SENTIMENT ANALYSIS OF PROTEIN BAR CONSUMERS FOR EFFECTIVE DIGITAL MARKETING IN IRISH RETAIL SECTOR – A BEHAVIOURAL APPROACH

Niranjan Ashok Hallur

18175996

M.Sc. in Management

NATIONAL COLLEGE OF IRELAND

SUBMITTED TO NATIONAL COLLEGE OF IRELAND

MAY, 2021

Niranjan Ashok Hallur
18175996

Abstract

A good marketing design promotes successful conveyance of information about the product or a service to its customers. It is considered as a medium through which organisations acquire their customers and capture potential markets. A conventional marketing strategy aims to bring a product to its customers through promotions, public relations and advertising. However, conventional marketing is not effective in communicating with the consumers on other wider segments like convenience seekers, relationship seekers and it lacks in defining the online audiences. Digital marketing targets these issues affordably. It allows organisations to communicate with wider sections of customer through digital media. This study is focused towards improving the efficiency of digital marketing approaches for protein bars segment in the Irish Retail Market through data mining and sentiment analysis. Research also highlights the market analysis on the variety of protein bars being sold in the Irish retail market and also the customer analysis highlighting the number of consumers that are currently purchasing form various online platforms. Reviews and comments from these platforms were mined to identify the customer opinions. A Sentiment Analysis algorithm was used to determine the emotional polarity towards the various protein bars. The analysis indicated higher levels of positivity for local brands. The study also sheds light on digital marketing approach based on behavioural personalisation for the indicated results. The accuracy of the algorithm was tested at 93.45% and the research indicates, behavioural personalisation through sentiment analysis can help contribute to e-commerce community in the process of inventory management.

Declaration

Submission of thesis and dissertation

NATIONAL COLLEGE OF IRELAND

RESEARCH STUDENTS DECLARATION FORM

NAME: NIRANJAN ASHOK HALLUR

Student Number: 18175996

Degree for which thesis is submitted: M.Sc. IN MANAGEMENT

Materials submitted for Award

a) I declare that the work submitted here is written by me only

b) I hereby declare that all the extracts taken from other secondary sources are properly cited and such information have been specifically acknowledged

c) My thesis will be included in an electronic format in the College Institutional Repository TRAP (Thesis reports and projects).

d)***Either*** ******I declare that no material contained in the thesis has been used in any other submission for an academic award.

Or * I declare that the following material is contained in the thesis formed part of a submission for the award of **M.Sc. IN MANAGEMENT.**

Signature of the research student: NIRANJAN ASHOK HALLUR

Date: 05thMAY, 2019

Submission of Thesis to Norma Smurfit Library, National College of Ireland.

Student name: **NIRANJAN ASHOK HALLUR**

Student number: **18175996**

School: **NATIONAL COLLEGE OF IRELAND**

Degree to be awarded: **M.Sc. IN MANAGEMENT**

Title of Thesis: **SENTIMENT ANALYSIS OF PROTEIN BAR CONSUMERS FOR EFFECTIVE DIGITAL MARKETING IN IRISH RETAIL SECTOR**

One hardbound copy of your thesis will be lodged in the Norma Smurfit Library and will be available for consultation. The electronic copy will be accessible in TRAP (http://trap.ncirl.ie/), the National College of Ireland's Repository. In accordance with the normal academic library practice, all theses lodged in the National College of Ireland Institutional Repository (TRAP) are made available on open access.

I agree to a hardbound copy of my thesis being available for consultation in the library. I also agree to an electronic copy of my thesis being made publicly available on the National College of Ireland's Institutional Repository TRAP.

Signature of Candidate: NIRANJAN ASHOK HALLUR

For completion by the School;

The aforementioned thesis was received by _____

Date:_____

This signed form must be appended to all hard bound and electronic copies of your thesis submitted to your school.

Acknowledgement

I would like to acknowledge my recognitions to various people who have travelled with me for the past few months, as I worked on this thesis. Primarily, I owe an enormous gratefulness to my supervisor, Prof. Naill O'Brien for his continuous encouragement and support. I offer my sincere appreciation for the learning opportunities provided by him. I could not have imagined having a better mentor for my research.

The completion of this thesis would not have been accomplished without the support of my friends who gave insights on the technical aspect of this thesis. I am greatly indebted to the Faculty of Graduate Studies at National College of Ireland for constant guidance, support and supervision and for providing essential resources required for this research.

Finally, I would like to acknowledge my deep gratitude to my parents for supporting me financially throughout the journey of my research work and for their unwavering love and my sincere appreciations to everyone who willingly helped me with all their abilities in completion of this research.

Table of Contents

Index of Figures

Index of Tables

1. Introduction

Marketing is an important component that contributes to the success of any business. It essentially includes the process of understanding the requirements of consumers and thereby recognising the target market that the organisation can serve best. Effective implementation marketing contributes to the development of a compelling value proposition through which businesses can entice and develop valued customers. The impression of today's modern marketing is not only to promote the sales but also to engage and build deeper relationships with customers that help brands to become significant part of consumers' lives (Armstrong & Kotler, 2016). However, the tried and proven conventional marketing practices are changing among the marketers as they start to adopt varying technologies of the digital age.

The traditional techniques of marketing are being redefined to accommodate the requirements of growing consumers with the help of technology and mass media communication networks playing major role in digital marketing. These new and updated techniques are bridging the gap between the organisations and their consumers. Moreover, the attention towards Information and Communication Technologies was also accompanied by globalisation. Fast connectivity and the widespread network fuelled by today's internet, enhanced the process of convergence by facilitating global communication channels between the organisations and its customers. Extending these techniques to current e-commerce models not only subsidised the business activities globally but also it added to the value chain of firms across various business sectors (Kraemer, et al., 2006).

It is essential for small business ventures to have precise marketing approaches. Businesses concerning products such as protein bars is an overlooked part of the wider retail market. The extensive use of digital platforms and the involvement of ICT in sales and business propositions of these products' id creating abundance of business data. This study is focused towards understanding the opinions of different segments of customers through a data mining approach called sentiment analysis. The research objectives and the rationale behind the study are given in a later section to justify the value of current research.

1.1. Research Context and Rationale behind the study

Organisations and marketers are defining a new set of customer relationship strategies and tools using the internet, smartphones, social media etc. for constant engagement with of consumers irrespective of brand experiences, conversations and communities. An efficient marketing strategy will reflect on the market share, customer equity and profits of the company. Customer engagement and customer value are essential for a good marketing strategy.

Following the changes brought by the information and communication technology, the process of digital marketing has established into one of the most common and effective ways of communication for organisations to know their consumers. It enhanced the conventional ways of product marketing by including internet resources, telecommunication networks and other global channels for inter and intra firm relations. Moreover, the complexity behind executing these techniques was simplified as the customers across the world are able to have seamless access to various products and services via entertainment media offered by different online platforms. Consumers and suppliers are able to buy and sell respectively on e-commerce

platforms like Amazon. Organisations are creating brand and product related content for marketing purposes using high speed communication channels and long-term evolution technologies like 5G internet, in the aims to target end devices such as smartphones, tablets and computers.

Not only have the consumers benefitted but also the suppliers are able to use technology as a medium to analyse and identify the demand characteristics of a product. (Kim & Rossi, 1994) have identified the relation between the redundancy of a type of product to its demand. The existence of similar type of products from different suppliers or of different brands may cause biased opinions in the purchase choices of customers irrespective of the extensive marketing. Customers tend to have uncertainty in the purchase process when there is redundancy in products and product features (Kang & Zhou, 2016). However, the purchase behaviours of consumers are changing along with technology creating new demographics of consumers. Identifying these new demographics has become one of the challenging factors contributing to the success of the companies especially in retail industry. Hence, it is deemed necessary for companies to define new digital marketing techniques to align with the customer requirements and to capture new demographics.

1.2. Research Objectives

The study investigated the retail sector in the Republic of Ireland with respect to Protein bars. The protein bar market is a large yet elusive in retail sector as they are considered as a healthy snack. The customers for these products are not regular but unique with certain health concerning goals. With changing lifestyles, consumers are changing their dietary routines and behavioural patterns towards snacks and other retail products. The evaluation of current market situation regarding these products is addressed. The study is also focused on identifying the emotions of consumers in association with already existing products that are being sold on e-commerce platforms. Evaluating these emotions provide direct insights about the influence of sentiments (emotions) in customer purchase behaviour and their socio-economic progression.

Listed following are the objectives defined for this research;

- Identifying and collecting text data from websites such as Amazon.com, ebay.com, etc. and other company specific websites like My Protein and Fulfil bars.
- Mining multi-domain text data to extract comments, ideas, opinions and feelings towards a defined product
- Establishing divergence in customer emotions through data mining
- Identifying the influence of customer emotions in purchase patterns of the product

1.3. Research Purpose and Value of Study

The current study is aimed at identifying the customer opinions about different protein bars available in Irish retail market. As mentioned above, text data such as comments and ideas with respect to available products in the market will be considered for data mining. With regard to the protein bar markets, there are numerous products from different manufacturers such as ***Optimum Nutrition by Glanbia, Barbells Functional Foods' Barbell Protein Bars, My Protein's Protein Bars, Quest Nutrition's Quest Bars, Origin Bars, Muscletech's NitroTech bars etc.*** competing in the retail sector in the Republic of Ireland. As protein bars are sold in

retail as snack bars, other relatively cheaper products such as cereal bars and nut bars can also be considered as competitors. Moreover, some products are being imported into Irish markets, resulting in tight competition. These products are being sold in various retail outlets like Supervalu Stores, Marks and Spencer, Spar etc. Apart from these outlets, some sellers are capturing the market with the help of e- commerce platforms such as Amazon, eBay, and their own websites.

The purpose of this research is to find the relevance of human emotions in a product purchase as they can contribute to the organisational decisions. The data pertaining to protein products that comprise a particular segment of retail sector are considered appropriate to establish the research. Protein bars are selected particularly as the study concerning the driving factors behind the product sales is significantly limited. Moreover, the growth in demand for these products in growing in different dimensions especially in e-commerce as the consumers are looking for better dietary alternatives. The purpose of this study is also directed to justify the fact that internet is not just a medium to buy and sell products, rather it is a platform to improve the competitive efficiency in the market. The study adds value to business community by offering operational insights about the consumers such as promotions to improve customer satisfaction and thus contributes to the customer value.

For products with high demand, there exists a need to identify the consumer requirements. The hypothesis of the study is that ***"customer opinions can be identified through sentiment analysis and the opinions thus extracted can be used for effective Digital Marketing of a product (Protein Bars) in Ireland."*** A data driven application such as *Emotion AI*[1]can be utilised to analyse datasets of multi domain comments extracted from Amazon.com and Supplier's websites.

The primary stimulus for the study is the role of digital marketing in e-commerce which will be discussed in later sections. The literature review below emphasises the importance of data and the data mining practices involved. Factors influencing the consumer perceptions towards products and their purchase behaviours were observed based on the current market trends. To understand the status of the market along with market characteristics such as the various products available, existing brands and other market analytics are included in the later sections of Literature Review. By identifying the research gaps, research questions were then formulated and the methodology utilised is based on pre-existing data mining techniques and new algorithms such as sentiment analysis. The process involved in the evaluation is mentioned in the Methodology section. This technique was selected due to its efficiency with respect to text data. Results from evaluating the collected sample of data is presented and the limitations and pit falls of this approach were identified in the discussion section of this research

[1] Shashank Gupta, *Sentiment Analysis: Concept, Analysis and Applications*, towardsdatascience.com (Online Source) Accessed on: 22-10-2020
Available at: https://towardsdatascience.com/sentiment-analysis-concept-analysis-and-applications-6c94d6f58c17

2. Literature Review

The advent of the digital revolution has introduced various marketing techniques and it altered the conventional ways of marketing to the very core. By communicating with the consumers electronically via the Internet and Communication Technologies (ICT) and other conventional communication channels, corporations are creating enormous amounts of data concerning the products, consumers and the way consumers respond. Such data can be subjected to analysis to identify customer requirements. The collected data is generally large and unstructured and is stored in web scale data warehouses. Analysing such data to draw assumptions is challenging and involves complex data processing techniques that are compute intensive, time consuming and they consume expensive resources. Complex data analytic practices such as data mining and machine learning algorithms can be used to identify the consistencies, predictabilities and concealed details in the dimensional model (Linoff & Berry, 2011). A readable data model with structured rows and columns is known as a dimensional model which on analysis can generate insights on customer behaviour to with respect to a particular type of product.

