SENTIMENT ANALYSIS OF PRODUCTS REVIEW:

A FIXED FEATURE-BASED APPROACH

BY

TAPASY RABEYA

ID: 183-25-704

This Report Presented in Partial Fulfillment of the Requirements for the Degree of Master of Science in Computer Science and Engineering

Supervised By

Professor Dr. Md. Ismail Jabiullah
Professor
Department of CSE
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

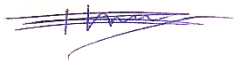
DHAKA, BANGLADESH

JANUARY 2022

APPROVAL

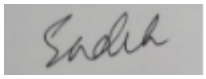
This Project titled "**Sentiment Analysis of Products Review: A Fixed Feature-Based Approach**", submitted by Tapasy Rabeya to the Department of Computer Science and Engineering, Daffodil International University, has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of M.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on

BOARD OF EXAMINERS



Chairman

Dr. Touhid Bhuiyan
Professor and Head
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University



Internal Examiner

Md. Sadekur Rahman (SR)
Assistant Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Internal Examiner

Moushumi Zaman Bonny
Assistant Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

External Examiner

Dr. Shamim H Ripon
Professor
Department of Computer Science and Engineering
East West University

DECLARATION

We hereby declare that, this project has been done by Tapasy Rabeya under the supervision of Professor Dr. Md. Ismail Jabiullah, Professor**, Department of CSE** Daffodil International University. We also declare that neither this project nor any part of this project has been submitted elsewhere for award of any degree or diploma.

Supervised by:

Professor Dr. Md. Ismail Jabiullah
Professor
Department of CSE
Daffodil International University

Submitted by:

Tapasy Rabeya
ID: 183-25-704
Department of CSE
Daffodil International University

ACKNOWLEDGEMENT

First we express our heartiest thanks and gratefulness to almighty God for His divine blessing makes us possible to complete the final year project/internship successfully.

We really grateful and wish our profound our indebtedness to Professor Dr. Md. Ismail Jabiullah, Professor, Department of CSE Daffodil International University, Dhaka. Deep Knowledge & keen interest of our supervisor in the field of NLP to carry out this project. His endless patience ,scholarly guidance ,continual encouragement , constant and energetic supervision, constructive criticism , valuable advice ,reading many inferior draft and correcting them at all stage have made it possible to complete this project.

We would like to express our heartiest gratitude to Dr, Touhid Bhuiyan, Professor, and Head**,** Department of CSE, for his kind help to finish our project and also to other faculty member and the staff of CSE department of Daffodil International University.

We would like to thank our entire course mate in Daffodil International University, who took part in this discuss while completing the course work.

Finally, we must acknowledge with due respect the constant support and patients of our parents.

ABSTRACT

E-commerce has become one of the most commodious methods of shopping because of the technological revolution. Research shows internet shopping is much preferred by people rather than the traditional mode of shopping. Millions of user-generated comments are posted daily on the web, and analysis of these opinions could be more directive towards the customer's and manufacturers. That makes the Sentiment analysis of online reviews one of the most sought-after research topic. This paper portrays our experimental work on domain-specific feature-based sentiment analysis of product review. In this paper, we worked with some fixed predefined core features of a product for presenting the customer's acceptance of the principal attributes of a product so that the manufacturer can improve the basic features quality. We have proposed a feature-oriented sentiment prediction scheme. That analyses the generated expressions from the textual reviews of a product for predicting sentiment and assigns scores for our predefined features to present a net sentiment profile of a product of all parameters. With 92% accuracy our sentiment detection scheme is proved to be an effective ways to highlight the core attributes that are seems to be the most to the purpose to the customer and manufacturer.

TABLE OF CONTENTS

CONTENTS	PAGE
Board of examiners	i
Declaration	iii
Acknowledgements	iv
Abstract	v
List of Figures	viii
List of Tables	ix
CHAPTERS	
CHAPTER 1: INTRODUCTION	**1-3**
1.1 Introduction	1
1.2 Motivation	2
1.3 Rationale of the study	2
1.4 Research Question	2
1.5 Expected Output	3
1.6 Report Layout	3
CHAPTER 2: LITERATURE REVIEW	**4-8**
2.1 Related Works	4
2.2 Research Summary	5
2.3 Comparative Analysis	6
2.4 Scope of the Problem	7

2.5 Challenges 8

CHAPTER 3: RESEARCH METHODOLOGY **9-16**

3.1 Research Subject and Instrumentation 9

3.2 Data Collection Techniques 9

3.3 Methodology and Data Analysis 10

3.3.1 Word Corpus 11

3.3.2 Feature Selection 11

3.3.3 Data Pre-processing 12

3.3.4 Generating Expression 13

3.3.5 Rule Based Approach for Sentiment Analysis 14

3.4 Statistical Analysis 15

CHAPTER 4: IMPLEMENTATION **17-18**

4.1 Experimental Results 17

4.2 Discussion 18

CHAPTER 5: CONCLUSION **19-20**

5.1 Summary 19

5.2 Conclusion 19

5.3 Future Scope 19

REFERENCES **21**

LIST OF FIGURES

FIGURES	PAGE NO
Figure 3.1: Raw data	9
Figure 3.2: Block diagram of our sentiment analysis technique	10
Figure 3.3: Sentiment Profile of Champ Diaper Pail	16

