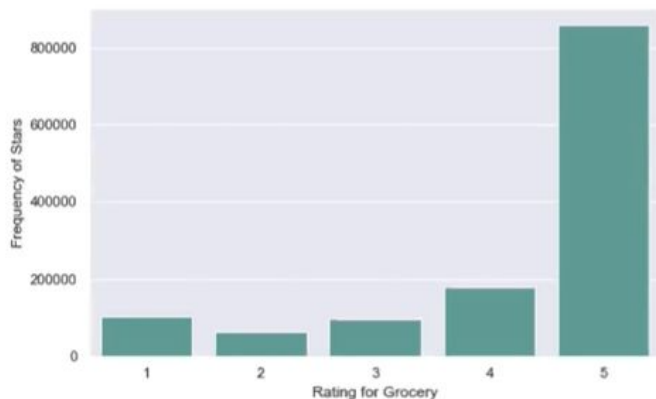

Reviewing the Amazon Review Scale: Sentiment Analysis

Team 1: Zhiqing Li, Zhourong Li, Pruthvi Reddy,
Shreyans Kothari

Introduction

Zhiqing Li



review_headline	rating	review_text
Storage on Wheels Cart	3	The cart is fine and works for the purpose for which I bought it. (Farmers's Markets, etc) but it stinks like hell. Even after having it in the open air for sometime, it still smells. I made the mistake of putting it in my car and now I can't get the smell out. Other than that it's fine for the price.
Tastes a Bit like Cough Syrup	3	My wife picked some of this up on sale. I usually drink Crystal Light so she though this would be a decent alternative. The taste isn't the greatest. If all you are looking for is a caffeine fix then this would work. But the cough syrup flavor wasn't my favorite.
Interesting bitters	3	I had a martini at a local distillery that used a bit of wormwood bitters. I bought these wanting to recreate the drink at home. These aren't quite the same. They have a bit more of an anise flavor than the other bitters did. That being said, I like them!
Good, but too small of a bag!	3	It was the correct flavored product, but it was in a tiny bag. I did convert the size from metric prior to buying it, but it was still a lot smaller than expected. Good if you have a craving for sour starburst that you can't find anywhere else. But it would be a bag worth \$2-3 to buy at the store in the US (if they carried it!). So the rare flavor has a cost I guess!
A very faint green tea taste.	3	These are excellent kit-kats, don't get me wrong. However, I guess I was just expecting a stronger green tea taste. It's very faint.

review_headline	rating	review_text
Decent flavor, value...meh.	4	I bought these on sale (2 for 4) at my local supermarket, so at 3 these are probably overpriced. The price is decent when comparing to MIO, however. The flavor I bought was actually cherry limeade, which tastes similar to the Great Value powder by the same description and the taste is the same. As far as "fantasy fruit", web reviews say it has a pretty indistinct fruit flavor but it tastes good. I can't comment much on the caffeine content since I'm fairly immune to caffeine at this point. However, if you're stepping down from caffeinated diet cola to flavored water, as I am, you're probably going to want something like this to ease the transition. This fits the bill well. My main gripe has to do with value. I assumed, since this was billed as "24 servings" that it would actually flavor 24 x 16.9-ounce bottles of water - but they must be referring to a different amount of water because at that rate, you would NOT be able to taste much of anything besides water. At best, you might get 10-12 uses out of a bottle if you want actual flavor. Compare that to Great Value powder, which sells for \$2 per 10-pack. Both are similar in price but I find that the powder can flavor two bottles of water. The liquid enhancer requires a large amount to get any actual flavor. Since there is no precise measurement, it is hard to tell when you've used too little or too much. The one advantage liquid flavoring has over powder is that it doesn't fly everywhere in a stiff breeze. When you get the red powder on your clothes (or anything porous), it is very difficult to get out. So if you're camping or running a fan, the liquid wins.
They're dice!	4	And they're pretty nice! One set looks like ice, and you get a set of four thrice. Kind of overpriced.
overwhelming smoky flavor	4	I use like a 1/3 of a teaspoon in a spicy cup of soup I make at times and the smoky flavor is so overwhelming that it's the only thing I end up tasting. And mind you, this is coming from a person that really adores smoky flavors... but really, the smokiness is just too much and makes every other flavor practically disappear from your soup/dish.
Great heat!	4	I haven't used this in a recipe for wings yet, but I have added to a beans and rice meal and it added a great heat! Once I make some wings, I'll add my findings. If it is as good as I am hoping, I'll change the review to 5 stars.
burning ring of fire!	4	Yes, this stuff is hotter than hell, literally hotter than hell. If you want to win the spiciest chili contest use a tablespoon of this. The flavor itself