2.1. Characteristics of Data

The data available from the internet contain different variables and datatypes. A data type can be defined as a specific kind of data that is unique by its value and the operations performed over it. In relation to the computational data, it is classified into primitive and non-primitive data types. Numeric and non-numeric values such as Boolean values are classified as primitive whereas the character values such as strings and arrays are classified non–primitive. To analyse data through computational techniques, it is essential to classify the values according to the data types. However, statistically the data variables can be classified as either qualitative or quantitative. Qualitative data deals with the nominal and ordinal data which is categorical in nature, whereas the Quantitative data deals with the interval and ratios in the data which is numerical in nature (Whitley & Ball, 2001). The data needs to be classified in accordance to the objectives of the analysis.

To simplify, the process of characterising the data, concerning the business entities can be understood by a closed loop marketing procedure in which the operational data acts as the data source for analysis that can help in marketing program creation and testing. This Cycle is known as *The Virtuous Cycle* of business opportunities (Zahra, 2008). The results generated rejuvenated the later cycles of analysis. The steps in the Virtuous Cycle can be identified by the figure below

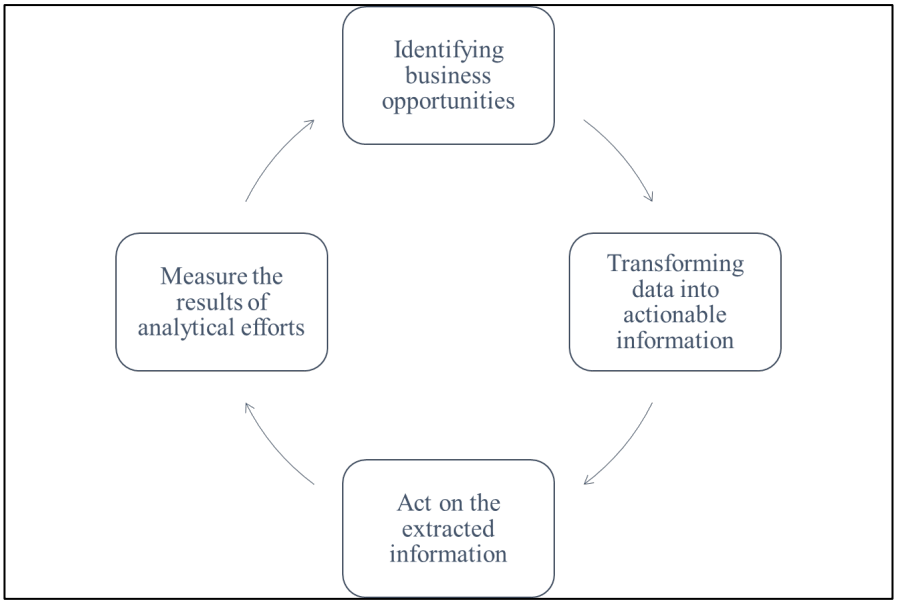


Figure 1 The Virtuous Cycle of Business Opportunities (Zahra, 2008; Linoff & Berry, 2011)

2.1.1. Identifying the Data for Business Opportunities

Data is being generated by the consumers around the world. This data is being collected and used for various purposes. The initial stage in data processing is to identify the business opportunities that the data can provide. This will provide the insights to form objectives for analysing the data. Analytical and statistical approaches to identify business opportunities is a time-consuming process and hence it is essential to work on the possible results, to identify the precise objective (Wang, et al., 2013). Business processes such as planning for new product, Direct marketing planning, understanding customer segments, evaluating the market, allocating market budgets etc. are a few examples where data analysis can contribute to the existing business efforts. This improves the efficiency of business executives by allowing them to take more educated decisions (Saura, 2020). To avoid the analytical and statistical efforts, it is important to assess the impact of actions taken so as to review the value of data mining efforts.

Figure 1 represents the availability of data in different forms from various sources. It can be seen that the data come from multiple systems in different formats. (Adi, 2009) and (Linoff & Berry, 2011)suggested that it is essential to identify the precise data sources and aggregate them together as they contribute to the success factors of a business. Preliminary, extraction and mining of data deals with issues such as mismatched keys and identities across databases, inconsistent systems, overwritten records etc. It is expensive and time consuming for companies to Solve these issues and process the data.

2.1.2. Data Transformation for appropriate Information

It was observed that data in general is available in abundance from different sources. It can be seen from Figure 2 that data required for marketing can be mined from sources such as Data marts, operational systems, marketing summaries etc. In some cases, the data is stored for auditing and other long-term purposes into large web scale data warehouses. This type of data is historical data to which more content will be continuously added, resulting in a constant change of format and content overtime. Data concerning customer relations, sales and business

management indicated information about customer segments, customer lifestyles and other external information. Other transaction data between business entities can indicate missing and incomplete fields if the data is sorted. Moreover, the main objective of analysing the data is to convert it into actionable results. (Linoff & Berry, 2011) suggests that the effective evaluation lies in interpreting the data to make business sense but not in the complexity of algorithm used. Following are the list of complexities that can be observed while processing the data;

- Bad or incomplete data formats such as the incomplete addresses of customers
- Unclear data fields such as mismatched dates of delivery in different systems
- Lack of functionality, such as the inefficiency of middlemen in the intra-firm transactions.
- Legal complications like giving legal reasons for the usage of data (data protection)
- Time complexities such as expending more time on processing data than marketing itself.
- Organisational factors like complexities associated with changes in regular practices

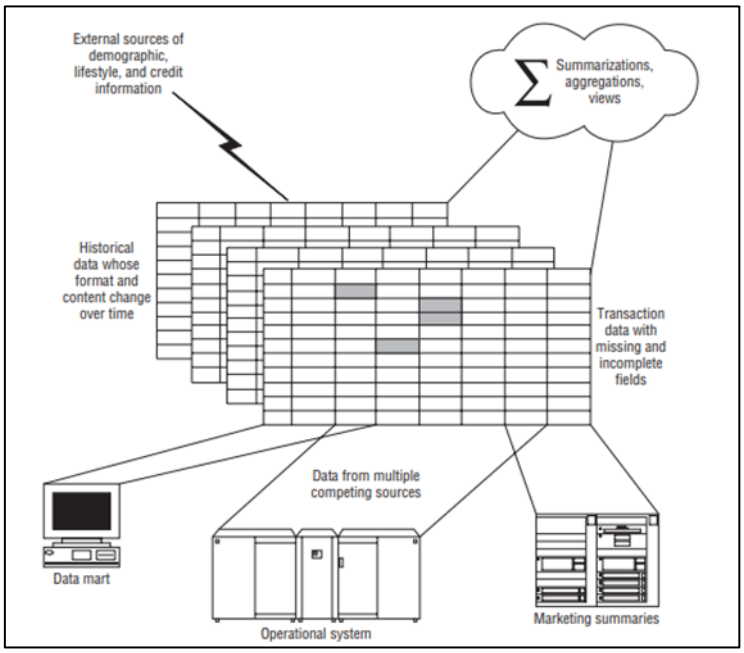


Figure 2 Various sources and forms of data (Linoff, 2011)

2.1.3. Act on the Information

As mentioned earlier, acting on the retrieved information through data transformation is the main objective of data mining as it benefits the decision-making process for companies. Better informed decisions will have better outcomes (Bijakšić, et al., 2018). In some cases, the business decision can be an already perceived one but the data supporting it acts as a factor for the confidence for taking the decision. Characteristically, the decisions or the actions towards the business data are inclined with the current business trends. Some of the business actions are observed as follows;

- Incorporating results into automated systems for online customers
- Reaching out to customers via telemarketing, emails and other online channels
- Sorting customer service priorities

- Regulating the inventory levels

2.1.4. Measuring the Results

The significance of measuring the results has been emphasised by (Markić, et al., 2016; Linoff & Berry, 2011). Though companies acquire these results from various sources and data analyses, less consideration is being given than it justifies as it has no direct return on investment. Moreover, individuals tend to compare and learn from the information they gather through word-of-mouth surveys adding to the efficiency of plans. This approach not only works for the individuals but also for the organisations that are trying to reflect on their marketing plans. Commonly, these efforts are measured according to the financial procedures and these are vital for companies. However, they should be measured along with the modelling efforts in both financial and marketing.

Additionally, the outcomes of data analysis can be used to create new business processes that can communicate with the customers and affect the relationship between companies and customers. (Armstrong & Kotler, 2016) emphasised that it is essential to assess the influence of companies on its consumers which will define customer value. Nevertheless, the data needs to be identified beforehand to process and measure the results. It is relatively difficult to classify the unstructured data to a structured data and hence the data warehouses are populated with unstructured data, pooled from various resources and interpreted as textual (word based) data (Liu, 2010). In relation to the marketing process, it was perceived through research that the information extracted by analysing the data about products and consumer behaviours can draw assumptions about their opinion towards the products available in market. To emphasise, the comments and reviews that consumers leave about their purchases on e-commerce and social media platforms can be interpreted to find customer opinions and their motivation for purchasing (Alessia, et al., 2015; Greaves, et al., 2013).

2.2. The importance of Textual Data

The initial indications of processing text data were emphasised by Alan Turing through the work of his Turing Machine in early twentieth century. The objective of his works was targeted towards answering the question, “Can Machines Think?” The initial iterations of his work were to control machinery for analysing what was then considered to be an immeasurable amount of data to simple translations from one language to another in the aims to win *Jeopardy* (Turing, 1950). Though the philosophical aspects of his work were a topic of discussion when the work was published, the technology of today’s generation is able to achieve more than what Alan Turing considered. Language can be considered as one of the vital components to establish communication between entities and it is important to identify its role in the business world (Harnad, 2009).

Pragmatically, large amount of business data is being stored across world wide web in the form of text. Various sources that the text can be extracted from are as follows;

- Consumer comments on websites and social media platforms can provide information about competitors.
- Notes made by business entities such as customer service executives, sales executives can convey information concerning factual interactions between customers.

- Articles and reports that may provide fact frameworks
- News articles and stories that explain the status of market

The opinions of consumers are recorded as comments in the form of text. Text data commonly has certain file formats for user accessibility. But data warehouses not only contain text but also content of other media formats. Hence, the file format for it will not be determined in data warehouses as it adds complexity in saving to the warehouse management system (Losiewicz, et al., 2000). The process of extracting relevant text data from large datasets is known as 'Text Mining'. Text Mining techniques implement various approaches including probabilistic approaches such as Topic Modelling and non-probabilistic approaches such as dimensionality reduction. Topic Modelling refers to a probabilistic framework that determines word clusters and document clusters in a given dataset. Whereas the dimensionality reduction is generally used along with other language processing techniques for reduced dimensionality in analysing datasets. Both the approaches have different outputs and functionalities. Some text mining techniques may utilise both approaches for fine grained analysis.

2.2.1. Natural Language Processing (NLP):

Natural Language Processing or NLP is defined as an interdisciplinary field that bridges Computer Science and Linguistics. It uses data mining and machine learning algorithms to achieve the end goal. It enables computers to analyse and draw assumptions from human languages[2]. Analysis of text data is handled by natural language processing. Natural language processing is an area of application which can be used to manipulate and understand raw natural language text for various analytical purposes. The prominence of NLP techniques developed overtime due to the increase in scope of world wide web and digital libraries (Eg. Data warehouses) (Chowdhury, 2003). Moreover, it attempts to understand the text by building assumptions of what the given data sample is actually saying. Sentiment analysis is among the key components of Natural Language Processing which is capable of analysing computational linguistics[3] and texts. Organisations and suppliers in retail sector can target such data to analyse and extract consumer needs. Hence, digital marketing can be greatly varied by including such software tools.

