Semantic Topic Analysis: Amazon Product Reviews

By Clifford Cheng
Galvanize SFO 33

Luxury Beauty

Agenda

Supervised ML

Semi-Supervised ML
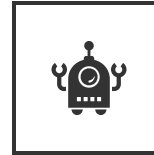
Data Overview

NLP: High | Low Rating Prediction

NMF: Semantic Product Analysis

Business Use Cases

Data Overview

Data Overview

Overall Rating

Upvotes

558,924 Reviews
Weighted Reviews Based on sqrt of Upvotes

Review Text

Product ID Number
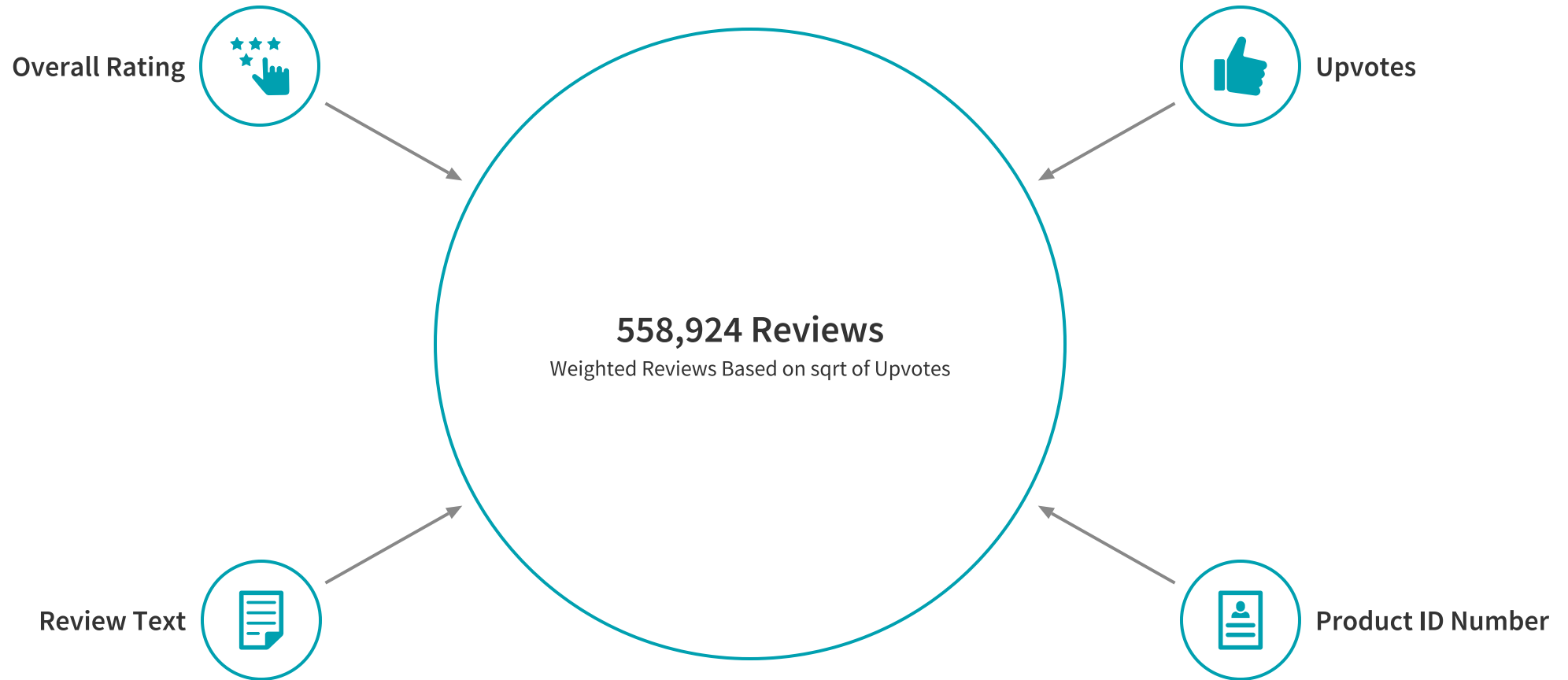
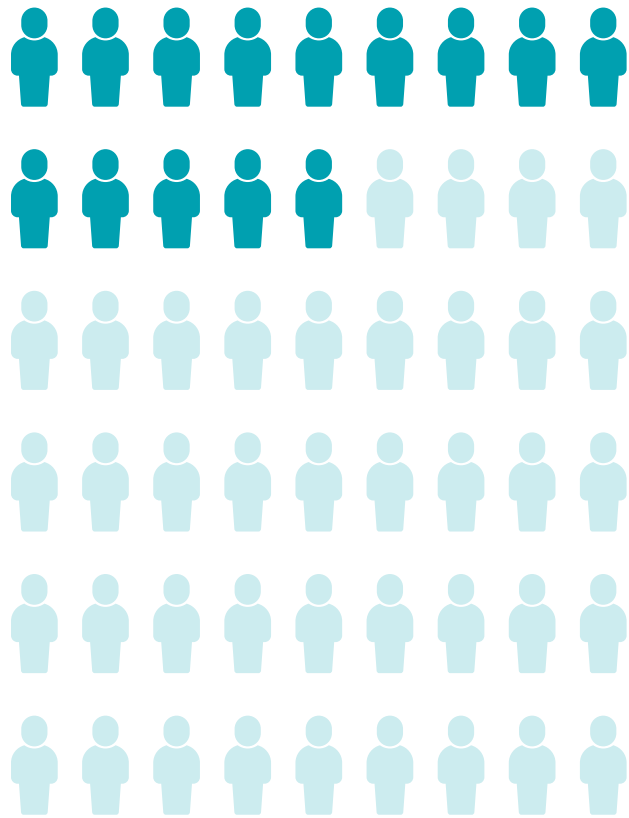# NLP: High | Low Rating Prediction

