

ASPECT MINING ON CONFECTIONERY PRODUCTSRupashini P R¹Rahul Aanand S²Akilan K³

¹ Guided by Assistant Professor Department of Computer Science and Engineering, Bannari Amman Institute of Technology, India

^{2,3} Student, Department of Computer Science and Engineering, Bannari Amman Institute of Technology, India

ABSTRACT

With the rapid growth in E-Commerce, reviews for popular products on the web have grown rapidly. Customers are often forced to wade through many online reviews. In order to make an informed product choice, these comments are useful for other users to decide in buying a product. The manual analysis of such a huge number of reviews is practically impossible. Scanning all these reviews would be tedious, time consuming, boring, and fruitless. It would be good if these reviews could be processed automatically and customers are provided with limited generalized information. To solve this problem, an automated approach to mine the aspects related to the products is required. Opinion Mining plays a major role to summarize customer reviews and make it easy for online customers to determine whether to purchase the products or not. The proposed work performs aspect-level opinion mining by extracting product aspects from reviews on E-Commerce sites and produces a summarized report of the most frequently discussed product aspects regarding the number of appearances in positive, negative, and neutral reviews. This project focuses on the aspect based opinion mining of confectionery product reviews. Given a set of reviews of a product, the model generates products' aspect profile summarizing its important features automatically.

Keywords:

Aspect Term Extraction, Sentiment analysis, POS tagging, Rating.

INTRODUCTION

In aspect-based opinion mining (ABOM), an entity's features or aspects are extracted, and opinions on those traits or aspects are determined. It is a text classification technique that developed from named entity extraction and sentiment analysis (NER). As a result, ABOM combines aspect extraction with opinion mining. Despite the fact that opinions about entities are useful, opinions regarding their components are more specific and enlightening. People's opinions on products have always been an important piece of information when buying a product online. With the rapid growth in E-commerce, reviews for popular products in the internet have grown rapidly. Customers are often forced to wade through many online reviews in order to make an informed product choice. These reviews are beneficial for other users to make a decision on buying a product. The manual analysis of such a huge number of reviews is very nearly impossible. Scanning all these reviews would be tedious, time-consuming, boring, and fruitless. It would be good if these reviews could be processed automatically and customers are provided with limited generalized information.

To solve this problem, an automated approach of aspect mining related to the products is required. Opinion Mining plays a major role to summarize customer reviews and make it easy for online customers to determine whether to purchase the products or not. The proposed work performs aspect-level opinion mining by extracting product aspects from reviews on e-commerce sites and produces a summarized report of the most frequently discussed product aspects regarding the number of appearances in positive, negative, and neutral reviews. This project focuses on the aspect based opinion mining of confectionery product reviews. Given a set of reviews of an item for consumption, the model generates products' aspect profile summarizing its important features automatically. This project focuses on the aspect-based opinion mining of confectionery product reviews. Given a set of reviews of a product, the model generates products' aspect profile, summarizing its important features automatically. Our project aims to provide this feature to users for improved decision making on confectionery products.

OBJECTIVES

This project focuses on the aspect-based opinion mining of confectionery product reviews. Given a customary of reviews of a product, the model generates products' aspect profile, summarizing its important features automatically. The project aims to provide this feature to customers for improved decision making on confectionery products.

PROBLEM IDENTIFICATION

People's thoughts on products have always been an important piece of information when buying a product online. With the rapid growth in E-commerce, reviews for popular products on the internet have grown rapidly. Customers are often forced to paddle through many online reviews in order to make a conversant product choice. These comments are useful for other consumers to make a decision on buying a product. The manual analysis of such a huge number of reviews is basically impossible. Scanning all these reviews would be tedious, time-consuming, boring, and fruitless. It would be good if these reviews could be processed automatically and customers are provided with limited generalized information.