Studies carried out by (Cambria & White, 2014) states that the unlike NLP human language processing captures every word and activates a torrent of relative responses and sensory experiences. These relative concepts and episodes of sensory experiences facilitate the effortless execution of NLP tasks such as disambiguation of word sense, textual entailment etc. It was also observed that most of the current language processing methods based on syntactic analysis of text. This process as mentioned by (Cambria & White, 2014), '*Jumping Curves*' supports organisational decisions for business management and marketing prediction. In relation to this approach, (Atalay, et al., 2019; Nasukawa & Yi, 2003) suggested that syntactic analysis using natural language processing is highly efficient in capturing favourability for

[2] See KC Karnes, Introduction to Natural Language Processing (NLP) | CleverTap, clevertap.com (Online Resource), accessed on: 19-04-21

[3] See *Computational Linguistics*, Stanford Encyclopaedia of Philosophy (2014); (Online Resource), accessed on: 15-12-20; Computational Linguistics (Stanford Encyclopaedia of Philosophy)

persuasive marketing. These approaches incline with current research objectives and the techniques based on sentiment analysis can be used to identify the purchase patterns in consumer data.

2.3. Consumer Purchase Behaviour vs Consumer Opinion

Consumers Behaviour refers to the characteristics of consumer interaction with the product (Bordoloi & Bhardwaj, 2020). Whereas the Consumer Opinions are essentially made after having experienced the product or before purchasing the product. Opinions of a potential consumer are influenced vastly by marketing strategies and value of the product in a certain market. Data relating to consumer behaviour on e-commerce platforms include aspects such as time spent on a web page, no. of clicks, time and dates of purchase, etc. Relatively the customer behaviours differ from the purchase behaviours as they deal with characteristics like size and recurrence of purchase. Moreover, these purchase patterns can further be classified as Pre and Post purchase behaviour, Complex buying behaviour, Habitual buying behaviour (Bordoloi & Bhardwaj, 2020; Wind & Thomas, 1980). The pre and post purchase buying behaviours define the characteristics of consumers before and after a purchase has been made. Whereas the complex behaviour and habitual behaviour deal with the complexities in perceptions of brand choices. Moreover, customer intents were observed to be relative to their decision-making process. For example, working class millennials tend to purchase groceries and household supplies during the day on weekends and in evenings on weekdays. Similarly, online shoppers exhibit different behaviour such as spending time on a type of product.

The criterion for customer decision meaning in a purchase is subjective to many factors. However, the stimulus for consumers to make a purchase is triggered before they actually purchase a product. The choices towards the range of products they have is influenced by the marketing techniques employed by the companies. Consumers tend to purchase products of brands that they are aware of (well marketed) (Armstrong & Kotler, 2016). Moreover, consumerism can also be a key contributor for decision making in today's purchase behaviours of consumers. Consumers are observed to have loyalty towards brands, post purchase and they tend to develop resistance towards products from newer brands. Accordingly, the decision making process varies between nominal, limited and extended decision making in consumerism (Kotler, 1971). Five stages in decision making can be seen as;

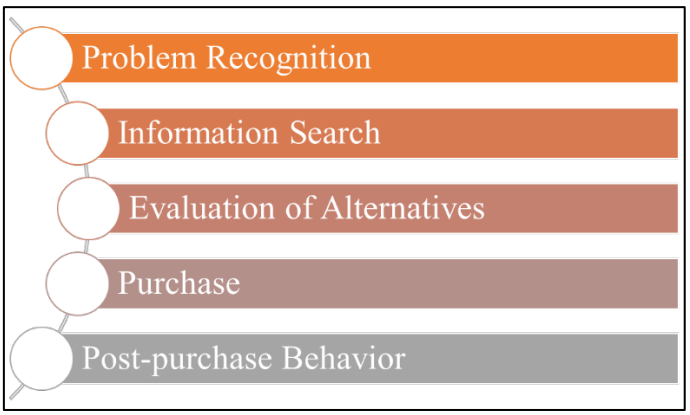


Figure 3 Stages of decision making

These aspects though they do not directly contribute to consumer opinions but they can be leveraged for digital marketing purposes and to identify consumer requirements. The marketing practices of companies and consumer behaviours complement each other. (Linoff & Berry, 2011) emphasised on identifying the data according to the business opportunities it can reflect on. In a business paradigm with collaborative and innovative consumers, customer opinions are vital, considering the degree of personalisation being expected in all the market products. However, the need for companies to acknowledge customer opinions depends on various factors such as the industry, involvement of product, degree of market saturation, economic practicality etc. Customer opinion give information about the emotions they have towards a particular product. For example, if we consider a comment on a snack bar like, "***The taste of these bars is absolutely phenomenal, but they are on expensive side.***" It can be interpreted that the customer is happy about the taste of the bar but he may not be satisfied by the pricing. Insights like these can be used take more informed business decisions.

In relation to current research objectives, free text data like comments from online platforms is considered as the data source. This may lead to the questions that "*what if the company had no data concerning purchases? – like comments, promotions and ideas*." The data can be extracted from various sources. Commercial intents can be identified from microblogs web content and features of competing products; visual comparison of data (Greaves, et al., 2013; Kumari & Singh, 2016). In cases of limited data, companies can analyse the results in conjunction with the market insights to act on the data. This is justified in The importance of Textual Data.

2.4. The Role of Digital Marketing

It was seen that the digital marketing heavily relies on the 5Ds; Digital devices, Digital Platforms, Digital media, Digital Data and Digital Technology (Chaffey & Ellis-Chadwick, 2019). A gradual evolution was observed in the way companies communicate with consumers about products through digital marketing. Modern consumers are observed to be developing various requirements with changes in the quality of lifestyle and their requirements are observed to be constantly changing due to the redundancy in products and by the influential digital marketing strategies. (Ushnish & Michael, 2021) emphasised on the constantly changing customer requirements and proposed a novel solution using Block chain techniques to identify the requirements.

(Alghizzawi, 2019) identified the role of digital marketing in the behaviour of customers. The study observed consumer behaviour by following digital marketing tends in three categories; social media platforms, mobile apps and electronic word of mouth (eWOM). Behavioural patterns towards certain type of product can be understood by following this approach. However, this fails to concentrate on the opinions that consumers have towards the products, evidently. It was observed that mobile app marketing through promotions and advertisements in mobile applications is quite effective in communicating with consumers. Conventional marketing strategies along with digital marketing strategies can help broaden the degree of communication between firms and also between firms and its audience. (Hoffman & Novak., 1997) explained this as reconstruction of market function. The study suggested the extension of this paradigm in One-to-many commination model as the web presents wider

communication prospects. Figure 4 represents the traditional One to Many marketing model with customers accessing firm related content from a digital/non-digital medium.

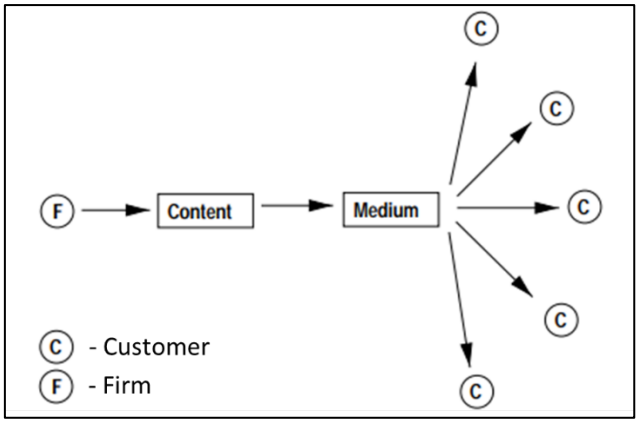


Figure 4One to Many marketing model for mass communication (Hoffman & Novak., 1997)

Consumers are affected by the changes in technology over time and the modern consumers are termed as Cyber Consumers (Wind & Mahajan, 2002). These consumers have defined themselves a set of rules such as quality thresholds, aftersales service etc., to get their required qualities in the products and services they buy and the information they seek. However, it is subjective to the type of product and service. These consumers redefine the manner of digital marketing by challenging the suppliers with the aforementioned distinctive requirements. This phenomenon however was observed as a consequence of early adaptation of digital marketing strategies.

Consumer requirements in such constantly changing segments fluctuate depending on how much they are being influenced by the Internet. An example for such change is that the gluten free products were not in such high demand before when compared to today but with the changes in customer lifestyle, their requirements are also changing accordingly. This could be the result of influence by social media marketing or word of mouth marketing or just a lifestyle change. Social Media channels such as Facebook, Instagram etc. are creating a bridge between firms and consumers.

Large numbers of social media platform users are being able to follow different product related and brand related posts, commercials and promotions and companies from small scale start-ups to large scale enterprises are able to communicate with the consumers through the internet. Moreover, companies are following new innovative digital marketing techniques to accommodate these requirements. The rise of *Influencer Led marketing* is a new way of digital marketing observed in recent trends (Campbell & Farrell, 2020). On platforms such as YouTube, Instagram, SnapChat etc. are few of the platforms on which the consumers are being influenced by different online celebrities. Moreover, the ‘cyber consumers’ are willing to try and experience new products and services that are marketed on the internet and this can be considered as a result of new digital marketing techniques.

2.4.1. Conventional Marketing Vs Digital Marketing

Marketing can be defined as a process of engaging consumers and managing profitable relationships with customers. The process by itself define goals to attract new customers and

to retain and grow current customers by ensuring customer value and delivering customer satisfaction (Armstrong & Kotler, 2016). Traditional marketing can be seen in various market segments. Its execution can be seen from the products in super markets, advertisements on TVs, promotions on magazines etc. Advertising and selling can only be seen as a part of marketing process which contributes to a much broader subject, *marketing mix.* A set of marketing tools used to engage, satisfy needs, and build relationships with customers constitute marketing mix.

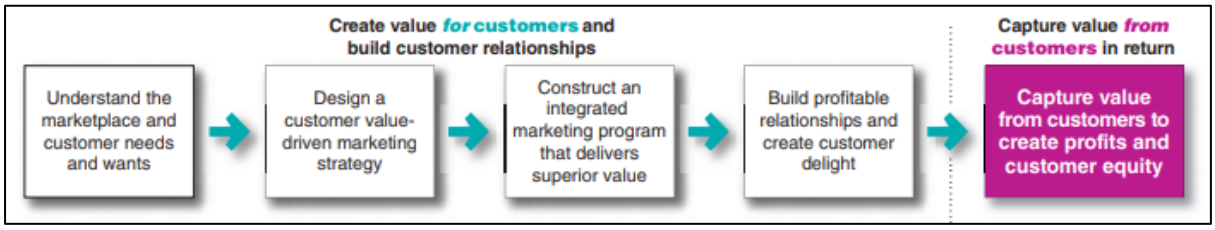


Figure 5 The marketing process (Armstrong & Kotler, 2016)

(Armstrong & Kotler, 2016)Figure 3 represents a simple five step model to create and capture customer value. The first four steps of the process explain how companies understand customers and create customer value to build stronger relationships. The final step yields returns to the companies through customer value. By creating customer value, companies can capture value from customers in the form of customer equity. Companies capture customer value by employing conventional marketing techniques that vary from customer segmentation, designing the right market plans, placement of products, employing marketing mix etc. However, the companies interact with consumers through traditional methods such as radio advertising, banners, broadcasting media, sponsorships, print ads, etc. The strength of traditional marketing is that it is capable of targeting local customers (Salehi, et al., 2012). Effectiveness of traditional marketing can be explained by taking Boston Market as an example. The CEO of Boston Market, George Michel is known to make regular visits to the company owned restaurants and work in the dining room to engage with the customers. It helped the company to understand their requirements and needs.