LIST OF TABLES

TABLES	PAGE NO
TABLE 2.1: Comparative analysis with other	6
TABLE 3.1: Lexicon categories	11
TABLE 3.2: Pre-processing by removing	12
TABLE 3.3: Pre- processing by splitting up	12
TABLE 3.4: Generated expression	13
TABLE 3.5: Sentiment Profile of champ diaper pail	15
TABLE 4.1: Sentiment prediction result	17
TABLE 4.2: Number of true positive, true negative, false positive, false negative	17

CHAPTER 1

Introduction

1.1 Introduction

Online shopping has become a well-liked looking technique ever since the net has declared a takeover. The increase in technology provides sensible opportunities to the vendor to succeed in the client in a lot of quicker, easier and in economic manner. Online shopping is rising in no time in recent years [1]. According to a survey, 81% of Internet users have looked up a product online at least once. [2].

Due to the wide range of online application, analyzing the sentiment of online reviews and other user-generated content are sought-after research concern now a days. Sentiment analysis examines all comments according to various granularity, sentence level, document level, and function-based sentiment analysis.

Aspect-level or feature-based sentiment analysis performs an in depth analysis in business for analyzing their client feedback knowledge, so that they will learn extra concerning their clients and emphasis better quality products to meet their needs. In feature-based sentiment analysis, every users comment is processed to determine the polarity associated with each feature or attribute [3].

An evaluation on smartphones portrayed that, nearly 50% users of smartphone do not use their phones to their full potential [4]. But over the time as technological advancement we have been experiencing different new features of smartphone, like: numerous built-in cameras, 5G-compatible handsets, foldable gadgets, as well as holographic screen projections and an expanded AI focus. An research on smartphone owners in the US and UK aged 10-64 showed that, these fancy new features may produce the biggest buzz, however enhancements to the core phone features are actually likely to drive additional sales[5].

Every product is designed and manufactured with some basic attributes and different brand merely provide some form of differentiation against their key competitors. For example: Smartphone; there are always some common attribute that most smartphones provide, and user usually look for the features while owning a phone. Like camera, storage capacity, Speed, battery backup, display and so on.

1.2 Motivation

To enable a product manager to understand to understand customer emotions in their marketing campaigns, Sentiment analysis is a powerful marketing tool. It is an important factor when it comes to product and brand recognition, customer loyalty, customer satisfaction, advertising and promotion's success, and product acceptance. Understanding the psychology of consumers can help product managers and customer success managers to alter their product roadmap with greater precision. The term emotion-based marketing is a broad umbrella phrase that encompasses emotional customer responses, such as "positive," "negative," "neutral," "negative," "uptight," "disgust," "frustration," and others. Understanding the psychology of customer responses can also increase product and brand recall.

1.3 Rationale of the study

Due to the wide range of online application, analyzing the sentiment of online reviews and other user-generated content are sought-after research concern now a days. There are different types of sentiment analysis; fine-grained sentiment analysis, emotion detection sentiment analysis, aspect-based analysis and intent analysis. In this work we have done a combinational work of emotion detection sentiment analysis and aspect-based analysis for better sentiment analysis output. As the emotion detection sentiment analysis is a sophisticated way of identifying the emotion in a piece of text, where Lexicons and machine learning are used to determine the sentiment and aspect-based analysis usually emphasis on the features of a product, so that the product's owner can understand how customers feel about specific attributes of the product. Our main aim after this empirical work is to come up with a better sentiment, so that we can make proper benefit for the customers review.

1.4 Research Question

Research Question 1: Does every user generated comment have sentiment such as positive or negative?
Research question 2: How to identify the basic features of a product?
Research question 3: Is it really worthy to analyze the sentiment based on fixed features?

1.5 Expected Output

In this paper, we worked with the core features of a product of a specific domain, like a baby diaper pail. Rather than focusing on the detailed features of a product, the work is intended to find out the customer's acceptance of the principal attributes. And present the sentiment profile through a graph. So that the product's company can emphasize the product's basic attributes quality.

1.6 Report Layout

The rest of the report is structured as follows. In chapter 2 the review works of our work, scope of the problem and challenges are presented. The Research methodology of our experimental work including data collection procedure, Statistical analysis are described in chapter 3. Chapter 4 describes the implementation and brief discussion of the analysis. Chapter 5 describes the comparative analysis of our work and finally the summary of the empirical research, conclusion and future scope are described in chapter 6.

CHAPTER 2

Literature Review

2.1 Related Works

NLP-based approach and Statistical model-based approach for creating rules are two major approaches for performing sentiment analysis. In this connection, Liu et al. First, apply text mining to extract attributes from the reviews and aggregate user reviews. Then for writing rules they used NLP approach. Here, reviews from two new apps were used: Widget apps in the Brain & Puzzle category and game apps in the Personalization category. In total they extracted six hundred textual reviews for their analysis.

SAS® Sentiment Analysis Studio 12.1 was used for their sentiment analysis, they found the SAS® Sentiment Analysis Studio 12.1 very useful as it provides the facility of defining the synonym list of products and features. They did not spell check and no stop list was used in the study. Finally, they come up with a precision of 86 percent and 94 percent respectively on the positive and the negative directory for testing the widget app dataset. And respectively the precision of 94 percent precision and 90 percent precision on positive and negative directory for testing game app dataset. [6]

An experimental work has been done on movie reviews, a domain specific feature-based sentiment analysis. They analyzed reviews of 100 Hindi movies 10 reviews for each and reviews are flagged manually by them to evaluate the performance of the formulation of aspect-oriented algorithms. The result was satisfactory as their proposed scheme performed more exact and centered opinion profile than the straightforward document-level assumption examination [7].