LITERATURE SURVEY

1. Nidhi Mishra, Sonal Meenu Singh worked on the explosion of Web 2.0, platforms such as blogs, ecommerce sites, peer-to-peer networks and social media, consumers have a broad platform and unlimited power to share their experiences in the form of reviews. With this, bulks of reviews are available for a single product, that needs to be analyzed, processed and mine. Opinion Mining or Sentiment Analysis is a Natural Language Processing and Information retrieval task that defines the customer's views or opinions through positive, negative or neutral sentiments. Aspect based Opinion Mining deals with aspects of the features. In this paper an aspect-based opinion mining system is proposed that classify reviews as positive and negative. The system also deals with two aspects in a review. Experimental results using reviews of mobile phones show an accuracy of 75% as compared to other methods. [1].
2. Anetta Barska, Julia Wojciechowska-Solis investigated on the dynamically developing trend of sustainable consumption is manifested, among others, by the growing interest in organic products on the part of consumers. The aim of this article was to identify the behavior of Polish consumers in the market of organic products and to establish a link between their environmental awareness and willingness to buy organic products. The authors hypothesized that there is a relationship between consumer awareness of the concept of sustainable consumption and the consumption of organic products. Consumer awareness means making conscious choices based on the knowledge expressed in the attitudes and, sometimes, preferences of the food brand. The research was conducted using a proprietary survey questionnaire. A total of 1067 respondents participated. A statistical analysis was performed by using Statistica 13.1 PL software, which includes descriptive statistics, the discriminant function analysis, and regression analysis. Motives were identified that are of crucial importance to the consumer deciding to purchase organic products. [2].
3. M.T. Mastura, H.N. Salwa, S.M. Sapuan, R.A. Ilyas and M.Y.M Zuhri conducted a study on the consumer packaged goods (CPG) markets greatly rely on packaging. Packaging allows products to be delivered to consumers in excellent condition and spoilage can be avoided. Packaging also plays an important role in extending shelf life and minimizing wastage for unpreserved foods such as fresh meat and fish, ready-prepared meals, and salads. Other products, including electronic equipment such as a mobile phone, as well as other domestic appliances such as vacuum cleaners, kettles, and toasters, all rely on packaging to protect them against the potential damage encountered in the distribution chain. [3].
4. Adisa Azapagic, Antonios Konstantas, Laurence Stamford's study has introduced a framework for the assessment of eco-efficiency at both product and sectoral levels based on life cycle thinking. Its application can provide valuable insights into a broad range of eco-efficiency indicators. The framework also incorporates multicriteria decision analysis to help aggregate the indicators into a single eco-efficiency score, while enabling a transparent analysis of how preferences for different indicators affect outcomes. The application of the framework has been demonstrated by estimating the eco-efficiency of the confectionary and the frozen desserts sectors in the UK. At the product level, the estimated eco-efficiencies range from 0.001 to 0.84, with cupcakes being most and semi-sweet biscuits least efficient. A general trend is observed showing that premium products, such as premium ice cream or cupcakes, tend to be considerably more eco-efficient than low-value items. [4].
5. Bing Liu, Jiajun Bu, Chun Chen, Guang Qiu & Jiajun Bu did an article that focuses on two important tasks in opinion mining, namely, opinion lexicon expansion and target extraction. We propose a propagation approach to extract opinion words and targets iteratively given only a seed opinion lexicon of small size.

The extraction is performed using identified relations between opinion words and targets, and also opinion words/targets themselves. The relations are described syntactically based on the dependency grammar. We also propose novel methods for new opinion word polarity assignment and noisy target pruning. In the evaluation, we compared our new approach with others on standard testing data sets. The results show that our approach outperforms other state-of-the-art methods in these two tasks. In the future, we plan to first focus on improving the precision of opinion word extraction by working on opinion word pruning methods. We will then also try to learn syntactic relations automatically from large corpuses using pattern mining techniques to improve the relation coverage. [5].

