FOM University of Applied Sciences for Economics & Management

Study Centre Essen

Project in the module

Big Data Analysis Project

to obtain the degree of a

MSc. in Big Data and Business Analytics

On the topic

Text Analysis in Online Sensory Marketing

Project Guide:
Prof. Dr. Buchkremer Rüdiger

Submission Date:

Submitted By:

1. Pukhraj Danaram suthar (543289)
2. Taral Sanjay Sarvagod (543298)
3. Hitesh Ashok Kale (543263)
4. Sagar Prabhu Panchal (543293)
5. Shreyashi (543297)
6. Pratik Keshav Dhumal (543272)
7. Sagnik Patra (550496)

Table of Contents

List of figures

List of Abbreviations

NLP (Natural language processing)
OSMI (Online Sensory Marketing Index)
Polarity
Sentiment Polarity Distribution
Neutral Zero Polarity
Boxplot
Textblob
Sentiment analysis
NTLK
Stop words
Lemmatization
Count vectorization
Pyplot

1. Introduction

The sensory market analysis is the now a days tends for the online business model. The marketing campaign mostly have strategic planning based on sensory analysis. Sensory market analysis is nothing but the powerful tactics to make user or customer sense of the product in all sense like hearing seeing smelling etc. But in online sensory technique only hearing and seeing methods are more evident and commonly possible. Through simulation technique the entire product can be presented digitally which feel like in real physically presence of product in terms of appearance and function. All the online business model in retail or any other sector using this online sensory marketing technique to grow the customer base which ultimately results into the revenue generation and gain competitive advantages in business. The best example of using sensory marketing and leveraging the opportunities is Starbucks coffee. It's not only a coffee place its more than that people use to have important meets at this place for different reasons. Also in retail industries are utmost important service like restaurant saloon etc. these sector also showing great importance in sensory marketing techniques and leveraging the opportunities. Sense like Sounds, images, smells, and even tastes can all affect the way that your target audience thinks about you. As we think about the terms with the science of sensory marketing, businesses can get help to gain high brand loyalty and discover the concepts that truly convince customers to buy.

The obvious question is the how is data analysis is helps in the sensory marketing the answer to this questions is there are many way the data analysis is going to prove as backbone for the online sensory marketing techniques. By analyzing the customer behavior pattern companies can take competitive advantages. In past time the companies have used color code and logos for sensory marketing but now a day's availability of new tool and techniques like simulations results into the advancement of the sensory marketing. The availability of data in large volume also help to understand user's personalization base approach which give deeper understanding of consumer area of interest. There are many free website available which gives OSM data. The types of data for OSM in media format as well as textual information. Any type of data can be harvest to analyze the OSM. As we know everything has its cons and pros limitation along with advantages also challenges. The biggest challenge for the any time of OSM data is the veracity of data the truth in the data otherwise the results may alter different ways. The OSM index can be used to find the success and we can measure the success of the online sensory marketing campaign.

1.1 Importance of topic

"Text Data analysis in Online Sensory Marketing"

Today's increase in volume of data generation and its handling tools and technology enable the sensory marketing techniques more evident in growth of business the importance of not only text data but also all other format of data are helpful in creating strong customer base for any online business models. With respect to the online sensory marketing the role of text data is utmost importance from text data analysis results in depth understanding of customer behavior analysis and helps to understand buying patterns of the customers Which

again helps in the formulated of new business strategy for the customer retention. In our proposed work we have disturbed the work into different teams. My team have got text data analysis in online sensory marketing index. We have got the data from Data collection teams about OSM index. The types of the data in sensory marketing is of many types such as-

- Color
- Image
- Text
- Graphics
- Video
- Lights

In this types of data we as a group have been responsible for the analysis of the text format data. Any type of headline in an advertisement that is coupled with images and the right fonts can have a significant impact.

1.2 Problem statement

"In the era of digitalization and industry 4.0 how text data analysis techniques used in an online sensory marketing to leverage the opportunities. Also the question is like, how can you tap into the sensual side of marketing to improve the potential and sustainability of business?"

1.3 Research question

"How text data analysis helps in online sensory marketing?"

As online sensory marketing technique might not be the new but increase in digitalization globally results in the creating the new opportunities in online business models hence it is obvious that increase in data leads to the increase in more possibilities. How companies can grow with the help of online sensory marketing techniques through text data analysis and availing new opportunities.

1.4 Expected outcome

"As a proof of concept we have perform data visualization and sentiment analysis to find out some meaningful insights."

As the expected output from text analysis team we have find some meaningful insights from data that we get from data collection teams. We have use python programming language to create data visualization graphs. Also we have done the sentiment analysis about text reviews from data and we have use textblob python library for the sentiment analysis.

1.5 Schedule planning

As a group we have put cumulative efforts to understanding respective concept from scratch to the finalizing results. We had gone through deep literature analysis of the online sensory marketing and how the online sensory marketing is proves an essential tool for the marketing wizards for the companies. We had design the schedule for the task allocation and better approach for the solution that we have as an output. Aa schedule planning help us to work as a team in efficient manner and also we had organized some scrum meeting for discussion and planning solution for the analysis. As we have planning milestone chart for carried out this task we helps us to take corrective feedback and suggestion from the professor guide.

Milestone chart

Here we completed this project in 5 milestone phases.

1. Introduction: - In introduction we did background research about financial frameworks. How financial frameworks works worldwide, Different financial frameworks around the globe? In this phase we found the Problem statement, research question and research design.

2. Modifications: - In this phase we received feedbacks from professor about the problem statement, research question and research design and did the all required corrections for project.
3. Data Understanding: - It was an important part of any analysis project. We as team worked on understanding of the data attributes. The data collection team has provides the support for the source information.
4. Data Mining: - In data mining we had presented some meaningful insights about the data that we had analyze. Also the results about the sentiment analysis is shown and derived with the help of python library called as textblob.
5. Final presentation: - Live demo, Final documentation and presentation.

This project management methods helps us to trace the our ways to handle the various phase efficiently we are as team coordinate among us and achieved this goal in most fitted manner it proves a handy tool for us to finish project on time.

2. Theoretical foundation

2.1 What is Text data Analysis?

Text analysis seeks to extract machine-readable data from unstructured text in order to allow data-driven content management. Text Analysis necessitates the creation of specialized text mining pipelines to overcome the ambiguity of human language and achieve high accuracy for a given area. Text analysis is the process of analysing texts to obtain machine-readable facts. Text Analysis is used to produce structured data from unstructured text. Slicing and dicing piles of unstructured, diverse documents into easy-to-manage and comprehend data bits is how the process works. **The ambiguity of human languages is the primary issue in text analysis.** A tailored text mining pipeline that combines or reflects these details is required to achieve high accuracy for a certain domain and document types.[1]

Information retrieval, lexical analysis to investigate word frequency distributions, pattern recognition, tagging/annotation, information extraction, and data mining techniques such as link and association analysis, visualization, and predictive analytics are all part of text analysis. The overall purpose is to convert text into data that can be analysed using natural language processing (NLP), various algorithms, and analytical methodologies. The interpretation of the information received is a crucial part of this procedure. Scanning a set of natural-language documents and either modelling the document set for predictive categorization or populating a database or search index with the information gathered is a typical application. When it comes to text mining, the document is the most important component. We describe a document as a textual data unit that can be found in a variety of collections. [2]

[1] Cp. https://www.ontotext.com/knowledgehub/fundamentals/text-analysis/
[2] Cp. https://en.wikipedia.org/wiki/Text_mining

Knowledge graphs for Text data analysis: - Knowledge graphs and modern Text Analysis technology are heavily intertwined. To enable a more accurate interpretation of the text, big graphs give background knowledge, human-alike concept and entity awareness; the outcomes of the analysis are semantic tags (annotations) that link references in the text to specific ideas in the graph. Facts derived from the text can be added to enrich the Knowledge Graph using these tags, which provide structured metadata that allows for improved search and analytics.

