

SUMMARY REPORT FOR TASK 21

Author: Sam Derry

Date: 6th March 2024

Project: Capstone Project - NLP Applications

DESCRIPTION

This report presents a summary of the results of a sentiment analysis. The task involved analysing the sentiment of a sample set of customer reviews for Amazon products. The dataset comprises 28,332 products reviews including basic product information, rating, review text, and more for each product. The analysis required that the dataset review text be cleaned prior to processing.

PREPROCESSING

The dataset was pre-processed by selecting only the product review text, and removing any records with a missing product review. Prior to determining the sentiment for each review, the text was converted to lowercase, stripped of any leading whitespace characters, and cleaned of identifiable stop words to try and reduce the quantity of redundant content from the review.

RESULTS

The sentiment analysis was performed by taking the review text and using the spaCy model to determine the sentiment value, between -1 and 1. This value was categorised as either positive, neutral or negative based on value ranges as per the following table.

Sentiment	Category
> 0.333	Positive
> -0.333 and < 0.333	Neutral
< -0.333	Negative

The analysis comprised searching the dataset for a sample of 10 product reviews in each category, which are tabulated below. Each review was manually evaluated for its sentiment category and compared to the NLP model's sentiment. This was intended to discover the accuracy of the model.

N.B. The analysis only found 9 product reviews with a negative sentiment category.

The 'Observed Sentiment' column contains the manually determined sentiment category, subject to the personal view of the report author. The category is presented in **Green** where the value matches the model's sentiment, and **Red** where there is a mismatch.

POSITIVE REVIEWS

The following reviews had a sentiment value greater than 0.333.

	Review	Observed Sentiment
1	Well they are not Duracell but for the price i am happy.	Positive
2	Seem to work as well as name brand batteries at a much better price	Positive
3	Bought a lot of batteries for Christmas and the AmazonBasics Cell have been good. I haven't noticed a difference between the brand name batteries and the Amazon Basic brand. Just a lot easier to purchase and have arrive at the house and have on hand. Will buy again.	Positive
4	Well if you are looking for cheap non-rechargeable batteries that last quite a while then these are perfect. Nothing more to say.	Positive
5	AmazonBasics AA AAA batteries have done well by me appear to have a good shelf life. I'll buy them again.	Positive
6	I find amazon basics batteries to be equal if not superior to name brand ones. Can't believe I didn't start buying them sooner! The packages are large and the price is great too.	Positive
7	Use it for my fish tank's light at night and works great, I love how you can easily switch it off and on if you want it on while guests are there.	Positive
8	just got em so I can't really comment on how good the do the job, good price, quick delivery but have only put two into one of my keyboards but they can go up to a year so who can say after three days	Neutral
9	we have many things that need aa battery they are great	Positive
10	Thankful that I was able to find on Amazon for a great price and even better shipping. Arrived in perfect condition and did exactly what I needed it to. Great purchase and would purchase again.	Positive

NEUTRAL REVIEWS

The following reviews had a sentiment value greater than or equal to -0.333 and less than or equal to 0.333.

	Review	Observed Sentiment
1	These batteries are very long lasting the price is great.	Positive
2	ive not had any probleme with these batteries have ordered them in the past been very pleased.	Positive
3	These do not hold the amount of high power juice like energizer or duracell, but they are half the price.	Neutral
4	When I first started getting the Amazon basic batteries I really liked them. With recent purchases, they do not seem to last like they had, or maybe a mixed-bag (inconsistent with some lasting better than others). I have not done any tests, but feel some other brands may last longer. However, the price is hard to beat.	Neutral
5	I don't know if I would buy thus brand again seems like they don't last as long as Duracell	Negative
6	In my opinion these did not last anywhere near as long as Duracel in things like LED candles (which is crazy) and trail cameras. Cameras were not exposed to cold temps more or less than other batteries. WE buy in bulk for the north house. Amazon Basics is great for things like sheets and beeding and towels. In my opinion the battery life, in the large package of aaa and aa size we purchased were lacking.	Negative
7	They don't last as long as the brand name but are good enough considering they are much cheaper.	Neutral
8	Bought these batteries for my Christmas gifts the month of (december) only lasted like 2months toys now need replacement batteries . I also used some for my doorbell and just now needs replacement batteries. Tv Remote control is still working but these batteries don't last very long...	Negative
9	They last as long as Duracell batteries in my Xbox one controlllers and none of them exploded like some of the reviews said.	Positive
10	These Amazon batteries did the job although I gave 4star only because I had a few I would say a hand full of batteries that were not as strong or were pretty weak but out of a box of 48 batteries, I will definitely buy again for this priceIm pretty well satisfied.Thank you!	Positive