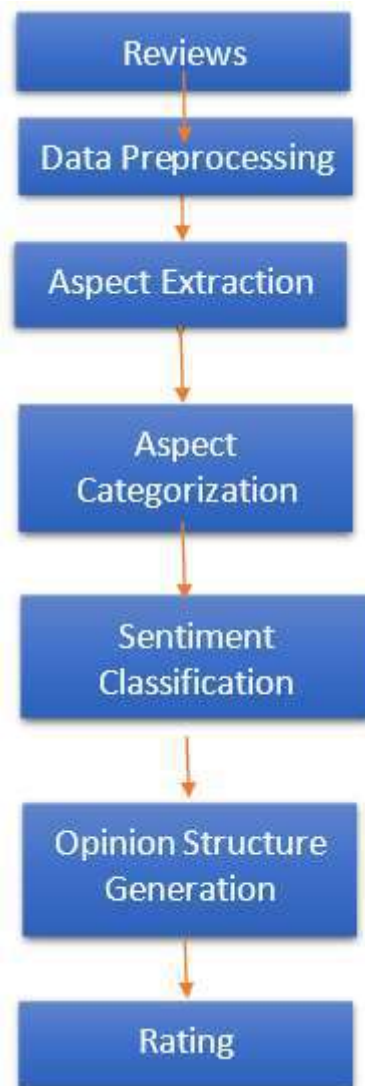
6. Fatemeh Hemmatian, Sohrabi Mohammad Karim states that opinion mining is considered as a subfield of natural language processing, information retrieval and text mining. Opinion mining is the process of extracting human thoughts and perceptions from unstructured texts, which with regard to the emergence of online social media and mass volume of users' comments, has become to a useful, attractive and also challenging issue. There are varieties of researches with different trends and approaches in this area, but the lack of a comprehensive study to investigate them from all aspects is tangible. In this paper they represent a complete, multilateral, and systematic review of opinion mining and sentiment analysis to classify available methods and compare their advantages and drawbacks, in order to have better understanding of available challenges and solutions to clarify the future direction. For this purpose, they present a proper framework of opinion mining accompanying with its steps and levels and then completely monitor, classify, summarize and compare proposed techniques for aspect extraction, opinion classification, summary production and evaluation, based on the major validated scientific works. In order to have a better comparison, they also propose some factors in each category, which help to have a better understanding of advantages and disadvantages of different methods. [6].
7. Bo Pang, Lillian Lee done a survey that covers techniques and approaches that promise to directly enable opinion-oriented information-seeking systems. Our focus is on methods that seek to address the new challenges raised by sentiment aware applications, as compared to those that are already present in more traditional fact-based analysis. We include material on summarization of evaluative text and on broader issues regarding privacy, manipulation, and economic impact that the development of opinion-oriented information access services gives rise to. To facilitate future work, a discussion of available resources, benchmark datasets, and evaluation campaigns is also provided. [7].

METHODOLOGY

This Aspect-based opinion mining (ABOM) involves extracting aspects or features of an entity and eliciting opinions about those aspects. It is a text classification method that evolved from sentiment analysis and named entity extraction (NER). ABOM is therefore a combination of viewpoints and opinion mining. While entity views are useful, entity views are more detailed and comprehensive.

To solve this problem, an automated approach to mine the aspects related to the products is required. Opinion Mining plays a major role to summarize customer reviews and make it easy for online customers to determine whether to purchase the products or not. The proposed work performs aspect-level opinion mining by extracting product aspects from reviews on e-commerce sites and produces a summarized report of the most frequently discussed product aspects regarding the number of appearances in positive, negative, and neutral reviews. This project focuses on the aspect-based opinion mining of confectionery product reviews. Given a set of reviews of a product, the model generates products' aspect profile summarizing its important features automatically. The system performs this task in several steps as data collection, pre-processing by removing stop words (words which does not perform any role in classifying the opinion, for example "I", "Me", "OMG", etc.), tokenizing reviews, pos tagging, aspect extraction, orientation and then identifying opinion words using aspect mining and classifying the product, according to the aspects by giving a score to each aspect, whether it is good, bad or a neutral product.

- DATA COLLECTION
- DATA PREPROCESSING
- POS TAGGING
- ASPECT TERM EXTRACTION
- SENTIMENT ANALYSIS
- INGREDIENTS ANALYSIS



DATA COLLECTION

To Many reviews are gathered from the web in order to determine the polarity of the sentences depending on their elements. The vast majority of consumer reviews are available on numerous places on the Internet. The reviews are gathered through e-commerce websites.

DATA PREPROCESSING

The review text that we extracted from the e-com sites had a lot of unclean data, so we created a cleaning script for the dataset. It removed unnecessary characters, hyperlinks, symbols, excess spaces, and other patterns of text that could not be processed by our system. From the cleaned dataset, we extracted the review text description for our analysis.

POS TAGGING

POS TAG	DESCRIPTION	EXAMPLES FROM ENGLISH LANGUAGE
JJ	Adjective	great, new, high, big, local
NN, NNS	Noun, Singular, or Mass	home, order, price, time
NNP	Noun, plural	Deliveries, services
VB	Verb/ Verb, past participle/ Verb, past	is, make, said, agreed, placed
RB	Adverb	already, really, still, early, now, yet
DT	Preposition or subordinating conjunction	with, in, over, of, by

ASPECT TERM EXTRACTION:

In Aspect-Based Sentiment Analysis, which tries to extract explicit aspect expressions from online user reviews, Aspect Term Extraction (ATE) is a crucial sub-task. So, what is the aspect term? Aspect terms are the target of opinions.