Hypothesis



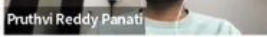
Hypothesis 1

Despite having a 5-point scale for rating, the Amazon rating system is a binary rating system: 5-stars are associated with positive reviews, and stars less than 5 are associated with negative reviews.

Hypothesis 2

The Amazon reviews with 3-4 stars would have a negative sentiment value attached to it.

Sentiment Analysis



- NLP technique
- Detects polarity: classifies text into positive / negative categories
- Obtain insights from linguistic data
- Aims to measure the attitude, sentiments, evaluations, and emotions of a speaker/writer based on the computational treatment of subjectivity in a text.
- It is also very useful in subjectively gauging Brand Sentiment, Voice of Customer, Monitoring Social Media and Customer service etc.

Preprocessing

- Removing URL's.
- Removing HTML tags.
- Chat word conversions.
- Emojis and Emoticons conversion to text.
- Prepared 3 columns - without stopwords, stemmed, lemmatized

Original: I just love this product-- Cafix Coffee Substitute Crystals . It dissolves completely, has a great taste and is caffeine free. What more could one want?

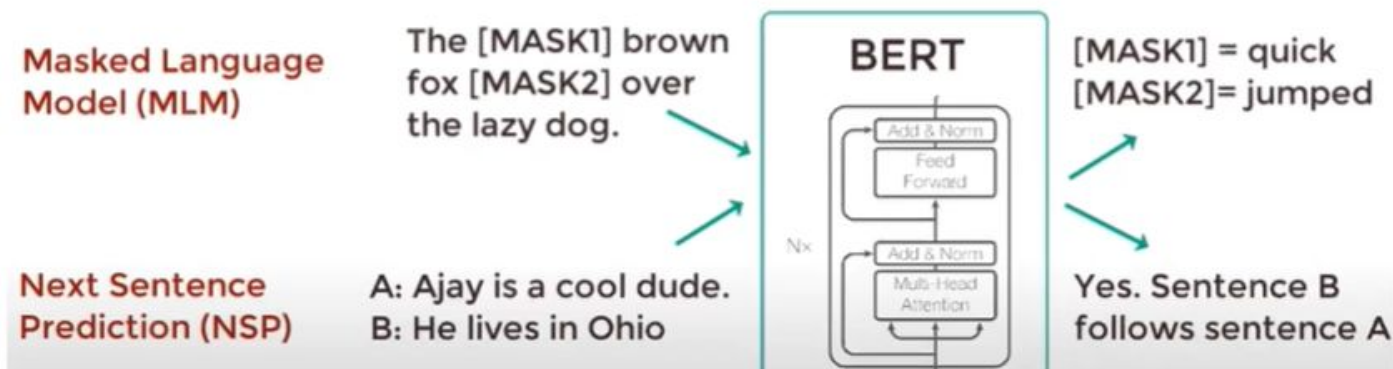
Stopwords: i love product-- cafix coffee substitute crystals . it dissolves completely , great taste caffeine free. what could one want?

Stemmed: i love product-- cafix coffe substitut crystal . it dissolv complet , great tast caffein free. what could one want?

Lemmatized: i love product-- cafix coffee substitute crystal . it dissolve completely , great taste caffeine free. what could one want?