Using a method based on machine learning Fang Luo, Cao et al. has done text sentiment analysis. And for textual representation, a vector space model of sentiment is used for solving the problem of lack of data. But their method is failed in word level sentiment analysis, is well performed in both document level and sentence level [8].

Sometime Sentiment analysis become very challenging task, when same word used for expressing different sentiment. A single punctuation may change a sentences sentiment. For example, “He loves this mobile?” is different from “He loves this mobile!”. Different emotions can be expressed by the same sentence with different punctuation. Focused on this two research was done using Chinese review. First they proposed a new weighting algorithm and a performance-based vector model for sentiment analysis of Chinese product reviews. In the review text, they considered both the modifying relationships between words and punctuation marks and used adverbs of the degree to indicate the intensity of mood. Lizhen et al. has proposed a feature based vector model having six tuple, specially

focused on adverbs of degree and punctuations for expressing the intensity of polarity more accurately. Data sets from three different domain were used here [9] [10].

A fuzzy based product review analysis was performed by Indhuza et al. to extend the feature based classification approach which was evaluated with SFU corpus. They have collected and used 2000 user generated product comments from different websites. The general accuracy of their system was 85.58%. The representation of input data really varies the accuracy of their proposed system. For performing the classification no labeled data set for training is required for the feature-based approach. To improve the performance focus on entity recognition, co-reference resolution and domain dependency is needed [2].

A feature based sentiment analysis was done with twitter reviews. At first they select some features from the tweets and a sentiment analysis is conducted on tweets only containing those features. This proposed method avails 40% more precision compared to features extracted directly from tweets. They have done this analysis for multiple languages [11]. Another feature based analysis performed by proposing the use of SVM in new dimension as a classifier and they resolve the negation problem [12]. Another feature based analysis was done in Arabic language for extracting and weighting features and detecting sentiments from Arabic reviews [3].

A back tracking approach has been introduced for detecting sentiment of Bengali text based on a hypothesis that people most of the time express their actual emotion in the last of the sentence. For that they have proposed a backtracking algorithm and got more than 70% accuracy. Their proposed methods performance varies from the input data [13]. Another Sentiment analysis of YouTube song comment was done by using mentioned back tracking approach with slight change in the algorithm [14].

2.2 Research Summary

Research is an organized way to find solutions of existing problems or problems that nobody has worked on before. It can be used for solving a new problem or it can be the expansion of past work on any particular field. Our Research is on detecting sentiment polarity from the user generated comments, that is associated with NLP(Natural Language Processing).AI(Artificial Intelligence) is challenging the human being to exceed human beings performance. There's been lots of work that has already done to detect sentiment as well as emotion using texts or documents from various languages. We have studied lots of paper related to detecting sentiment and emotion from text, lyrics, sentence etc. They used different methods and among them we have chosen a rule based method for sentiment analysis. For that reason we have used a strong corpus that will match the words according to the sentiment. At first an expression will be generated from the comments and then the

sentiment will be predicted, whether it is positive or negative or neutral. If it's positive then a point will be assigned for the specific feature and finally the sentiment profile of that product will be presented through a graph.

2.3 Comparative Analysis

All the three papers mentioned below are concentrated on feature-based sentiment analysis. TABLE 2.1 presents a comparative analysis of our work with three other papers that lay with the same research focus.

TABLE 2.1: Comparative analysis with other

Attributes	Paper 1	Paper 2	Paper 3	Proposed method
Analyze with Fixed features	NO	NO	NO	YES
Use POS tagging for Feature Extraction	YES	YES	NO	NO
Expression generation form text	NO	NO	NO	YES
Use new approach to predict sentiment	NO	NO	YES	YES
Use of precision and recall	YES	YES	YES	YES

Paper 1 *Mining Opinion Features in Customer Reviews
Paper 2 *Sentiment analysis of movie reviews: A new feature-based heuristic for aspect-level sentiment classification
Paper 3 *Feature-based Sentiment Analysis Approach for Product Reviews

In our paper, we worked with some core features of a product unless analyzed all the features of a product. For that background analysis was required to determine the fixed features. Paper 1 had identified two kinds of features from analyzing users' comments: the most frequent and infrequent features. Paper 2 has made an extensive search for identifying the aspects in different film awards, movie review sites, and film magazines and worked out a list of aspects. Paper 3 has used a feature-based six-tuple vector model which contains both features and opinions that describe reviewers' opinions on the features.

We have not used POS tagging for feature extraction. As we analyzed with fixed features so that we have added feature words in our corpus. We have four kinds of words in our

corpus: positive, negative, negation, and feature word. Paper 1 and 2 have used POS tagging for extracting features from the user comments. Paper 1 additionally used a human tagger to evaluate the discovered features, by reading manually all the reviews and producing a manual feature list for each product. Paper 3 has proposed an algorithm for extracting feature words from a sentence.

For analyzing sentiment we had an intermediate step called generating expression. We generate an expression for every sentence. We have denoted our words with a single alphabet to generate expression. Like- P for Positive word, N for negative word, and B for negation word.