Marketing has come a long way from being just selling and advertising. Digital marketing is similar to traditional marketing but it leverages the connectivity offered by internet and world wide web. Unlike conventional marketing, digital marketing is capable of reaching out to global consumers. It mainly employs marketing techniques such as email marketing, search engine optimisation, content marketing, social media marketing, influencer led marketing etc. Marketers are adapting to the changing marketing trends and are developing new approaches. The companies interact with the consumers through various developments such as imaginative and interactive websites, mobile applications, online videos, social media content and influencers (Chaffey & Ellis-Chadwick, 2019). Digital marketing can be effective among the segment of customers who spend time more on digital media or the customers who choose online shopping over offline shopping. Hence, digital marketing can also be called online marketing.

Consider an average consumer who wants to shop online, there is a fairness of probability that the consumer thinks firstly of Amazon despite the existence of other popular vendors. To

emphasise this, Amazon being the pioneer to start the virtual book store (online shop for books) in 1995, quickly they became a one stop shop for essentially anything a consumer could imagine to purchase. (Armstrong & Kotler, 2016) explains this as the company being customer driven to its core. But in the words of Jeff Bezos, it is creating the right values for customers. Amazon boasts its strength in numbers and it reasons that each one of its 237 million customers should have a personalised store. Hence, Amazon was the first to analyse purchase history and browsing patterns to provide personalised experiences. Amazon not only sells goods on its website but also it engages consumers to create personalised customer relationships.

Area	Traditional Marketing	Digital Marketing
Target Customers	Relatively easier to target local customers	It extends reachability to global consumers from around the world
Marketing approach	Conventional approach is a personal approach as it directly communicates with consumers personally and through different means	Physical approaches are not required as it mostly employs online marketing techniques. However, physical marketing compliments the overall marketing strategy
Consumer interaction	It requires less interaction between the supplier and the consumer as the promotion mediums are capable of communicating directly to the consumers.	Various digital platforms and social media websites are in existence. Mobile applications, analytical tools, search engine optimising tools etc. are capable of passively collecting customer data.
Cost	Marketing costs are significantly high as they involve physical mediums for communication	Marketing costs are relatively less as media content can be created online. However, it is subjective to market requirements

Table 1 Differences between Conventional and Digital Marketing[4]

[4] See Traditional Marketing vs Digital Marketing | 10 Amazing Key Differences (educba.com) accessed on: 10-01-21

2.4.2. Trends in Digital Marketing

Digital marketing technique	2017	2018	2019
Content marketing	20.3	20	15.1
Big data	20.2	14	14.4
Artificial intelligence and machine learning	-	14	15.1
Marketing automation	10.3	9	14.8
Mobile marketing	9.2	9	5.9
Social media marketing	8.8	10	6.8
Conversion rate optimisation and improving website experiences	5.9	5	7
Internet of Things marketing application	5.4	3	3.4
Search engine optimisation (SEO)	4.1	4	2.7
Wearables	3.4	2	1.3
Paid search marketing	3.3	2	2.2
Online PR	2.7	3	2.8
Communities	2.6	3	3.2
Display	1.6	1	1.5
Partnership	1.5	2	1.8
Other	1.9	1	1.7

Source: compiled by the authors using: *(Smart Insights, 2017; 2018; 2019)*

Table 2Trending Digital Marketing Trends 2017-2019

Changes in the views of public, patterns in the use of media and lifestyles of consumers are initially noted by the marketing industry. Smart Insights[5] conducted an analysis of the various digital marketing trends. The analysis was distinguished by (Kotane, et al., 2019). The changes in trends can be seen in table 1. It can be observed that different marketing strategies such as Big Data, Content marketing and mobile marketing have downward trends by the year 2019. Though the analysis was made for three consecutive years, Artificial Intelligence and Marketing automation were observed to be gathering potential in the coming years (Smart Insights, 2019). Moreover, marketing based in artificial intelligence was not used before 2017 see, Table 1. By 2021, Artificial Intelligence is among the top digital marketing techniques along with Programmatic Advertising, Chatbots, Influencer Marketing and Personalisation. [6]

(Palamarova, 2018) conducted a comparative study to analyse the digital marketing trends in European countries including Ireland. The study suggested that the digital marketing trends in Ireland, along with other countries are in line with the global trends. Surveys from various consumers indicated that Social media is one among the most highly promoted platforms for effective marketing. However, companies are constantly adapting to the ever-changing consumer behaviours. Identifying the purchase behaviours of consumers has become one of the vital components in market research and the growth in digital marketing has influenced the buying behaviour patterns in consumers (Bordoloi & Bhardwaj, 2020).

[5]See, About Smart Insights | Smart Insights accessed on: 25-11-20
[6] See, 42 Digital Marketing Trends You Can't Ignore in 2021 (singlegrain.com) accessed on: 08-03-21

2.5. Analysis and Personalisation of Consumer Behaviour

Conveying messages personally is observed to be one of the nest ways of communication. Irrespective of the audience, communicating the information with respect to terms that are relevant will establish strength and clarity in the communication. Personalisation is considered to be one of the critical components in designing marketing strategies through Customer Relationship Management. Customer relationship management can be considered as a preeminent medium to engage customers. CRM allows companies to directly communicate with consumers. It allows the organisations to learn from its consumers and produce products and services not only according to their requirements but also tailors them to give unique and personalised experience (Kingsnorth, 2019). Personalisation can be achieved at different levels; they can be seen as;

1. True Personalisation
2. User Defined
3. Behavioural

True personalisation can be used to define the concept itself. It directly sends relevant messages to consumers by tailoring them according to the individual. This is achieved by techniques such as using the individual's names or personally crafted messages like "*we know you like French vanilla with cinnamon rolls. We hope you'll like this combo offer !!!*" However, treating individuals precisely the same as the other individuals because of some common factors is not efficient for everyone. Reducing the concept further suggests two methods, one is to let the consumers give information about their requirements which gives user defined personalisation and the other method is to learn from the consumer behaviour about their requirements which defines the behavioural personalisation.

Behaviour personalisation can be defined as a variable that is capable of representing a consumer or an individual's behaviour from the indicators analysed from different data sources. Concerning the current research, the indicators are the behavioural traits of the consumer and the text data processed from "*the web*[7]"can be considered as the data source. The information thus collected can be used along with data mining applications such as Natural Language Processing (NLP) to make marketing decisions. However, behavioural personalisation is challenging, considering the data that is being generated from various sources. Interpreting the available data can be the initial phase to make the precise assumptions from the data. Hence, data processing, machine learning, artificial intelligence, predictive analysis and other data analysis solutions as such are being applied for analytics (Saura, 2020). However, it is the responsibility of companies to make relevant assumptions.

2.6. Sentiment Analysis and Opinion Mining

As mentioned in section 1.2, consumer behaviour towards a product and their opinion towards a product are two different entities but they complement each other. Consumers in retail markets generally develop a purchase pattern towards products. For example, working class millennials tend to purchase groceries and household supplies during the day on weekends and

[7] Referring to the resources gathered from different sources like comments, articles, blogs etc.

in evenings on weekdays. Similarly, online shoppers exhibit different behaviour such as spending time on a type of product.

Sentimental analysis is among the highly mentioned topics in Machine Learning, Artificial Intelligence and marketing management research communities. As research suggests, the success of a product greatly relies on the opinions of customers. Moreover, identifying the customer requirements can promote customer loyalty and brand success. (Markić, et al., 2016; Bijakšić, et al., 2018) suggested that conventional techniques can be used collect data and recognise customer opinion and attitudes concerning a product. The study defined the role of sentimental analysis in marketing. As marketing strategies need to be constantly reviewed by analysing customers, sentiment analysis can provide a solution for continuous study of customer attitudes. Other studies defined sentiment analysis as a practice to find emotions, feelings and attitudes from text data (Liu, 2010). Therefore, it was categorised as a progression of Natural Language Processing[8].

Applications for sentiment analysis were not only observed in marketing but also in Medicine (Greaves, et al., 2013) used sentimental analysis to find the opinions in patients by targeting the text posts and comments as data sets. Considering the emotions, (Alessia, et al., 2015) analysed the emoticons in text posts to discover the polarity of feelings. Since the analysis is to identify the emotions, not all the opinions will be positive but the negative opinions also play a crucial role in product and service development and it helps in translating informative conversations to relative (Pathak & Pathak-Shelat, 2017). Considering packaged food and snacks, for example, snack bars (Breakfast Bars, Nut Bars, Cereal Bars etc.) external factors such as packaging, the taste and quality of the product have been considered to identify the consumer choices (Pinto, et al., 2017). Considering protein bars, there are various local and international brands competing in the retail sector in Republic of Ireland. These products are sold on various online platforms targeting the "*Cyber Consumers*". There exists a need for new and upcoming local manufacturers such as Origin Bars[9] and FulFil[10] bars to establish the polarity of emotions concerning products of other brands to effectively communicate with the customers and accommodate their requirements.

2.7. Industry Review and Situational Analysis

To support the current research, a brief analysis of current market trends regarding the designated product - protein bars and retail market has been carried out. This section of the study is to understand the current market situation and get insights into the marketing techniques of existing suppliers. As a product protein bars are placed as nutritional snacks and are being sold as a retail product. As justified in section 1.3, this product is selected for the

---

[8] See, *Natural Language Processing,* IBM Cloud Education; Online Source, accessed on 09-12-20 What is Natural Language Processing? | IBM

[9] See Origin Bars - *Kerry company aims to take a bigger bite of snack market with protein bars* Irish Examiner. Retrieved on 7 July 2020; https://www.irishexaminer.com/breakingnews/business/kerry-company-aims-to-take-a-bigger-bite-of-snack-market-with-protein-bars-833180.html

[10] See *Nutrition Clear thought - Interview: Fulfil Nutrition*, Clearwater International. Retrieved on 7 July 2020; https://www.clearwaterinternational.com/publications/nutrition-clearthought/interview-fulfil-nutrition

study as the market study concerning this product is limited and potential for the growth of local manufacturers adds value to the techniques emphasised in this study.

a. Market Trends:

Protein bars are being consumed as a healthy alternative to normal snack bars. Research indicated a potential for global snack bar market growth by 2026. And these products are being sold in fitness supplement store, departmental stores, grocery stores and super markets as a retail product. A growth rate of 3.7% in sales of retail products have been observed till 2nd quarter of 2020 and it is expected that this trend may grow with a growth rate of 2.5% in the coming years. With respect to protein, it is not only being sold in the form of bars but also in other forms such as Whey Proteins, Protein Shakes, Cakes, etc., targeting a wider section of the market. Mintel's Global New Products Database, Board Bia indicated an increase of 498% in the manufacturing high protein content products both in UK and Ireland[11] (Binchy, 2019).

b. Market Profile:

The retail sector has been perhaps the greatest supporter of the Irish exchequer, making 23% of expense and tax receipts altogether and it is the major market sector in Ireland. Looking at the food and drink industry, it stands apart as a distinctive sector for the Irish market producing a turnover of €27.5 billion. The key insights of Ireland's food and drink industry can be seen as follows;

- Supplier of produce for food and grocery sectors and supports other market sectors
- Covers a larger export network extending to 180 countries globally with 53% exports from internal producers.
- around 71% of resources are procured in Ireland and 50% of the related services are bought from native work force.
- The sector's broad footprint extending to global markets is contributing to the country's economy[12]

The fitness industry and the food & drink industry are interlaced as the governments around the world are attempting to take measures to target the current crisis[13]. Origin bars, Fulfil bars, and Quest bars are the major players in the local retail market and they are domestic companies with greater market share in Ireland. Other international manufacturers such as Grenade Bars, Glanbia's Optimum Nutrition and PhD Performance Nutrition are competing along with the local manufacturers. Catering consumer requirements of different customer segments with a wide product portfolio is observed to be the mode of operation for these brands.

---

[11] See Global Food and Drink Trends, Mintel's Global New Products Database, Mintel.com, 2021 Accessed on: 20-04-2021

[12] See Food and Drink Ireland - Report generated by Food and Beverages (FDI) sector exposing the market profile. Accessed on 28 June 2020; https://www.fooddrinkireland.ie/Sectors/FDI/FDI.nsf/vPages/Food_Industry~sector-profile!OpenDocument

[13] See Colin Gleeson, Bord Bia unveils measures to tackle Covid-19 pandemic (irishtimes.com), Accessed on: 20-04-2021

2.7.1. Customer Analysis

Kotler indicated that the trends in markets and the marketing strategies vary with the variations in demographics. Since, protein bars are consumed by different segments of consumers it is essential to identify and define the demographics (Kotler & Armstrong, 2016). A brief customer analysis has been included for the study in the aims to recognise the customer base. Given below are the insights on customer impression on the Irish retail industry.