BERT: Sentiment Analysis

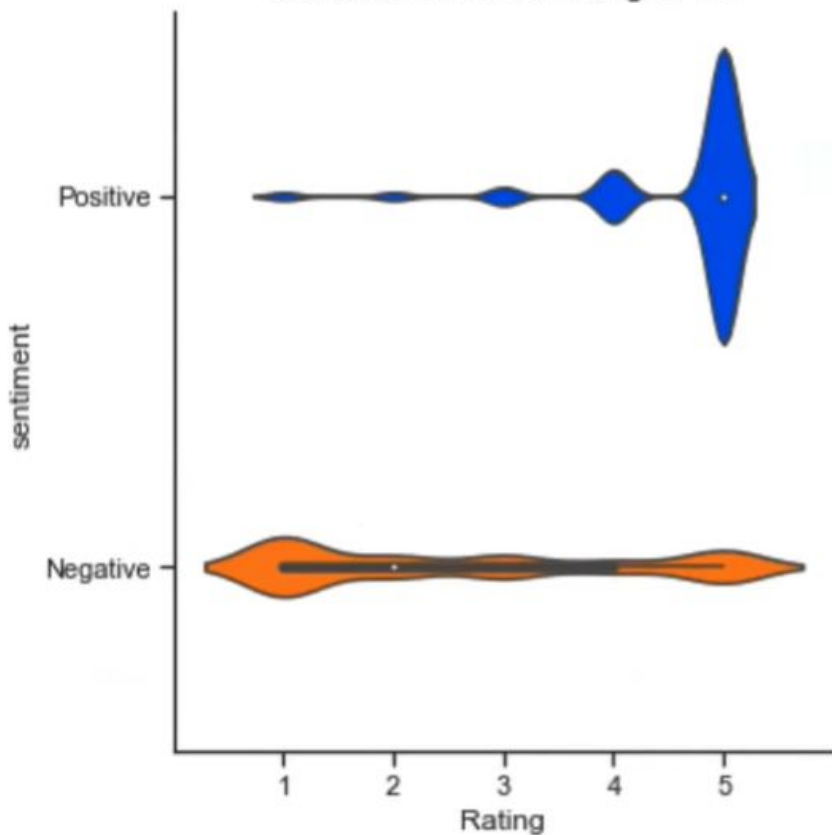
- BERT is a Masked Language Model.
- Uses Bi-directional encoders for transforming the text.
- We used BERT-uncased pretrained model.
- Fine Tuned it with a dataset of classified reviews from IMDB.
- Predicted sentiment for 125k amazon reviews.
- Training Accuracy: 0.99 ; Validation Accuracy: 0.886