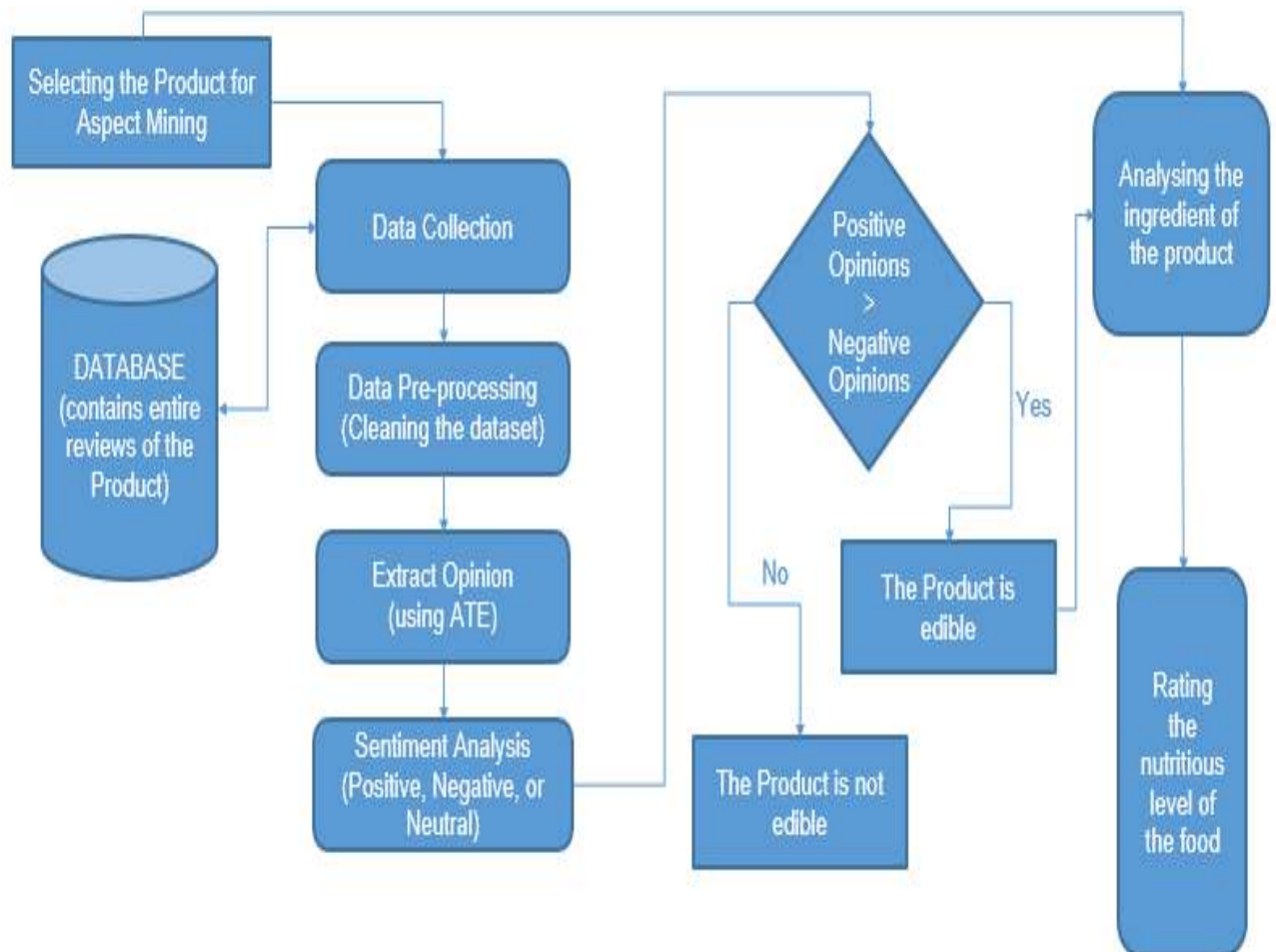
SENTIMENT ANALYSIS:

Sentiment analysis is a method of natural language processing (NLP) that detects whether a piece of text is good, negative, or neutral. It is sometimes referred to as opinion mining or emotion artificial intelligence. A sentiment analysis tool can comprehend what customers are saying, how they are saying it, and what they truly mean from both an individual's and the public's perspective by evaluating text and statistics.