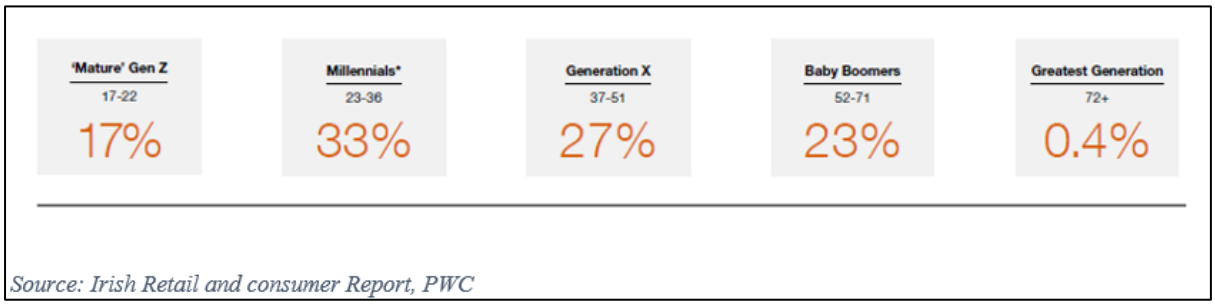


Figure 6 Breakdown of millennials based on age

The above given figure represents the percentage of customers from different age groups contributing to the market through online and offline sources. These millennials are classified into three major age groups

- Young millennials (18-26)
- Core millennials (27-31)
- Mature millennials (32-36)

It was observed that a mix of young and core millennials between ages 18 and 34 make up the higher market share as this age group contain individuals of various lifestyles such as students, working class individuals, athletes, young professionals, etc. Moreover, the purchase patterns of these consumers highly depend on the reason behind the purchase. The percentage of store purchases were recorded as 35, 42 and 39 from young millennials, core and mature millennials respectively. The remaining part of consumers from each age group accounts for the online purchases from various ecommerce platforms. This indicates that a larger percentage of consumers from each age group tend to buy online. However, this consumer base can be broken down based on the purchase reasons and perceptions they have towards products.

It can be observed from Figure 7 that a higher concentration consumers of around 26% tend to have a perception that protein bars are a healthy snack and their purchase reflects the records. Based on the reasons for purchase, the next higher percentage of consumers perceive protein bars as an on the go snack and the next two lesser cohorts of people tend to purchase with a perception of pre/post workout protein for gym going and just as a protein supplement for everyday requirements. Reasons such as meal replacement, taste, for sports activity, energy boost etc. are some of the less significant reasons. Moreover, consumers totalling 37% of all consumer base, consider it as a product for healthy lifestyle. Stable consumers are observed to purchase product in the aim to reach fitness goals. From the observations it can be assumed that the current customer base is populated majorly with consumers of athletic lifestyles.

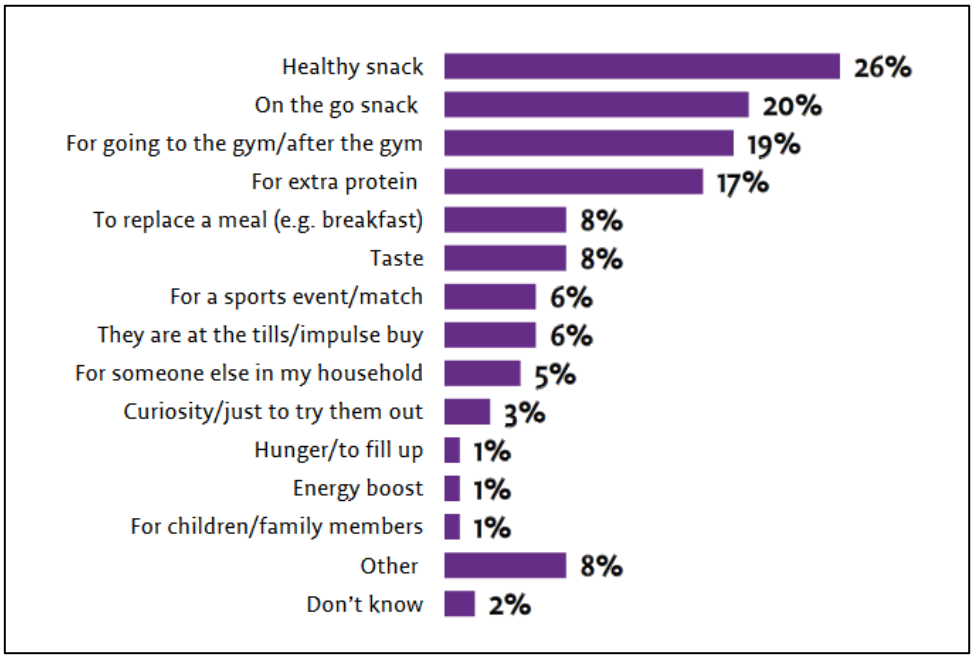


Figure 7Purchase rate of consumers based on reason for consumption (Safefood, 2019)

Current customer Index:

- The population of around 11.92% is between the ages of 15 to 24
- The age group of 25 to 54 covers 42.86% of the population
- 43% of population in Ireland engages in at least one sport, as indicated by Irish Times

Customer Segments who use online platforms frequently tend to purchase multi-packs. These consumers develop opinions based on post purchase experience and factors like the taste of the product, quality, ingredients, nutritional value etc. These opinions are expressed on social media platforms, seller's websites, e-commerce websites etc. Consumers are thus generating abundance of data by leaving various comments, opinions and ideas regarding the products. Organisations can use this data to identify the consumer behaviour on e-commerce platforms and also consumer opinion.

In relation to millennials (23-36) and generation X (37-51) consumers who cover 33% and 27% of the retail markets respectively (Irish Retail and Consumer Report, PWC, 2019), they incline to have larger online foot prints and this customer segment is relatively more technically inclined. Consumer reports suggest that they spend more time online and it presents an opportunity to promote products on Social media and online platforms. Conventional marketing techniques will not be able to capture this market segment considerably unlike other market segments. But Digital marketing techniques can be used to capture other customer segments along with the above-mentioned demographics (Chaffey & Ellis-Chadwick, 2019; Kingsnorth, 2019). Hence there is a need for companies to consider strategic digital marketing. The term Digital Marketing can be considered as an umbrella term for various online marketing techniques. (Kannan & Li, 2017) defines digital marketing as a technology-enabled process

through which companies communicate with consumers to cooperatively communicate, build and deliver value for the stakeholder.

2.8. Gap Analysis

The literature review provided a thorough understanding of different aspects that data mining and sentimental analysis can target. Previous research was explored in accordance with questions that can rise with respect to the influence of data mining techniques over marketing and the benefits of data mining in marketing the product 'protein bars.' Moreover, aspects of marketing, its impact in consumer opinions, product personalisation and achieving them through sentiment analysis were summarised. However, some of the methodologies such as targeting wider segments of consumers through marketing and true personalisation were not able to align with some of the research objectives. Limitation in traditional marketing techniques were identified. An industry review of the Irish retail sector provided insights on the strength of consumers using online platforms for product purchases, considering the pandemic. The study indicated the need to leverage the information gap between these consumers and the companies. Bridging the gap through data mining can help the companies to make better educated decisions towards customers

2.8.1. Research Problem and Questions

Marketing and digital marketing greatly rely on the customer data. The present day technology is generating abundance of customer data from various sources online. Conventional marketing strategies are not capable of processing such data to extract customer insights. The research problem is defined as ***"Traditional marketing techniques lack potential in identifying the customer requirements, opinions and attitudes, there exists a need to categorize what the consumers think about the products to effectively communicate and identify consumer needs to define better digital marketing strategies."*** The following research questions are formulated based on the literature review;

How to use customer information on online platforms to identify opinions towards protein bars in the Irish retail sector?

Other identified business problems are given below.

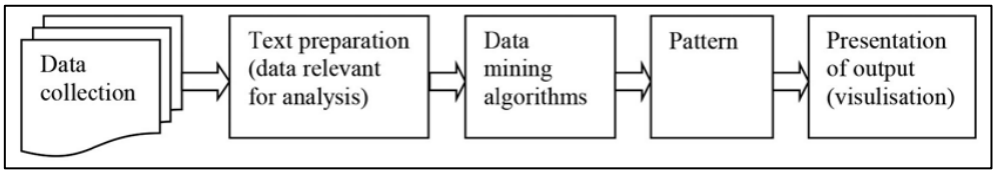
- Which products in the inventory has higher and lower reviews to differentiate into superior and inferior products respectively?
- What are consumers' post purchase opinions and their purchase patterns?
- Is it possible to compare positive and negative opinions of products?
- How digital marketing can be influenced through sentiment analysis?

3. Methodology

As suggested by previous research, there are various ways to analyse data and extract the opinions of customers. Traditional techniques such as generating information through questionnaires on product quality, service quality and conversations have been discussed in literature which are not as effective in identifying the opinions of online consumers. However, these techniques can be used as a medium for data collection. The methodology for the analysis follows a structural pattern to analyse the text data. The research is concentrated on the analysis

of user opinions on e-commerce websites to gain insights into emotional polarity towards different protein bars in Irish Market.

Emotional Polarity can be defined as a technique to realise the emotions of abstract polarities (Adi, 2009). This area mainly deals with the Emotion Theory based in Natural Language Processing principles. Abstract polarities in this case are the emotions that the consumers express towards products. In relation to the data requirements, the data is collected considering local and international protein bars available in Irish Retail Market. As mentioned above, imported products like Optimum Nutrition bars, NitroTech bars, Barbells Protein bars can be considered along with FulFil bars, Origin bars which are based in Ireland.



Source: (Bijakšić, et al., 2018)

Figure 8Flow of sentiment analysis algorithm

The above presented process flow represents the step by step execution of a typical sentiment analysis algorithm. The Data sets defined are mined for relevant variables that are required. The unstructured data which is generally multi-dimensional in context will be prepared. The data preparation constitutes the collection of required data from the large datasets. Thus formed data is subjected to mining algorithms depending on the datatype and variables that is needed. The algorithms find patterns in the data and present a visual description of the analysis. For example, people who find the product as a 'healthier snack' would post positive comments on how it can be a healthier option for a snack. The positive comments and the negative comments define emotional polarity towards the product. The analysis is visualised with the percentage of emotional polarity against the defined variables such as the taste, nutritional value, ingredients, etc.

3.1. Data Collection

Following the hypothesis, the dataset requires Protein Bars available in Ireland as a primary variable. Considering the retail sector in Ireland, there are various protein bar manufacturers and they are being appreciated by various demographics of customers (Board Bia). As the market review suggests large number of buyers are making purchases from online retail vendors via e-commerce platforms. Amazon being one of the largest sellers in Ireland, they do not have a dedicated Irish website. However, they ship products to Ireland through various distribution channels.

Product reviews are gaining significance with the evolution of retail stores to online stores (e-commerce). Customers are posting their post purchase experiences in the form of reviews directly the websites hosting these products. In relation to text data mentioned by (Linoff & Berry, 2011), all these review and comments can be considered as rich data sources for analytical purposes. It grants an opportunity for companies to gain insights from the market reactions towards a specific product.