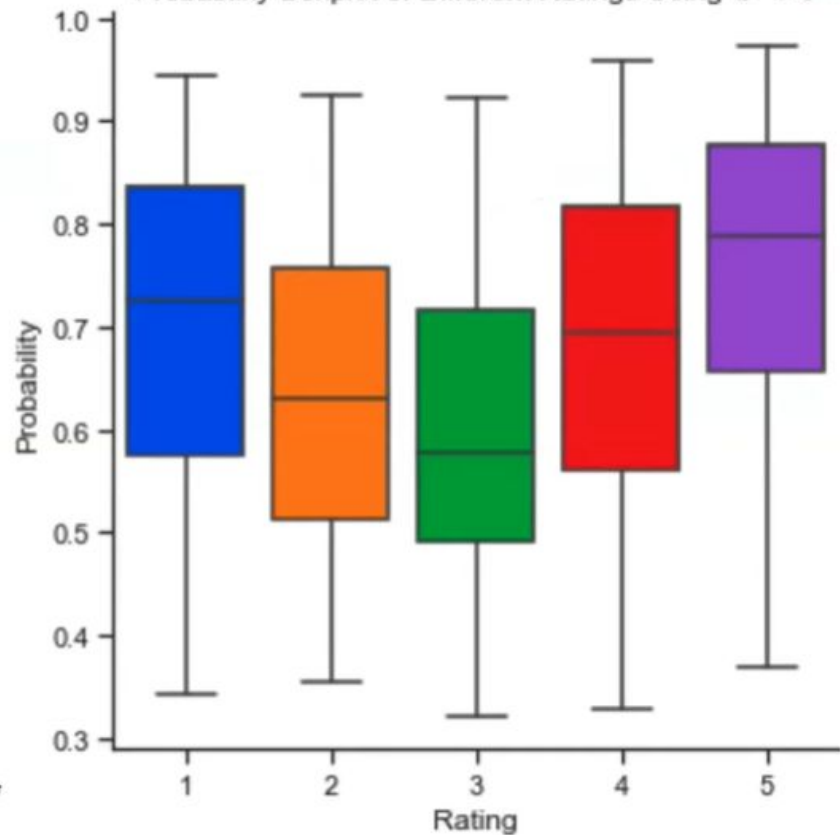
GPT-3: Sentiment Analysis

- Generative Pre-trained Transformer 3(GPT-3) is an autoregressive language model uses deep learning to produce human like text.
- Pre-trained using 175 billion machine learning parameters.
- Demonstrate that a language model trained on enough data can solve NLP tasks that it has never encountered.
- GPT-3 studies the model as a general solution for many downstream jobs **without fine-tuning**.
- Language Translation, Classification, Text Generation, Conversation, etc.

Sentiment Distribution Using GPT-3



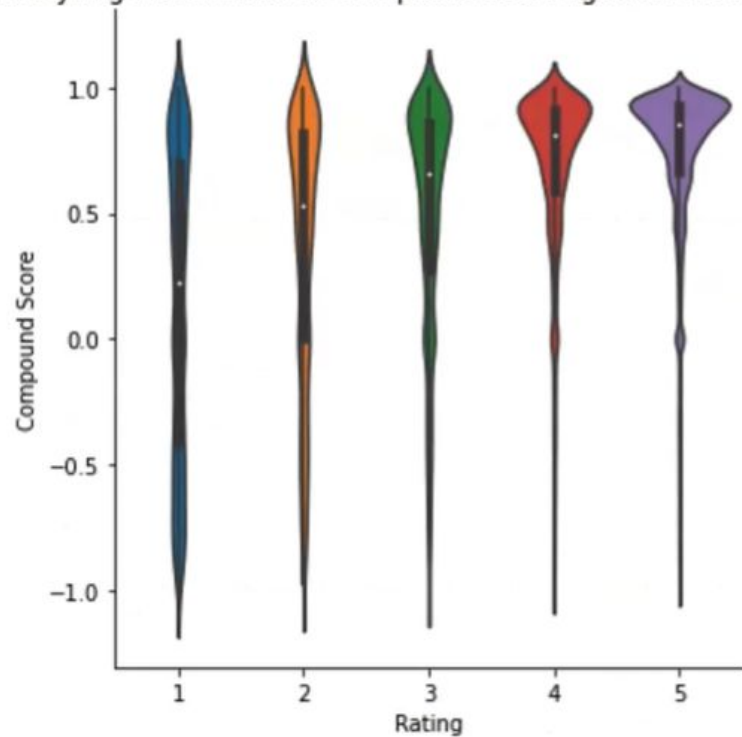
Probability Boxplot of Different Ratings Using GPT-3