INGREDIENTS ANALYSIS:

Using the open food facts library, the ingredients of the product is extracted by entering the bar code of the product, which is already stored in the database. The ingredients nutrition is analyzed and given a score depending on its nutrition whether it is very good, good, bad, or very bad by rating A, B, C, D respectively.

MODEL AND SYSTEM ARCHITECTURE



RESULTS AND DISCUSSION**DATA COLLECTION**

```

class AmazonScraper:
    review_date_pattern = re.compile('(?:Jan(?:uary)?|Feb(?:ruary)?|Mar(?:ch)?|Apr(?:il)?|May|Jun(?:e)?|Jul(?:y)?|Aug(?:ust)?)')
    product_name_pattern = re.compile('https://(?:www.amazon.in/|.+)\product-reviews')
    def __init__(self):
        self.session = requests.Session()
    def scrapeReviews(self, url, page_num, filter_by='recent'):

    try:
        if re.search('^+(?=/)', url)==None:
            return "end"
        review_url = re.search('^+(?=/)', url).group()
        review_url = review_url + '?reviewerType=all_reviews&sortBy={0}&pageNumber={1}'.format(filter_by, page_num)
        print('Processing {0}...'.format(review_url))
        response = self.session.get(review_url)

        product_name = self.product_name_pattern.search(url).group(1) if self.product_name_pattern.search(url) else ''

        if not product_name:
            print('url is invalid. Please check the url.')
            return
        else:
            product_name = product_name.replace('-', ' ')

        soup = BeautifulSoup(response.content, 'html.parser')
        review_list = soup

        reviews = []

        product_reviews = review_list.find_all('div', {'data-hook': 'review'}) # return reviews

        for product_review in product_reviews:
            if product_review==None:
                return "end"
            if product_review.find('a', {'data-hook': 'review-title'})==None:
                return "end"
            review_title = product_review.find('a', {'data-hook': 'review-title'}).text.strip()
            review_body = product_review.find('span', {'data-hook': 'review-body'}).text.strip()
            rating = product_review.find('i', {'data-hook': 'review-star-rating'}).text
            username = product_review.a.span.text
            user_profile = 'https://amazon.in/{0}'.format(product_review.a['href'])

            reviews.append(UserReview(product_name, review_title, review_body, rating, username))
        return reviews
    except Exception as e:
        print(e)
        return None

```

This is the code for collecting data that is the set of reviews of the product for reviewing. It connects with amazon review scrapper and gets all the reviews posted by people as a single set.

DATASET

The dataset created from review scrapper model for Pediasure Product

```
In [3]: df.head()
```

```
Out[3]:
```

	product_name	review_title	comment	rating	username
0	Pediasure Specialized Nutrition Children Choco...	Taste	Good	5.0 out of 5 stars	Udita Bhatt
1	Pediasure Specialized Nutrition Children Choco...	Good product	Kids like the product. Very good flavor	4.0 out of 5 stars	Bharat Parmar
2	Pediasure Specialized Nutrition Children Choco...	Good	Good	5.0 out of 5 stars	Anita
3	Pediasure Specialized Nutrition Children Choco...	Essential	Very good product for growing kids	5.0 out of 5 stars	Abhishek Bakshi
4	Pediasure Specialized Nutrition Children Choco...	Best price to buy again	Good flavour & taste , I like it	5.0 out of 5 stars	Ravi

Now the reviews of the pediasure product have been collected. The next step is to take these reviews for ATE.

ASPECT MINING**DATA PREPROCESSING: Stop words removal**

```
[17]: def preProcessing (inputFileStr, outputFileStr, printResult):
inputFile = open(inputFileStr,encoding="utf8").read()
outputFile = open(outputFileStr,"w",encoding="utf8")

cachedStopwords = nltk.corpus.stopwords.words("english")
cachedStopwords.append('OMG')
cachedStopwords.append(':-')
cachedStopwords.append('READ')
cachedStopwords.append('...')
cachedStopwords.append('...')
cachedStopwords.append('...')
result = ' '.join[word for word in inputFile.split() if word not in cachedStopwords]]
if(printResult):
    print('Following are the stopwords')
    print(cachedStopwords)
    outputFile.write(str(result))
outputFile.close()

preProcessing("Review.txt", "preProcessing.txt", "new.txt")
```

