

A Project Report on

**Understanding of Complaints and Praises of Woohoo Gift card – Google Reviews**

Submitted in partial fulfilment for award of degree of

**PGDM**

In **Business Analytics**

Submitted by

**HARSHA GV**

R19MBA83

Under the Guidance of

**Jay B.Simha**

Chief Mentor

REVA Academy for Corporate Excellence

**REVA University**

Rukmini Knowledge Park, Kattigenahalli,

Yelahanka, Bangalore – 560064

**<Month, Year>**



# Candidate’s Declaration

I, HARSHA GV hereby declare that I have completed the project work towards the first year of Diploma in Business Analytics at, REVA University on the topic entitled “Understanding of Complaints and Praises of “Woohoo Gift card” – Google Reviews” under the supervision of Jay B.SIMHA and chief mentor. This report embodies the original work done by me in partial fulfilment of the requirements for the award of degree for the academic year 2021.

Place: Bengaluru Name of the Student:

HARSHA GV

Date: Signature of Student



# Certificate

This is to Certify that the Project work entitled “Understanding of Complaints and Praises of “Woohoo Gift card” – Google Reviews” carried out by HARSHA GV with R19MBA83, is a bonafide student of REVA University, is submitting the first year project report in fulfilment for the award of PGDM in Business Analytics during the academic year 2021. The Project report has been tested for plagiarism, and has passed the plagiarism test with the similarity score less than 15%. The project report has been approved as it satisfies the academic requirements in respect of PROJECT work prescribed for the said Degree.

Signature of the Guide Signature of the Director

Name of the Guide Jay B.Simha Name of the Director Shinu

Guide Director

External Viva

Names of the Examiners

1. <Name> <Designation> <Signature>
2. <Name> <Designation> <Signature>

Place: Bengaluru

Date:



# Acknowledgement

I would like to thank each and every one part of the Reva for giving me the opportunity to complete my PGDM course and Project report on “Understanding of Complaints and Praises of “Woohoo Gift card” – Google Reviews”

I convey my sincere gratitude for the support provided by Hon’ble Chancellor, Dr. P Shayma Raju, Hon’ble Vice Chancellor, Dr. M. Dhanamjaya, Registrar, Dr. N. Ramesh, Director Shinu, Chief Mentor Jay B.Simha, Trainers Ratnakar, Mithun, Pannendra, IBM Trainers, Program office Pullikeshi, Ravi, coordinators Tousif, Hasitha, Bhumika, Finance - Shobha , IT, Security, Cafeteria, Cleaning staff, and BA07 classmates.

Place: Bengaluru

Date:



# Similarity Index Report

This is to certify that this project report titled <Title of the project> was scanned for similarity detection. Process and outcome is given below.

Software Used:

Date of Report Generation:

Similarity Index in %:

Total word count:

Name of the Guide:

Place: Bengaluru Name of the Student: HARSHA GV

Date: Signature of Student

Verified by:

Signature

Dr. Shinu Abhi,

Director, Corporate Training

# List of Abbreviations

|  |  |  |
| --- | --- | --- |
| **Sl. No** | **Abbreviation** | **Long Form** |
| 1 | SVM | Support Vector Machine |
| 2 | NRC | The National Research Council Canada's (NRC) Emotion Lexicon |
| 3 | TFIDF | Term Frequency -Inverse Document Frequency |

# List of Figures

|  |  |  |
| --- | --- | --- |
| **No.** | **Name** | **Page No.** |
| Figure No.6.1 | Flow chart of business understanding | 16 |
| Figure No.8.1 | Word cloud – Overall Woohoo reviews | 22 |
| Figure No.8.2 | Word cloud – Complaint Woohoo reviews | 24 |
| Figure No.8.3 | Word cloud – Praise Woohoo reviews | 26 |

# List of Tables

|  |  |  |
| --- | --- | --- |
| **No.** | **Name** | **Page No.** |
| Table No.2.1 | Literature reviewed | 11 |
| Table No.6.1 | Feature Variables details | 17 |
| Table No.7.1 | Data Variables details | 19 |
| Table No.7.2 | Sample data | 19 |
| Table No.8.1 | Data Variables and options | 20 |
| Table No.8.2 | Trend 2 class Sentiment | 28 |
| Table No.8.3 | Trend 5 class Sentiment | 29 |
| Table No.8.4 | Cross Tab – Validation of Lexicon | 30 |
| Table No.8.5 | Trend analysis- Word count and Length- 2 class | 31 |
| Table No.8.6 | Trend analysis- Word count and Length- 5 class | 31 |
| Table No.8.7 | Features by 5 class Sentiment | 32 |
| Table No.8.8 | Trend analysis- Features- Complaint | 33 |
| Table No.8.9 | Trend analysis- Features- Praises | 34 |
| Table No.9.1 | Model approach | 35 |
| Table No.10.1 | Model output | 36 |
| Table No.10.2 | Classification report – 2 Class | 37 |
| Table No.10.3 | Confusion Matrix – 2 Class | 37 |
| Table No.10.4 | Classification report – 5 Class | 39 |
| Table No.10.5 | Confusion Matrix – 5 Class | 39 |
|  |  |  |

# Abstract

This is a short study on analysing the sentiments towards the brand woohoo gift, earlier studies analysed text only with 2 scale sentiments negative only and not negative. 5 scale sentiments neutral, praise, positive only, complaint and negative only, gives better view on feeling of the customers. Complaints and Praise subset provides us more informative and in-depth understanding on the customers feeling. Our analysis shows that praises and complaints have noticeable difference when compared with negative only or positive only reviews.

Keywords :- Text Mining, Sentiment analysis

# Table of Contents

[Candidate’s Declaration 2](#_Toc47857456)

[Certificate 3](#_Toc47857457)

[Acknowedgement 4](#_Toc47857458)

[Similarity Index Report 5](#_Toc47857458)

[List of Abbreviations 6](#_Toc47857458)

[List of Figures 6](#_Toc47857459)

[List of Tables 6](#_Toc47857460)

[Abstract 7](#_Toc47857461)

[Chapter 1: Introduction 9](#_Toc47857462)

[Chapter 2: Literature Review 11](#_Toc47857463)

[Chapter 3: Problem Statement 12](#_Toc47857464)

[Chapter 4: Objectives of the Study 13](#_Toc47857465)

[Chapter 5: Project Methodology 14](#_Toc47857466)

[Chapter 6: Business Understanding 15](#_Toc47857467)

[Chapter 7: Data Understanding 16](#_Toc47857468)

[Chapter 8: Data Preparation 17](#_Toc47857469)

[Chapter 9: Data Modeling 18](#_Toc47857470)

[Chapter 9: Data Evaluation 19](#_Toc47857471)

[Chapter 10: Deployment 20](#_Toc47857472)

[Chapter 11: Analysis and Results 21](#_Toc47857473)

[Chapter 12: Conclusions and Recommendations for future work 22](#_Toc47857474)

[Bibliography 23](#_Toc47857475)

[Appendix 24](#_Toc47857476)

[Plagiarism Report 25](#_Toc47857477)

[Publications in a Journal/Conference Presented/White Paper 25](#_Toc47857478)

[Any Additional Details 26](#_Toc47857479)

# Chapter 1: Introduction

**Background**

Gift cards, also known as gift certificates, gift vouchers, or gift tokens, are generally issued by a retailer or bank. In addition, gift cards are used as a developmental strategy for promoting businesses, attract new customers, increase brand awareness, reduce fraud, and improve business sales. Moreover, a gift card can be utilized as an alternative source for money at a particular store or business store for purchases related to goods or services.