Ratings Distribution (558,924 Reviews)

25%

75%
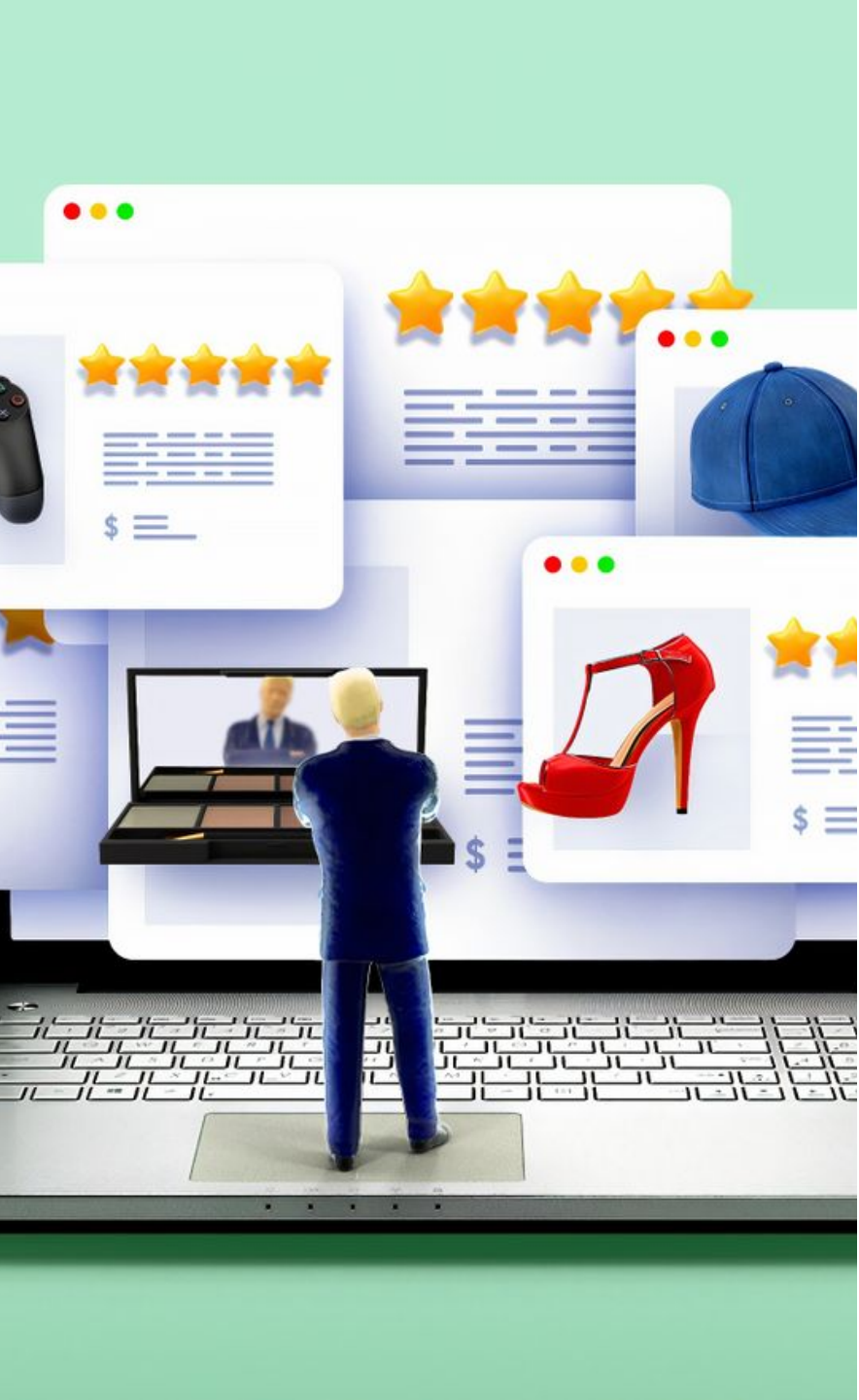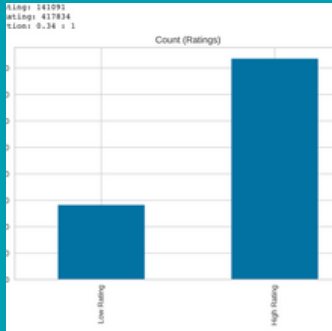
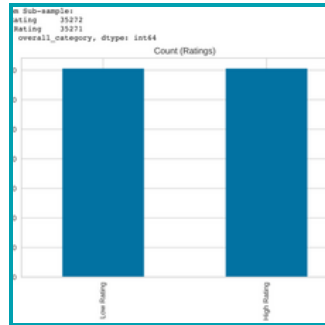1-3 Stars (Low Rating)

