Information Retrieval Mid Project Review

"Aspect Based Product Review Mining"

Group No. 9

Authors: *Bhaswanth Gudimella, MT22025* *Kushagra Mittal, MT22108*
Vamshi Krishna Vannekala, MT22086 *Vaibhav Nauriyal, MT22129*
Ujjwal Garg, MT22085 *Shreeyash Khalate, MT22069*

1. Updated Problem Formulation:

What is the problem identified in the project?

The objective of the project is to develop a system that analyzes large volumes of generalized customer reviews on Amazon products of various categories and extracts the different aspects/features mentioned in the reviews like quality, design, usability, durability. The dynamic extraction of aspects can lead to extracting everything that is labelled as 'NOUN' by the NLTK POS tagging technique. Thus, an external interference is needed to determine which of the retrieved aspects are product features and which are not. All aspects that correspond to product features can then be evaluated with their opinion reviews to calculate a particular rating for each aspect.

The goal of our proposed system remains to provide more detailed evaluation of a product's performance with respect to its different features.

2. Updated Literature Review:

Is there any related work?

<Research paper references given below>

The research paper[1] in this area, propose methodologies to classify a review as positive and negative, where they come up with two key innovations : metadata substitutions and variable length features. They also Coupled the product name substitution with the best substring algorithm yielded 85.3 percent accuracy, higher than the 84.6 percent accuracy of bigrams in classification. Another research paper[2] introduces OPINE, an unsupervised information extraction system which mines reviews in order to build a model of important product features, their evaluation by reviewers, and their relative quality across products. OPINE solves the opinion mining tasks outlined above and outputs a set of product features, each accompanied by a list of associated opinions which are ranked based on strength. Another proposed technique[3] is an aspect-based product recommendation model that identifies aspects and sentiments from user reviews and incorporates them into a collaborative filtering approach to improve the accuracy of recommendations.

Analysing the importance of aspects review for any product, a user gives their opinion on a particular product may it be positive or negative, and the authors call the analysis of these opinions as 'Sentimental analysis'. There are two machine learning techniques to perform this analysis, they are Lexicon based and SVM. The existing system uses only unigram Lexicon system where only words like 'nice', 'bad', 'great', etc. are taken into account. But this method fails to include the inverted words like 'not'. When the review is 'not bad' and if the word 'not' is excluded then the review becomes negative, to overcome this a bi-gram Lexicon system is proposed. Also, previous proposed systems fail to compute the rating of individual features of a product just based on the customer review. Some platforms like Flipkart have implemented a similar approach but only for limited products. We will be implementing this for large group of products from different categories like electronics, books, DVDs, and clothing.

Thus, through our survey, we get an overview of the aspect-based opinion mining field, including tasks such as aspect extraction and sentiment classification, methods for addressing these tasks, and applications in various domains.

3. Updated Baseline Results:

Techniques used and retrieved results

First 3 steps remain same as mentioned in previous Baseline Submission

1. *Data Preprocessing:*

The first step in the project majorly involved data pre-processing steps like collecting the scattered data of different product categories and merging them into one single dataset,
Using NLTK to lowercase the categorical columns in the dataset, removing stop words, removing punctuations and non-alphabetic symbols, etc.
Then we calculated the initial sentiment behind every product review based on the user rating.

2. *Sentiment Analysis:*

We computed the compound sentiment score of each review using SentimentIntensityAnalyzer() and 'vader_lexicon'. A new rating and corresponding sentiment for each product is calculated based on this compound sentiment score.

This can further include computing polarity, subjectivity of the text in order to assign an aspect score to each aspect of a particular product.

3. *Aspect-Opinion Extraction and Aspect-Rating Calculation:*

This step further included more data preprocessing steps. First we remove the relative pronouns (that, which, who, whom, whose, etc.). So the statement like ‘it is that good’ becomes ‘it is good’. Then we remove adverbs so they are not detected as aspects further. We then join the related nouns so the phrase ‘battery life’ is detected as one single aspect ‘battery_life’.

We also handle negated words in the text like ‘not good’ by replacing them with their closest antonym. Thus ‘not good’ becomes ‘bad’ and ‘not terrible’ may become ‘okay’ and so on.

