cards:

1: "Gift Card Purchase",

2: "Gift Card Review and Balance",

Magazine Subscription:

1: "Recipes and Cooking",

2:"Money, Home, and Information",

Beauty:

1: "Nail Polish and Soap Review",

2: "Hair Product and Toothpaste",

Conclusion

Advantages

- ~90% sentiment analysis on full text reviews

- ~8 topics extraction in each category

- Unsupervised sentiment analysis on sentence-based corpus

Further Research

- Class imbalance

- Bag-of-words

- Sentence based topic modeling

- More categories

Example

overall	reviewerID	Product id	sentence	Highest_Similarity	Highest_Similarity_Topic	Sentiment	Sentiment_Category
5	A2G90R2ZU6KU5D	B00006L9LC	Got this shampoo as a solution for my wife's dandruff problem.	0.8525694	Scented Soap and Hair Care	0	Neutral
5	A2G90R2ZU6KU5D	B00006L9LC	She got rid of any dandruff signs after 3-4 uses.	0.77942353	Scented Soap and Hair Care	0	Neutral
5	A2G90R2ZU6KU5D	B00006L9LC	As long as she is happy, I am happy too!	0.56335485	Soap, Bar, and Fragrance	0.51666667	Positive
3	A2NEJIX5NIH2ZZ	B000YFSR5G	They fit great.	0.111194819	Running Shoes Comfort and Size	0.6	Positive
3	A2NEJIX5NIH2ZZ	B000YFSR5G	But they fade bad	0.37034148	Running, Toe, and Color	-0.7	Negative

Beauty
Fashion