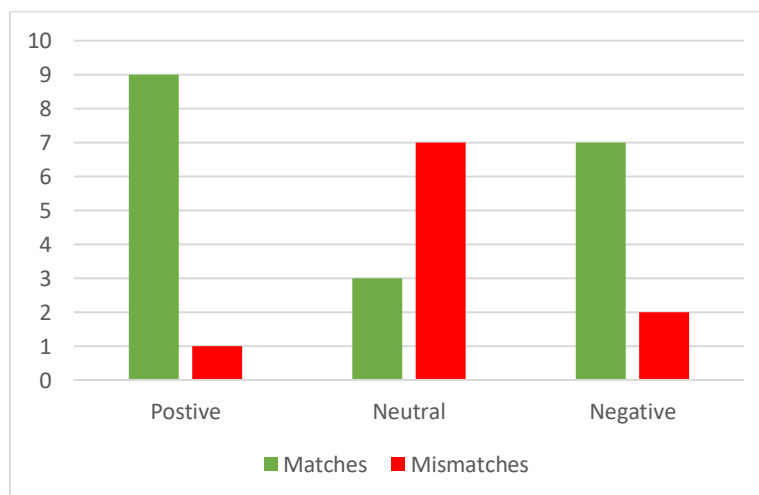
NEGATIVE REVIEWS

The following reviews had a sentiment value less than -0.333.

	Review	Observed Sentiment
1	I order 3 of them and one of the item is bad quality. Is missing backup spring so I have to put a pcs of aluminum to make the battery work.	Negative
2	Bulk is always the less expensive way to go for products like these	Neutral
3	they seemed to not last as long as other name brand names disappointed!	Negative
4	Sorry to say I've been disappointed with these. They didn't last as long in my remotes and RC transmitters as even other store brand alkaline batteries.	Negative
5	Mediocre performance	Neutral
6	Maybe mine is a bad batch, had a serious leak on one of them, these batteries seem to last as long as duracells but duracells never leaked in my experience > Batteries were inexpensive, work well in multiple devices.I hate running out of triple A batteries!	Negative
7	At least half of this batch had no charge, or at least so little that they died within an hour of being put in a walkie-talkie. Very disappointed	Negative
8	I've been using these on my roku remote and the lifespan is terrible. About two weeks, when Duracell's have been 1-2 months	Negative
9	Awful poor performance. My mouse ran for several weeks to several months from Duracells but have never gotten more that several days from these. Will not buy again.	Negative
10	-	-

INSIGHTS

The results show a strong correlation between the model's and the author's observed sentiments for positive and negative categories. However, there is poor correlation for neutral sentiments.



The reasons for this are due to the complexity of sentiment analysis, as well as design choices made in this task such as the categorisation thresholds.

There are instances where the review indicates multiple sentiments, which would be difficult for the model to categorise. For example, positive review number 8 contains the text “can't really comment” and “quick delivery” and the model has assigned a positive category. In the author's view the review sentiment should be neutral

because the reviewer states that they have just received the product and cannot judge the product, and the speed of delivery isn't relevant to information about the product.

Some reviews contain words that appear to mislead the model. For example, neutral review number 9 contains the text "exploded" which would have negative connotations when reviewing batteries, but in context is "none of them exploded" which is positive.

If this project was to be repeated, some tuning of the threshold values could be attempted to achieve an improved correlation between the model and observations. Additionally, further cleaning of the data could be explored to remove misleading words or phrases, and reviews that are determined to be highly opinion-based rather than fact-based could be excluded.