Difference between Text Analysis, Text mining and Text Analytics:

The terms text analysis and text mining are interchangeable. The scientific discipline that underpins text mining is known as information extraction. What is Information Extraction? is a great post to read. A collection of common Text Analysis jobs is provided.

All of these concepts apply to partial Natural Language Processing (NLP), which has the purpose of retrieving specified information from a text in the most practical way possible rather than fully understanding it. This entails striking a fair balance between the time and effort required to design and maintain the analytical pipeline, as well as its computational cost and performance (e.g., how much memory it requires and how long it takes to process a single document). Text Analytics and Text Analysis are two terms that are used interchangeably. While these may appear to be synonyms to the untrained ear, there is a slight difference worth noting from the perspective of practice and experience. Text analysis aids in the translation of a text into data language. Text Analytics kicks in to help make sense of these data when Text Analysis "prepares" the content.

The word "text analysis" refers to the process of computing and analysing texts. While Text analytics is a set of techniques and methodologies for converting textual material into data, which may then be mined for insights, trends, and patterns.

2.2 Online Sensory Marketing

What is Sensory marketing?

In the marketing area, sensory marketing is widely used, and it may explain the majority of our purchasing decisions. It has an impact on consumer perceptions of new food courts and restaurants springing up around every corner. Visual, aural, olfactory, gustative, and tactile marketing are all examples of sensory marketing. Sensory branding is a sort of marketing that appeals to all of your senses when it comes to your brand. It connects with customers on an emotional level through the senses. By appealing to the senses, brands can deceive buyers' emotional associations.[3] Sensory marketing is a potent strategy that elicits significant responses by appealing to at least one of the five senses. In-store retailers have proven to be successful in appealing to customers' senses. For example, the aroma of a Starbucks store is distinct and recognizable. How might marketers elicit these same emotions through their online

[3] Cp. https://www.researchgate.net/publication/309154767_Analyzing_the_Impact_of_Sensory_Marketing_on_Consumers_A_Case_Study_of_KFC

store? The answer is to understand how to properly duplicate in-store sensory experiences in order to fully engage customers and increase their confidence in the quality and accuracy of things purchased online. A sensory experience is a person's perception of goods, services, or other essentials in a service process as an image that challenges the mind and senses.

Because we are now in the recovery phase, there is so much more at stake, the epidemic has caused firms to reconsider their marketing strategy. Marketers must ignore the jargon and get back to basics by appealing to the five senses of customers to drive the path to buy while seeking for fresh methods to reinvent themselves and embrace a whole new type of customer.

Sensory marketing has always been connected with engaging in-store customers; but what if marketers could use the senses of sight, hearing, smell, taste, and touch to engage online customers?[4]

Sense-Appealing E-Commerce

Although touch cannot be replicated online, there are a variety of techniques to engage the other senses in order to creatively appeal to customers and improve the online purchasing experience.

Sight

For internet retailers, highly visual cues are extremely crucial. Consumers make the majority of their shopping decisions based on what they see on screen, trusting that what they receive at their door is the same as what they saw online. Brands can utilize high-quality dynamic photos that reveal the appropriate colour and texture of cloth to provide buyers a realistic perspective of products. Shoppers who want to get a better feel for an item before they buy will appreciate detailed zooming and 360-degree capabilities.

Sound

Sound and music should portray the appropriate mood for a firm's distinct brand image, taking into account factors such as genre and tempo. When individuals access the website, this should not be the default soundtrack. Sound, on the other hand, should be employed intentionally as an additional rather than a disruptive aspect. Visa, for example, uses a variety of digital features to generate a sense of security and consistency, including a combination of animation, sound, and haptic sensation as users complete a transaction. According to the company's research, 81 percent of customers think businesses who employ sound or animation cues to signal those online actions have taken effect have a better reputation.

[4] Cp. https://www.marketingmag.com.au/hubs-c/how-digital-sensory-marketing-is-key-to-appealing-to-todays-consumer/

Smell and Taste

Customers' emotional levels can rise by 28 percent when they smell and taste a product. Because customers do not yet have access to the long-awaited smell-o-vision or taste-o-vision capabilities, marketers must rely on descriptive words and graphics to replicate the product's scent or flavour.

Touch

It's tough to engage the sense of touch in a digital context, which has resulted in widespread webrooming and showrooming. These strategies make it more difficult for retailers to convert impulse buyers while yet allowing customers to have their chosen shopping experience. Consumers are more likely to engage in mental simulations when IoT is combined with AR and VR. To put it another way, imagine yourself doing something specific, such as using a product. Multi-gesture apps and direct touch effects, such as pinching and scrunching a material, should be considered by brands to enhance online conversions.[5]

2.3 Online Sensory Marketing Technologies

Brands can use technology to take advantage of digital sensory marketing and make the most of their ecommerce channel. The following are the most up-to-date technologies and capabilities for delivering sensory marketing in a digital setting.

Video that can be purchased

Video is a potent and more popular vehicle across web channels in today's hyper visual, digital-driven society. Brands can take this asset even further by making videos shoppable, allowing them to show clips alongside a sequence of product photos. Clickable videos bring products to life while maintaining an interactive and enjoyable experience.

Artificial Intelligence

AI is a must-have technology for a range of ecommerce applications, including customer chatbot services and giving personalized services to online shoppers. It may also automate media assets on a website, such as image and video tagging and transcription, allowing brands to optimize visual material for a shopper's sense of sight. Marketers can provide visual experiences much faster and more accurately than previously by removing manual processes and replacing them with AI.

Augmented Reality (AR)

AR delivers a unique digital experience for people by allowing them to interact with brands in real time via their mobile devices. For example, Amazon allows users to

[5] Cp. https://www.marketingmag.com.au/hubs-c/how-digital-sensory-marketing-is-key-to-appealing-to-todays-consumer/

imagine furniture in their own space, Sephora lets users to upload a selfie and "virtually try on" various things available in the store, and Quay Australia allows users to try on glasses/sunglasses to see what best suits their face shape.

Virtual Reality (VR)

Virtual Reality (VR) is another advanced sensory technology that allows clients to have an immersive experience with the goods, making online buying easier and more accessible while also appealing to several senses at the same time. Many firms are adopting virtual reality to build virtual tours, such as Diesel, which has launched its HypeRoom, which allows customers to view, spin, and interact with products in 360-degree displays online. Using Google Cardboard glasses, the North Face developed a 360-degree virtual tour of Yosemite National Park.