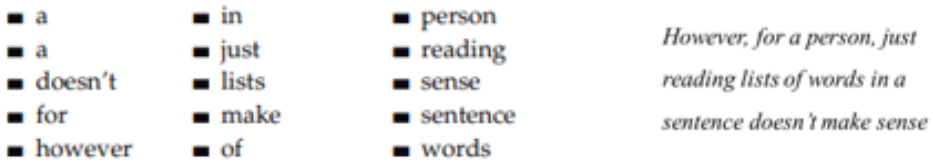
Initial iterations of data collection were believed to have been implemented to collect data in two different datasets, primary and secondary data sets. However, this was not feasible as the primary dataset included data to be collected from surveys. Though this approach is efficient for smaller datasets, it lacks the information required in a wider perspective. The complexity to draw impressions from smaller datasets is relatively higher and time consuming. And the data collected through surveys is not effective for text mining. Hence a single dataset is defined with comments and reviews collected from different websites. This dataset is capable of being used to analyse larger aspects of the product with respect to the research objectives. A website called "discountsuplements.com" was selected to extract the customer reviews. This website has a significant market share in Ireland and it sells not only the products that are made in Ireland but also supplements that are imported from different countries. This site is an example of supplement sellers along with sellers like SO nutrition, HP nutrition, etc.

Amazon has a relatively larger footprint, not only in Ireland but also around the world and it provided an opportunity to target protein bars that are being exported to other global markets. The data required for this research is taken from the all the reviews of consumers recorded over a data span of 18 years primarily on Amazon. The dataset is created based on the products from different manufacturers but are sold on Amazon. Moreover, the dataset contains all the raw data concerning all the products that are sold on Amazon. Since this data is unstructured, further preparation is required to extract relevant text data for training the data model. Training the algorithm gives capacity to draw assumptions on the whole data from a provided sample

3.2. Text Preparation

The process of data mining is to find and exploit relevant patterns in any form of data. Text is considered as one of the key forms of data as it can be processed with less complexity unlike other forms of data. Moreover, it is expensive and time consuming to deal with other forms of data. Research suggests that unstructured data can be efficiently converted into structured data with the help of text mining. It can be used to add derived columns to a model set as a process of converting the data into structured features. Extracting the derived variables to form columns is a process of identifying specific patterns in text. For example, a house address can be used to identify if someone is living in it by looking at the house number. To mine such information from raw data, characters such as H.No., Apt.No., Apt., # etc. can probably refer to a house or an apartment number. However, analysing text data to find patterns have several flavours. Sentiment analysis is one of them which is capable of analysing the emotions of a writer.

Preparing and representing raw data in a readable format is an initial step of text mining. The datasets can be stored with the raw data but it does not necessarily help the text mining systems or computers to understand them. There are various approaches to understand the raw data. The "*understanding*" approach is considered suitable as the analysis requires larger datasets that are not possible to analyse using "*bag of words*" approach. The selected approach enables to understand the dataset and what each word in it signifies. A bag of words approach divides a gives paragraph or a set of words in a document. It's hard to make sense of a set of words. Consider following list of words for example,



The list represents the sentence adjacent to the list. The list of words in the sentence are arranged alphabetically. Although, the precise meaning of the sentence cannot be obtained from the list the collection of words can provide some information. This list of does not provide information about any particular topic but the vocabulary used to define ideas can provide information on the subject. It may deliver some information about the sentiments but for individual words and it is not efficient form larger data samples.

3.3. Sentiment Analysis as Data Mining Application

It was observed from the literature review that there are various approaches to analyse text data. Processing and visualising the data to make it readable one aspect of data processing but drawing assumptions from such data in the primary objective of data mining. Sentiment analysis is a subsidiary of text mining. The goal of this application is to automatically examine data sources and identify how individuals feel about a subject, in this case protein bars. As mentioned in section 2.8.1 this data mining approach is capable of assigning sentiment score to the given data. This score ranges between 1, 0, -1 representing negative, neutral and positive emotions. Tracking sentiments of consumer over a period of time helps organisations to understand the market situation.

Sentimental analysis is efficient for published data similar to the dataset of reviews considered for analysis. The approach of sentimental analysis involves following processes;

- Clean the text
- Finding a topic of interest
- Determining the sentiment
- Application of sentiments

Cleaning of data is required to clear the anomalies such as incomplete sentences, redundant names, similar reviews on different platforms etc. Cleaning the data helps in reducing the complexities in the interpretation of data. Once the readable data is obtained, it is essential to identify the topic of interest. Considering the selected dataset, large number of variables are included. Irrelevant variables such as names of insignificant products are included in the dataset. In relation to the research objectives, the topic of interest is the reviews on different protein bars. By analysing the reviews, the sentiment behind the review can be interpreted. For example, if the review says "*gross – I like all kinds of Power Bars, Odwalla, Balance Bars, Premier Bars. Just don't like the taste of Chocolate - Chocolate Chunk.*". The sentiment behind it is interpreted based on the attitude represented by the psycho-social dictionary[14]. The review can then be categorised to have a negative sentiment as the word "*gross*" will be marked as

[14] Psycho-social dictionary is a collection of thousands of words with various nuances of meanings, particularly related to human attitudes.

negative in the psycho-social dictionary (Linoff & Berry, 2011). These sentiments are applied to the data model to gather insights from the selected sample. As mentioned above, the reviews will be categories into negative, neutral and positive reviews.

3.3.1. Evaluation model

It can be explained by Figure 10 that the data that is prepared has comments, Ideas, etc. will be analysed to extract three polarities about the products. One is to find the consumers that are happy for a particular product, the second is sad and the third is the customers with neutral opinions.

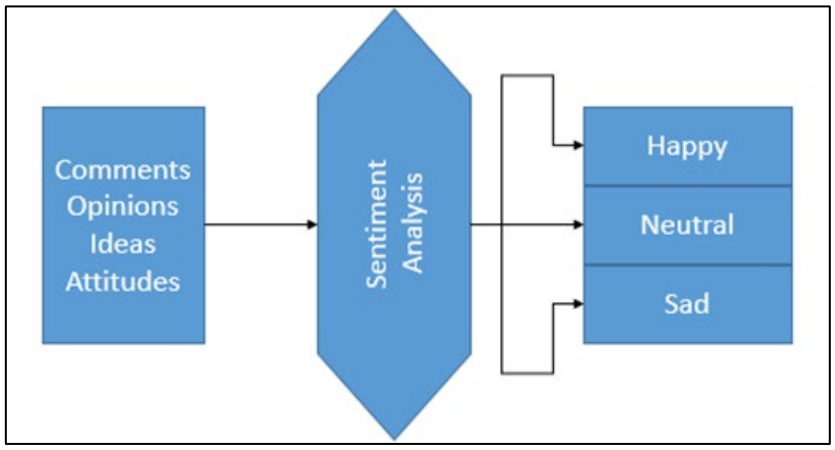


Figure 9 Execution flow of the data

The opinions thus extracted and further be analysed to identify the characteristics that triggered the emotions. This helps in designing better marketing and digital marketing plans targeted toward niche customer segments such as "Cyber Consumers" (Wind & Mahajan, 2002). The future studies on this analysis can be inclined to define effective communication strategies with audience based on the opinions extracted for protein bars.

As a data analytics application, this provides a scientific credibility to a qualitative analysis. The algorithm is capable of generating opinions in real time. Hence it adds value to the research especially in effective digital marketing as the sentiments can be leveraged to make marketing decisions. Following the steps involved in the process of sentiment analysis, the data set with reviews and comments of consumers is considered for analysis. A sentiment analysis application is considered to mine the provided data. For this research, Python based FastText[15] library for text classification is used to define the data model and a sentiment analysis application based on support vector machine is considered. With the help of the algorithm, the interpreted sentiments can be applied to the data sample to find the emotional polarities.

3.4. Applications in Digital Marketing

The aim of analysing the data is to identify customer sentiments to create personalised experiences. As emphasised by (Kingsnorth, 2019), personalisation is effective in serving the customers with relevant messages at a given point of time. True personalisation is one of the digital marketing approaches to provide tailored experiences to customers. In a way sentient

[15] See, fastText an open source library for text classification Accessed on: 25-04-2021

analysis acts like a user defined personalisation, but customers do not voluntarily give their opinions. As the customer opinion are being extracted through data mining it is more inclined as behavioural personalisation. The results of analysis can be processed through a model which can make real time business decisions. Similar to the above-mentioned sections of methodology, behavioural personalisation collects information about the behavioural traits of customers in this case, the data collected is customer reviews.

Following the marketing strategies proposed by (Kotler & Armstrong, 2016), this approach can be used to build better customer relationship management in preparing integrated marketing plans. It was suggested that CRM is vital for modern marketing. To achieve research objective, the three steps approach – understanding the market, designing value driven strategies for marketing and building a marketing program can be followed. However, the entire marketing programs are subjective to organisational goals and the type of product. Through data processing, the research aims to create digital marketing approach through behavioural personalisation in CRM. Considering the example from section 3.3, if the individual reviews the product as "*gross*" it can be used to create personal emails like "*we heard that you didn't like our chocolate chunk bar. Would you be interested in trying out other flavour in the combo created especially for you?*"

3.4.1. Assumptions:

However, it is essential for the companies to not only interpret data but also to create assumptions from the data. Following assumptions can be made in relation to the current study.

- It is assumed that a sample size of 30000 examples taken from given data is enough to represent the population of reviews
- It is assumed that the data from text model will be efficient to train the algorithm with an anticipated 70% accuracy
- It is assumed that personalised messages to consumers by analysing their sentiment towards protein bars is efficient in establishing better customer satisfaction.

4. Evaluation

The evaluation model for this research is developed based on the sentiment analysis models emphasized by (Liu, 2010; Markić, et al., 2016). Following the steps involved in the data mining process, the data is selected and processed to make text models. The algorithm is executed on a computer with AMD A8 processor with 16 GB of RAM. Software environment required for processing and visualization is Python with NumPy, Matplotlib[16] libraries. These libraries are used to visualize the data. Microsoft Excel is used to create a sample.

4.1. Data

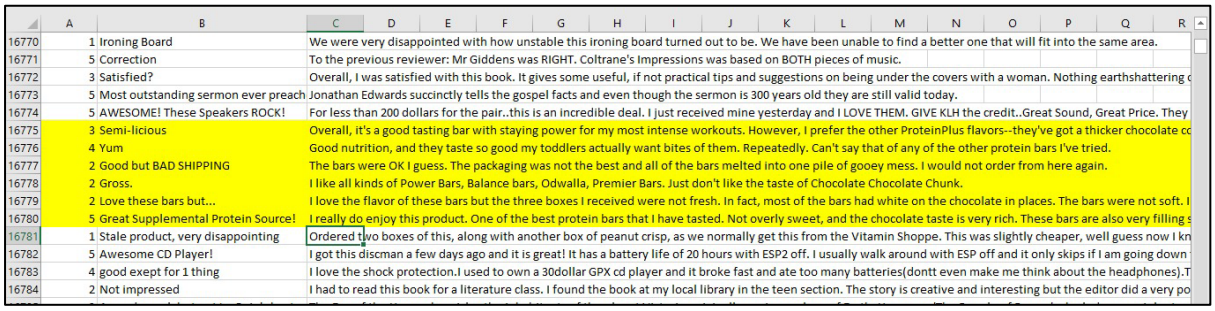
Based on the considered dataset as mentioned in section 3.2, raw data was extracted from the customer reviews of Amazon.com. The dataset is taken from previous studies of (Zhang, et al.,

[16] A Python library, Capable of inserting plots with the help of general purpose GUIs

2015). This data set was considered as it offers millions of observations. The data set was considered suitable as it has large number of observations that have been recorded over time.

The dataset is modelled using Microsoft excel based on the required product types such as the products that are essentially considered as retail products. The full score data set has produced approximately 35 million customer reviews. Most of the observations in the data are based on the customer satisfaction with the product. These observations comprised of customer reviews that include ratings, user and product information as plain text. The dataset is populated by (Zhang, et al., 2015) and used as a text classification benchmark. In total the sample has 3,000,000 training samples and 650,000 testing samples. The extensibility of this dataset makes it suitable for sentiment analysis. As mentioned above, the training sample is used to apply the sentiments to the whole dataset with respect to the outcomes of test sample.

