Prediction of Success Factors of FMCG Commercials using Signal Processing and Machine Learning Algorithms

by

Nazmus Saquib
17101469
Kaniz Fatema Mithun
17101124
Jarin Tasnim
17101091
Pushpol Sen
16301181

A thesis submitted to the Department of Computer Science and Engineering
in partial fulfillment of the requirements for the degree of
B.Sc. in Computer Science

Department of Computer Science and Engineering
Brac University
June 2021

Declaration

It is hereby declared that

1. The thesis submitted is my/our own original work while completing degree at Brac University.

2. The thesis does not contain material previously published or written by a third party, except where this is appropriately cited through full and accurate referencing.

3. The thesis does not contain material which has been accepted, or submitted, for any other degree or diploma at a university or other institution.

4. We have acknowledged all main sources of help.

Student's Full Name & Signature:

Nazmus Saquib
17101469

Kaniz Fatema Mithun
17101124

Jarin

Jarin Tasnim
17101091

Pushpol Sen

Pushpol Sen
16301181

Approval

The thesis/project titled "Prediction of Success Factors of FMCG Commercials using Signal Processing and Machine Learning Algorithms" submitted by

1. Nazmus Saquib (17101469)
2. Kaniz Fatema Mithun (17101124)
3. Jarin Tasnim (17101091)
4. Pushpol Sen (16301181)

Of Spring, 2021 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of B.Sc. in Computer Science.

Examining Committee:

Supervisor:
(Member)

Zavid Parvez
Digitally signed by Zavid Parvez
DN: cn=Zavid Parvez, o=Brac University, ou=CSE, email=zavid.parvez@bracu.ac.bd, c=US
Date: 2021.06.09 20:23:46 +10'00'

Dr.Mohammad Zavid Parvez
Associate professor
Department of Computer Science and Engineering
Brac University

Program Coordinator:
(Member)

Md.Golam Rabiul Alam
Associate professor
Department of Computer Science and Engineering
Brac University

Head of Department:
(Chair)

Sadia Hamid Kazi, PhD
Chair person and Associate professor
Department of Computer Science and Engineering
Brac University

Abstract

In the highly advanced and competitive business world, today's marketing strategy has become very difficult. To satisfy the needs and necessities of consumers, advanced marketing research methods are required to recognize consumer's preferences. For the purpose of product or service promotion to the mass people, companies are spending big budgets on TVCs (TV Commercials). Since the organizations are leaning towards digitized advertising at a rapid pace, effective research is required to improve the system. All the TVCs can not influence the viewers in the same way. There must be some factors that effectively increase the success rate of a TVC. For these factors to find out and make TVCs resource efficient, intensive research work in this field has been of primary priority to upgrade the industry. This kind of research is able to make a huge impression to maximize the outcome of any advertisement. Only a few literature were found on predicting TVC success factors but none of them used textual data as well as brain signal at the same time to find the factors that are most important according to our study. Although advertisement success has been a game changer for the FMCG industry lately, there is enough room to work in this industry. Since advertisements on FMCG (Fast Moving Consumer Goods) spend a comparatively bigger budget than others and it has a better influence on daily purchases of the majority of people, we decided to work on this industry particularly. In this research, we have used both the subjective and objective measurements as our dataset. Textual data taken from the interviewees as well as their brain signal extracted by EEG machine has been applied to implement the algorithms on. To predict the vital factors, we have implemented some supervised machine learning along with some deep learning algorithms like ANN and MLP to pull out our outcome. Among all of the features that we have worked on, it is found that 'relevant message' is the most important factor in an advertisement to convince a viewer. Including 'relevant message' we have taken all other crucial factors in consideration and found out the importance of the factors to make an advertisement successful. Executing machine learning methods, we have achieved highest 96% of accuracy and executing deep learning, highest 93% of accuracy was achieved. The results proves the crucial relationships between the features that we used and the advertisement success.

Keywords: EEG, Advertisement, FMCG, Emotion, Purchase Behaviour

Acknowledgement

First of all, we express our humble gratitude to the Almighty Allah for assisting us and directing us in the right direction.