4+ Stars (High Rating)

NLP: High | Low Rating Prediction
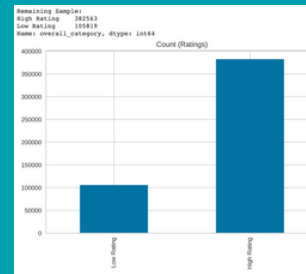

100%

Data
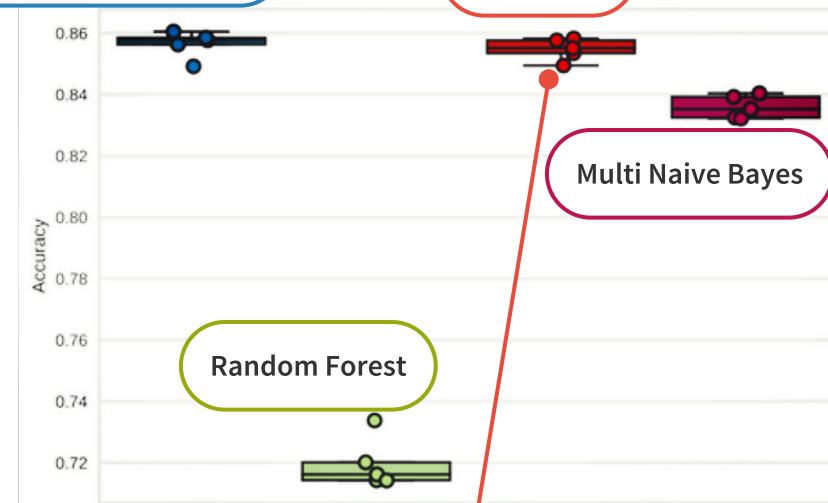
25%

Subsample

75%

Remaining Sample

Logistic Regression

Linear SVC

Multi Naive Bayes

Random Forest

Remaining Sample (f1-score):
High Rating: 91%
Low Rating: 73%

NMF: Semantic Product Analysis

High Ratings (417,834 Reviews)