Area of your work – Woohoo Gift Card India \_ Google reviews

Company “Qwikcilver solutions Pvt Ltd” brand Woohoo gift card has more than 75% Market share in India, the Indian market is expected to grow to 9.34Billion by 2024. and Global industry is expected to grow by 2076 billion by 2027.

**Current status**

The need for study of google reviews given for the Woohoo Gift card application in the google app store. As the brand is the number one Gift card application in India, even though it has been in India for more than 10 Years, Still the consumers are not fully aware of the brand or concept of digital and physical gift cards. To cater to the customers better and grow faster analyze the google reviews.

**Why this study**

The scope of the study is to use the Woohoo google reviews to do Text analysis and understand the Customer sentiment complaints and praises.

In the Text analysis, social media and data available in the public domain studied to helps us understand the customer behavior, expectations, needs, and where the brand is failing to fulfill.

The study will help the organization and the gift card industry to understand from the sentiment analysis what are the factors that drive customers to buy the product, be loyal or move away from the brand due to bad delivery of service.

Sentiment analysis research is very effective in predicting positive and negative polarity ratings at different granularity at the word count used, length review written, product features performance over a period of time.

# Chapter 2: Literature Review

Referred the Linguistic Understanding of Complaints and Praises in User Reviews,

written by Kavitha Ganeshan, the intuition is that praise sentences and complaints tend to be more informative than plain positive only or negative only sentences. This paper thus tries to understand the properties of such text that considered as complaints and praises.

Summary of this Literature

The study has the following analysis done: -

1. In the paper they defined the Negative only, Complaint, Positive only and Praises.
2. For the defined sentiments they have shown the Average number of words, and Average length Analysis.
3. Noun and adjective usage: For example, a negative only sentence such as ‘the screen is bad’ or a complaint such as ‘the screen is not clear’ both have nouns (‘screen’) and adjectives (‘bad’ and ‘clear’). Both the noun and adjectives play a role in indicating negative sentiment.
4. Paste tense analysis: By each sentiment count of Past tense words used.
5. Negation analysis: For example, consider the negation in the following sentence: this lasts all night and feels really great on my skin not oily cakey or heavy”.
6. Intensifier usage: For example, to express appreciation on some restaurant service one may say ‘The service was extremely fast and the food was super delicious!’.

In relation to our project , considered Woohoo gift card Google reviews to understand the sentiments , Average number of words , Average review Length and compare the sentiments with product features . To help the management to work towards improving the product service and delight the customers.

**Research Gaps**

There are no research gaps, considered the study/research paper as reference and replicating it as case study to analyze the customers sentiments.

**List of 5 literature reviewd**

Table 2.1: **literature reviewd**

|  |  |  |  |
| --- | --- | --- | --- |
| **Author** | **Year** | **Study** | **Details of the analysis done** |
| Bo Pang and Lillian Lee | 2008 | Opinion mining and sentiment analysis | For the same movie there can be different opinion from 2 people  a) we will analyse individual and statistical accuracy and grade given by them.  b) sentiment polarity classification - overall positive or negative.  c) transforming each review into a feature vector to analyse which feature is important. |
| Hyun Duk Kim, Kavita Ganesan, Parikshit Sondhi, and ChengXiang Zhai. | 2011 | Comprehensive review of opinion summarization (opinion mining survey). | opinion summarization are built upon some closely related  research areas such as sentiment classiﬁcation, text summarization,  opinion summarization are built upon some closely related  research areas such as sentiment classiﬁcation, text summarization,  opinion summarization are built upon some closely related  research areas such as sentiment classiﬁcation, text summarization  Analysed  1) Aspect / feature-based opinion  2) Sentiment Prediction  3) Summary generation  a) statistical summary  b) Text selection |
| Kavita Ganesan and ChengXiang Zhai | 2012 | Opinion-Based Entity Ranking | Key words in the reviews are being used to associate with the product features. Conducting sentiment analysis using Praise words(good) and intensifiers(astonishingly).  Key words in the reviews are being used to associate with the product features. Conducting sentiment analysis using Praise words(good) and intensifiers(astonishingly). |
| Kavita Ganesan and  Guangyu Zhou | 2016 | Linguistic Understanding of Complaints and Praises in User Reviews | Analysing the reviews on 4 scale Praise, Postive only , Complaints and Negative only. a) Sentence length analysis.  b) Noun and adjective usage  c) Past tense analysis  d) Negation analysis  e) Intesifer usage |
| Quratulain Rajput, Sajjad Haider, and Sayeed Ghani | 2016 | Lexicon-Based Sentiment Analysis of Teachers’ Evaluation | Analysed students’ feedback towards the teachers. Sentiment analysis 3scale, positive, negative and neutral.  a) Polarity tagging  b) word frequency  c) word attitude  d) Overall attitude  c) word cloud visualisation. |

# Chapter 3: Problem Statement

**Business Problem**

Gift card industry is available in India for more than a decade. But the product is yet to be completely accepted by customers. In order to serve the customers better and understand the concern areas. Current Penetration is around 5 to 15 %, goal is to **Increase the penetration to 25%** by analysing the customer complaints and Praises using the Google reviews given by customers for the past 7 years.

**Analytics Solution**

For a product to perform and meet the Expected/promised service from the beginning purchase to end service redeem of a gift card needs to be provided, there are multiple factors end-to-end work.

The problem is worth the study as it can help the existing as well any other Gift card companies on how to run the business ensuring that the customer's expectations/needs are met and kept high. Do and Don’ts to delight the customer. This would help in attract, retain and grow customers by working towards improving the customer journey and delivery process.

# Chapter 4: Objectives of the Study

The purpose of the capstone project is to understand the Customers Sentiments towards complaints, Negative only, Positive only and praises and the This would help in attract, retain and grow customers by working towards improving the customer journey and delivery process.

**Primary and Secondary objectives**

**Primary**

Increase the Penetration from current 5- 15% to 25%.

**Secondary**

Finding the areas of concern which needs to be addressed immediately

1. Application
2. Customer service
3. Delivery service
4. Product not meeting customer requirement.

**Expected outcome.**

1. Understand the sentiments trend analysis.
2. How efficiently the Lexicons are able to predict the sentiments.
3. In building the optimized model, having better accuracy in Predicting the Sentiments.
4. Understand the Avg words analysis and Average Length analysis.
5. Understand the improvement areas using SWOT analysis.
6. Based on the outcome, brand can work towards decreasing the complaints and Negative only sentiment. Increase in praise and positive only sentiment.