4.1.1. Cleaning the data



	A	B	C
16770	1	Ironing Board	We were very disappointed with how unstable this ironing board turned out to be. We have been unable to find a better one that will fit into the same area.
16771	5	Correction	To the previous reviewer: Mr Giddens was RIGHT. Coltrane's Impressions was based on BOTH pieces of music.
16772	3	Satisfied?	Overall, I was satisfied with this book. It gives some useful, if not practical tips and suggestions on being under the covers with a woman. Nothing earthshattering
16773	5	Most outstanding sermon ever preach	Jonathan Edwards succinctly tells the gospel facts and even though the sermon is 300 years old they are still valid today.
16774	5	AWESOME! These Speakers ROCK!	For less than 200 dollars for the pair..this is an incredible deal. I just received mine yesterday and I LOVE THEM. GIVE KLH the credit..Great Sound, Great Price. They
16775	3	Semi-licious	Overall, it's a good tasting bar with staying power for my most intense workouts. However, I prefer the other ProteinPlus flavors--they've got a thicker chocolate co
16776	4	Yum	Good nutrition, and they taste so good my toddlers actually want bites of them. Repeatedly. Can't say that of any of the other protein bars I've tried.
16777	2	Good but BAD SHIPPING	The bars were OK I guess. The packaging was not the best and all of the bars melted into one pile of gooey mess. I would not order from here again.
16778	2	Gross.	I like all kinds of Power Bars, Balance bars, Odwalla, Premier Bars. Just don't like the taste of Chocolate Chocolate Chunk.
16779	2	Love these bars but...	I love the flavor of these bars but the three boxes I received were not fresh. In fact, most of the bars had white on the chocolate in places. The bars were not soft. I
16780	5	Great Supplemental Protein Source!	I really do enjoy this product. One of the best protein bars that I have tasted. Not overly sweet, and the chocolate taste is very rich. These bars are also very filling
16781	1	Stale product, very disappointing	Ordered two boxes of this, along with another box of peanut crisp, as we normally get this from the Vitamin Shoppe. This was slightly cheaper, well guess now I kn
16782	5	Awesome CD Player!	I got this discman a few days ago and it is great! It has a battery life of 20 hours with ESP2 off. I usually walk around with ESP off and it only skips if I am going down
16783	4	good exept for 1 thing	I love the shock protection.I used to own a 30dollar GPX cd player and it broke fast and ate too many batteries(dontt even make me think about the headphones).T
16784	2	Not impressed	I had to read this book for a literature class. I found the book at my local library in the teen section. The story is creative and interesting but the editor did a very po

Figure 10 Representation of test model

The above figure represents the sample of test model drawn from the selected dataset of raw text. The highlighted section shows a few reviews given on some of the protein bars. Three columns can be seen representing the rating given between 1 to 5, the review title and the review text. The review text can be analysed to identify the emotional polarity. However, the text data from the above sample needs to be classified into a set of variables to identify anomalies such as redundancies, empty fields, incomplete statements in the document.

With the help of Mitplotlib library, the text data extracted for the model is classified. The text data is classified using type of product as a classifier. The classifier then picks the data from the provided sample and classifies according to the product into a set of defined variables to form a structured data format (table). Along with the variables mentioned above - rating, review title and review text, other attributes such as the product ID, Amazon Standard Identification Number, Name of the product, category, manufacturer, keys associated, reviewer city and province are defined based on the full dataset. A sample of text classification, visualised in the Mitplotlib user interface is represented in Figure 11.

	id	name	asins	brand	categories	keys	manufacturer	reviews.text	reviews.title	reviews.userCity	reviews.userProvince
0	AVqkIhwDv8e3D1O-lebb	All-New Fire HD 8 Tablet, 8 HD Display, Wi-Fi,...	B01AHB9CN2	Amazon	Electronics,iPad & Tablets,All Tablets,Fire Ta...	841667104676,amazon/53004484,amazon/b01ahb9cn2...	Amazon	This product so far has not disappointed. My c...	Kindle	NaN	NaN
1	AVqkIhwDv8e3D1O-lebb	All-New Fire HD 8 Tablet, 8 HD Display, Wi-Fi,...	B01AHB9CN2	Amazon	Electronics,iPad & Tablets,All Tablets,Fire Ta...	841667104676,amazon/53004484,amazon/b01ahb9cn2...	Amazon	great for beginner or experienced person. Boug...	very fast	NaN	NaN

Figure 11 Representation of text classification

4.1.2. Identifying the topic of interest

The topic of interest is the subject of analysis. In this case it is "Protein Bars." Out of the raw data present in dataset, there may be a number of comments and reviews regarding the protein bars. Though the prepared text model is readable and understandable, the data is huge and can be tedious for human interpretation. Hence a descriptive analysis of data is needed. This approach gives an opportunity to have a brief idea of what to expect from the text model. It gives descriptive data such as mean standard deviation etc.

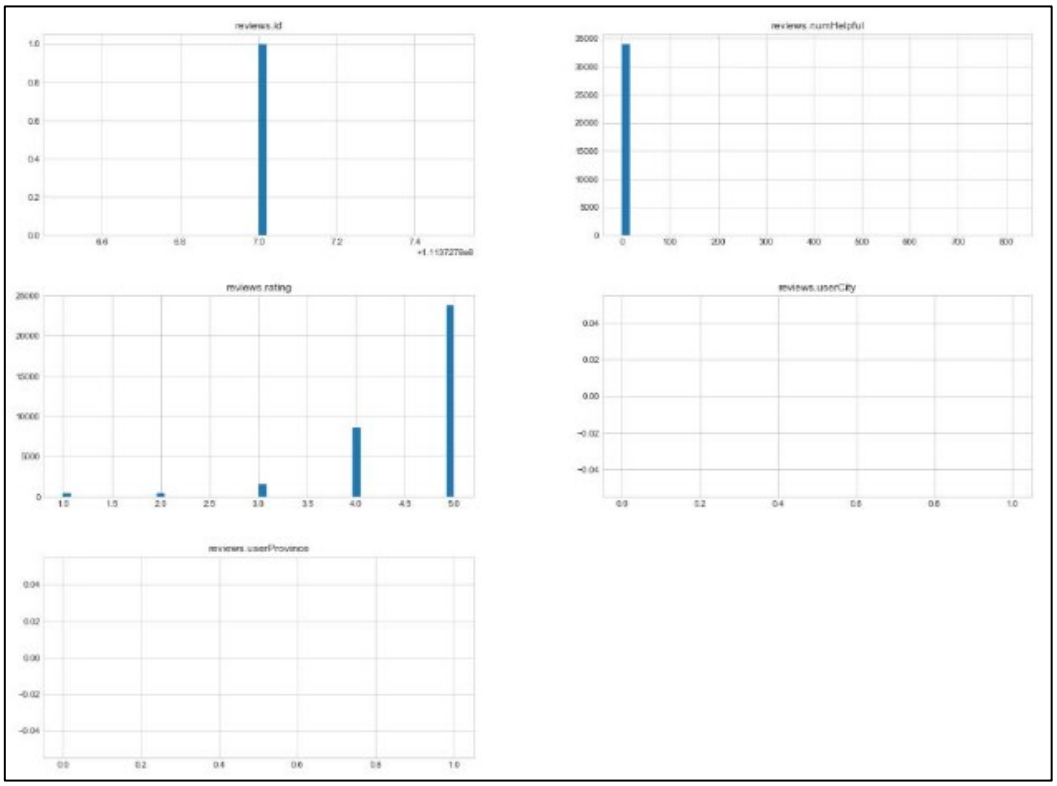
	reviews.id	reviews.numHelpful	reviews.rating	reviews.userCity	reviews.userProvince
count	1.0	34131.000000	34627.000000	0.0	0.0
mean	111372787.0	0.630248	4.584573	NaN	NaN
std	NaN	13.215775	0.735653	NaN	NaN
min	111372787.0	0.000000	1.000000	NaN	NaN
25%	111372787.0	0.000000	4.000000	NaN	NaN
50%	111372787.0	0.000000	5.000000	NaN	NaN
75%	111372787.0	0.000000	5.000000	NaN	NaN
max	111372787.0	814.000000	5.000000	NaN	NaN

Table 3 Descriptive statistics based in the topic of interest

The above given table represents the descriptive statistics drawn from the test model. The model contained approximately 600,000 observations. The descriptive analysis has been implemented to evaluate the strength of ratings based in the given product. It can be seen that the number of reviews that can be useful for sentiment analysis are 34131 and the number of reviews that can be leveraged based in the rating are 34627. Apart from the mean and std. deviation, the strength of review for the given product ID according to the rating can be interpreted that there is a higher percentage of 5-star rated reviews. However, the following assumptions can be made based on the above data;

- A low standard deviation is observed with a review score 4.58 on average, most reviews are positive
- A higher std. deviation was observed with an average of 0.6 for the no. of helpful reviews
- From std. deviation, it can be said that a number of 0-13 customers found the reviews helpful
- From ***max(reviews.numHelpful)***, it can be interpreted that 814 people found the most helpful reviews as helpful.

4.1.3. Visualising the distribution



The distributions are visualised using the plotting library. From the above representation, it can be assumed that the outliers are significant and there is an opportunity to consider the reviews that are helpful to more than 50 consumers. From the rating distribution, most of the examples were highly rated and the 5-star ratings are twice the size of other ratings combined.

4.1.4. Determining sentiment

Using the data that has been established in the above steps, a classifier is used to determine the sentiments from reviews. The classifier is defined to recognise the sentiment based on rating. The impressions of executing classifier over the variable *review.rating* are given in Figure 12. The accuracy of the model is recorded at 93.45% with 27701 training samples and 12526 words in the given training sample

```
4349     Positive
30776    Positive
28775     Neutral
1136     Positive
17803    Positive
7336     Positive
32638    Positive
13995    Positive
6728     Negative
22009    Positive
11047    Positive
22754    Positive
5578     Positive
11673    Positive
19168    Positive
14903    Positive
30843    Positive
5440     Positive
28940    Positive
31258    Positive
Name: Sentiment, dtype: object
```

Figure 12 Sentiment determination

5. Conclusion

The research model was designed to accommodate the defined research objectives. Primarily, the research was directed to understand the importance of marketing approaches for retail markets. The industry review of Irish retail market suggested a higher proportion of consumers from varying demographics are visiting online shopping platforms for their retail purchases. By studying different digital marketing techniques combined with data mining, it was observed that the data being generated by consumers in the form of their attitudes and opinions towards products in online platforms highlighted rich information that can be used for marketing purposes. Since the considered data is being generated continuously on various ecommerce websites, a text mining approach was appropriate. Text mining in association with sentiment analysis highlighted the significance of opinions in customers. Now these opinions can be mined to use with personalization strategies to create better and efficient communication technique.

Form the analysis, it was evident that products with less reviews are not able to justify the statement "Lower rated products are inferior". However, the products that have relatively higher ratings could justify that they are superior products and they are sold in high numbers. For example, it was observed that Fulfil bars have higher ratings based on several different factors. This brand though it initially started manufacturing its products in Ireland, it was shifted to the UK for various business reasons and they are observed to have highest market share in Ireland. These bars continue to sell in high numbers when compared to less significant products such as Origin Bars which indicated relatively lower ratings. Moreover, Origin Bars are new in the market and they are sold in limited retail outlets including e-commerce websites. From the analysis it can be said that Fulfil bars are a superior product to Origin Bars. However, the analytical approach used is not efficient in recognizing the consumer patterns in less data and more input data can be considered to draw assumptions on lower rated products.

On analysis, the dataset was observed to be positively skewed because of relatively lower observations towards low rated products. Successful analysis of the data using Sentiment Analysis was executed despite it skewedness. The analysis was able to determine and classify the reviews into Positive, Negative and Neutral sentiments. The analysis was also anticipated to prove that the concentration of positive reviews is associated with the sales as the customers interpret those reviews as useful for their pre purchase experience. The descriptive analysis of the data showed significantly higher standard deviation indicating that a total of 814 customers in the given dataset found the most helpful reviews as useful. These can be the customers who get influenced from reviews. For example, "*Yum - Good nutrition, and they taste so good my toddlers actually want bites of them. Repeatedly. Can't say that of any of the other protein bars I've tried*." This comment is taken from the dataset. The analysis suggests that there is a probability that a consumer reading this review may find it useful and make a purchase.