We would like to convey the deepest appreciation to Dr. Mohammad Zavid Parvez, Associate professor, Department of Computer Science and Engineering, Brac University, who has supported us in developing our ideas and directing us in bringing it into practice. We are dedicating this thesis to our parents for their continuous support and encouragement. Finally, we would like to convey our special thanks to Mr. Shahan Jamil Bhuiyan who motivated us from the beginning of this research and helped us with his valuable suggestions.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

Advertisements have become a very important part of our lives. Above all, new items may acquire a positive impression by a legitimate commercial.Advertisements always have been one of the most powerful means of communication with the users of the products.perfectly. In our Media, we see different advertisements of different types. Media are also various types.

1.1 Motivation

Advertisements have long been one of the most effective strategies of reaching out to product users. Above all, a legitimate commercial can give new things a positive impression. The marketing environment has widened and become sophisticated. As a result, a company's marketing expenditure must be carefully allocated [23] . Year by year companies are increasing their advertising budget as they see the outcome of the spending. Covid situation has made the situation quite more favorable for the advertisers [37]
The figure here depicts the digital ad spending budget by year and predict for the next years. The expenditure quite sums the whole situation of the advertisement industry. As FMCG [22] (Fast Moving Consumer Goods) industry is more dependent on TVCs, running a research on advertisement of this industry could bring up important values to improve the tactics in the future. So we aim to work on determining the success factors of advertisements and show the impacts of some specific attributes on which we suggest to focus on while creating commercials.

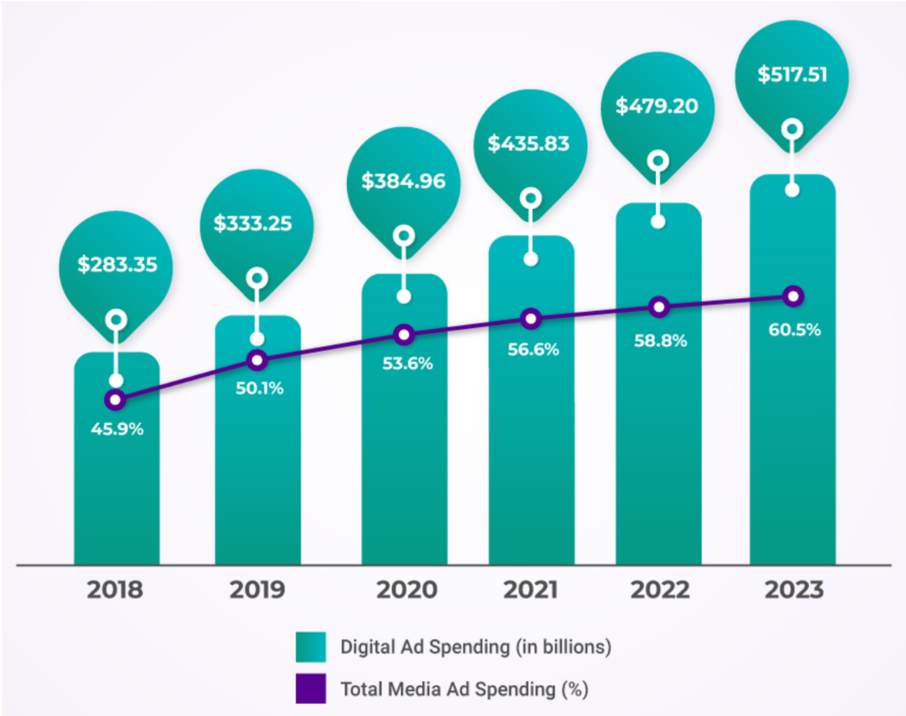


Figure 1.1: World wide digital ad spendings

1.2 Problem Statement

We all know that marketing is extremely important for companies. The marketing plan is crucial for start-up companies. They can readily launch their products in markets after they gain public attention. A perfect organization's principal goal is to achieve this. Every owner of a start-up aspires to take their businesses to the next level. Sales may suffer greatly from poor advertising, and the company may be forced to close without the perfect advertisements. They suffer a lot of controversies as a result of this.

In today's economic environment, our young people want to run their own businesses rather than work for others. As a result, they form their own organizations. Some people start their careers while still in school. They put their money and efforts on the line to start their own businesses. Entrepreneurs are what they are called. New ideas and perspectives are valued by all, and entrepreneurs play a significant part in this field. Though this is regrettable in such a way that many entrepreneurs have failed to make a profit. The inability to maintain a proper advertising policy is the primary cause of the loss. They put their all into their work, yet they don't seem to be achieving much progress.

A successful advertisement can change people's minds about a company or its products. As a result, we come up with the idea of identifying the essential elements of the marketing process, notably the marketing variables. We also propose which elements are required for a successful advertisement. The public responses have been used to identify these factors.