Further we extract Nouns and their related Adjectives from each review. Thus Nouns are our aspects and their corresponding adjectives become the description(opinion) of that aspect. Based on the description of each aspect of the product, a corresponding sentiment score is calculated using AFINN tool between -5 and +5. Based on this calculated sentiment score, a new rating is assigned to each aspect of the product out of 5.

For each of the aspect extracted from the text, a sentiment score is assigned to their respected descriptions (adjectives). Based on this sentiment score a rating (out of 5) is assigned to each aspect. The new rating for the product will be the average of ratings of all aspects extracted from that product review.
Example, The text 'The battery life of the phone is good but the camera quality is not that good.' gives the following output –

{'battery_life': 'good', 'camera_quality': 'bad'}
{'battery_life': 3.0, 'camera_quality': -3.0} (Sentiment Score between -5 to +5)
{'battery_life': 4, 'camera_quality': 2} (rating out of 5)
Sentient Score by Aspects: 0.0 (-1 to +1)
Rating by Aspect Sentiment Score: 3.0 (out of 5)

4. ***Updated Results:***

We were successful in extracting the aspects and their corresponding adjectives in the review text. All the negated words in the text were handled. We were able to eliminate all adverbs and relative pronouns and were able to join related nouns in the sentence.

The previous accuracy received for the sentiment detection was 80%

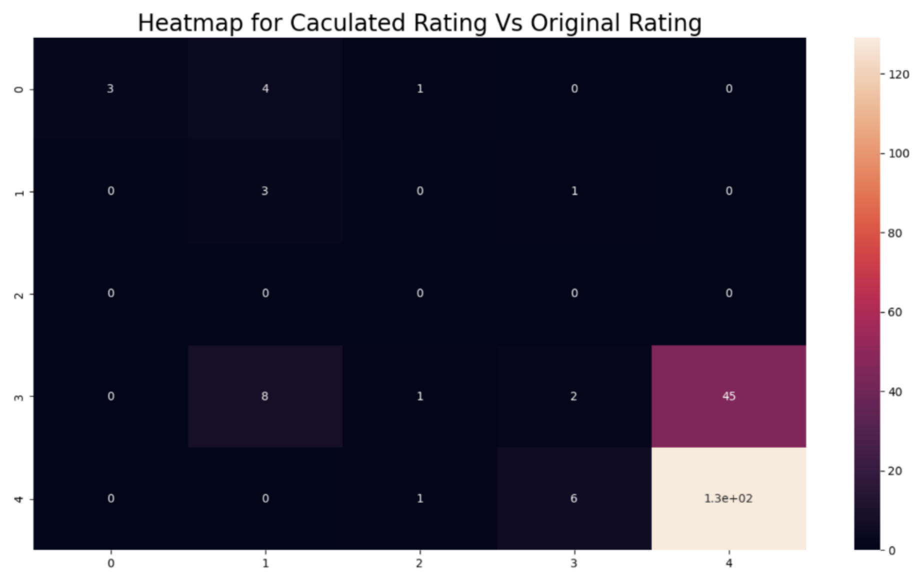
```
Calculated Sentiment Accuracy  0.800848040970291
```

To improve the accuracy we detected the contradicting rating-review pairs where the rating is excellent but the submitted review is negative or vice-versa.

The updated accuracy for same data is 95.38%

```
Calculated Sentment Accuracy  0.9538113430290708
```

Updated confusion matrix for calculated rating Vs original rating is –



Updated Classification Report for new rating vs original rating is –

```
Classification report:
              precision    recall  f1-score   support

         1.0       0.72      0.46      0.57      3933
         2.0       0.20      0.29      0.24      2370
         3.0       0.00      0.00      0.00         0
         4.0       0.26      0.09      0.14     15588
         5.0       0.73      0.91      0.81     39098

    accuracy                           0.65     60989
   macro avg       0.38      0.35      0.35     60989
weighted avg       0.59      0.65      0.60     60989
```