It Makes Sense for Online Businesses

During the pandemic, the limited effect of physical shopping venues should not limit brands. With creative technology, they can still bring multi-sensory clues to life through consumers' cellphones or laptops. Ecommerce merchants can employ digital sensory marketing to build strong, engaging online experiences this holiday season and beyond by appealing to the senses with these capabilities.[6]

2.4 Natural Language Processing

Natural language processing (NLP) is an area of computer science concerned with computers' capacity to understand text and spoken words in the same way as people do. Computational linguistics—rule-based human language modelling is combined with statistical, machine learning, and deep learning models in NLP. These technologies, when used together, allow computers to process human language in the form of text or speech data and 'understand' its full meaning, including the speaker or writer's intent and sentiment.[7] NLP is used to power computer algorithms that translate text from one language to another, respond to spoken requests, and summarise large volumes of data quickly—even in real time. Natural language processing (NLP) refers to a computer program's capacity to read spoken and written human language, often known as natural language. It's a part of AI (artificial intelligence) (AI).

How does natural language processing work?

NLP allows computers to comprehend natural language in the same way that people do. Whether the language is spoken or written, natural language processing uses artificial intelligence to absorb real-world data, interpret it, and make sense of it in a way that a computer can understand. Computers have programs to read and microphones to gather audio, much as people have diverse sensors such as ears to hear and eyes to see. And, just as people have a brain to absorb the information, computers have a software to do the same. The input is translated to code that the computer can

[6] Cp. https://www.cmswire.com/digital-marketing/how-to-use-sensory-marketing-online/
[7] Cp. https://www.ibm.com/cloud/learn/natural-language-processing

interpret at some point during the processing. Natural language processing is divided into two stages: data pre-processing and algorithm development.[8]
Data pre-processing is the process of preparing and "cleaning" text data so that machines can examine it. Pre-processing transforms data into usable form and highlights textual elements that an algorithm can use. This can be accomplished in a variety of ways, including:

Tokenization: When text is split down into smaller parts to work with, this is what happens.

Stop word removal: This is when common words are deleted from a document, leaving just the unique words that provide the most information about it.

Stemming and lemmatization: To process words, they are reduced to their base forms.

Part-of-speech tagging is a technique for identifying certain parts of speech. This is when words are classified according to their part of speech, such as nouns, verbs, and adjectives.

An algorithm is created to process the data once it has been pre-processed. There are a variety of natural language processing methods, however the following two are the most widely used:

Rules-based system: This system employs language rules that have been carefully crafted. This method was employed in the early stages of natural language processing and is still used today.

Machine learning-based system: Statistical approaches are used in machine learning algorithms. They learn to do jobs based on training data provided to them, and as more data is processed, they alter their approaches.[9]

2.5 Literature Analysis

Analysing the impact of Sensory Marketing On Customers: Study of KFC
This paper was published in general of Us-China. Publicity Administration, April 2016 states how senses are useful in a growing business in the marketing strategy around emotions and the fundamentals for the companies which differentiate themselves from the competitors.
This paper is based on the consumer's buying behaviour, especially for KFC customers. For this project, the data was collected by asking a hundred random questionnaires which were belief as preliminary required data regarding the senses. The motive behind asking the questions in the survey was how senses are useful in marketing and growing business. It was an IBM SPSS text analytics survey and was analysed using descriptive statistics and chi-square test. The analysis and findings were giving the results of experiencing music and experiencing sensations in the KFC outlet.

[8] Cp. https://searchenterpriseai.techtarget.com/definition/natural-language-processing-NLP
[9] Cp. https://searchenterpriseai.techtarget.com/definition/natural-language-processing-NLP

It stated that 47 per cent of people found the music as relaxing, in the outlet 9.7% found it exciting and 32.1% found it enjoyable. Whereas last 10.4% founded nothing. Well, in terms of experiencing sent 9% found it relaxing, 43% found it exciting and 31% founded enjoying whereas the remaining founded nothing. This shows that experiencing sent and experiencing music plays an important role in making a business using senses for KFC in their KFC restaurants.

Employing Sensory Marketing as a Promotional Advantage for creating brand differentiation and brand loyalty. This paper was published in Pakistan Journal of Commerce and social science. The data collected was by questionnaires, distributed to 520 people from which 478 questioners were returned back. And the data analysis, done by assessment of normality test where results were obtained by analysing mean, standard deviation, variance, skewness and kurtosis of the data. The results were also obtained by internal reliability test. And the significance of the model and NOVA. The taste which was used shows the effect of independent ways variables of the sensor marketing on the dependent variable by brand differentiation, the result shows the model is significant. This paper provides two major contributions. Firstly, the sensor perception item set was developed, which established a skill development relying on literature, search expert, interviews, feedbacks and several tastings. It was developed by keeping in mind, the five sense perceptions, like visual acoustic, haptic, Olfactory and gustatory. This paper shows the solution as a holistic measurement of sensory perception. And secondly, it provides the significant relationship between the brand value and the sensory marketing strategy. The results, show the inside, how the brand uses sensory marketing, as a positive experience for the customers in the context of different products and around different industries.

Digital Century Marketing: Integrating New Technologies into Sensory Online Experience

The paper is written on important senses touch and smell. And the aim of the paper is how we can implement the experience of the original product with the help of devices and technologies by sitting at home. The first technology is known as haptic enabling Technologies. It is a new form of physical contact, which states the technology which can make you feel the real experience of the product, which you are surfing online. And the second one is a technology used to experience the smell of a of a product by just viewing its image with the help of chemicals and devices. This technology makes it feel like almost real life experience.

3. Methodology

3.1 Data Descriptions

We have created our own dataset by scraping through different web portal which include attribute like BRANCHE, COMPANY, TEXT.
BRANCHE: This attribute contains branch of the company e.g. Food, Technology, automobile etc.
COMPANY: This attribute contains the company names of following branches.
TEXT: This attribute contains Text data for companies.

3.2 Data Pre-processing

What is Data Pre-processing?
In the actual world, data is frequently incomplete: it lacks attribute values, specific attributes of relevance are missing, or it simply contains aggregate data. Errors or outliers make the data noisy. Inconsistent: having inconsistencies in codes or names.
Following are the steps involve in textual data pre-processing:
Step 1: Import the libraries
Step 2: Import the dataset
Step 3: Checkout the missing values
Step 4: NLP pre-processing (Removing Special Character, lowercase, stemming, removing stop wards)

Step 1: Import libraries

```
import pandas as pd
import numpy as np
from textblob import TextBlob
from sklearn.feature_extraction.text import CountVectorizer
import warnings
warnings.filterwarnings('ignore')
import matplotlib
import matplotlib.pyplot as plt
import seaborn as sns
matplotlib.rcParams['figure.figsize'] = (10.0, 6.0)
import plotly.graph_objs as go
import cufflinks
pd.options.display.max_columns = 30
from IPython.core.interactiveshell import InteractiveShell
from chart_studio import plotly as py
import plotly.figure_factory as ff
InteractiveShell.ast_node_interactivity = 'all'
from plotly.offline import iplot
cufflinks.go_offline()
cufflinks.set_config_file(world_readable=True, theme='pearl')
from bokeh.io import output_notebook
output_notebook()
```

[5] ✓ 0.1s Python

Figure (2): Import Libraries

This is how we utilize the import keyword in Python to import libraries, and these are the most commonly used libraries by Data Scientists. (Jupyter Notebook was my tool of choice.)

NumPy

NumPy (Numerical Python) is an excellent tool for doing scientific computations and simple and complex array operations.
The library has a lot of useful features for working with n-arrays and matrices in Python. It facilitates the processing of arrays that store values of the same data type and simplifies array math operations (including vectorization). In reality, using the NumPy array type to vectorize mathematical operations improves speed and reduces execution time.

