

NLP/Unsupervised Learning

Project Write-up

May 28, 2021

Abstract

Athletic apparel retailer Lululemon will be entering the athletic shoe market in 2022. Though not an exhaustive list, Nike, Adidas, Under Armour, New Balance, and Asics will be some of their primary competitors in this segment. My project looks at Amazon reviews for these competitor's shoes, with similar expected purpose, and identify topics of most interest to consumers in order to develop a focus and potential marketing campaign. As a majority of Lululemon's client base is female, the focus will only be on women's shoes in this analysis.

Data

Over 16,000 customer reviews were scraped from Amazon.com using Google's web scraper. The file Shoes_NLP.csv created in Excel is resulting file read into python for analysis.

Tools

Exploratory data analysis – Excel

NLP algorithms – Non-negative Matrix Factorization (NMF) was used.

Visualizations – Matplotlib, Excel, and Wordcloud

Presentation – PowerPoint

Communication

I will be presenting results to the NLP / Unsupervised Learning cohort on Friday, May 28, 2021. A PDF of the presentation, NLP_Lulu.pdf, is also included on my personal GitHub repository.