# Chapter 5: Project Methodology

**Conceptual Framework**

1. Text Cleaning.
2. Labelling the Reviews 2 class (Positive and negative) and 5 class (complaints, Negative only, Positive only and praises) to understand the sentiments.
3. Identifying the best lexicon NRC VS Vader which give better results.
4. Test the 2 class and 5 class sentiments with Naive Bayesian and SVM (support vector Machine) models.
5. Analyse the Average word count and Average length of the reviews
6. Analyse the complaints and Praises to generate insights
7. Analyse the product features contributing to sentiments performance over the period of time to generate insights.

**Research Design**

Leverage on Text analytics using TDSP process.

* 1. Business Understanding
  2. Data Acquisition and understanding
     1. Use web scrapping of woohoo brand Google reviews to collect customer reviews/ comments.
     2. Text cleaning
     3. Labelling 2 class and 5 Class sentiments
     4. Feature data set (Identifying and labelling features)
     5. Exploratory
     6. Document matrix (Bag of Words (BoW) using Count Vectorizer and TFIDF)
  3. Modelling – Model training and Evaluation for 2 class and 5 class sentiments
  4. Deployment
  5. Customer Acceptance

The tools used for this project for solving the problem in analyzing the sentiment mentioned below.

1. Python tool for cleaning the data, implement the Lexicon sentiments and Build Naive Bayesian Model.
2. MS Excel for reviewing, doing data review, count sentiment, word count, review length count, trend analysis and for coding the Product features.

# Chapter 6: Business Understanding

Gift Card industry complete process... to understand the business explained below

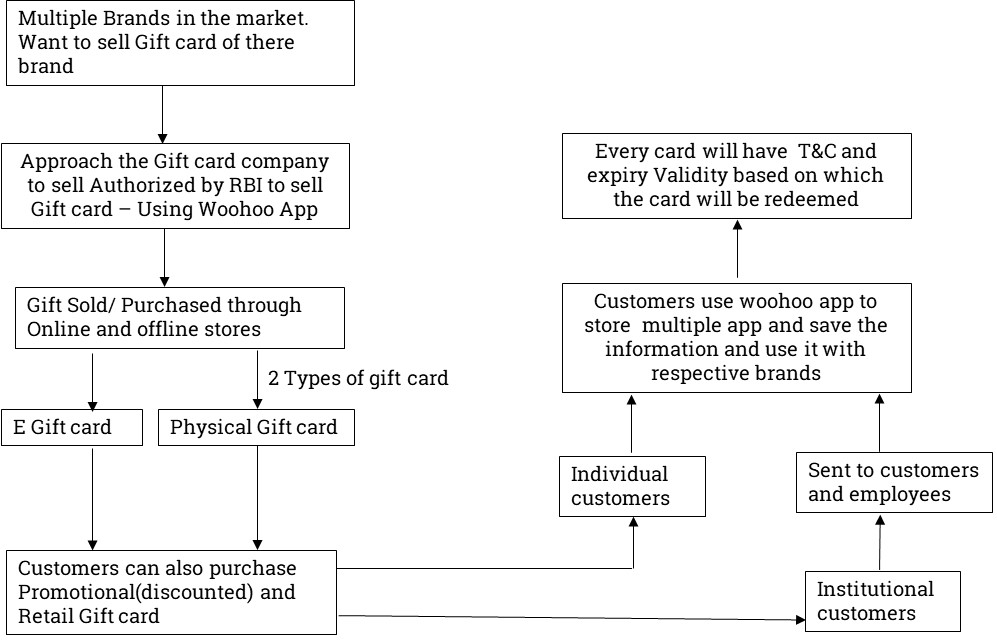


Fig 6.1: Flow chart of business understanding

As per business understanding identified the product features. This would help us to determine if the company is providing better services to the customers.

Table 6.1: Feature Variables details

|  |  |  |  |
| --- | --- | --- | --- |
| # | Variables | Product features | Details |
| 1 | feature\_negative\_1 | Customer Support not good | Customer Support / Chat Support Service/ Awaiting Response/Help me issues |
| 2 | feature\_negative\_2 | Delivery of service not good | Delivery Service Issues / Receiving, Activation, Redeem, Refund, OTP |
| 3 | feature\_negative\_3 | App/ Web Issues | Web App Issues / Login / UI / How to use/ App Issues |
| 4 | feature\_negative\_4 | Payment Issues | Payment Issues / Add other Payment Brands |
| 5 | feature\_negative\_5 | Promotion/offers Issues | Promotional offer Issues / less offers / Cashbacks / Add brands |
| 6 | feature\_negative\_6 | Validity Issues | Validity Issues |
| 7 | feature\_negative\_7 | Not Trusting the Brand | Un trustable / Frauds/Waste /Cheat / Fake/ not good /Scam |
| 8 | feature\_positive\_1 | Speed of delivery good | Faster buying process |
| 9 | feature\_positive\_2 | App/ Web good | Web App Good/ Nice /Great / Awesome |
| 10 | feature\_positive\_3 | Promotion/offers good | Best Offers Deals / Discounts |
| 11 | feature\_positive\_4 | Price is good | Price is good / Save Money |
| 12 | feature\_positive\_5 | Customer Support good | Customer Service is good |
| 13 | feature\_positive\_6 | Easy to use | Easy to Use |
| 14 | feature\_positive\_7 | Multiple Brands to choose good | Buy Wide range of brands |
| 15 | feature\_positive\_8 | Delivery of service good | Delivery service is Good |
| 16 | feature\_positive\_9 | Payment Good | Easy Payment / Multiple Payment modes |

**Challenges**

1. The technology and customers behavior will always tend to change at faster pace. Required robust technology, which can cater to all kind of customers individual and institutional customers.
2. Promotional/cash back offers can be provided only for some duration, so that customers get the experience of the product service. Customers always expect for offers which cannot be provided.
3. Customers are not fully aware about the Expiry date, terms and conditions so that customers redeem gift card within the time frame.

**Monetary Impact**

Increases the overall revenue by 50 crores.

<https://www.tofler.in/qwikcilver-solutions-private-limited/company/U72200DL2006PTC360078>

# Chapter 7: Data Understanding

Extracted the data 2872 records for the Woohoo Gift card from the google play app store. For the past 7 years data.