Pandas

In the data science life cycle, Pandas (Python data analysis) is a requirement. Along with NumPy in matplotlib, it is the most popular and commonly used Python package for data research. It is frequently used for data analysis and cleansing, with about 17,00 comments on GitHub and a community of 1,200 contributors. Pandas offers quick, versatile data structures, such as data frame CDs, that make working with structured data simple and natural.[10]

Warnings

Warning messages are often used to warn the user of a condition in a program that does not require raising an exception and terminating the program. When a program uses an obsolete module, for example, a warning should be issued.[11]

Step 2: Import the Dataset

Figure (3): Import Dataset

We import our data-set using Pandas, and the file I used here is. csv format [Note: You don't have to deal with CSV files all of the time; you can also deal with Html or Xlsx (Excel file)].
However, because of their minimal weight, we employ CSV files to access and use quickly. You can see that we utilize the head method (This function retrieves the first n rows for the object based on position) after importing the dataset. It's useful for

[10] Cp https://www.dataquest.io/blog/15-python-libraries-for-data-science/
[11] Cp https://hackernoon.com/what-steps-should-one-take-while-doing-data-preprocessing-502c993e1caa

quickly determining whether your object contains the correct type of data. It returns 5 rows by default.)

Step 3: Checkout the missing values

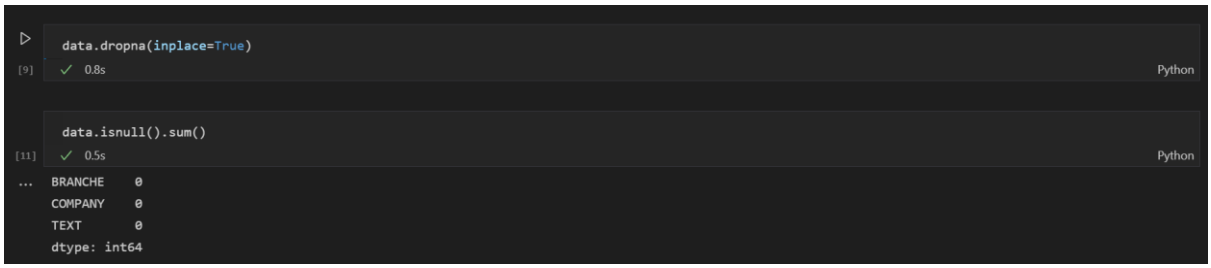


Figure (4): Missing Value Imputation

The concept of missing values is critical to grasp in order to manage data effectively. If the researcher does not appropriately handle the missing numbers, he or she may make an incorrect inference about the data. The researcher's results will differ from those where missing values are present due to inappropriate management.

Drop the Missing value technique

This method can be used on a feature that contains numeric data, such as the year column or the goal column for the home team. The feature's mean, median, or mode can be calculated and replaced with missing values. This is an approximation that can contribute to the data set's variance. However, the loss of data can be mitigated by using this strategy, which produces better results than removing rows and columns. A statistical strategy to addressing missing values is to replace with the above three approximations. This strategy is also known as data leakage during training. Another option is to use the deviation of nearby numbers to estimate it. If the data is linear, this works better.

Step 4: NLP pre-processing (lowercase, stemming, removing stopwards)

Figure (5): Text Pre-processing

Lowercase

Although sometimes missed, one of the simplest and most effective forms of text preparation is to lowercase all of your text data. It can be used to solve most text mining

and NLP problems, and it can be very useful when your dataset isn't very huge. It also considerably improves predicted output consistency.

Stemming
The process of eliminating inflection in words (for example, disturbed, difficulties) to their root form is known as stemming (e.g. trouble). In this situation, the "root" may not be a true root word, but rather a canonical variant of the original term.
Stemming is a rudimentary heuristic procedure that chops off the ends of words in the hopes of appropriately converting them to their root form. Because the ends were just hacked off (ughh, how vulgar!), the terms "trouble," "troubled," and "troubles" might actually be changed to troublinstead of trouble.

Stopword Removal
Stop words are a group of words that are frequently employed in a language. Stop words in English include "a," "the," "is," "are," and others. The idea behind stop words is that by deleting low-information terms from a document, we can concentrate on the crucial words.
For example, if your search question is "what is text preprocessing?" in the context of a search engine, you want the search system to prioritize surfacing results that discuss text preprocessing over documents that discuss what is. This can be accomplished by blocking the analysis of all terms from your stop word list. Stop words are frequently used in search engines, text classification software, topic modeling, and topic extraction, among other applications.

3.3 Data Visualization

Data visualization is the practice of attempting to comprehend data by presenting it in a visual context in order to reveal patterns, trends, and connections that might otherwise go undetected.
Python has a number of excellent graphing packages with a variety of useful features. Python provides a great library for creating interactive, live, or highly customizable charts.

Matplotlib
The most well-known Python data visualization library is matplotlib. Despite the fact that it is almost a decade old, it is still the most extensively used plotting package in the Python community. It was created to seem like MATLAB, a proprietary programming language that was established in the 1980s.[12]

Seaborn
In just a few lines of code, Seaborn uses matplotlib to create stunning graphs. Seaborn's default styles and color palettes, which are supposed to be more aesthetically pleasant and modern, are the main distinction. You'll need to know matplotlib to alter Seaborn's defaults because Seaborn is built on top of it.

[12] Cp https://www.kaggle.com/getting-started/108792

Plotly

You may be familiar with Plotly as an online data visualization platform, but did you realize you can use it from a Python notebook? Plotly, like Bokeh, specializes in interactive plots, but it also has several charts that aren't found in most libraries, such as contour plots, dendograms, and 3D charts.