Following are the stopwords
['i', 'me', 'my', 'myself', 'we', 'our', 'ours', 'ourselves', 'you', 'you're', 'you've', 'you'll', 'you'd', 'your', 'yours', 'yourself', 'yourselves', 'he', 'him', 'his', 'himself', 'she', 'she's', 'her', 'hers', 'herself', 'it', 'it's', 'its', 'itself', 'they', 'them', 'their', 'theirs', 'theirs', 'themselves', 'what', 'which', 'who', 'whom', 'this', 'that', 'that'll', 'these', 'those', 'am', 'is', 'are', 'was', 'were', 'be', 'been', 'being', 'have', 'has', 'had', 'having', 'do', 'does', 'did', 'doing', 'a', 'an', 'the', 'and', 'but', 'if', 'or', 'because', 'as', 'until', 'while', 'of', 'a', 't', 'by', 'for', 'with', 'about', 'against', 'between', 'into', 'through', 'during', 'before', 'after', 'above', 'below', 'to', 'from', 'up', 'down', 'in', 'out', 'on', 'off', 'over', 'under', 'again', 'further', 'then', 'once', 'here', 'there', 'when', 'where', 'why', 'how', 'all', 'any', 'both', 'each', 'few', 'more', 'most', 'other', 'some', 'such', 'no', 'nor', 'not', 'only', 'own', 'same', 'so', 'than', 'too', 'very', 's', 't', 'can', 'will', 'just', 'don', 'don't', 'should', 'should've', 'now', 'd', 'll', 'm', 'o', 're', 've', 'y', 'ain', 'aren', 'aren't', 'couldn', 'couldn't', 'didn', 'didn't', 'doesn', 'doesn't', 'hadn', 'hadn't', 'hasn', 'hasn't', 'haven', 'haven't', 'isn', 'isn't', 'ma', 'mightn', 'mightn't', 'mustn', 'mustn't', 'needn', 'needn't', 'shan', 'shan't', 'shouldn', 'shouldn't', 'wasn', 'wasn't', 'weren', 'weren't', 'won', 'won't', 'wouldn', 'wouldn't', 'OMG', ':-', 'READ', '...', '...', '...']

The data preprocessing is done to remove the stop words.

TOKENIZATION OF REVIEWS

```
In [18]: def tokenizeReviews(inputFileStr, outputFileStr, printResult):
tokenizedReviews = {}
inputFile = open(inputFileStr, "r", encoding="utf8").read()
outputFile = open(outputFileStr, "w", encoding="utf8")
tokenizer = nltk.tokenize.punkt.PunktSentenceTokenizer()
uniqueId = 1;
cachedStopWords = nltk.corpus.stopwords.words("english")
for sentence in tokenizer.tokenize(inputFile):
    tokenizedReviews[uniqueId] = sentence
    uniqueId+= 1
outputFile.write(str(tokenizedReviews))
if(printResult):
    for key,value in tokenizedReviews.items():
        print(key, ' ',value)
outputFile.close()

tokenizeReviews("preProcessing.txt", "tokenizeReviews.txt", "new1.txt")
```