<https://play.google.com/store/apps/details?id=com.giftbig.mobile&_branch_match_id=952083693414301500>

Table 7.1: Data Variables details

|  |  |
| --- | --- |
| **Variables** | **Options** |
| Rating | 1, 2, 3, 4, 5 |
| Year | 2015, 2016, 2017, 2018, 2019, 2020, 2021 |
| Reviews | Customer feedback for Woohoo gift card - source google play store |

Sample Data

Table 7.2: Sample data

|  |  |  |
| --- | --- | --- |
| **Rating** | **Year** | **Review** |
| **1** | **2021** | ##10103920## case I'd Worst poor customer service they are not giving any response after 4 day's of register complaint. They call me but disconnected the call without giving any proper answer. Now I have not any way to reach customer support so I tried this.... I will never change my review..... |

# Chapter 8: Data Preparation

Pre-processing

1. Going through the reviews, manually labelling of the sentiment by 2 class, 5 class and product features.

Table 8.1: Data Variables and options

|  |  |
| --- | --- |
| **Variables** | **Options** |
| senti\_2class (Labelled) | Not Negative, Negative |
| senti\_5class (Labelled) | Negative only, Complaints, Neutral, Positive only, Praise |
| feature\_negative\_1 | 0,1 |
| feature\_negative\_2 | 0,1 |
| feature\_negative\_3 | 0,1 |
| feature\_negative\_4 | 0,1 |
| feature\_negative\_5 | 0,1 |
| feature\_negative\_6 | 0,1 |
| feature\_negative\_7 | 0,1 |
| feature\_positive\_1 | 0,1 |
| feature\_positive\_2 | 0,1 |
| feature\_positive\_3 | 0,1 |
| feature\_positive\_4 | 0,1 |
| feature\_positive\_5 | 0,1 |
| feature\_positive\_6 | 0,1 |
| feature\_positive\_7 | 0,1 |
| feature\_positive\_8 | 0,1 |
| feature\_positive\_9 | 0,1 |

1. **Load the data into python and doing the preliminary analysis**
2. Total 2782 Records available.
3. There are Zero Missing values in the data
4. In the review’s sentiment (manually labelled the corpus) 57.9% Positive and 42.1% Negative
5. **Cleaning the data by the below process.**
6. Removing special characters, punctuation, square bracket, numbers, double spacing.
7. Convert all the upper case to lower case.
8. Tokenization: Tokenization is the act of breaking up a sequence of strings into pieces such as words, keywords, phrases, symbols and other elements called tokens. Created the tokens in a separate column in the data frame.
9. Stop Words removal: - A stop word is a commonly used word (such as “the”, “a”, “an”, “in”) that a search engine has been programmed to ignore, both when indexing entries for searching and when retrieving them as the result of a search query. Post tokenization remove the stop words.
10. Stemming and Lemmatization: - Stemming and lemmatization are methods used by search engines and chatbots to analyse the meaning behind a word. Stemming uses the stem of the word, while lemmatization uses the context in which the word is being used.
11. **Document Matrix.**

Building the Document matrix (mathematical matrix that describes the frequency of terms that occur in a collection of documents) using the Bag of words (commonly used in methods of document classification where the (frequency of) occurrence of each word is used as a feature for training a classifier).

For our study used 2 Document matrix.

1. Bag of words using count vectorizer (involves counting the number of occurrences each word appears in a document)
2. Bag of words using TFIDF (Term Frequency — Inverse Document Frequency”. This is a technique to quantify a word in documents, usually compute a weight to each word which signifies the importance of the word in the document and corpus.)

**Word cloud - Overall**



Fig 8.1: Word cloud – Overall Woohoo reviews

Some the key words noticed in the Word cloud.

* Woohoo
* app
* Gift card
* Money offer
* Nice app
* payment
* Best app
* Worst app
* Customer care
* Cashback
* Redeem
* Application
* Good app
* Time even
* Buy gift
* Amazing
* Give

**Word cloud - complaint**

****

Fig 8.2: Word cloud – Complaint Woohoo reviews

Some the key words noticed in the Word cloud.

* Worst app
* Payment
* Didn’t
* Give
* Don’t
* Fake app
* Time
* Error response
* Fraud app
* Refund
* Redeem
* Worst service
* Customer care
* Call
* Received
* Waste
* cashback

**Word cloud - Praise**

****

Fig 8.3: Word cloud – Praise Woohoo reviews

Some the key words noticed in the Word cloud.

* Great
* Good app
* Best app
* Really
* Awesome
* Best
* Nice
* Offer
* Gift card
* Woohoo
* Got cashback
* Love
* Happy
* Easy
* Super
* Amazing
* Instant discount

**Descriptive Analytics**

The following topic explained below: -

1. Trend analysis- 2 class Sentiment
2. Trend analysis- 5 class Sentiment
3. Lexicon Sentiment Analysis- 2 class
4. Cross Tab – Validation of Lexicon
5. Trend analysis- Word count and Length- 2 class
6. Trend analysis- Word count and Length- 5 class
7. Features by 5 class Sentiment
8. Trend analysis- Features- Complaint
9. Trend analysis- Features- Praises
10. **Trend analysis- 2 class Sentiment**

Table 8.2: **Trend** **2 class Sentiment**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Number** | **2015** | **2016** | **2017** | **2018** | **2019** | **2020** | **2021** | **Total** |
| Negative | 453 | 96 | 94 | 127 | 97 | 166 | 138 | 1171 |
| Not Negative | 539 | 125 | 60 | 101 | 89 | 376 | 321 | 1611 |
| Grand Total | 992 | 221 | 154 | 228 | 186 | 542 | 459 | 2782 |
|  |  |  |  |  |  |  |  |  |
| **Percentage** |  |  |  |  |  |  |  |  |
| Negative | 46% | 43% | 61% | 56% | 52% | 31% | 30% | 42% |
| Not Negative | 54% | 57% | 39% | 44% | 48% | 69% | 70% | 58% |
| Grand Total | 36% | 8% | 6% | 8% | 7% | 19% | 16% | 100% |

In the Year 2021 the company has the highest Positive Rating of 70% compared to previous year. But when observed for 2021 it has only 16% weighted when compared with reviewers of the total 7 years data.

In the 2 class sentiment overall, had 58% Not Negative and 42% Negative.

1. **Trend analysis- 2 class Sentiment**

Table 8.3: **Trend** 5 **class Sentiment**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Number** | **2015** | **2016** | **2017** | **2018** | **2019** | **2020** | **2021** | **Total** |
| Complaint | 251 | 44 | 49 | 76 | 46 | 80 | 78 | 624 |
| Negative only | 202 | 52 | 45 | 51 | 51 | 86 | 60 | 547 |
| Neutral | 37 | 7 | 3 | 3 | 4 | 10 | 10 | 74 |
| Positive only | 381 | 84 | 47 | 87 | 78 | 277 | 218 | 1172 |
| Praise | 121 | 34 | 10 | 11 | 7 | 89 | 93 | 365 |
| Grand Total | 992 | 221 | 154 | 228 | 186 | 542 | 459 | 2782 |
|  |  |  |  |  |  |  |  |  |
| **Percentage** |  |  |  |  |  |  |  |  |
| Complaint | 25% | 20% | 32% | 33% | 25% | 15% | 17% | 22% |
| Negative only | 20% | 24% | 29% | 22% | 27% | 16% | 13% | 20% |
| Neutral | 4% | 3% | 2% | 1% | 2% | 2% | 2% | 3% |
| Positive only | 38% | 38% | 31% | 38% | 42% | 51% | 47% | 42% |
| Praise | 12% | 15% | 6% | 5% | 4% | 16% | 20% | 13% |
| Grand Total | 36% | 8% | 6% | 8% | 7% | 19% | 16% | 100% |

Overall: -Majority of the reviews are towards “Positive only” 42% which is good, but the company has to work towards converting the “positive only” to “Praise”. and reduce the negative only and complaints around 20% to positive only.

