



# Amazon Products Sentiment Analysis

Ahsun Asad, Anisha Kishore, Khadija King

# Background

- In today's age of social media, companies receive an overwhelming amount of information
- One major use case is customer product reviews: a large company can easily receive millions of product reviews in total
- These reviews can impact how a company prioritizes how to respond to customer complaints and analyze customer feedback
- Sentiment Analysis is the process of identifying and categorizing opinions to determine whether the user has a positive, neutral, or negative opinion based on their language compared to their ratings system



# Database

- As of 2022, Amazon had over 250 million product reviews
- We used an Amazon product database that specifically looked at various Amazon products such as the Amazon Kindle, Echo, Fire Tablet, etc.
- The database included user reviews of these products rated 1-5 stars
- Our goal was to parse through the reviews, in order to create our Sentiment Analysis model
- Sample reviews we parsed through are below

0,3,I thought it would be as big as small paper but turn out to be just like my palm. I think it is too small to read  
,5,This kindle is light and easy to use especially at the beach!!!,Great light reader. Easy to use at the beach  
,4,"Didnt know how much i'd use a kindle so went for the lower end. im happy with it, even if its a little dark",Great  
,5,"I am 100 happy with my purchase. I caught it on sale at a really good price. I am normally a real book person, but  
,5,Solid entry level Kindle. Great for kids. Gifted for a kid of my friend and they love to use it to read more than t  
0,5,This make an excellent ebook reader. Don't expect much from this device except to read basic ebooks. The good thin  
,5,"I ordered this for my daughter, as I have the black paperwhite, and I love it. I read quite a bit and the larger b  
0,4,I bought my Kindle about 2 months ago and the battery is already dead and will not charge,not good quality  
,5,"amazon kindle is always the best ebook, upgrade every new model",best ebook

# Coding Process

- We standardized the data of Amazon product reviews prior to modeling it:
  - We initially cleaned up the model by removing unnecessary rows, and translating the text into numbers to turn the written reviews into something that could be analyzed by our model in a predictive manner.
  - From there, we padded our sequence for training and testing, which we did after normalizing the features. After this, we initialized our model with two hidden layers, then compiled and trained our model before finally evaluating the model and plotting the accuracy results.
- We looked at various reviews on a 5-star scale and made a judgment on what number of stars would correspond to a specific sentiment:
  - 4 stars or higher is positive
  - 3 stars is neutral
  - 2 stars or lower is negative
- The model would then seek out key words to determine what reviews had a specific sentiment, and predict whether a review was negative/neutral/positive

```
def assign_sentiment(rating):  
    if rating >= 4:  
        return 'positive'  
    elif rating == 3:  
        return 'neutral'  
    else:  
        return 'negative'  
  
df['sentiment'] = df['Reviews_Ratings'].apply(assign_sentiment)  
df['sentiment_num'] = df['sentiment'].map({'positive':1, 'negative':2, 'neutral':0}).values
```



# Predictive Behavior, Optimization, and Performance

- The goal of our model was to have 75% classification accuracy or 0.80 R-squared
- We ended up with 0.05% classification accuracy
- Looking back, part of the problem with our model was that we had been looking at every single word in a review for our sentiment analysis
- For example, the word “thought” could be present in a neutral, negative, and positive review (“I thought it was terrible vs. I thought it was amazing”)
- One way to optimize our model’s performance would be to have a stricter guideline for the key words our model is seeking out for predictive behavior



# Conclusion

- With how prominent social media and online reviews are, Sentiment Analysis plays a key role in determining customer responses when companies have large volumes of data, such as millions of product reviews
- Although our model was limited to reviews of Amazon products, it could also be used across several different companies and scenarios
  - Various retailer product reviews over time
  - Breaking down which customer demographics have strong reactions to certain products
  - Prioritizing customer service issues
- We do plan on further working on the model and improving it over time to increase our classification accuracy and R-squared
- Once improved, the model would prove highly valuable for monitoring social media reactions and analyzing customer feedback