25%

75%
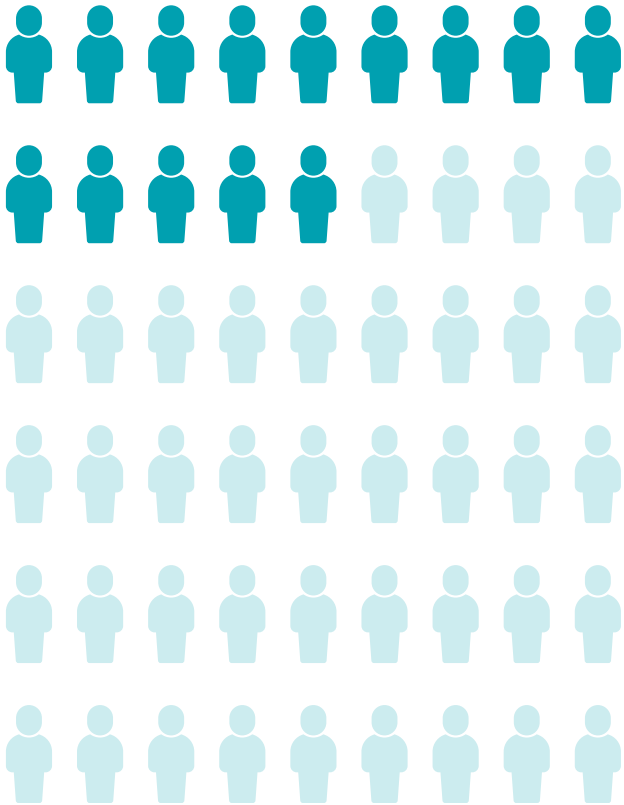
1-3 Stars

4+ Stars

NMF Components: High Rating Groupings

① **Product Feels Great**

['stuff great', 'great good', 'feels', 'great stuff', 'great works', 'great price', 'feels great', 'great great', 'works great', 'great']

② **Best Product**

['owned', 'best far', 'best best', 'best price', 'best stuff', 'best market', 'far best', 'far', 'market', 'best']

③ **Excellent Quality**

['excellent works', 'excellent excellent', 'excellent quality', 'excellent service', 'delivery', 'quality', 'excellent great', 'service', 'excellent price', 'excellent']

④ **Amazing Product**

['great amazing', 'absolutely amazing', 'absolutely', 'feels', 'amazing great', 'amazing works', 'feels amazing', 'works amazing', 'stuff amazing', 'amazing']

⑤ **Family Loves Product**

['loved', 'daughter loves', 'happy', 'daughter', 'gift', 'husband loves', 'husband', 'wife loves', 'wife', 'loves']

⑥ **Product Highly Recommended**

['recommend highly', 'great recommend', 'definitely', 'highly recommended', 'recommended', 'definitely recommend', 'great highly', 'highly recommend', 'highly', 'recommend']

⑦ **Great Customer Service**

['fast shipping', 'price great', 'great great', 'shipping', 'worth price', 'worth', 'fast', 'good price', 'great price', 'price']

TSNE Projection of 5000 Documents

Great Customer Service
['shipping', 'great awesome', 'awesome stuff', 'works awesome', 'fast']

Family Loves Product
['daughter loves', 'happy', 'daughter', 'gift', 'husband loves', 'husband', 'wife loves', 'wife', 'loves']

Favorite Product
['favorite great', 'absolute', 'absolute favorite', 'far favorite', 'far', 'favorite']

Excellent Product
['excellent works', 'excellent excellent', 'excellent quality', 'excellent service', 'delivery', 'quality', 'excellent great', 'service', 'excellent price', 'excellent']

Best Product
['owned', 'best far', 'best best', 'best price', 'best stuff', 'best market', 'far best', 'far', 'market', 'best']

Amazing Product
Best Product
Excellent Product
Excellent Quality
Family Loves Product
Favorite Product
Feels Wonderful
Great Customer Service
Product Feels Great
Product Highly Recommended
Product Is Worth It
Product Works Great

Low Ratings (141,091 Reviews)

25%

75%

1-3 Stars

4+ Stars

NMF Components: Low Rating Groupings

1 **Decent Product**

['ok great', 'expensive', 'ok good', 'ok doesn', 'ok didn', 'ok better', 'better', 'works ok', 'works', 'ok']

2 **Product Does Not Work**

['didn results', 'care', 'didn care', 'notice', 'help', 'didn help', 'didn difference', 'didn notice', 'didn work', 'didn']

3 **Complete Waste of Money**

['absolutely', 'worth money', 'complete', 'save', 'save money', 'complete waste', 'work waste', 'waste money', 'money', 'waste']

4 **Product Does Not Last**

['doesn cover', 'great doesn', 'good doesn', 'doesn good', 'hold', 'doesn hold', 'stay', 'doesn stay', 'doesn work', 'doesn']

5 **Product Broke**

['leaked', 'opened', 'seal broken', 'package', 'return', 'seal', 'item broken', 'item', 'box', 'broken']

6 **Product Seems Fake**

['believe fake', 'item', 'thing', 'label', 'authentic', 'beware', 'fake fake', 'original', 'different', 'fake']

7 **Expected More From Product**

['wasn expected', 'work expected', 'smaller', 'lighter', 'better', 'expected better', 'lighter expected', 'smaller expected', 'good expected', 'expected']