2021: - Over the past 6 years the company has been able to improve the scores for “Positive only” for 2021 with 47%, and improved the praise to 20 %.

1. **Lexicon Sentiment Analysis- 2 class**

Sentiment analysis using the lexicons

1. Lexical Affinity. This approach trains probability from linguistic corpus. It not only detects obvious affect words but also assigns sentiment to arbitrary words.
2. NRC Lexicon Sentiment analysis: (The National Research Council Canada's (NRC) Emotion Lexicon

In the NRC lexicon we are able to understand the emotional side of the reviews along with positive and negative sentiments. 8 emotions are anger, anticipation, disgust, fear, joy, sadness, surprise and trust.

1. Vader Lexicon Sentiment analysis (Valence Aware Dictionary and sentiment Reasoner) Emotion Lexicon

In the Vader sentiment analysis, we are able to understand the Positive and negative sentiments. We rate it as Negative, positive and neutral the compound score >0 is Not negative and <0 is Negative.

We have done cross tab of the 2 Class sentiment with lexicon output to validate know which lexicon give better results.

1. **Cross Tab – Validation of Lexicon**

Table 8.4: Cross Tab – Validation of Lexicon

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Number | Lexical Affinity | | NRC | | Vader Sentiment | |  |
| Sentiment | Negative | Not Negative | Negative | Not Negative | Negative | Not Negative | Total |
| Negative | 512 | 659 | 726 | 445 | 536 | 635 | 1171 |
| Not Negative | 8 | 1603 | 141 | 1470 | 11 | 1600 | 1611 |
| Total | 520 | 2262 | 867 | 1915 | 547 | 2235 | 2782 |
|  |  |  |  |  |  |  |  |
| Percentage |  |  |  |  |  |  |  |
| Negative | 43.7% | 56.3% | 62% | 38% | 46% | 54% | 42% |
| Not Negative | 0.5% | 99.5% | 9% | 91% | 1% | 99% | 58% |

NRC Lexicon is providing better results in identifying the sentiments.

Negative is 62 % which is better compared to other lexicons Vader with 46% and Lexical Affinity 44%.

Not Negative is 91 % is better than other lexicons, even though the percentage for Vader with 99 % and Lexical Affinity 99.5 % is more, NRC is better as it is giving better results for negative sentiment also.

1. **Trend analysis- Word count and Length- 2 class**

Table 8.5: Trend analysis- Word count and Length- 2 class

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Average of Words count** | | |  | | | | | |
| Average | 2015 | 2016 | 2017 | 2018 | 2019 | 2020 | 2021 | Average |
| Negative | 22.5 | 23.3 | 22.0 | 22.0 | 19.6 | 21.2 | 22.8 | 22.09 |
| Not Negative | 9.9 | 10.2 | 7.7 | 6.7 | 4.9 | 9.4 | 10.1 | 9.27 |
| Grand Total | 15.7 | 15.9 | 16.4 | 15.2 | 12.5 | 13.0 | 13.9 | 14.67 |
| **Average text Length** | | |  | | | | | |
| Negative | 122.06 | 125.61 | 118.04 | 120.17 | 106.26 | 117.39 | 122.96 | 119.96 |
| Not Negative | 53.93 | 54.55 | 41.82 | 35.37 | 25.62 | 50.19 | 52.91 | 49.73 |
| Grand Total | 85.04 | 85.42 | 88.34 | 82.61 | 67.67 | 70.77 | 73.97 | 79.29 |

* Reviews average of words count is 14.67 words are used and 79.29 is the average text length of the reviews.
* 2021 negative reviews average of words count is around 23 and the average text length is 123.
* 2021 not negative reviews average of words count is around 10 and the average text length is 53.
* customers write more for Negative reviews, compared to Not negative.

1. **Trend analysis- Word count and Length- 5 class**

Table 8.6: Trend analysis- Word count and Length- 5 class

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Average of words count** |  |  |  |  |  |  |  |
| **Average** | **2015** | **2016** | **2017** | **2018** | **2019** | **2020** | **2021** | **Average** |
| Complaint | 29.09 | 35.98 | 30.80 | 28.96 | 26.76 | 29.04 | 30.58 | 29.70 |
| Negative only | 14.42 | 12.56 | 12.36 | 11.71 | 13.06 | 13.86 | 12.65 | 13.41 |
| Neutral | 3.30 | 4.57 | 1.00 | 4.00 | 1.50 | 3.20 | 4.70 | 3.43 |
| Positive only | 7.08 | 6.30 | 4.45 | 4.94 | 4.59 | 6.39 | 6.23 | 6.27 |
| Praise | 20.74 | 21.12 | 25.10 | 21.27 | 9.71 | 19.34 | 19.63 | 20.08 |
| Grand Total | 15.67 | 15.90 | 16.42 | 15.24 | 12.52 | 12.98 | 13.89 | 14.67 |
|  |  |  |  |  |  |  |  |  |
| **Average text Length** |  |  |  |  |  |  |  |
| Complaint | 158.39 | 192.82 | 164.41 | 158.21 | 147.00 | 163.00 | 165.73 | 161.94 |
| Negative only | 76.91 | 68.75 | 67.56 | 63.49 | 69.51 | 74.97 | 67.37 | 72.07 |
| Neutral | 17.54 | 20.43 | 3.33 | 19.33 | 5.75 | 17.90 | 20.50 | 17.12 |
| Positive only | 38.50 | 32.01 | 23.45 | 25.44 | 23.96 | 33.59 | 32.03 | 33.13 |
| Praise | 113.67 | 117.26 | 139.70 | 118.27 | 55.43 | 105.51 | 105.34 | 109.63 |
| Grand Total | 85.04 | 85.42 | 88.34 | 82.61 | 67.67 | 70.77 | 73.97 | 79.29 |

Complaints has the highest average for both average of words count 29.70 and average text length 162.

For the Positive only and neutral has the Lowest average for both average of words count < 6 and average text Length < 33.

**Conclusion:**

Customers tend to write more when they are either very happy (Praise) or very unsatisfied(complaint)

On an average over the years the reviews average of words count and average text length has reduced from 16 to 14 and 85 to 74.