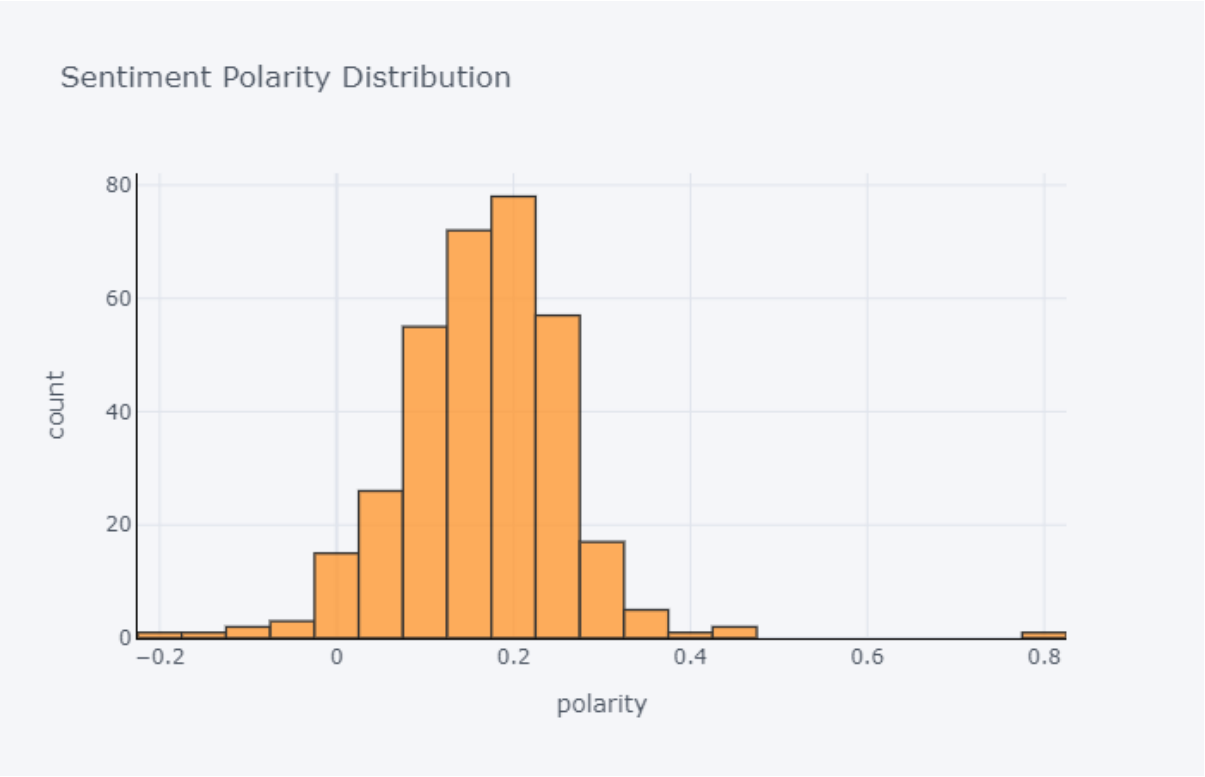


Figure (6): Sentiment Polarity Distribution

In above chart we have plot the bar graph for sentiment polarity count where x axis represents the polarity which ranges from -1 to 1 and y axis represent the count of each polarity value.

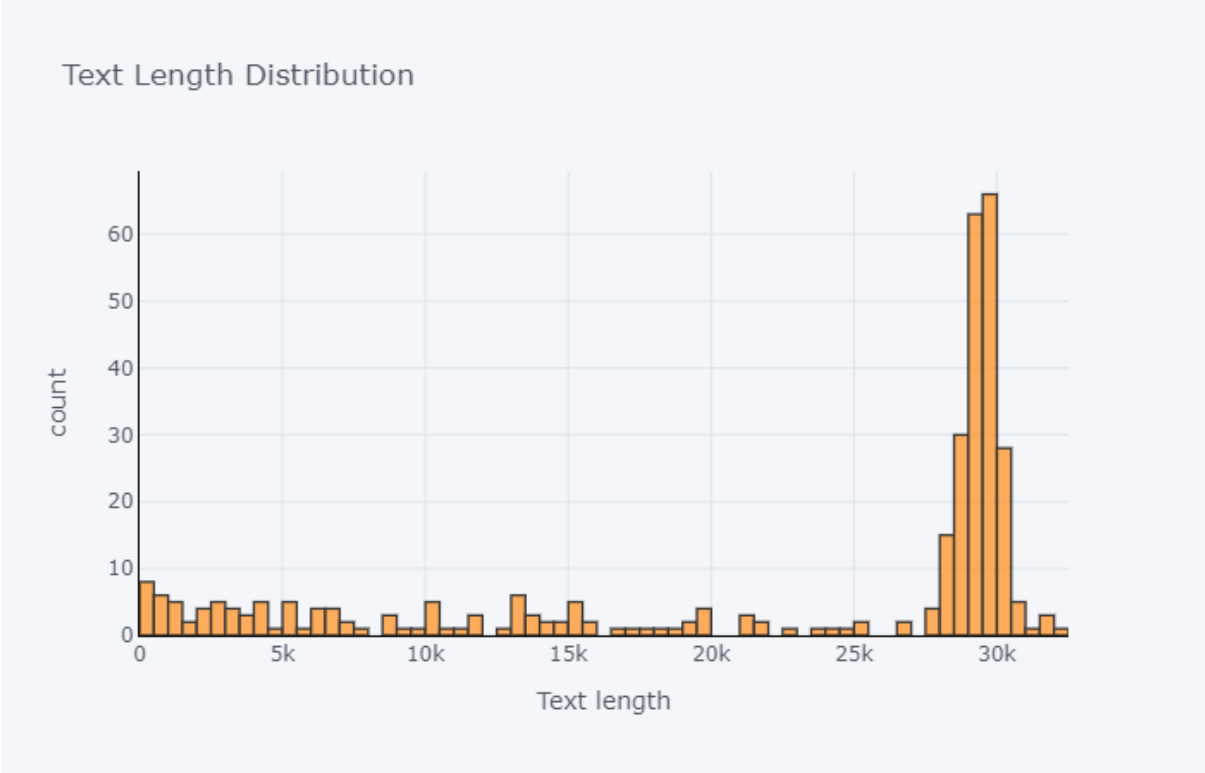


Figure (7): Text Length Distribution

In above chart we have plot the bar graph for Text Length Distribution for each row. This graph shows the text length count for each row. Where x axis represents the text length count and y axis represent the count.

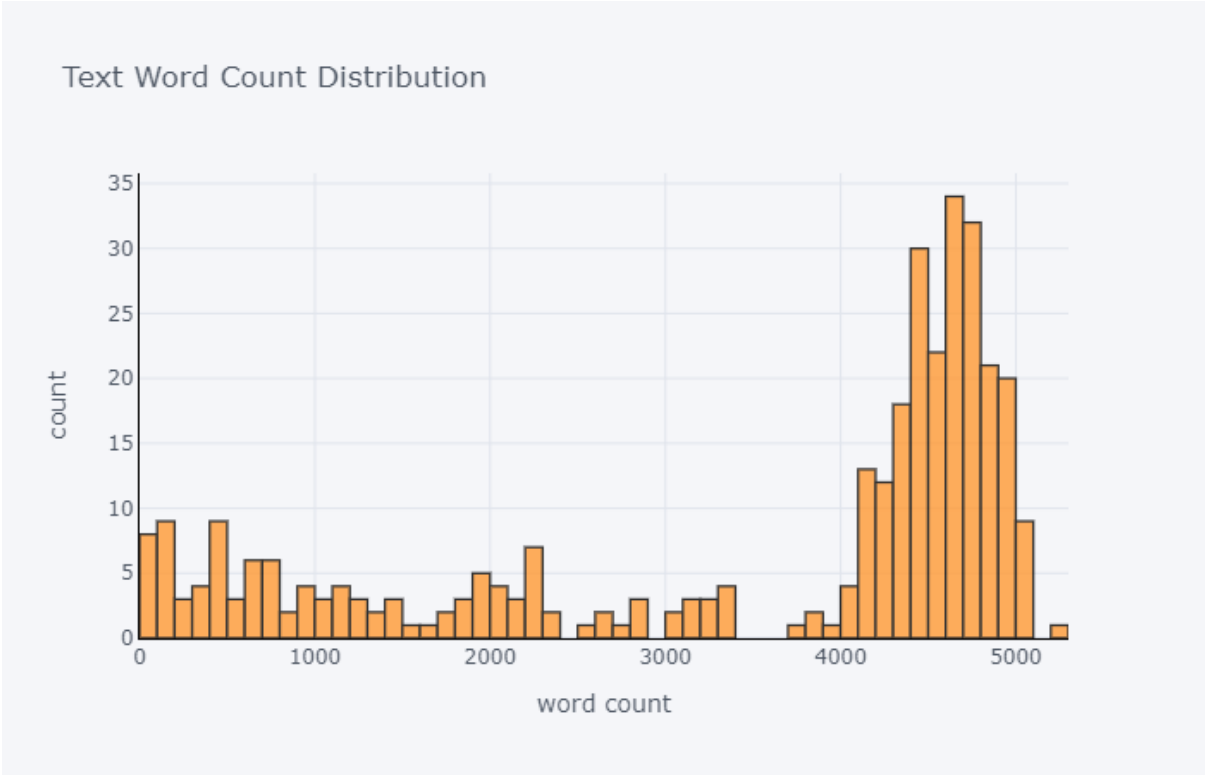


Figure (8): Text Word Distribution

In the above chart we have plot the bar graph for text word count distribution. This graph represent the word count distribution with respect to each row. Where x axis represents the word count and y axis represent the count.