8 **Extremely Disappointed**

['package', 'extremely', 'extremely disappointed', 'original', 'opened', 'work disappointed', 'return', 'box', 'item', 'disappointed']
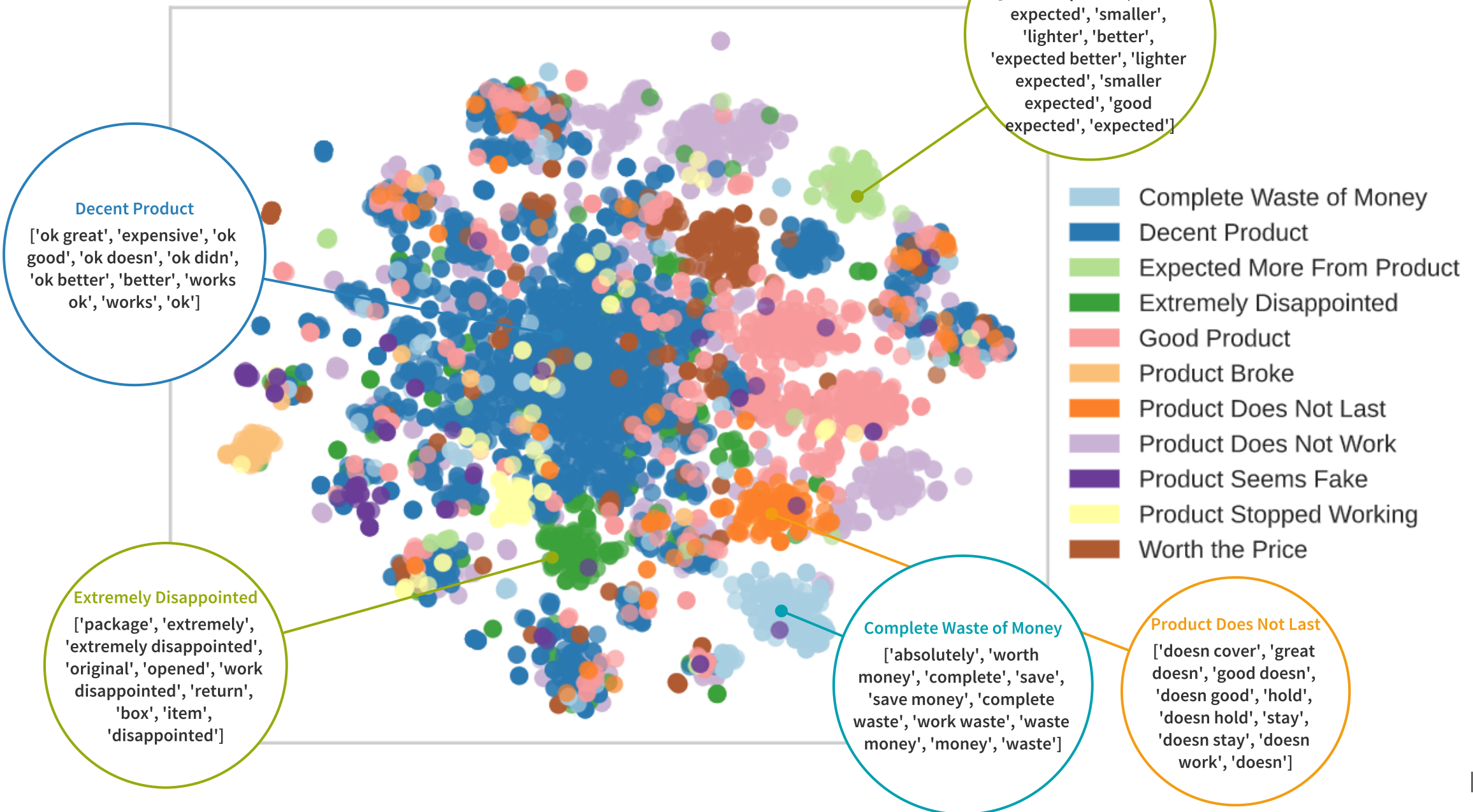
TSNE Projection of 5000 Documents

Expected More From Product

['wasn expected', 'work expected', 'smaller', 'lighter', 'better', 'expected better', 'lighter expected', 'smaller expected', 'good expected', 'expected']

Decent Product

['ok great', 'expensive', 'ok good', 'ok doesn', 'ok didn', 'ok better', 'better', 'works ok', 'works', 'ok']