5. ***Output:***
   The final retrieved output for an input text like "The battery life of the phone is good but the camera quality is not that good." is –

```
Original Text :  The battery life of the phone is good but the camera quality is not that good.


After Removing Pronouns :  The battery life of the phone is good but the camera quality is not good .
After Removing Adverbs(but keeping negations) :  The battery life of the phone is good but the camera quality is not good .
After Joining Related Words :  The battery_life of the phone is good but the camera_quality is not good .
Negated words in the text are :  ['good']


Aspects: ['battery_life', 'camera_quality']
Adjectives: ['good', 'bad']
Aspects With Description: {'battery_life': 'good', 'camera_quality': 'bad'}
Aspects With Sentiment Scores (-5 to +5):  {'battery_life': 3.0, 'camera_quality': -3.0}
Aspects With Ratings (Out of 5):  {'battery_life': 4, 'camera_quality': 2}


Normal Sentiment Score  0.7003
Rating by Normal Sentiment Score:  5
Sentient Score by Aspects:  0.0
Rating by Aspect Sentiment Score:  3.0
```

The first 5 rows of updated dataframe after extracting aspects and calculating ratings is :

	index	product_name	review_text	rating	initial_sentiment	normal_sentiment_score	normal_sentiment	rating_by_normal_sentiment	aspect_ratings
0	0	sphere: books: michael crichton	sphere michael crichton excellant novel certai...	5.0	positive	0.9231	positive	5	{'novel': 3, 'misssion': 3, 'mile': 3, 'spacec...
1	1	healing from the heart: a leading surgeon comb...	dr oz accomplished heart surgeon field cardiac...	4.0	positive	0.93	positive	5	{'transplantation': 4, 'epilogue': 4, 'medicin...
2	2	mythology: dc comics art of alex ross 2007 cal...	gorgeous artwork comic books contains extraord...	5.0	positive	0.9118	positive	5	{'books': 4, 'artwork': 3}
3	4	guns, germs, and steel: the fates of human soc...	going short sweet review many others commented...	5.0	positive	0.9912	positive	5	{'review': 4, 'others': 2, 'value': 3, 'though...
4	5	world's fairs and the end of progress: an insi...	since worlds fair fanatic read almost everythi...	4.0	positive	0.9633	positive	5	{'worlds': 4, 'read': 4, 'book_author': 3, 'er...

5.1 Web Interface and Backend API Calls:

The below *Figure 1* shows API call to fetch the product name and aspect sentiment. This API is developed using Django framework.