POS TAGGING

```
In [19]: def posTagging(inputFileStr,outputFileStr,printResult): #Tagging each word
inputFile = open(inputFileStr,"r",encoding="utf8").read()
outputFile=open (outputFileStr,"w",encoding="utf8")
inputTuples=ast.literal_eval(inputFile)
outputPost={}
for key,value in inputTuples.items():
    outputPost[key]=nltk.pos_tag(nltk.word_tokenize(value))
if(printResult):
    for key,value in outputPost.items():
        print(key, ' ',value)
outputFile.write(str(outputPost))
outputFile.close()

posTagging("tokenizeReviews.txt", "posTagging.txt", "new1.txt")
```

```
1 [(['Taste', 'NN'), ('Good', 'JJ'), ('product', 'NN'), ('Good', 'NNP'), ('Essential', 'NNP'),
('Best', 'NNP'), ('price', 'NN'), ('buy', 'VB'), ('Too', 'NNP'), ('good', 'JJ'), ('Super', 'NN
P'), ('Good', 'NNP'), ('Taste', 'NNP'), ('Good', 'NNP'), ('Beneficial', 'NNP'), ('kids', 'NN
S'), ('12yrs', 'CD'), ('Great', 'NNP'), ('discount', 'NN'), ('Different', 'NNP'), ('flavours',
'VBZ'), ('Vanilla', 'NNP'), ('rec', 'NN'), ('d', 'MD'), ('Good', 'VB'), ('Colour', 'NNP'), ('S
uperb', 'NNP'), ('Price', 'NNP'), ('high', 'JJ'), ('market', 'NN'), ('Good', 'NNP'), ('Good',
'NNP'), ('product', 'NN'), ('Good', 'NNP'), ('Worst', 'NNP'), ('Flavour', 'NNP'), ('quality',
'NN'), ('Nice', 'NNP'), ('Taste', 'NNP'), ('Good', 'NNP'), ('Best', 'NNP'), ('children', 'NN
S'), ('Good', 'NNP'), ('quality', 'NN'), ('products', 'NNS'), (',', ','), ('result', 'NN'), ('o
k', 'JJ'), ('Nice', 'NNP'), ('product', 'NN'), ('Amazon', 'NNP'), ('good', 'JJ'), (',', ','),
('it', 'PRP'), ('s', 'VBZ'), ('delivery', 'NN'), ('it', 'PRP'), ('...', ':'), ('K
ids', 'NNP'), ('love', 'VB'), ('Ok', 'NNP'), ('Good', 'NNP'), ('Nice', 'NNP'), ('Good', 'NNP'),
('Ok', 'NNP'), ('Very', 'NNP'), ('tasty', 'JJ'), ('😂😂', 'NN'), ('Damage', 'NNP'), ('Produc
t', 'NNP'), ('Good', 'NNP'), ('health', 'NN'), ('Very', 'RB'), ('good', 'JJ'), ('👍', 'NN'),
('good', 'JJ'), ('Super', 'NNP'), ('pediasure', 'NN'), ('7+', 'CD'), ('chocolate', 'NN'), ('400
gm', 'CD'), ('Nice', 'NNP'), ('Is', 'NNP'), ('bad', 'JJ'), ('flavour', 'NN'), ('No', 'NNP'),
('discount', 'NN'), ('given', 'VBN'), ('Healthy', 'NNP'), ('Not', 'RB'), ('suited', 'VBD'), ('D
aughter', 'NNP'), (',', ',')]
2 [(['Inside', 'NNP'), ('package', 'NN'), ('teared', 'VBD'), ('A', 'NNP'), ('good', 'JJ'), ('p
```

Pos tagging is the process of marking up a word in a text (corpus) as corresponding to a particular part of speech, based on both its definition and its context.

ASPECT EXTRACTION:

```
In [20]: def aspectExtraction(inputFileStr,outputFileStr,printResult): #Finding list of aspectation words
inputFile = open(inputFileStr,"r",encoding="utf8").read()
outputFile = open(outputFileStr,"w",encoding="utf8")
inputTuples=ast.literal_eval(inputFile)
prevWord=''
prevTag=''
currWord=''
aspectList=[]
outputDict={}
for key,value in inputTuples.items():
    for word,tag in value:
        if (tag=='NN' or tag=='NNP'):
            if (prevTag=='NN' or prevTag=='NNP'):
                currWord=prevWord + ' '+word

            else:
                aspectList.append(prevWord.upper())
                currWord= word
        prevWord=currWord
        prevTag=tag

    for aspect in aspectList:
        if (aspectList.count(aspect)>1):
            if (outputDict.keys()!=aspect):
                outputDict[aspect]=aspectList.count(aspect)
    outputAspect=sorted(outputDict.items(),key=lambda x: x[1],reverse = True)
    print(outputAspect)
    outputFile.write(str(outputAspect))
    outputFile.close()

aspectExtraction("posTagging.txt", "aspectExtraction.txt", "new1.txt")
```

```
[('PEDIASURE', 4), ('NICE', 4), ('PRODUCT', 4), ('HEALTH', 4), ('TASTE', 3), ('A', 2), ('GOOD', 2), ('VERY', 2), ('CHILD', 2), ('NICE PRODUCT', 2)]
```

SENTIMENT ANALYSIS

```
[21]: def orientation(inputWord):
      wordSynset=wordnet.synsets(inputWord)
      if(len(wordSynset) != 0):
          word=wordSynset[0].name()
          orientation=sentiwordnet.senti_synset(word)
          if(orientation.pos_score()>orientation.neg_score()):
              return True
          elif(orientation.pos_score()<orientation.neg_score()):
              return False
```

```
IDENTIFY_OPINION_WORDS('opinion_words.txt', 'opinion_words.txt', 'opinion_words.txt', 'opinion_words.txt')
...
PRODUCT : Positive => 58.7 Negative => 28.65
HEALTH : Positive => 55.77 Negative => 26.92
TASTE : Positive => 53.19 Negative => 23.4
<_io.TextIOWrapper name='identifyOpinionWords.txt' mode='w' encoding='cp1252'>
```

By identifying the opinion words, it will be easy to classify the reviews in which way we want it.