Example: BPBP is the expression of "The bags never fit right and then it never spun right."

The bags never fit right and then it never spun right

B P B P

Finally, sentiment will be predicted by analyzing these expressions. Paper 1, 2 & 3 don't generate any kinds of expression.

We have used a new and simple rule-based approach for predicting the sentiment of a comment. Paper 2 used SentiWordNet based approach to determine every sentence sentiment polarity, on the other side paper 3 used a new weighting algorithm for sentiment analysis of Chinese reviews, and paper 1 used their proposed algorithm.

All of the papers used precision and recall for the performance calculation of their empirical work.

2.4 Scope of the Problem

Sentiment analysis from text is incipiently a content based classification which expatiate the concept from Natural language processing including Machine Learning as well. Sentiment analysis is an interesting field of study. Nowadays, it has been adding values to the industry as because sentiment analysis bases its results on factor that are so inherently humane, it is bound to become one of the fundamental drivers of many business decisions in future. Improved accuracy and consistency in text mining techniques can helps to overcome the current problems. Currently, as the next wave of knowledge discovery, text analysis is achieving high commercial values. Because of the wide range of online application, analyzing the sentiment of online reviews and other user-generated content are become a great research concern now a days. With the technological advancement and increasing user's dependency on web marketing are creating great scope for the researcher to contribute more. As because Sentiment Analysis is a quite challenging task, since the

expression in natural language is complicated and changes from time to time, no method can exhaustively deal with every situation.

2.4 Challenges

For some comments our proposed algorithm failed to predict any sentiments. As our system generate expressions from the lexicons, so if a comment does not contain any positive, negative or negation word, our system failed to predict any sentiment. But sometimes the comments express very strong sentiment.
For example - “Holds a LOT of diapers--a full week of toddler diapers!”
This comment strongly reveal the capacity of the Diaper pail but our system fail to predict anything though it could not generate any expression from it.
Another example- “The Champ uses regular store bags.”
This comment is explain the simplicity of this diaper pail in case using bags. But as the sentence is not contains any sentiment word our system fail to predict sentiment.

CHAPTER 3

Research Methodology

3.1 Research Subject and Instrumentation

Our research subject is sentiment analysis which is a process of computationally identifying and categorizing opinions expressed in a piece of text, especially in order to determine whether the writer's attitude towards a particular topic, product, etc. is positive, negative, or neutral. Sentiment analysis (or opinion mining) is a natural language processing (NLP) technique, it is often performed on textual data to help businesses monitor brand and product sentiment in customer feedback, and understand customer needs. Our research goal is to present a sentiment profile of a product so that, people who wants to buy a particular product don't need to scroll down all the users review. The sentiment profile will show the sentiment statistics of different features.

3.2 Data Collection Techniques

Ding, Liu and Yu, WSDM-2008 had produced a data set containing more than 1000 user's generated comments from online platform based on some specific domain [15]. They used this data set for improving the performance of a lexicon based approach. In their data set there were data from 9 different domain. We have chosen data from one domain for our analysis. Figure 3.1 shows the raw comments used in analysis.

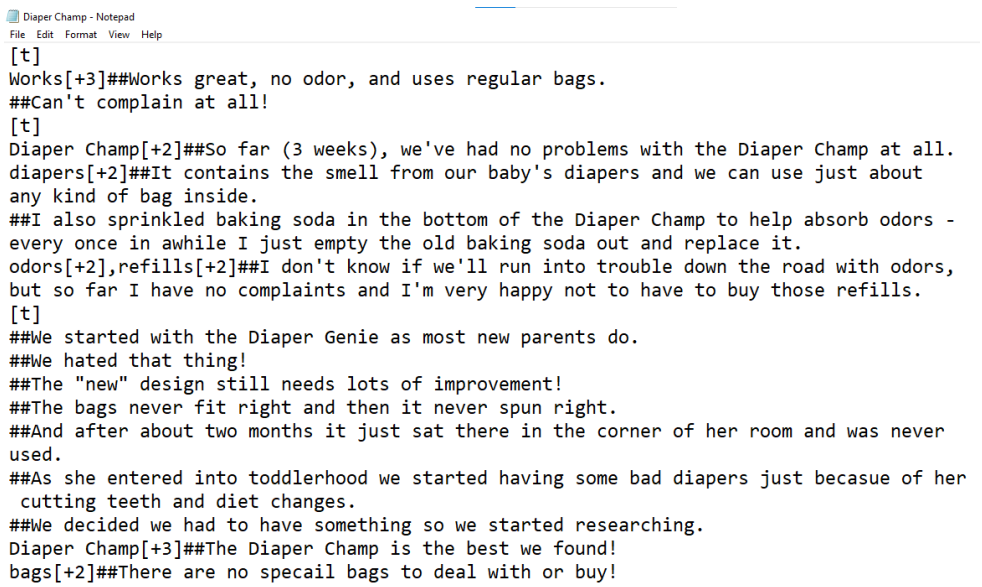

Diaper Champ - Notepad
File Edit Format View Help
[t]
Works[+3]##Works great, no odor, and uses regular bags.
##Can't complain at all!
[t]
Diaper Champ[+2]##So far (3 weeks), we've had no problems with the Diaper Champ at all.
diapers[+2]##It contains the smell from our baby's diapers and we can use just about
any kind of bag inside.
##I also sprinkled baking soda in the bottom of the Diaper Champ to help absorb odors -
every once in awhile I just empty the old baking soda out and replace it.
odors[+2],refills[+2]##I don't know if we'll run into trouble down the road with odors,
but so far I have no complaints and I'm very happy not to have to buy those refills.
[t]
##We started with the Diaper Genie as most new parents do.
##We hated that thing!
##The "new" design still needs lots of improvement!
##The bags never fit right and then it never spun right.
##And after about two months it just sat there in the corner of her room and was never
used.
##As she entered into toddlerhood we started having some bad diapers just becasue of her
cutting teeth and diet changes.
##We decided we had to have something so we started researching.
Diaper Champ[+3]##The Diaper Champ is the best we found!
bags[+2]##There are no specail bags to deal with or buy!