Average of words count**: -**

Over the past 7 years the customers are writing very less words to express their feeling having “Praise” in year 2021 is 20 and writing more for “complaint” in year 2021 is 31.

Average text length**: -**

Complaint Length of reviews is increasing over the past years for 2021 with 165 and for praise it has reduced to 105.

1. **Features by 5 class Sentiment**

Table 8.7: Features by 5 class Sentiment

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Features / 5 class sentiment** | **Complaint** | **Negative only** | **Neutral** | **Positive only** | **Praise** | **Grand Total** |
| Customer Support not good | 31% | 5% | 0% | 0% | 0% | 8% |
| Delivery of service not good | 63% | 11% | 0% | 0% | 0% | 16% |
| App/ Web Issues | 65% | 46% | 1% | 1% | 1% | 24% |
| Payment Issues | 19% | 8% | 0% | 1% | 0% | 6% |
| Promotion/offers Issues | 27% | 11% | 3% | 1% | 1% | 9% |
| Validity Issues | 2% | 0% | 1% | 0% | 0% | 1% |
| Not Trusting the Brand | 63% | 20% | 1% | 0% | 0% | 18% |
| Speed of delivery good | 0% | 0% | 0% | 1% | 14% | 2% |
| App/ Web good | 1% | 3% | 0% | 49% | 82% | 32% |
| Promotion/offers good | 0% | 0% | 0% | 4% | 42% | 7% |
| Price is good | 0% | 0% | 0% | 0% | 3% | 0% |
| Customer Support good | 0% | 0% | 0% | 1% | 5% | 1% |
| Easy to use | 0% | 0% | 0% | 4% | 35% | 6% |
| Multiple Brands to choose good | 0% | 0% | 0% | 3% | 21% | 4% |
| Delivery of service good | 0% | 0% | 0% | 1% | 25% | 4% |
| Payment Good | 0% | 0% | 0% | 0% | 6% | 1% |
|  | 624 | 547 | 74 | 1172 | 365 | 2782 |

Overall, 32 % of the customers feel the App/web good, 24 % of the Customers feel that there are issues in the App/ Web which needs to be fixed. 3rd highest is Not trusting the Brand having 18 %.

**Praise**

Under the Sentiments wise Praise stands significantly high with 82% customers feeling the app / web application is good.

Promotion/offers good 42%, Easy to use 35%, delivery of service is Good 25% and Access to Multiple Brands to choose for purchase 21% are the other key features customers lookout for while purchasing gift card. Brand should focus on to increase the Positive only and praise

**Complaint**

App/ Web Issues 65%, Delivery of service not good 63%, Not Trusting the Brand63%, Customer Support not good 31%, are the Key product features due to which the customers have given complaints and Negative only reviews.

Brand should focus on reducing the scores and improving the service.

1. **Trend analysis- Features- Complaint**

Table 8.8: Trend analysis- Features- Complaint

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Feature /Year** | **2015** | **2016** | **2017** | **2018** | **2019** | **2020** | **2021** | **Total** |
| Customer Support not good | 18% | 34% | 33% | 28% | 37% | 48% | 54% | 31% |
| Delivery of service not good | 54% | 59% | 69% | 64% | 63% | 68% | 79% | 63% |
| App/ Web Issues | 76% | 70% | 73% | 64% | 70% | 43% | 40% | 65% |
| Payment Issues | 8% | 14% | 49% | 28% | 15% | 20% | 36% | 19% |
| Promotion/offers Issues | 45% | 23% | 8% | 17% | 20% | 14% | 9% | 27% |
| Validity Issues | 3% | 5% | 2% | 0% | 7% | 1% | 1% | 2% |
| Not Trusting the Brand | 65% | 64% | 67% | 66% | 72% | 58% | 56% | 63% |
| Speed of delivery good | 0% | 0% | 0% | 0% | 0% | 1% | 0% | 0% |
| App/ Web good | 2% | 2% | 0% | 0% | 0% | 1% | 3% | 1% |
| Promotion/offers good | 1% | 0% | 0% | 0% | 0% | 0% | 0% | 0% |
| Price is good | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% |
| Customer Support good | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% |
| Easy to use | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% |
| Multiple Brands to choose good | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% |
| Delivery of service good | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% |
| Payment Good | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% |

**Complaint**

There are 4 areas on which the brand has to work on improving Delivery of service not good 79%, Not Trusting the Brand 56%, Customer Support not good 54% and Payment Issues 36%, which has increased for the past 7 years.

App/ Web Issues has got improved from earlier 76% to 40% and for promotional offers from 45% to 9%. Brand should focus on reducing the scores and improving the service.

1. **Trend analysis- Features- Praises**

Table 8.9: Trend analysis- Features- Praises

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Year** | **2015** | **2016** | **2017** | **2018** | **2019** | **2020** | **2021** | **Total** |
| Customer Support not good | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% |
| Delivery of service not good | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% |
| App/ Web Issues | 2% | 0% | 0% | 0% | 0% | 0% | 0% | 1% |
| Payment Issues | 1% | 0% | 0% | 0% | 0% | 0% | 0% | 0% |
| Promotion/offers Issues | 3% | 0% | 0% | 0% | 0% | 0% | 0% | 1% |
| Validity Issues | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% |
| Not Trusting the Brand | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% |
| Speed of delivery good | 5% | 18% | 20% | 9% | 0% | 8% | 32% | 14% |
| App/ Web good | 80% | 68% | 70% | 73% | 57% | 91% | 87% | 82% |
| Promotion/offers good | 57% | 53% | 30% | 55% | 29% | 37% | 26% | 42% |
| Price is good | 3% | 0% | 0% | 0% | 0% | 3% | 3% | 3% |
| Customer Support good | 5% | 3% | 20% | 9% | 29% | 0% | 6% | 5% |
| Easy to use | 28% | 50% | 30% | 36% | 43% | 56% | 17% | 35% |
| Multiple Brands to choose good | 18% | 18% | 20% | 9% | 14% | 35% | 13% | 21% |
| Delivery of service good | 17% | 12% | 50% | 27% | 43% | 16% | 47% | 25% |
| Payment Good | 7% | 3% | 0% | 0% | 0% | 4% | 8% | 6% |

**Praise**

There are 4 areas on which the brand has performed well in features App/ Web good 87%, Speed of delivery 32% and Delivery of service good 47%.

Promotion/ offers good 26%, Easy to use 17% and Multiple Brands to choose good 13% It is positively contributing to praises but it has good reduced for the past 7 Years. If the score is improved it will help in getting more praises.

# Chapter 9: Data Modeling

Modeling Process

1. Post cleaning of the data (please refer the slide data preparation) first create 2 document matrix. Bag of words using count vectorizer and Bag of words using TFIDF.
2. Split the data by Train and test for 2 different sample size 70-30 and 80 -20 split
3. Using the Train data, build the model and apply this to test data to predict and validate the accuracy.
4. For the 2-document matrix- used 2 models Naive Bayesian Model and Support Vector Machine Model.
5. In total 16 Model built, 8 models for 2 class sentiment and 8 models built for 5 class sentiment
6. Evaluate the model based on accuracy, Precision and recall.
7. 2 Class sentiment and 5 class sentiment for both models have been built to check the efficiency of the sentiments. Based on higher accuracy recommendation can be provided to clients to improve the services.