INGREDIENT ANALYSIS

```
In [27]: import openfoodfacts
product= openfoodfacts.products.get_product('8710428015877 (EAN / EAN-13)')#bar code of the product from website
print(product)
```



```

n [33]: In [33]: def ingredient():
            d={'A':"VERY GOOD IN NUTRITION",
              'B':"GOOD IN NUTRITION",
              'C':"AVERAGE UN NUTRITION",
              'D':"BAD IN NUTRITION",
              'E':"VERY BAD IN NUTRITION"}
            print("NUTRITION SCORE [A B C D E] ")
            print(product['product']['nutriscore_grade'].upper(),"=",d[product['product']['nutriscore_grade'].upper()])
            ingredient()

NUTRITION SCORE [A B C D E]
B = GOOD IN NUTRITION

```

As stated in above methodology rating of the products is done on a scale of 4 by accessing the food facts library which is an open source.

FUTURE SCOPE

Many online shopping sites include candy product reviews, star ratings, and reviews. With this we analyze information about the product from the customer's point of view and make suggestions to people who want to buy similar products. Today people care about what they consume. We can also find out about the ingredient in the product and warn about the potentially harmful ingredient.

ACKNOWLEDGEMENT

We wish to express our sincere thanks to Faculty guide Mrs. Rupashini P R (Assistant Professor II) for their constructive ideas, encouragements, excellent guidance and much needed technical support extended to complete our project.

CONCLUSION

Many e-commerce sites contain confectionery product reviews, star rating and followers of reviews. With this information we analyze about the product from customer perspective and the suggestion to people who want to buy similar products. Now a day people are concern of what they consume. We can also discover out about the ingredient containing the product and warn about the contain of any possible harmful ingredient.

REFERENCES

1. Nidhi Mishra, Sonal Meenu Singh (2016), "Aspect Based Opinion Mining for Mobile Phone". doi: [10.1109/NGCT.2016.7877474](https://doi.org/10.1109/NGCT.2016.7877474)
2. Anetta Barska, Julia Wojciechowska-Solis (2021), "Exploring the Preferences of Consumers' Organic Products in Aspects of Sustainable Consumption: The Case of Polish Consumer, Agriculture". doi:10.3390/agriculture11020138.
3. M.T. Mastura, H.N. Salwa, S.M. Sapuan, R.A. Ilyas and M.Y.M Zuhri. (2021), "Life Cycle Assessment of Bio-Based Packaging Products, Bio-Based Package". doi :10.1002/9781119381228.ch22.
4. Adisa Azapagic, Antonios Konstantas, Laurence Stamford (2020), "A Framework for Evaluating Life Cycle Eco-efficiency and an Application in the Confectionary and Frozen Desserts Sectors. Sustainable Production and Consumption". doi: 10.1016/j.spc.2019.12.006.
5. Bing Liu, Jiajun Bu, Chun Chen, Guang Qiu & Jiajun Bu, (2011), "Opinion Word Expansion and Target Extraction Through Double Propagation ". Computational Linguistics, 37(1), 9–27. doi:10.1162/coli_a_00034.
6. Fatemeh Hemmatian , Sohrabi Mohammad Karim (2017), "A Survey on Classification Techniques for Opinion Mining and Sentiment Analysis". doi:10.1007/s10462-0179599-6.
7. Bo Pang, Lillian Lee (2008), "Opinion Mining and Sentiment Analysis". Foundations and Trends in Information Retrieval, Vol. 2(1-2): pp. 1-135. doi:10.1561/15000000011.
8. Hsinchun Chen, David Zimbra (2010), "AI and Opinion Mining". IEEE Intelligent Systems, 25(3), 74–80. doi:10.1109/mis.2010.75.
9. Deepika Kumawat, Vinesh Jain (2015), "POS Tagging Approaches: A Comparison". International Journal of Computer Applications (0975 – 8887) Volume 118 – No. 6, May 2015.
10. Andrea Dominici, Enrico Marone, Fabio Boncinelli, Francesca Gerini (2021), "Determinants of online food purchasing: The impact of socio-demographic and situational factors". Journal of Retailing and Consumer Services.