Figure 3.1: Raw data

3.3 Methodology and Data Analysis

In this feature based sentiment analysis we worked with some fixed attributes of a specific product so that we can present a sentiment profile of the basic features of a products. Figure 3.2 presents the block diagram of our work.

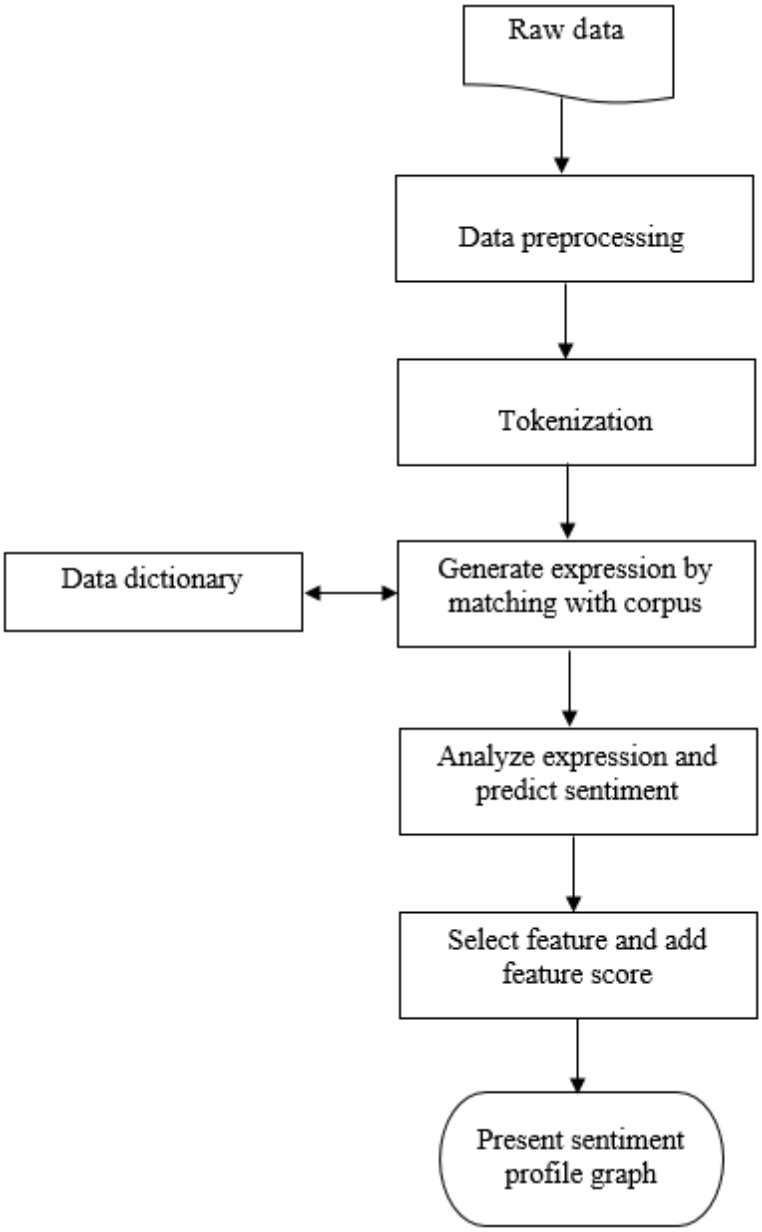


Figure 3.2: Block diagram of our sentiment analysis technique

3.3.1 Word Corpus

For our analysis we used four kinds of words. Positive, Negative, Negation and feature words. As we are working with fixed features so we need feature words for identifying the features. For positive and negative lexicon we have used Minqing Hu and Bing Liu's data set [16]. We have added some positive and negative word in the data set while analyzing the processed comments. And the negation and features words have been added to the dictionary for our analysis.

TABLE 3.1: LEXICON CATEGORIES

Positive	P	good, great, fantastic…
Negative	N	worst, clumsy, expensive..
Negation	B	not, never, no, nothing, nobody….
Feature word	F	odor, bag, size, capacity, price…

3.3.2 Feature Selection

For selecting the features of children diaper pail, first a general study was performed of the basic attributes for which a diaper pail is introduced. And we found some principal attributes of it. Firstly, a diaper pail should have the aptitude to block the odor from children diaper. *Some pail Refills* are designed for providing reliable protection from diaper odors which ordinarily tends to expensive for the user. So second attribute is the way of refilling the pail bag. Thirdly, the size and capacity of a pail. They determine how many dirty diapers it can hold, as it will determine how often it will need to be emptied. And finally the expenses of a pail use. Except these features rest of the comments will be considered under the "product's use" category. From our analysis six selected features are odor blocking, using of pail bag, size, capacity, price and product's use.