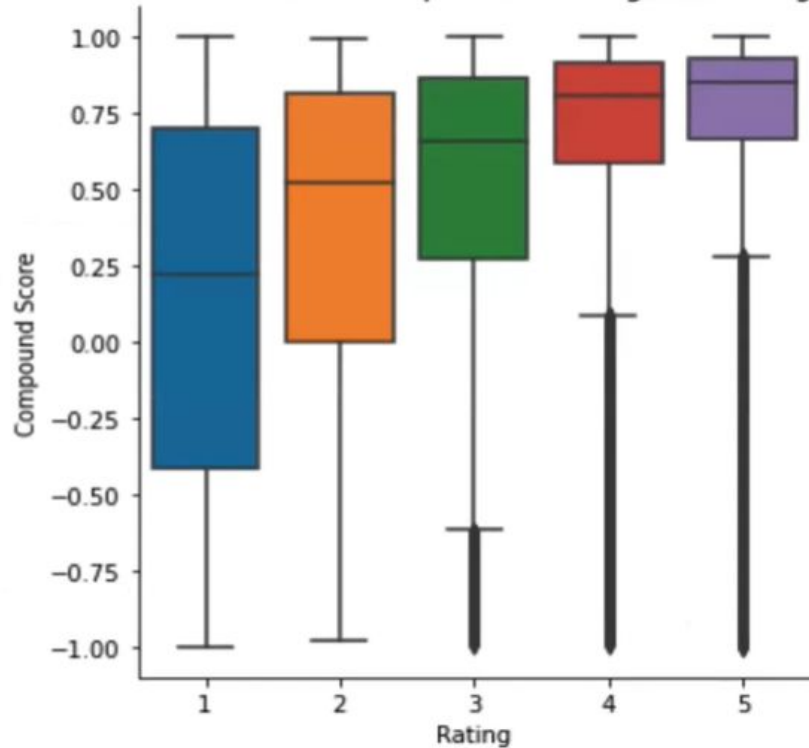
NLTK: Sentiment Analysis

- Pre-trained sentiment analyzer
- VADER (Valence Aware Dictionary and sEntiment Reasoner)
- VADER's `SentimentIntensityAnalyzer()` takes in a string and returns a dictionary of scores in each of four categories:
 - Negative
 - Neutral
 - Positive
 - Compound (computed by normalizing the above scores)
- 125K reviews
- Compound score range: [-1, 1]

Analyzing Product Review Compound Score against Product Rating



Product Review Compound Score against Rating



Analysis / Results

Review (5 Stars)	Model	Score
<p>This is my absolute, undisputed favorite tea right now. I love Darjeeling, but I'm not wildly fond of the lighter, first flush ones for being too delicate. This Darjeeling, especially when steeped a while, has a good tannic bite. It's bright and warm all at the same time, and pretty much explodes with that classic 'Darjeeling' flavor. It's not even remotely delicate, but neither is it hard-edged. It's sort of like good-looking men in bespoke suits -- strong but refined. I use boiling water, steep for 4-5 minutes, and with a large mug use one Splenda and just a tiny splash of milk. Then get out of my way, because with this tea, I can take on the world!</p>	BERT range [0,1]	0.99
	GPT-3	Positive
	NLTK (compound) range [-1,1]	0.975

Analysis / Results

Review (4 Stars)	Model	Score
<p>And they're pretty nice! One set looks like ice, and you get a set of four thrice. Kind of overpriced.</p>	BERT range [0,1]	0.96
	GPT-3	Negative
	NLTK (compound) range [-1,1]	0.8313

Analysis / Results

Review (3 Stars)	Model	Score
<p>The cart is fine and works for the purpose for which I bought it. (Farmers's Markets, etc) but it stinks like hell. Even after having it in the open air for sometime, it still smells. I made the mistake of putting it in my car and now I can't get the smell out. Other than that it's fine for the price.</p>	BERT range [0,1]	0.019
	GPT-3	Negative
	NLTK (compound) range [-1,1]	-0.7992

Analysis / Results

Review (2 Stars)	Model	Score
<p>Not the best chai taste. Pretty grainy. Is OK, at best. Have had better tasting ones.</p>	BERT range [0,1]	0.0039
	GPT-3	Negative
	NLTK (compound) range [-1,1]	0.3176

Hypothesis Conclusion

- Based on our models and analysis we mostly tend to reject both the hypothesis.
- The outcome is conditioned on the model itself, hence any improvement or specificity in the training of models would lead to better outcome.
- The fine-tuned BERT model also had a test accuracy of 0.88 only, other models might be underperforming.
- Reviews also have a specific nature, where people mentions both Pros and Cons and hence the entire text is unique in that sense to classify.