Topic Modeling Project Report for Timberland Products on Amazon

Ramya Prabhakar

Introduction: For the Topic Modeling project, I chose the brand **Timberland** since we ordered a pair of Timberland sandals for my son, Praniil, last year and he loved it so much that we ordered two more pairs this year!

Topics: The topics are listed below followed by a word cloud of the topics.

```
0: heel like kept b003jy80si
1: 34 slippers fit husband
2: wide width eee narrow
3: great slippers fit product
4: recommend would slippers highly
5: slipper comfortable great nice
6: size larger ordered small
7: years slippers last worn
8: little snug slippers bit
9: perfect fit happy husband
10: pair slippers comfortable nice
11: warm feet slippers comfortable
12: well made slippers comfortable
13: shoes house great comfortable
14: small run size slippers
15: 12 13 size fit
16: good comfortable fit look
17: quality high price excellent
18: looks classy slipper great
19: pair second years first
20: slippers leather comfortable product
21: gift great loves christmas
22: terry evans well atlin
23: best slippers ever owned
24: loves husband slippers wears
```

Topics with brief description

Topic 0: Most reviews about the slippers; heel area concerns but kept the toes warm.

Topic 1: Most reviews about slippers and how they fit their husband well. Interesting to note the special character &#34 appears in almost every review and hence the number 34 appears in the topic!

Topic 2: Reviews related to wide (EEE) footwear. Some people liked the fit but several mentioned they weren't as wide as they were supposed to be.

Topic 3: Hideaway Moccasin slippers. Not a single negative comment. This product fit well and kept all buyers happy!

Topic 4: These slippers kept most buyers so happy that they would recommend them! Also bought on special occasions like Christmas, Father's day, birthdays etc.

Topic 5: Great value for money, very comfortable slipper, most buyers found them nice

Topic 6: Referred to a product that ran almost 1 size small and most buyers recommended buying a size or a size and half larger. Looks like the feedback was taken seriously and the product is no longer available on Amazon!

Topic 7: Most buyers used this product for years and ordered a replacement once they were worn out. Some of the buyers noted a decline in the quality over the years.

Topic 8: The slippers fit a bit snug at first but then fit better with time

Topic 9: All praises for these slippers that fit their husband perfectly and made them happy!

Topic 10: Another product that fit well and extremely comfortable. A couple of negative reviews made it in here.

Topic 11: These slippers are very comfortable and also keep the feet rather warm

Topic 12: Very well made, comfortable slippers

Topic 13: Great shoes to wear around the house and as with the other products, very comfortable

Topic 14: These slippers run a size or two small and users suggest sizing up

Topic 15: These reviews mainly address concerns regarding the fit of sizes 12 and 13. People felt size 12 wasn't true size 12 and size 13 fit them better. It was mainly written by larger shoe size buyers.

Topic 16: Good shoes that fit well, comfortable to wear and look great. Some reviews also mentioned a good value for the price

Topic 17: Excellent slippers of high quality and the pricing seems perfect

Topic 18: Great slippers that look classy. More happy reviews!

Topic 19: This is an interesting topic as it contains reviews that compare and contrast various products that were bought first and second times. While most reviews tended towards being positive, some of the reviews noted a decline in quality.

Topic 20: This one seems to be a misnomer! Most reviews are about how these slippers/shoes are not as comfortable as they expected them to be.

Topic 21: Great Christmas gift loved by recipients.

Topic 22: Great pair of terry atlin slippers that kept all reviewers happy

Topic 23: Best slippers that they ever owned

Topic 24: Husbands love wearing these slippers and have been very happy with them

Preprocessing and Model Steps

1) Extract the asins for the brand Timberland using the first data set

2) Extract the reviews for all the asins extracted in step 2.

3) Running k-means clustering for 50 clusters after extracting the Timberland reviews and the asins related to the same revealed a list of topics that seemed rather similar. The clusters were visualized using LDA and based on the same, the cluster values were changed a couple of times and having 25 clusters seemed to give distinct topics that adhere to the thumb rule of clustering, which is to ensure homogeneity within and heterogeneity between clusters.

4) Topic Classification of all the reviews into the final 25 topics derived in step 3.

5) Investigate non - English reviews

6) Perform sentiment analysis.

Marketing and Product Insights:

1. Attributes that people like about our products

Most topics allude to the fact that the products are high on comfort and have loyal customers that return for the same product regularly.

2. Attributes that people do not like about our products

Topic 2 specifically highlights the issues with the wide slippers, EEE in particular. Explains how it is not as big as it is supposed to be. Interesting to note that the product with the most negative reviews is not available anymore. This shows great response from the makers.

Topic 20 mentions that slippers/shoes are not as comfortable as they expected them to be. It would be interesting to dig deeper to understand why this might have occurred and take corrective actions as required.

3. Purchase occasions

Topics 4 and 21 talk about great gifting products that were well received. So this brand might do well to have some special promotions during the holiday seasons.

4. Product Improvement Ideas

Topic 19 contains concerns about the decline in quality over the years and is something that the wri

A couple of topics express concerns over sizing of specific sizes such as size 12, 13. This is another area of improvement.

5. Pricing Suggestions

Topic 16 and 17 suggest that customers are happy with the quality of the products especially at the given price points. This is good feedback and the marketing team could use this fact to their advantage

6. What's not in the data

It was surprising to note that there were not many mentions of kids' footwear in the reviews. Either the parents were too busy to write them or weren't buying them that often.
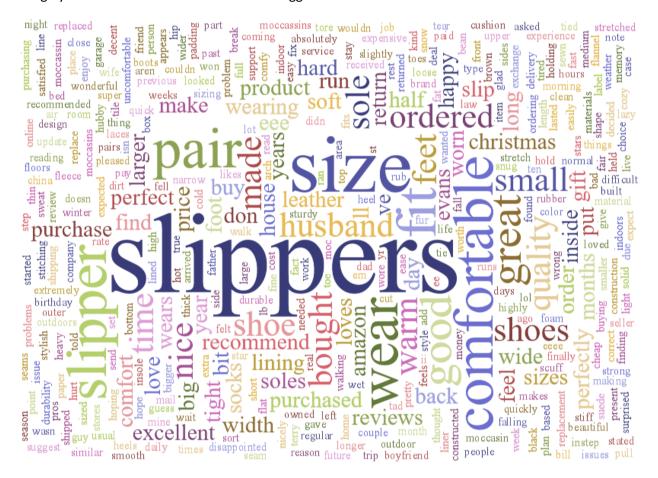
7. Two points that stood out

The first point that stood out for me was the fact that most reviewers were females that had purchased the products for their husbands or their fathers or their sons. Made me wonder if men did not buy these products often enough, or they just did not write reviews as often as women did. This might be a worthwhile point to consider in the marketing strategy, since segmentation would typically target men for men's footwear.

The second one is the fact that most of the products that were mentioned in the negative reviews are no longer available on Amazon.

Diving further into the data: It is interesting to note that most of the topics sound positive with very few complaints.

This triggered the interest in running some sort of sentiment analysis on the reviews. Tried textblob and then worked on Vader to arrive at the compound sentiment scores that revealed the fact that most of the reviews were positive with just over 600 negative reviews out of a total of 28325 reviews, which is roughly around 2%. The word cloud below suggests the same.

Additionally, conversion of comments from languages other than English was attempted using Googletrans. However, ran into issues with the same and just focused on the sentiment analysis.

Future work: It would be good to classify the negative reviews and identify problem areas to be addressed from a product manager's perspective.