3.2.3 Data Pre-processing

For our experiment, we mainly focused on single-line comments. For long text comments, firstly we minimize the lines that don't contain any sentiment support word. Table I demonstrates the processed user comments after removing the part containing no sentiment word. In the first comment, "We started with the Diaper Genie as most new parent's do." this line is not associate with any kinds of sentiment. That's why we just kept only "We hated that thing!" this line.

TABLE 3.2: PRE-PROCESSING BY REMOVING

Raw data	**Processed data**
1. We started with the Diaper Genie as most new parents do. We hated that thing!	We hated that thing!
2. The bags never fit right and then it never spun right. And after about two months it just sat there in the corner of her room and was never used.	The bags never fit right and then it never spun right.
3. We decided we had to have something so we started researching. The Diaper Champ is the best we found!	The Diaper Champ is the best we found!

And secondly, split up a long text comment into two comments. TABLE II demonstrates how we split up a long text comment. The first comment in the table is been divided into two comments.

TABLE 3.3: PRE-PROCESSING BY SPLITING UP

Raw data	**Processed data**
1. The diaper champ couldn't be easier to use. And the biggest plus is that you don't have to buy any expensive refills.	1(a) The diaper champ couldn't be easier to use.
	1(b) The biggest plus is that you don't have to buy any expensive refills.

2. Diaper Champ is by far the best diaper pail! It holds a lot of diapers, not to mention the odor! It's easy to use (just a flip of the lid, and the diaper is gone!). It's also easy to empty out.	2(a) Diaper Champ is by far the best diaper pail!
	2(b) It holds a lot of diapers, not to mention the odor!
	2(c) It's easy to use (just a flip of the lid, and the diaper is gone!).
	2(d) It's also easy to empty out.

3.3.4 Generating Expression

When an input text is given first an expression will be generated from the given text. In order to generate expression from the input sentences we have denoted the positive, negative and negation words as follow.

Positive ⟶ P
Negative ⟶ N
Negation ⟶ B

And sentiment prediction will done by analyzing the expression. TABLE IV shows the generated expression and selected features from the sentences.

TABLE 3.4. GENERATED EXPRESSION

Input Text	**Expression**	**Feature**
The bags never fit right and then it never spun right	BPBP	Bag
I'm very happy not to have to buy those refills.	PB	Refill
Love that you don't have to purchase expensive refills	PBN	Refill, Price
Inexpensive as baby products go!	P	Price

3.3.5 Rule Based Approach for Sentiment Analysis

Take an expression for analysis

If the expression contain only P
then declare positive
else if P is followed by B and B is followed by B or none
then declare positive

else if P is followed by N and N is followed none
then declare positive

else if P is followed by B and B is followed by P
then declare positive
else if P is followed by B and B is followed by N
then declare positive

If the expression contain only N
then declare negative
else if N is followed by B and B is followed by B or none
then declare negative
else if N is followed by P and P is followed by none
then declare positive

else if N is followed by B and B is followed by N
then declare negative
else if N is followed by B and B is followed by P
then declare negative

If the expression contain only B
then declare negative
else if B is followed by P
then declare negative
else if B is followed by N
then declare negative

else if a string is long
then find the mentioned string above to reform a string to find the result

3.4 Statistical Analysis

Statistical analysis is one of the core parts while doing any kind of experiment. It uses statistics as a research methodology. It makes easier to get approximate solutions for any research or experiment when the actual process to sort out the result is very highly complicated or unknown in its true form. The research which we have experimented consists of lots of complexity. As a result, by applying statistical analysis, it makes quite simple to sort out an approximation of a satisfactory outcome.

For receiving sentiment profile of a product, our proposed aspect-level sentiment analysis algorithm is a unique way. The Resulting sentiment profile is extremely useful for users, informative and easy to understand. Once sentiment of a text is predicted then it will score a feature. For every feature there are two types of scoring. Positive scoring and negative scoring. TABLE 3.5 presents the table representation of Sentiment Profile of Champ Diaper Pail and Figure 3.3 Presents the graphical representation.

TABLE 3.5 SENTIMENT PROFILE OF CHAMP DIAPER PAIL

Features	**Number of Positive comments**	**Number of Negative comments**
Odor	63	51
Bag refill	57	27
Size	0	3
Capacity	3	0
Price	27	0
Product’s use	105	45