Extremely Disappointed

['package', 'extremely', 'extremely disappointed', 'original', 'opened', 'work disappointed', 'return', 'box', 'item', 'disappointed']

Complete Waste of Money

['absolutely', 'worth money', 'complete', 'save', 'save money', 'complete waste', 'work waste', 'waste money', 'money', 'waste']

Product Does Not Last

['doesn cover', 'great doesn', 'good doesn', 'doesn good', 'hold', 'doesn hold', 'stay', 'doesn stay', 'doesn work', 'doesn']

Complete Waste of Money
Decent Product
Expected More From Product
Extremely Disappointed
Good Product
Product Broke
Product Does Not Last
Product Does Not Work
Product Seems Fake
Product Stopped Working
Worth the Price

MADE WITH
beautiful.ai

Business Use Case

NLP: High | Low Rating Prediction

"**I love this... and more importantly they have beautiful colors. I'm getting more. I bought 2 previous to this but they weren't delivered so I sent this one to my office and voila!**"

Classification: **High Rating**

"**I do not understand the positive reviews... so clumpy and streaky and weird. its the least nice... and one of the most expensive.**"

Classification: **Low Rating**

- Sentiment Analysis for sites without ratings such as twitter

Product Analysis

1 "The polish does not cover well... not matter what you do. Don't Spend your money on this."

Classification: **Complete Waste of Money**

2 "I had high hopes for this product. Unfortunately..."

Classification: **Product Does Not Work**

3 "I like the dark blue color. However even with a base coat it left a dark stain on my nails..."
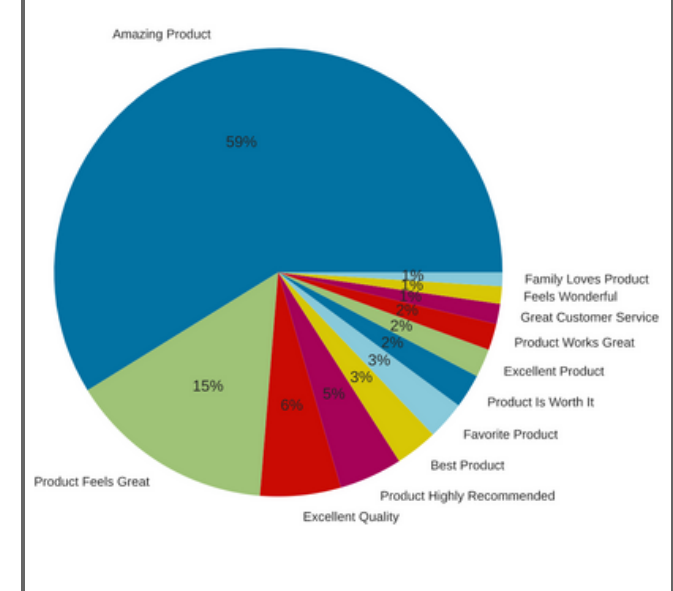
Classification: **Decent Product**

4 "Sorta thin but works as expected for the price"

Classification: **Product Is Worth It**
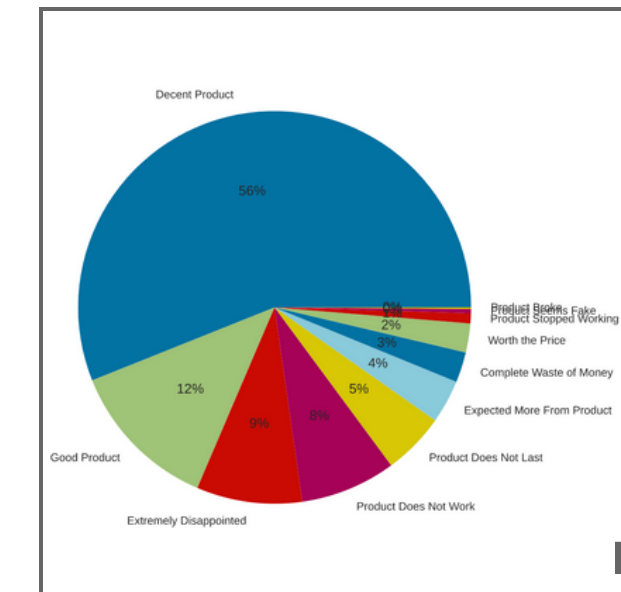
5 "wife loves this brand"

Classification: **Family Loves Product**

CND Vinylux Weekly Nail Polish .5 oz

4.1 Stars

$10.50

Thank you!

Clifford Cheng
SF 33