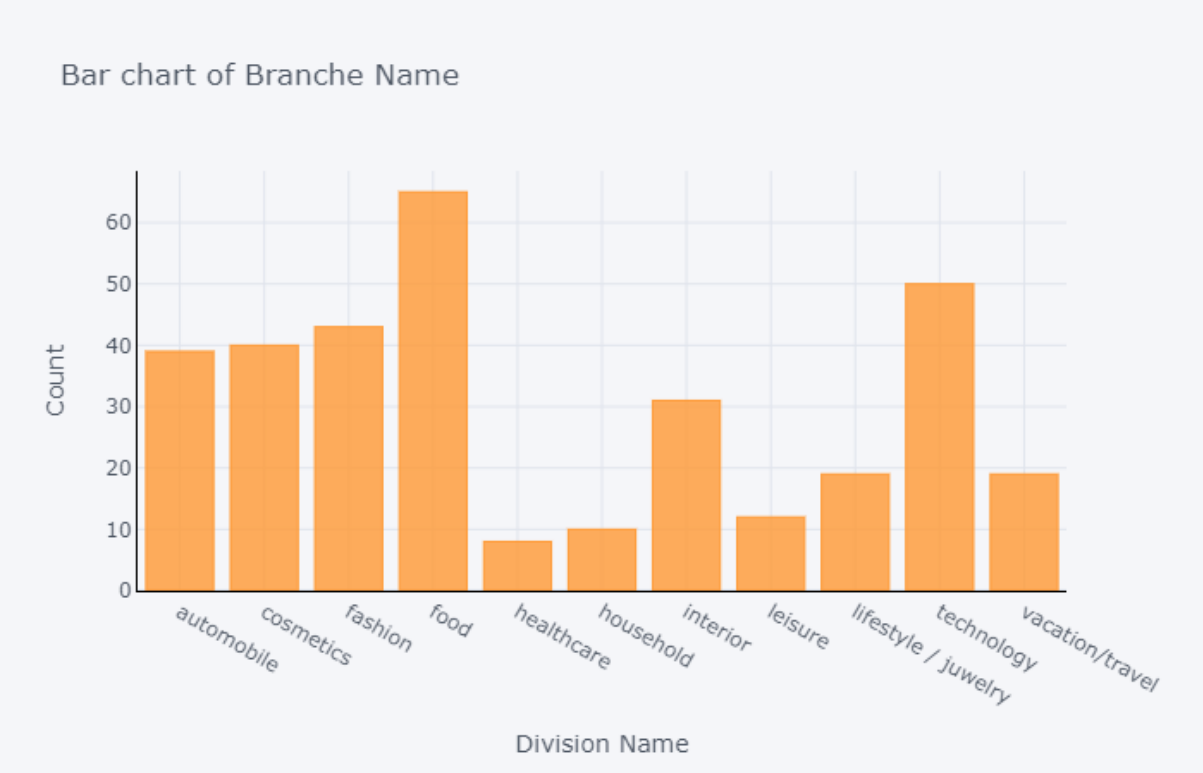
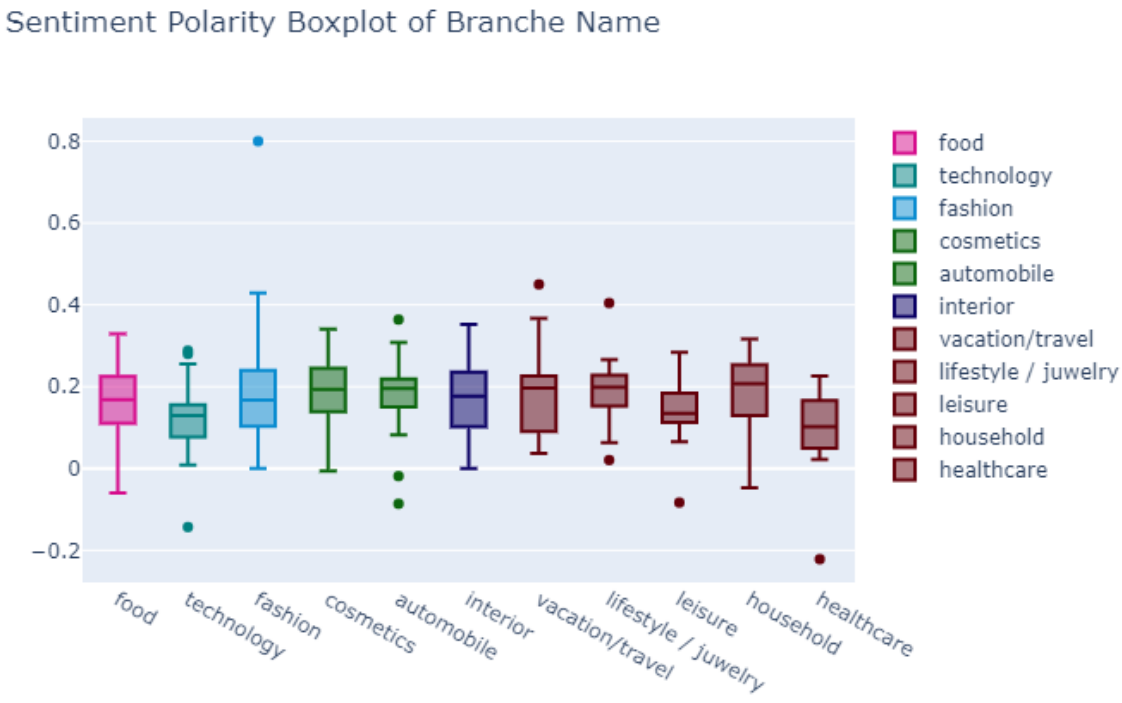


Figure (9): Branch Name Count

In the above chart we have plot the bar graph for Branch Name Count. This graph represent the Company Branch count where x axis represent the Branch name and y axis represent the count.

Figure (10): Sentiment Polarity Boxplot of Branch Name



The above chart we have plot box plot for Sentiment Polarity distribution for Branch Name.