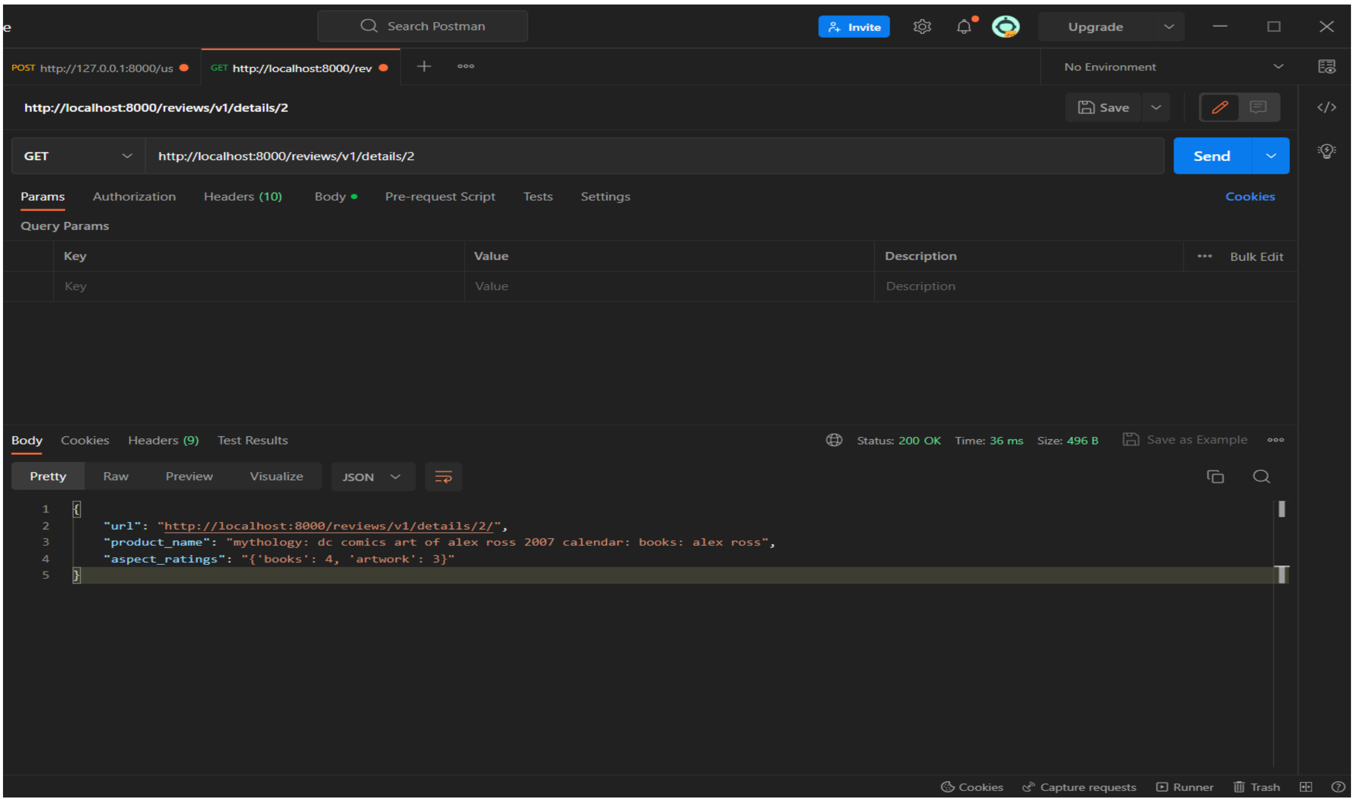


Figure 1 : GET request (API call) via Postman

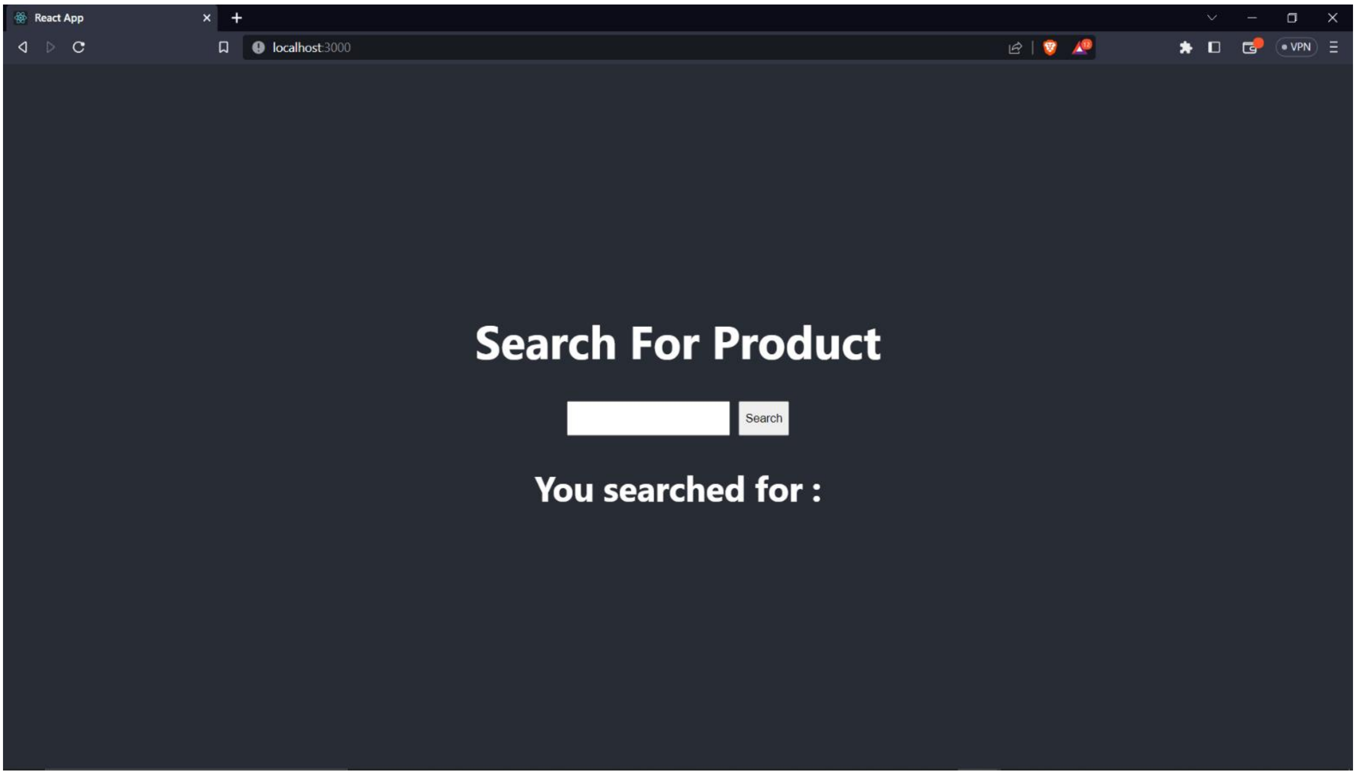


Figure 2 : React App Web Interface

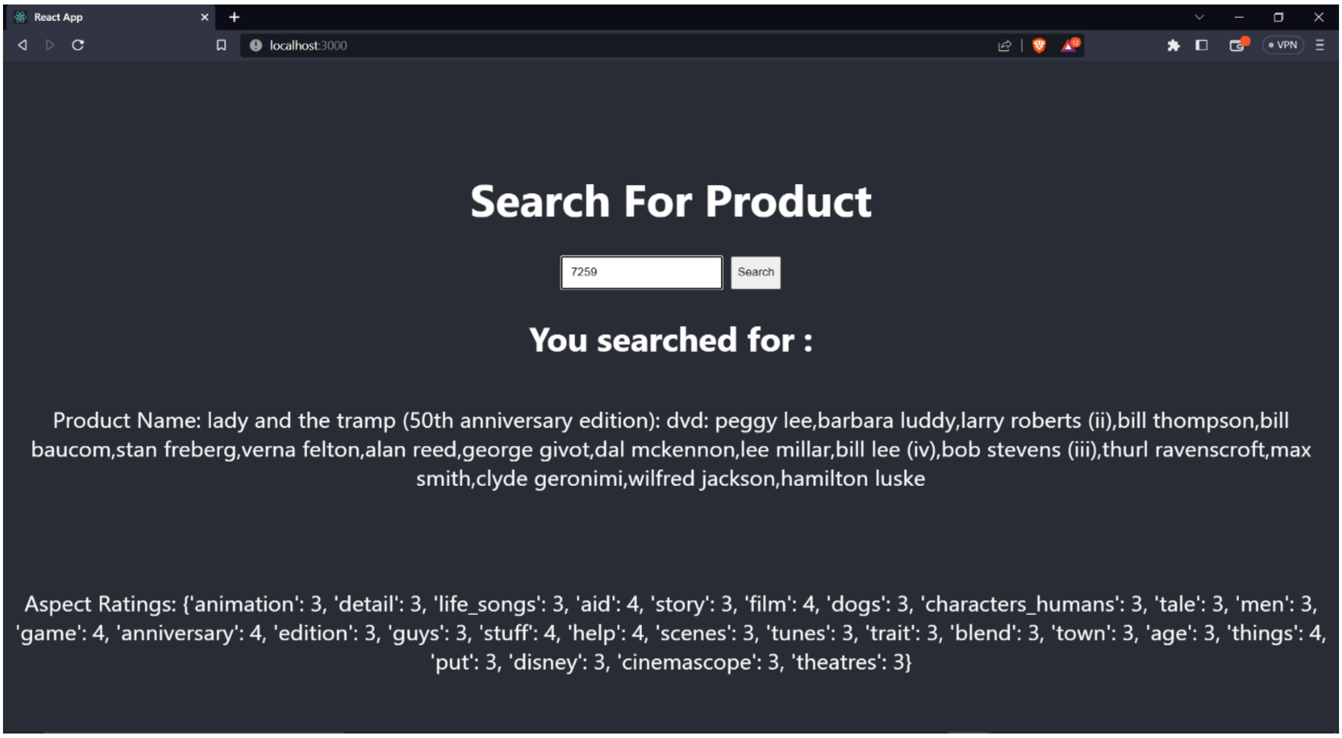


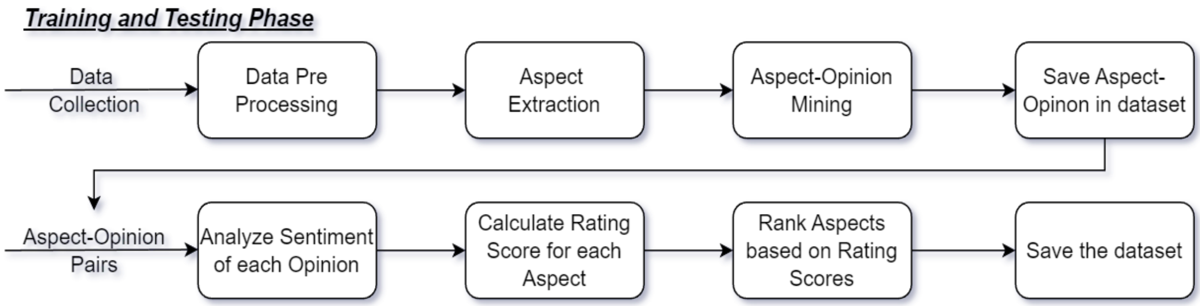
Figure 3 : Input and Output for a specific product

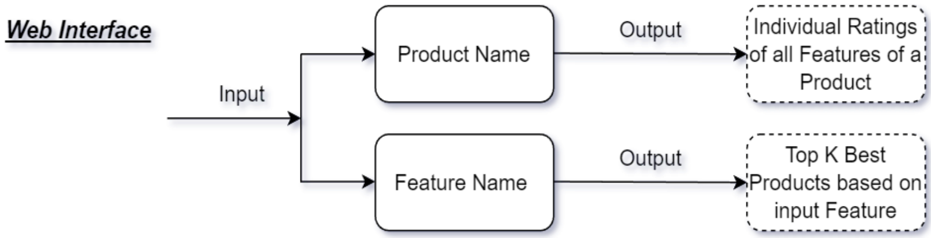
The above figures show the react app to search for a given product using product id and then displays the product name along with extracted aspects and associated sentiments from the product reviews.

4. <u>Proposed methodology:</u>

1. Data collection: Collect a large dataset of customer reviews for the target product or products.
2. Preprocessing: Clean and preprocess the data by removing noise, such as irrelevant words, and normalizing the text, such as converting all text to lowercase.
3. Aspect extraction: Extract relevant aspects or features of the product that are being discussed in the reviews. This can be done using methods such as rule-based parsing, unsupervised machine learning algorithms or NLTK POS tagging technique.