Table 9.1: Model approach

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Model table | 2 class | 2 class | 5 class | 5 class |
|  | 70-30 | 80-20 | 70-30 | 80-20 |
| Naive Bayesian - Bag of words - Count Vectorizer |  |  |  |  |
| Naive Bayesian - Bag of words - TFIDF |  |  |  |  |
| SVM - Bag of words - Count Vectorizer |  |  |  |  |
| SVM - Bag of words - TFIDF |  |  |  |  |

# Chapter 10: Data Evaluation

Modeling Output - comparison

Table 10.1: Model output

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Model output | 2 class | 2 class | 5 class | 5 class |
|  | 70-30 | 80-20 | 70-30 | 80-20 |
| Naive Bayesian - Bag of words - Count Vectorizer | 93% | 94% | 69% | 68% |
| Naive Bayesian - Bag of words - TFIDF | 93% | 94% | 70% | 71% |
| SVM - Bag of words - Count Vectorizer | 92% | 94% | 73% | 74% |
| SVM - Bag of words - TFIDF | 94% | 95% | 70% | 71% |

**2 Class sentiment**

SVM - Bag of Words (BoW) using TFIDF, with 80 train and 20 test Split for the 2-class sentiment has given higher accuracy of 95%.

**5 Class sentiment**

SVM - Bag of Words (BoW) using Count Vectorizer, with 80 train and 20 test Split for the 5-class sentiment has given higher accuracy of 74%.

**2 Class sentiment Model- Results**

SVM - Bag of Words (BoW) using TFIDF, with 80 train and 20 test Split

Classification report

Table 10.2: Classification report – 2 Class

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | precision | recall | f1-score | support |
|  |  |  |  |  |
| Negative | 92% | 95% | 93% | 226 |
| Not Negative | 97% | 94% | 95% | 331 |
|  |  |  |  |  |
| accuracy |  |  | 95% | 557 |

Confusion Matrix

Table 10.3: Confusion Matrix – 2 Class

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | Predictions | |  |
|  |  | Negative | Not Negative |  |
| Actual | Negative | 215 | 11 | 226 |
| Not Negative | 19 | 312 | 331 |
|  |  | 234 | 323 |  |

2 Class sentiment

SVM - Bag of Words (BoW) using TFIDF, with 80 train and 20 test Split for the 2-class sentiment has given higher accuracy of 95%.

This model is good as it has the precision of identifying the Negative 92% with the recall of 95%.

For the Not negative it has a precision of 97% and recall of 94%.

As precision and recall are more than 92 % for both sentiments, the F1 score is also good.

**5 Class Sentiment Model- Results**

SVM - Bag of Words (BoW) using Count Vectorizer, with 80 train and 20 test Split

Classification report

Table 10.4: Classification report – 5 Class

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | precision | recall | f1-score | support |
|  |  |  |  |  |
| Complaint | 72% | 75% | 73% | 115 |
| Negative only | 66% | 61% | 64% | 111 |
| Neutral | 50% | 25% | 33% | 12 |
| Positive only | 81% | 90% | 85% | 252 |
| Praise | 62% | 43% | 51% | 67 |
|  |  |  |  |  |
| accuracy |  |  | 74% | 557 |

Confusion Matrix

Table 10.5: Confusion Matrix – 5 Class

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | Predictions | | | | |  |
|  |  | Complaint | Negative only | Neutral | Positive only | Praise |  |
| Actual | Complaint | 86 | 25 | 0 | 1 | 3 | 115 |
| Negative only | 26 | 68 | 1 | 14 | 2 | 111 |
| Neutral | 0 | 1 | 3 | 8 | 0 | 12 |
| Positive only | 3 | 7 | 2 | 227 | 13 | 252 |
| Praise | 5 | 2 | 0 | 31 | 29 | 67 |
|  |  | 120 | 103 | 6 | 281 | 47 |  |

5 Class sentiment

SVM - Bag of Words (BoW) using Count Vectorizer, with 80 train and 20 test Split for the 5-class sentiment has given higher accuracy of 74%.

This model is good as it has the precision of >66% in identifying the Negative sentiments Complaint and negative only with recall >61%.

For the Not negative it has the precision of >62% in identifying Positive sentiments- Praise and Positive only. and recall of 90% for Positive only and 43% for Praise.

# Chapter 11: Deployment

Present this to the company management to take appropriate steps to improve the performance. Trained and tested the model and got 89 % accuracy. After 6 months again retesting the delivery of service sentiments performance.

# Chapter 11: Analysis and Results

**Outcome of Lexicons Sentiment analysis.**

NRC Lexicon is providing better results in identifying the sentiments.

Negative is 62 % which is better compared to other lexicons Vader with 46% and Lexical Affinity 44%.

Not Negative is 91 % is better than other lexicons, even though the percentage for Vader with 99 % and Lexical Affinity 99.5 % is more, NRC is better as it is giving better results for negative sentiment also.

**Outcome of Model in predicting the Sentiments.**

* **2 Class Sentiment Model- Results**

SVM - Bag of Words (BoW) using TFIDF, with 80 train and 20 test Split for the 2-class sentiment has given higher accuracy of 95%.

This model is good as it has the precision of identifying the Negative 92% with the recall of 95%.

For the Not negative it has a precision of 97% and recall of 94%.

As precision and recall are more than 92 % for both sentiments, the F1 score is also good.

* **5 Class Sentiment Model- Results**

SVM - Bag of Words (BoW) using Count Vectorizer, with 80 train and 20 test Split for the 5-class sentiment has given higher accuracy of 74%.

This model is good as it has the precision of >66% in identifying the Negative sentiments Complaint and negative only with recall >61%.

For the Not negative it has the precision of >62% in identifying Positive sentiments- Praise and Positive only. and recall of 90% for Positive only and 43% for Praise.

**Strength:**

1. The brand has been able to improve its performance in Not Negative 54 % in 2015 increased to 70% in 2021.whereas for negative it has decreased 46% in 2015 to 30% in 2021
2. The brand has been able to reduce its Negative scores by 54 % in 2015 increased to 70% in 2021.
3. The brand has been able to increase its praise from 12 % in 2015 to 20% in 2021.
4. The brand has been able to reduce its complaints from 25% in 2015 to 17% in 2021.
5. 2021 not negative reviews average of words count is around 10 and the average text length is 53. Customers are writing less as they are happy.
6. Over the past 7 years, the customers are writing very few words  to express their feeling having “Praise” in the year 2021 is 20.
7. Praise average length of reviews is reducing over the past years for 2021 with 105.
8. For the “Praise” there are 4 areas in which the brand has performed well in features App/ Web good 87%, Speed of delivery 32% and Delivery of service good 47%.