Sentiment analysis proved to be an effective approach to identify post purchase opinions of cyber consumers, especially considering websites like Amazon. However, this approach lacks credibility when considering company websites. As the analysis suggests, it requires more data to draw assumptions. Websites like Amazon are global and consumers are asked to leave feedback about their purchase experience every time they buy something. This is not the same

with other websites. For instance, consumers from Ireland can purchase from Amazon via the UK but consumers from the UK may not buy from smaller websites like discountsupplements.ie for various reasons like supply chain limitations, shipping problems, less recommendations, etc.

Though the research model was designed only to analyze data from a well-established dataset of product reviews, it is capable of mining text data from multiple domains. Similar to this approach other algorithms like SpeedText can be used to design data models of larger text data sources like blogs and articles. This allows users to establish divergence in customer emotions. In conclusion, the analysis substantiated that behavioral personalization can be redefined through data mining approaches. Protein bar brands can analyze the attitudes their customers have towards different variables of the product such as the price, taste, nutritional value etc. Used along with personalization techniques, opinions mined from the analysis can be used to create personal messages for direct marketing. This proves the hypothesis that customer opinions can be analyzed to create effective digital marketing – through behavioral personalization.

6. Discussion

The results from analysis are aligned with the research questions to justify the hypothesis. The hypothesis was proved by highlighting few digital marketing techniques like true personalisation and behavioural personalisation and data mining techniques such as natural language processing and sentiment analysis. The primary objective of this research was to analyse how customers feel about protein bars in the Irish retail market. It can be seen from the evaluation that the research model was successful in extracting the sentiments of customers from a set of millions of observations. Though the approach was successful, a few limitations were noted based on the complexity of the implementation and application usability. Assumptions can be made based on the algorithm that the sentiment analysis approach can be effective for social CRM strategies (Chaffey & Ellis-Chadwick, 2019). The data from social media networks can be subjected to this analysis to draw information that can be used for business development. The proposed model can bring a paradigm shift in traditional CRM. However, personalisation approach can be considered as one of the effective ways for protein bars. Creating personalised emails can be effective to communicate with consumers. The study proves that behavioural personalisation can be executed with the help of opinion mining. Products such as Fulfil and Origin Bars can learn from their customers through opinion mining and use behavioural personalisation for better CRM and business management.

6.1.1. Limitations, Challenges and Future work

Since the proposed methodology works with customer data. The primary challenge that can be considered is the regulations with using the data. Though the regulations concerning the privacy of end users such as cookie policies were established, they have influenced the sales and marketing techniques to ensure customer safety especially while using their data. Apart from the data other challenges such as the constantly changing customer requirements and behaviours and the training of staff in application of these new and technologically inclined techniques. Some of these challenges may be expensive to implement but necessary for efficiency and productivity of company. The noted challenges are;

- Privacy and data protection
- Changing consumer behaviours
- Training
- Processing the data

Considering the limitations, sentiment analysis can be seen as a method of opinion mining as it allows to classify opinions according to emotional polarity. This procedure however can be applied to a spectrum of subjects such as brand, person or a product. Through the research, the opinions of the considered dataset were successfully interpreted. However, more data was required to process the observations with lesser ratings. This complexity generated limitation in processing the low rated products. The other limitation observed from the analysis is noise reduction in data. Since large datasets were analysed, the algorithms captured redundancies in data such as the same product IDs, Standard Identification numbers, etc. this anomaly resulted in a skew with ~93% of efficiency. In relation to the personalisation technique emphasised in the research model lack of information due to missing addresses and other details limits the model in defining precise communication medium. Moreover, the personalisation model collapses as some the consumers with missing data will get generic messages. Not only the data integrity but the accuracy is also essential as the probability of incorrect extracting information from consumer data is evidential. However, the execution model depends on the assumptions made by the organisation for the interpretation of data.

The future work can be projected towards data mining techniques such as Artificial Neural Networks to target more complex marketing applications. The protein bar industry can be benefitted from digital marketing approaches as relatively less promotions and marketing is involved with these products. Other digital marketing approaches that align with this data mining model such as transforming the user experience consumer service and reputation management can be included as research prospects.

7. References

Adi, T., 2009. A theory of emotions based on natural language semantics. In: *In Handbook of Research on Synthetic Emotions and Sociable Robotics: New Applications in Affective Computing and Artificial Intelligence.* s.l.:IGI Global, pp. 292-323.

Alessia, D., Ferri, F., Grifoni, P. & Guzzo, T., 2015. Approaches, tools and applications for sentiment analysis implementation. *International Journal of Computer Applications,* 125(3), pp. 26-33.

Alghizzawi, M., 2019. The role of digital marketing in consumer behavior: A survey. *Intrtnational Journal of Information Technology and Language Studies,* 3(1), pp. 24-31.

Armstrong, G. & Kotler, P., 2016. *Principles of marketing.* 16 ed. Harlow: Pearson Education Limited.

Atalay, S., Kihal, S. E. & Ellsaesser, F., 2019. *Using Natural Language Processing to Investigate the Role of Syntactic Structure in Persuasive Marketing Communication,* Frankfurt: SSRN.

Bijakšić, S., Markić, B. & Bevanda, A., 2018. Decision Support System for Marketing Metrics. *Economy, Finance and Business in Southeastern and Central Europe*, 03 May, pp. 741-755.

Binchy, G., 2019. *Understanding protein into the future,* Dublin: Bord Bia – Irish Food Board.

Bordoloi, A. & Bhardwaj, S. N., 2020. Digital Marketing: its Impact on Consumer Buying Behaviour. *Solid State Technology,* 63(5), pp. 5037-5042.

Cambria, E. & White, B., 2014. Jumping NLP curves: A review of natural language processing research. *IEEE Computational intelligence magazine,* 9(2), pp. 48-57.

Campbell, C. & Farrell, J. R., 2020. More than meets the eye: The functional components underlying influencer marketing. *Business Horizons,* 63(4), pp. 469-479.

Chaffey, D. & Ellis-Chadwick, F., 2019. *Digital marketing.* 7 ed. Harlow, England ; New York: Pearson.

Chaffey, D. & Ellis-Chadwick, F., 2019. *Digital marketing.* 7 ed. s.l.:Pearson UK.

Chowdhury, G. G., 2003. Natural language processing. *Annual review of information science and technology,* 37(1), pp. 51-89.

Greaves, F. et al., 2013. Use of sentiment analysis for capturing patient experience from free-text comments posted online. *Journal of medical Internet research,* 15(11), p. 239.

Harnad, S., 2009. The annotation game: On Turing (1950) on computing, machinery, and intelligence. In: R. Epstein & G. Peters, eds. *Parsing the Turing Test: Philosophical and Methodological Issues in the Quest for the Thinking Computer.* s.l.:Kluwer, p. 28.

Hoffman, D. L. & Novak., T. P., 1997. A new marketing paradigm for electronic commerce. *The information society,* 13(1), pp. 43-54.

Kang, Y. & Zhou, L., 2016. *Longer is better? A case study of product review helpfulness prediction.* Baltimore County, Association for Information Systems.

Kannan, P. & Li, H. (., 2017. Digital marketing: A framework, review and research agenda. *International Journal of Research in Marketing,* 34(1), pp. 22-45.

Kim, B.-D. & Rossi, P. E., 1994. Purchase frequency, sample selection, and price sensitivity: The heavy-user bias. *Marketing Letters,* 5(1), pp. 57-67.

Kingsnorth, S., 2019. True Personalization. In: *Digital marketing strategy: an integrated approach to online marketing.* London, Philadelphia, New Delhi: Kogan Page Publishers, pp. 205-207.

Kotane, I., Znotina, D. & Hushko, S., 2019. Assessment of trends in the application of digital marketing. *Scientific Journal of Polonia University,* 33(2), pp. 28-35.

Kotler, P., 1971. What consumerism means for marketers. *Harvard Business Review,* 50(3).

Kotler, P. T. & Armstrong, G., 2016. *Principles of Marketing.* 16 ed. s.l.:Pearson Education.

Kraemer, K. L., Dedrick, J., Melville, N. P. & Zhu, K., 2006. *Global e-commerce: impacts of national environment and policy.* 1 ed. New York: Cambridge University Press.

Kumari, N. & Singh, S. N., 2016. Sentiment analysis on E-commerce application by using opinion mining. *2016 6th International Conference-Cloud System and Big Data Engineering (Confluence)*, 14 Jan, pp. 320-325.

Linoff, G. S. & Berry, M., 2011. *Data mining techniques: for marketing, sales, and customer relationship management.* 3 ed. Indianapolis: Wiley Publishings, Inc..

Liu, B., 2010. Sentiment analysis and subjectivity. In: N. Indurkhya & F. J. Damerau, eds. *Handbook of natural language processing.* Cambridge, UK: Chapman & Hall/CRC , pp. 627-666.

Losiewicz, P., Oard, D. W. & Kostoff, R. N., 2000. Textual data mining to support science and technology management. *Journal of Intelligent Information Systems,* 15(2), pp. 99-119.

Markić, B., Bijakšić, S. & Bevanda, A., 2016. Sentiment Analysis of Social Networks As a Challenge To the Digital Marketing. *Ekonomski vjesnik/Econviews-Review of Contemporary Business, Entrepreneurship and Economic Issues,* 29(1), pp. 95-107.

Nasukawa, T. & Yi, J., 2003. Sentiment analysis: Capturing favorability using natural language processing. *Proceedings of the 2nd international conference on Knowledge capture*, 23 Oct, pp. 70-77.

Palamarova, P., 2018. Comparative Study in Digital Marketing Trends in Adventure Tourism: Bulgaria, Estonia, Ireland and Scotland.. *Journal of Emerging Trends in Marketing and Management ,* 1(1), pp. 242-251..

Pathak, X. & Pathak-Shelat, M., 2017. Sentiment analysis of virtual brand communities for effective tribal marketing. *Journal of Research in Interactive Marketing.*

Pinto, V. R. A. et al., 2017. Influence of package and health-related claims on perception and sensory acceptability of snack bars. *Food Research International,* Volume 101, pp. 103-113.

Safefood, 2019. *A survey of high-protein snack foods,* Cork: Safefood.eu.

Salehi, M., Mirzaei, H., Aghaei, M. & Abyari, M., 2012. Dissimilarity of E-marketing VS traditional marketing. *International journal of academic research in business and social sciences,* 2(1), p. 510.

Saura, J. R., 2020. Using Data Sciences in Digital Marketing: Framework, methods, and performance metrics.. *Journal of Innovation & Knowledge,* Volume 5.

Turing, A. M., 1950. Computing machinery and intelligence. *Mind,* 59(236), p. 433.

Ushnish, S. & Michael, K. H., 2021. Meeting Changing Customer Requirements in Food and Agriculture Through the Application of Blockchain Technology. *Frontiers in Blockchain,* 4(1), p. 5.

Wang, J. et al., 2013. *Mining new business opportunities: Identifying trend related products by leveraging commercial intents from microblogs.* Seattle, Association for Computational Linguistics, pp. 1337-1347.

Whitley, E. & Ball, J., 2001. Statistics review 1: Presenting and summarising data. *Critical Care,* 6(1), pp. 1-6.

Wind, J. & Mahajan, V., 2002. Digital marketing. *Symphonya. Emerging Issues in Management,* Volume 1, pp. 43-54.

Wind, Y. P. & Thomas, R. J., 1980. Conceptual and methodological issues in organisational buying behaviour. *European Journal of Marketing,* 14(5,6), pp. 239-263.

Zahra, S. A., 2008. The virtuous cycle of discovery and creation of entrepreneurial opportunities. *Strategic Entrepreneurship Journal,* 2(3), pp. 243-257.

Zhang, X., Zhao, J. & LeCun, Y., 2015. Character-level Convolutional Networks for Text Classification. *Advances in Neural Information Processing Systems (NIPS),* 04 Sep.28(1).