4. Sentiment analysis: Analyze the sentiment of each review with respect to each aspect. This can be done using supervised machine learning algorithms or lexicon-based approaches.
5. Rating calculation: Calculate a rating score for each aspect based on the sentiment analysis results.
6. Ranking: Rank the aspects based on their rating scores to identify the most important aspects of the product.

4.1 Proposed System Flow





4.2 Further possible improvements/Alternative approaches:

Deep Learning Architectures that can be used to perform aspect-based sentiment analysis.

1. **Recurrent Neural Networks (RNNs)**: RNNs are widely used in ABSA to analyse sequential data, such as sentences or paragraphs. They can be used to capture the context of each aspect in a sentence and predict the sentiment of the corresponding aspect.
2. **Convolutional Neural Networks (CNNs):** CNNs can be used to extract important features from the text and identify the sentiment of each aspect. They can also be used to classify text at different levels, such as word, sentence, or document level.
3. **Transformer-based Models:** Transformer-based models, such as BERT and GPT, have achieved state-of-the-art performance in many NLP tasks, including ABSA. These models can be fine-tuned on ABSA datasets to predict the sentiment of each aspect in a given text.
4. **Recursive Neural Networks (ReNNs):** ReNNs are a type of neural network that can recursively combine the representations of different aspects in a sentence to predict the overall sentiment of the sentence. This approach can be useful when there are multiple aspects in a sentence and their sentiments interact with each other.

Some of the above architectures can be used and we would also try out some of the above architectures.

5. Conclusion:

Accuracy for sentiment retrieval is improved from previous accuracy of 80% to 95.38%
Accuracy for new rating calculated compared to original rating is improved from 54% previously to 65%, although accuracy is a subjective measure here but still can be improved further by excluding extracted aspects which do not correspond to the product quality. Furthermore, neural networks can be used to extract more accurate aspects and their opinions. A static set of aspects for each product type can be considered to compare the relevancy of dynamically extracted aspects.