**Weakness:**

1. The brand needs to work towards reducing negative scores, an average Negative 30% minimum is scored every year(past 7 years) in spite of being in the industry for more than 10 years this needs to be brought to 10% as bad customers experience will affect the brand growth.
2. Overall Complaints 22% and Negative only 20% combined 42 % shows that the company has to work towards improving its customer’s experience.
3. 2021 negative average of words count is around 23 and the average text length is 123. Customers are writing more words to explain the problems that need to be addressed.
4. Over the past 7 years, the customers are writing more words to express their feeling having “complaint” in the year 2021 is 31.
5. Complaint average length of reviews is increasing over the past years for 2021 with 165.

**Opportunity:**

1. The category penetration is currently around 5 to 15%, which shows the brand has huge potential to tap into this market and gain more market currently 75%.
2. Push from the government for Digital payments and Covid restrictions is a good opportunity for the brand to increase its share.
3. For the complaints App/ Web Issues has got improved from earlier 76% to 40% and for promotional offers from 45% to 9%. Brand should focus on reducing the scores and improving the service.
4. For the “praise” there are 3 if the score is improved it will help in getting more praise. Promotion/ offers good 26%, Easy to use 17% and Multiple Brands to choose good 13% It is positively contributing to praises but it has well reduced for the past 7 Years.
5. India with a population of more than 130 crores and we have only 5 to 15 % penetration has a huge potential.

**Threat:**

1. For the complaints, there are 4 areas on which the brand has to work on improving Delivery of service not good 79%, Not Trusting the Brand 56% and Customer Support not good 54% which has increased for the past 7 years.
2. There are more than 9 Brands (combined share of 25%) in the gift card industry in India. Which results in huge competition.
3. Customers tend to get attracted when more promotional offers/cashback/discount is issued by other Gift card companies. Customers might shift to competitors.
4. Any new technological advancements might impact the growth of the company. Example: E Rupi was introduced by the government to India in order to reach the end-users government benefits and the money given can be redeemed only with the authorized Merchant sellers only.

# Chapter 12: Conclusions and Future Scope

**Conclusion**

**Why Woohoo Gift cards are studied.**

 The scope of the study is to use the Woohoo google reviews to do Text analysis and understand the Customer sentiment complaints and praises. The gift card industry is growing slowly in India.

**Problem Statement.**

Increase the penetration to 25% by analysing the customer complaints and Praises using the Google reviews given by customers for the past 7 years.

**Approach - Results**

* Analysing the Sentiments trends for 2 classes and 5 classes. Praise sentiment is increasing.
* Analysing the Words count and text length - Customers write more for Negative reviews, compared to Not negative.
* Analysing the Lexicon sentiment. NRC Lexicon is providing better results
* Analysing the Model to predict the sentiment outcomes - for 2 class model accuracy of 95% and 5 class model accuracy of 74%.

**Recommendation based on the above findings.**

The brand has to work on improving in 4 areas

1. Delivery of service needs to be improved, better user interface and highlighting the T&C.
2. Work towards improving brand trust.
3. Customer Support should be quick in responding to the queries and taking them to closure.
4. Payment gateway better handling of payment failures and money getting deducted but not reflecting in account needs to be addressed.

**Summary**

Overall, the Brand has made a significant improvement in improving the sentiments scores for the past 7 years from 2015 to 2021. 70% of the customers are happy with the Brand, it has to focus on 30% of the unsatisfied customers. Under the Sentiments Praise stands significantly high with 82% of customers feeling the app/web application is good.

Complaint – scored around 63% is due to App/ Web Issues, Delivery of service not good and Not Trusting the Brand.

**Future scope of study.**

1. Negation handling.
2. N-grams (unigram, bigram, trigrams).
3. Noun and Adjective Usage.
4. Paste tense analysis
5. Intensifier Usage
6. Taking reviews and comments from other social media platforms Facebook and twitter.

# Bibliography

1)

Title=Linguistic Understanding of Complaints and Praises in User Reviews, author=Ganesan, Kavita and Zhou, Guangyu, book title= Proceedings of NAACL-HLT, pages=109--114, year=2016.

<https://kavita-ganesan.com/complaints-and-praises/#.YQPVoI4zaHv>

2)

Title = Lexicon-Based Sentiment Analysis of Teachers’ Evaluation, Author: - Quratulain Rajput, Sajjad Haider, and Sayeed Ghani, Hindawi Publishing Corporation Applied Computational Intelligence and So, Computing Volume 2016, Article ID 2385429, 12 pages http://dx.doi.org/10.1155/2016/2385429

<https://downloads.hindawi.com/journals/acisc/2016/2385429.pdf>

3)

[Ganesan and Zhai2012] Kavita Ganesan and ChengXiang Zhai. 2012. Opinion-based entity ranking. Information retrieval, 15(2):116–150.

<http://kavita-ganesan.com/wp-content/uploads/2010/03/EntityRank.pdf>

4)

[Kim et al.2011] Hyun Duk Kim, Kavita Ganesan, Parikshit Sondhi, and ChengXiang Zhai. 2011. Comprehensive review of opinion summarization (opinion mining survey).

<https://www.researchgate.net/publication/50273740_Comprehensive_Review_of_Opinion_Summarization>

5)

[Pang and Lee2008] Bo Pang and Lillian Lee. 2008. Opinion mining and sentiment analysis. Foundations and trends in information retrieval, 2(1-2):1– 135.

<https://www.cs.cornell.edu/home/llee/omsa/omsa.pdf>

# Appendix

## Plagiarism Report[[1]](#footnote-1)

## Publications in a Journal/Conference Presented/White Paper[[2]](#footnote-2)

## Any Additional Details

Website Links

<https://www.aclweb.org/anthology/W16-0418.pdf>

<https://www.alliedmarketresearch.com/gift-cards-market>

<https://www.businesswire.com/news/home/20200921005319/en/India-9.34-Billion-Gift-Card-and-Incentive-Card-Markets-2015-2019-2024-with-Covid-19-Update-Q2-2020---ResearchAndMarkets.com>

<https://www.qwikcilver.com/>

<https://www.woohoo.in/>

<https://www.businesstoday.in/magazine/the-hub/story/digigifting-185772-2019-04-01>

<https://www.businesswire.com/news/home/20200921005319/en/India-9.34-Billion-Gift-Card-and-Incentive-Card-Markets-2015-2019-2024-with-Covid-19-Update-Q2-2020---ResearchAndMarkets.com>

<https://www.tofler.in/qwikcilver-solutions-private-limited/company/U72200DL2006PTC360078>

1. Turnitn report to be attached from the University. [↑](#footnote-ref-1)
2. URL of the white paper/Paper published in a Journal/Paper presented in a Conference/Certificates to be provided. [↑](#footnote-ref-2)