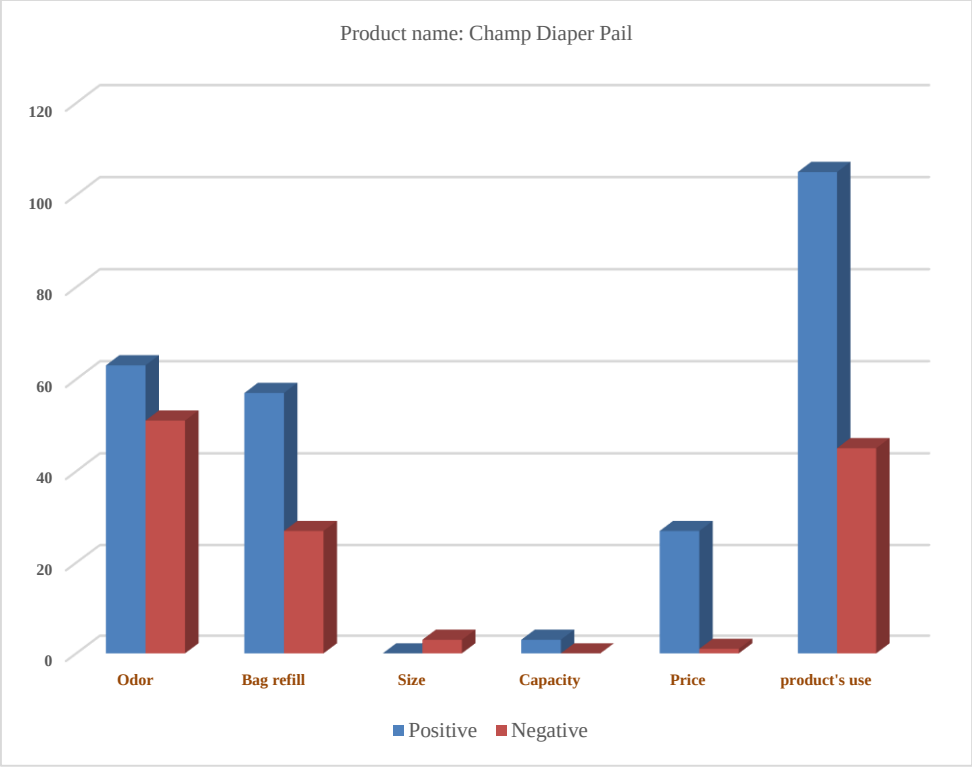


Figure 3.3: Sentiment Profile of Champ Diaper Pail

CHAPTER 4

Implementation

4.1 Experimental Results

Our fixed featured based analysis on specific domain is done using 354 user generated product reviews. We have used 4 types of lexicon for generating expression for sentiment detection purpose. Among 354 comments our algorithm perform 282 right prediction and 36 wrong prediction.TABLE VI & TABLE VII our evaluation result.

TABLE 4.1 SENTIMENT PREDICTION RESULT

Total Comments	Right Prediction	Wrong prediction	Failed to predict
354	282	36	36

TABLE 4.2 NUMBER OF TRUE POSOTIVE, TRUE NEGATIVE, FALSE PPASITIVE AND FALSE NEGATIVE

TP	TN	FP	FN
207	75	6	30

Here, TP is true positive which indicating the number of correctly classified sentences as positive, FP is false positive which indicating the number of incorrectly classified sentences as positive, TN is true negative which indicating the number of correctly classified sentences as negative and FN is false negative which indicating the number of incorrectly classified sentences as negative.

Recall, precision and F-measure are used for evaluating the performance of our algorithm, which are defined as follows:

$$Precision = \frac{TP}{TP+FP} \quad (1)$$

$$Recall = \frac{TP}{TP+FN} \quad (2)$$

$$F\text{-}Measure = \frac{2*Precision*Recall}{Precision+Recal} \quad (3)$$

With 97.2% precision and 87.3% recall, our Empirical results shows 92% accuracy of our proposed algorithms performance. Empirical results indicate that our sentiment detection scheme an effective ways to highlight the core features that are the most relevant to the customer and manufacturer. The performance of the algorithm varies from the input data representation

4.2 Discussion

For some comments our proposed algorithm failed to predict any sentiments. As our system deals with generated expression, so if a comment does not contain any positive, negative or negation word, our system failed to predict any sentiment. But sometimes the comments depict very strong sentiment.
For example-“Holds a LOT of diapers--a full week of toddler diapers!”
This comment strongly reveal the capacity of the Diaper pail but our system fail to predict anything though it could not generate any expression from it.

CHAPTER 5

Conclusion

5.1 Summary

In our work, we worked with the core features of a product of a specific domain, like a baby diaper pail. Every product is designed and manufactured with some basic attributes and different brand merely provide some form of differentiation against their key competitors. For example: Smartphone; there are always some common attribute that most smartphones provide, and user usually look for the features while owning a phone. Like camera, storage capacity, Speed, battery backup, display and so on. Rather than focusing on the detailed features of a product, the work is intended to find out the customer's acceptance of the principal attributes. So that the product's company can emphasize the product's basic attributes quality.

5.2 Conclusion

Sentiment analysis is complicated and a quite challenging task as the expression in natural language changes in no time, no method can exhaustively deal with every situation. In our work we focused on fixed features of a product so that user can only get a clear idea about the basic attributes from the products comment. To do so, an experiment is performed with a generic proposed algorithm. In this algorithm, for sentiment analysis first the system generates expression of the comments with three kinds of word from our dictionary and identify the feature word from the text. After applying our sentiment prediction algorithm for every predefined features, number of positive and negative comments are calculated to depict the sentiment profile graph. With 92% accuracy the algorithm used for feature based sentiment analysis is very simple and no previous training is required. It can be used to produce very detailed and useful sentiment profile of basic features of any product.

5.3 Future scope

1. Will deal with some sentences having no sentiment words but still containing strong sentiment.
2. Will deal with huge user's comments from different domain.
3. Will compare our simple algorithm performance with other established algorithms.
4. We have done our preliminary experiment on very small amount of data. We will work with large amount of data so that the result will be more accurate.