3.4 Sentiment Analysis

Sentiment analysis is the process of determining whether a sentence is positive, negative, or neutral. Text analytics' sentiment analysis system combines natural language processing (NLP) and machine learning techniques to assign subject, theme and category weighted sentiment scores to entities in a sentence or phrase. Sentiment analysis algorithms are primarily focused on defining the opinions, attitudes, and even emotions of a text corpus. The range of established emotions varies greatly from method to method. Standard analyzers define up to three basic polarity emotions (positive, negative, and neutral), while more advanced models are more limited. As a result, they can determine six "universal" emotions beyond the anode (e.g., anger, disgust, fear, happiness, sadness, surprise). The origin of Kansei analysis can be traced back to the 1950s, when Kansei analysis was mainly used for paper documents. However, today it is widely used for sentiment analysis by mining subjective information on Internet content, including texts, tweets, blogs, social media, news articles, reviews and comments. This is done using a variety of techniques, including NLP, statistics, and machine learning techniques.

These organizations use mined information to identify new opportunities and more effectively target targeted demographic messages. The Obama administration uses sentiment analysis to predict the public reaction to policy announcements. But before we can analyze emotional texts and mouths, we need to understand the parts that make them up. The process of dividing a document into components involves multiple sub-functions, including part-of-speech (PoS) tagging. Part-of-speech tagging is verbs, nouns, adjectives and adverbs. All lights are ignited by any person, any place, any person. Recently, scientists and scholars have done much times in the field of appreciation analysis. During the initiation period, the binary classification was expected to classify between positive and negative. One of the text mining applications is sentiment analysis. Most of the data is in text format, and over the last few years people have discussed more about NLP.

Improvement is an ongoing process, and many product-based companies are looking to leverage these character mining technologies to identify customer sentiment and improve their products. This information helps you understand end-user trends and needs that lead to customer satisfaction. Text mining is a huge concept, so this article is divided into two sub-chapter. The main focus of this article is to use Python to calculate two emotional bipolar and subjective scores. The polarity range is 1 to 1 (sound volume), which indicates whether the text contains positive or negative feedback. Most companies like to stop their analysis here, but in the second article, we want to create some labels from these scores and expand our analysis. Finally, you can train multiple label multiclass classifiers and predict future reviews. Sentiment analysis is an incredibly difficult task for humans. On average, in relation to sentiment analysis, comment-to-comment agreements (measuring how two (or more) labelers can make the same comment decision) are very low. And because machines learn from the data provided, sentiment analysis categorizers may not be as accurate as other types of sorters. All inflammations are fired in the portal vein. It is very difficult to analyze

emotions without confusion. However, the system cannot learn about the context unless explicitly stated. One of the problems that arises in context is the change in polarity.

Why Sentiment Analysis is Required?

In today's environment, which suffers from data overloads (although this does not mean better or deeper insight), companies can distract the collected customer feedback. But for just humans, it is not yet possible to manually analyze any kind of error or prejudice. Often the company with the highest intent is in the power vacuum of insight. We know that we need insight to inform our decisions. And you know it's missing. But you don't know the best way to get them. Sentiment analysis provides an answer to what is the most important question. Sentiment analysis can be automated, so decisions can be made based on a significant amount of data rather than always incorrect and simple intuition.

Types of Sentiment Analysis

People are having a very wide range of emotions like sad or happy, interested or uninterested so different sentiment analysis model are available to capture the variety of emotion's of people, so the different kind of sentiment analysis are positive, negative and neutral.

Fine -Grained:

This model helps us to derive polarity precision where we can conduct sentiment analysis across the polarity categories such like positive, very positive, negative, neutral or very negative. So if we consider the rating scale for instance scale from 1 to 5 where 1 will be considered as very negative and 5 as very positive, if for a scale 1 to 10 we can consider 1-2 as very negative and 9-10 as very positive, so this model helps us to extract the sentiments of people according to the scale.

Aspect-Based:

Compared to the previous model of fine grained which helps us to see the general polarity of customer reviews, Aspect based analysis delves deeper. It also helps us to dig deeper and determine the actual aspects of which individuals are talking about. So as an example a review is there of a customer like if you're an automobile manufacturing company and you bought a review stating the dashboard may be more attractive. With aspect based analysis we will easily determine that the reviewer commented something negative or positive about the dashboard of the car.

Emotion Detection:

Emotion Detection because the name suggests helps us to detect the emotions which incorporates anger, sadness, happiness, fear, frustration, worry, panic , tension etc. Typically it uses lexicons which may be a collection of words that convey certain emotions to people, some advanced classifiers also utilize robust machine learning algorithms. Recommendations are to use ml over lexicons because people express emotions in an exceedingly various of the way like for instance like "I am close to sink" this line expresses panic and fear but in an exceedingly similar way we will say "I will fall down for you" has a completely positive meaning but the keyword fall is related to panic or fear during this lexicon which can result in inaccurate emotion detection.

Intent Analysis:
Intent analysis determines the consumer's intent which may save company's time, money, effort because most of the days businesses find yourself by chasing the buyer that don't commit to buy anything anytime soon so if we will apply accurate intent analysis we are able to solve this issue. This model also helps us to spot the intent of the customer whether the customer will buy some products or is simply browsing around. So if the customer is willing to buy we will track them and target specifically by advertisements. If the customer doesn't intent to shop for any product we are able to save time on our resources and also our advertisement.

Subjectivity/Objectivity identification:
Subjectivity / objectivity identification involves classifying sentences and text fragments into one of two categories: subjective or objectivity. However, it should be noted that there are difficulties in performing these types of analysis. The main problem is that the meaning of words often varies depending on the context.

Procedure to perform sentiment analysis.
In today's modern-day sentiment analysis, the approaches are classified into three major categories well known as knowledge based, statistical based and hybrid. Using this three categories sentiment analysis are performed.
Knowledge Based: This type of approach includes the classification of text based on words that depicts emotion.
Statistical: This type of approach uses statistical method which utilizes machine learning algorithms such as latent semantic analysis and deep learning for accurate sentiment detection.
Hybrid: This type of approach leverages both knowledge based and statistical techniques for on point sentiment analysis.

Sentiment Analysis using TextBlob

Textblob is a python library for processing of textual data. It came into existence for Steven Loria, which provides a simple API for getting into common natural language processing tasks such as part of speech, noun phrase extraction or sentiment analysis. The difference that the TextBlob package applies to sentiment analysis is that it requires a predefined set of categorized words because it is rule-based. For example, these words can be uploaded from the NLTK database. You can also define emotions based on semantic relationships and the frequency of each word in the input sentence, resulting in more accurate output. Textblob output defines a polarity task which is float within the range [-1.0, 1.0] where -1.0 is a negative polarity and 1.0 is a positive polarity. Score can also turn equal to 0 which stands for a neutral evaluation of a statement as it doesn't contain any words from the training set.

The subjectivity or Objectivity identification task reports a float range of [0.0, 1.0] where 0.0 is a very objective sentence and 1.0 is very subjective. So, there are scattered examples of python interaction with textblob sentiment analyzer starting from a model based on different datasets for calculating tweet sentiments through the twitter API. The sentiment property is also named tuple of the form of sentiment for polarity and subjectivity. One of the most important things about the textblob is that it allows the user to choose an algorithm for implementation of the high level nlp tasks.

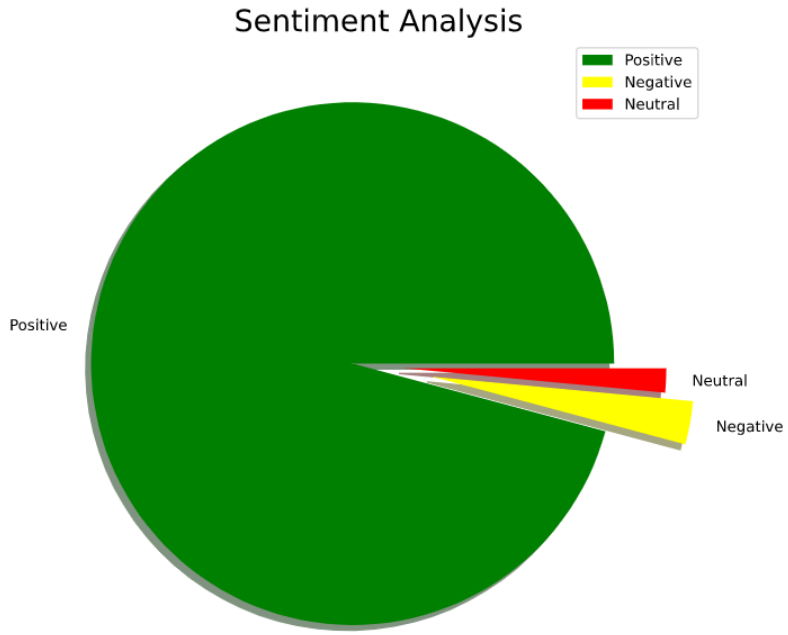


Figure (11): Sentiment Analysis

In the above graph it is clearly shown that majority of the companies has a positive sentiment.

4. Conclusion

By going through the data and analyzing it we can see the behavior of the data and discuss the value preposition of data and use it to build the model which will help in growing the business using Sensory data. The data which we received by reviews of the company has mostly positive reviews and can be used as key feature to build a OSMI model for business development.

References

1. *Aaker, J. L. (1997). Dimensions of Brand Personality. SSRN Electronic Journal, 1. Almeida e Silva, C. M., Okimoto, M. L. R. L., & Tanure, R. L. Z. (2012). Product perception from sensory stimuli: the case of vacuum cleaner. Work, 41, 1178–1184.*
2. *Amsteus, M., Olsson, H., & Paulsson, R. (2015). The scent of a successful venue:(In) congruent scent and consumer attitude towards a café. International Journal of Business and Social Science, 6(5), 232- 243.*
3. *Aslam, M. M. (2006). Are You Selling the Right Colour? A Cross-cultural Review of Colour as a Marketing Cue. Journal of Marketing Communications, 12(1), 15–30.*
4. *Bagozzi, R. P., & Yi, Y. (2011). Specification, evaluation, and interpretation of structural equation models. Journal of the Academy of Marketing Science, 40(1), 8–34.*
5. *Bell, S. and C.P. Bell: 2007, 'Future Sense: Defining Brands through Scent,' The Journal of the Marketing Society 38 (Autumn).*
6. *Bitner, M. J. (1992). Servicescapes: The Impact of Physical Surroundings on Customers and Employees. Journal of Marketing, 56(2), 57.*
7. *Bradford, A. (2017). Deductive reasoning vs. inductive reasoning. Live Science.*
8. *Medina, J. (2011).Brain rules: 12 principles for surviving and thriving at work, home, and school. ReadHowYouWant. com.*
9. *Celier, P. (2004). Le marketing sensoriel (Sensorial marketing). Mohammedia ENSET School.*
10. *Cox, D. F. (1969). The sorting rule model of the consumer product evaluation process, risk taking and information handling in consumer behavior. Boston, MA: Graduate School of Business Administration, Harvard University. Pp. 324-369.*
11. *Hirsch, A. R., & Gay, S. E. (1991). The effect of ambient olfactory stimuli on the evaluation of a common consumer product. Proceedings from 13th Annual Meeting of Association for Chemoreception Sciences, April.*
12. *Hul, M. K., Dube, L., & Chebat, J.-C. (1997). The impact of music on consumers' reactions to waiting for services. Journal of Retailing, 73(1), 87-104.*
13. *Hultén, B. (2011). Sensory marketing: The multi-sensory brand-experience concept. European Business Review, 23(3), 256-273. Hultén, B. (2012). Sensory cues and shoppers' touching behaviour: The case of IKEA. International Journal of Retail &*
14. *Distribution Management, 40(4), 273-289.*
15. *Hultén, B., Broweus, N., van Dijk, M., & Waxberg, C. (2008). Sinnesmarknadsforing (Marketing of the senses). Malmo, Sweden:*
16. *Imram, N. (1999). The role of visual cues in consumer perception and acceptance of a food product. Nutrition & Food Science,*
17. *99(5), 224-230.*
18. *Kahn, B., & Deng, X. (2010). Effects on visual weight perceptions of product image locations on packaging. In A. Krishna (Ed.),*
19. *Sensory marketing: Research on the sensuality of products (pp. 259-278). New York: Taylor and Francis Group.*

20. *Kellaris, J. J., & Kent, R. J. (1992). The influence of music on consumers' temporal perceptions: Does time fly when you're*
21. *having fun? Journal of Consumer Psychology, 1(4), 365-376.*
22. *Kellaris, J. J., & Kent, R. J. (1993). An exploratory investigation of responses elicited by music varying in tempo, tonality, and*
23. *texture. Journal of Consumer Psychology, 2(4), 381-401.*
24. *Kotler, P. (1974). Atmospherics as a marketing tool. Journal of Retailing, 49(4), 48-61.*
25. *Krishna, A. (2010). Sensory marketing: Research on the sensuality of products. New York: Routledge.*
26. *Krishna, A., & Morrin, M. (2008). Does touch affect taste? The perceptual transfer of product container haptic cues. The Journal*
27. *of Consumer Research, 34, 807-818.*
28. *Laird, D. A. (1935). How the consumer estimates quality by subconscious sensory impressions: With special reference to the role*
29. *of smell. Journal of Applied Psychology, 16, 241-246.*
30. *Lindstrom, M. (2005). Broad sensory branding. Journal of Product & Brand Management, 14(2), 84-87.*
31. *McCabe, D., & Nowlis, S. (2003). The effect of examining actual products or product descriptions on consumer preference.*
32. *Journal of Consumer Psychology, 13(4), 431-439.*
33. *Messaris, P. (1997). Visual persuasion: The role of images in advertising. Thousand Oaks, CA: Sage.*
34. *Milliman, R. E. (1986). Using background music to affect the behavior of supermarket shoppers. Journal of Marketing, 46(3), 86-91. Morrin, M., & Ratneshwar, S. (2003). Does it make sense to use scents to enhance brand memory? Journal of Marketing*
35. *Research, 40(1), 10-25.*

Declaration in lieu of Oath

I hereby declare that I produced the submitted report with no assistance from any other party and without the use of any unauthorized aids and, in particular, that I have marked as quotations all passages which are reproduced verbatim or near-verbatim from publications. Also, I declare that the submitted print version of this thesis is identical with its digital version. Further, I declare that this report has never been submitted before to any examination board in either its present form or in any other similar version. I herewith agree/disagree that this report may be published. I herewith consent that this report may be uploaded to the server of external contractors for the purpose of submitting it to the contractors' plagiarism detection systems. Uploading this report for the purpose of submitting it to plagiarism detection systems is not a form of publication.

26/08/2021

Sagar Prabhu Panchal	543293
Pukhraj Danaram Suthar	543289
Taral Sanjay Sarvagod	543298
Shreyashi	543297
Hitesh Ashok Kale	543263
Pratik Keshav Dhumal	543272
Sagnik Patra	550496