5. We want to integrate our system with text to speech converter in order to build something helpful for blind people

References

[1] Lopes, Privanka Nazreth, Tanava Prabhutendulkar, and Deepa Yogesh Naik. "A Comparative Study On Online v/s Offline Shopping of Eyewear." *Reflections-Journal of Management* 5 (2015).

[2] Indhuja, K., and Raj PC Reghu. "Fuzzy logic based sentiment analysis of product review documents." In *2014 First International Conference on Computational Systems and Communications (ICCSC)*, pp. 18-22. IEEE, 2014.K. Elissa, "Title of paper if known," unpublished.

[3] Abd-Elhamid, Laila, Doaa Elzanfaly, and Ahmed Sharaf Eldin. "Feature-based sentiment analysis in online Arabic reviews." In *2016 11th International Conference on Computer Engineering & Systems (ICCES)*, pp. 260-265. IEEE, 2016.Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, "Electron spectroscopy studies on magneto-optical media and plastic substrate interface," IEEE Transl. J. Magn. Japan, vol. 2, pp. 740–741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].

[4] smartphones-come-many-features-majority-used. available at <<https://blogs.oregonstate.edu/econ439/2014/01/17/ />>. last accessed on 10-19-2021 ar 12:00 PM.

[5] smartphone-features-consumers , available at <https://blog.gwi.com/chart-of-the-week//>> ,last accessed on 10-19-2021 ar 12:00 PM.

[6] Liu, Jiawen, Mantosh Kumar Sarkar, and Goutam Chakraborty. "Feature-based Sentiment analysis on android app reviews using SAS® text miner and SAS® sentiment analysis studio." In *Proceedings of the SAS Global Forum 2013 Conference*, vol. 250. 2013.

[7] Singh, Vivek Kumar, Rajesh Piryani, Ashraf Uddin, and Pranav Waila. "Sentiment analysis of movie reviews: A new feature-based heuristic for aspect-level sentiment classification." In *2013 International Multi-Conference on Automation, Computing, Communication, Control and Compressed Sensing (iMac4s)*, pp. 712-717. IEEE, 2013.

[8] Luo, Fang, Cheng Li, and Zehui Cao. "Affective-feature-based sentiment analysis using SVM classifier." In *2016 IEEE 20th International Conference on Computer Supported Cooperative Work in Design (CSCWD)*, pp. 276-281. IEEE, 2016.

[9] Lizhen, Liu, Song Wei, Wang Hanshi, Li Chuchu, and Lu Jingli. "A novel feature-based method for sentiment analysis of Chinese product reviews." *China communications* 11, no. 3 (2014): 154-164.

[10] Wang, Hanshi, Lizhen Liu, Wei Song, and Jingli Lu. "Feature-based Sentiment Analysis Approach for Product Reviews." *J. Softw.* 9, no. 2 (2014): 274-279.

[11] Erdmann, Maike, Kazushi Ikeda, Hiromi Ishizaki, Gen Hattori, and Yasuhiro Takishima. "Feature based sentiment analysis of tweets in multiple languages." In *International Conference on Web Information Systems Engineering*, pp. 109-124. Springer, Cham, 2014.

[12] Devi, DV Nagarjuna, Chinta Kishore Kumar, and Siriki Prasad. "A feature based approach for sentiment analysis by using support vector machine." In *2016 IEEE 6th International Conference on Advanced Computing (IACC)*, pp. 3-8. IEEE, 2016.

[13] Rabeya, Tapasy, Sanjida Ferdous, Himel Suhita Ali, and Narayan Ranjan Chakraborty. "A survey on emotion detection: A lexicon based backtracking approach for detecting emotion from Bengali text." In *2017 20th international conference of computer and information technology (ICCIT)*, pp. 1-7. IEEE, 2017.

[14] Rabeya, Tapasy, Narayan Ranjan Chakraborty, Sanjida Ferdous, Manoranjan Dash, and Ahmed Al Marouf. "Sentiment analysis of bangla song review-a lexicon based backtracking approach." In *2019 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT)*, pp. 1-7. IEEE, 2019.

[15] From, G. M. R. K. B. "Jump to: navigation, search (Ding et al., 2008)⇒ Xiaowen Ding, Bing Liu, and Philip S. Yu.(2008)."A Holistic Lexicon-based Approach to Opinion Mining." In: Proceedings of the International Conference on Web Search and Web Data Mining (WSDM 2008). doi: 10.1145/1341531.1341561 Subject Headings: Opinion Mining, Sentiment Analysis, Context-dependent Word, Product Mention, Lexicon-based Algorithm."2

[16] Hu, Minqing, and Bing Liu. "Mining and summarizing customer reviews." In *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 168-177. 2004.

183-25-704

ORIGINALITY REPORT

17%	11%	8%	6%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	raw.githubusercontent.com Internet Source	3%
2	A. A. Pinto, D. A. Rand. " Classifying structures on dynamical fractals: 1. The moduli space of solenoid functions for Markov maps on train tracks ", Ergodic Theory and Dynamical Systems, 2008 Publication	1%
3	Submitted to Daffodil International University Student Paper	1%
4	www.ijetcse.com Internet Source	1%
5	www.amazon.com Internet Source	1%
6	Submitted to Aristotle University of Thessaloniki Student Paper	1%
7	K Indhuja, Raj P C Reghu. "Fuzzy logic based sentiment analysis of product review documents", 2014 First International	1%