1.3 Aim and Objectives

Traditional marketing strategies have been converted to a strategic strategy, with major contributions from technology, in order to reach the organizational goal. If enough data is analyzed, current technology can almost predict anything. The success of advertisements is essential to some companies, and in certain cases the entire market. Before launching a product, companies conduct market research and come up with a strategy for marketing it to the target market. Our main goal is to use machine learning algorithms to predict the success rate of advertisements and analyze the factors. There have been some works on advertisements previously, but we are now focusing on the relationship between human emotion and advertisement. We have shown participants many advertisements and tried to find out whether they thought the advertisements were engaging and appealing or not. Again, we tried to observe if the advertisements had any influence on their purchase intent. There have been previous works on advertisements, but this time we are focusing on the relationship between human emotion and advertisement when making a purchase. We are working on a dataset that is based on the responses of a group of people in order to predict the success rate and analyze the factors.

There has been a survey form with inquiries about the advertisements that we presented to the people. The questions is designed in such a way that they cover nearly every aspect of the advertisements and can properly represent how the audience felt about them. Then, using SVM [32], we have figured out which segment of people prefers which part of the advertisement, and we'll figure out where to target that segment. We applied tree algorithms in our research paper like Random Forest, XGBoost. Tree algorithms are modern and they have a higher accuracy level. Then we applied MLP algorithm from deep learning in our datasets. These algorithms show us the various rates of purchasing the product as well as the general general opinion of it. The public's willingness to buy a product will be increased if a positive impression is made.

Chapter 2

Literature Review

Advertising is a huge and successful industry, and advertisers want their products or services to be seen as highly attractive and rewarding. In [10] they have worked on the significance of clustering of customers in the U.S. wine market. Throughout the most recent 15 years, the U.S. wine market has been developing consistently. The amount has increased from 2688 to 8862 from the year 1999 to 2016 of wineries. Most of the wineries are located in different regions of the U.S such as the Mid Atlantic region. The selling rate of local wines in those regions was very high until outside regions like Europe, South America, Oceania tried to start their business in the same wine markets. To avoid this uprising competition, it has become very important to elevate the local wine industry of the US in those regions.A Cluster Analysis has been done on consumers for understanding their purchase intent, perspective, and social demographic attitudes. 4 clusters have been found from the cluster analysis that has been done on wine markets. The 4 classes are Detractor, Enthusiasts, Neutral, and Advocators. The class 1 people are those who do not want to purchase local wine and about 67.4% of them have never purchased before. Class 2 people are those who are very interested in purchasing wine and they have a very high similarity of features with class 4 consumers. It is mentioned that around 74.5% of people in class 2 have already bought local wines. And the class 4 people are either supportive or defensive about purchasing local wine and about 60% of them have purchased before. But the class 3 people have a 50% purchase rate. This study on consumers will help to promote the local wine business in those regions.

In [5] quite similar work has been done but on the human resources of the hospital management industry. The aim and objective of this research are to find out whether management decisions are influenced by various things to approve a HRIS in the hospitals of Bangladesh. This paper incorporates two influential acceptance theories of the HOT-fit model and the TOE for understanding this issue.13 Factors in four dimensions are examined to detect their impact in hospital decisions on HRIS adoption. Implementing a method which is non-probability around five hundred fifty copies of surveys have been divided to HR employees from ninety two commercial medical clinics in our country. Responses that have been done by participants around 69.63% were valid responses. Depending on the Human resource management implementation phase, three groups were formed from the responses: adopters, prospectors, and laggards. This survey show 5 of the issues, which are Information Technology services, resources for top administration, personnel Information Technology skills, potential value, and competing constraints. The findings identify the five most important factors: IT architecture, senior management support, staff IT capabilities, purchase intention, and increased competition. All indicators were shown to be varied in different adopting groups, according to the study.

In [11] the author talks about emotionomics based on EEG signal, EMG, EDA, ECG, CNN, IoMT. He also talks about emotionomics based on emotion recognition system, affective computing, convolutional neural network. To understand human's affective state researchers have attempted to utilize some widely used AI method. For perceiving, realizing, and predicting a human's emotional state, specialists have analyzed human reactions, statements, social media platforms, and physiological signs. Researchers can observe emotions by the use of IoMT and clinical sensors. For understanding, the emotional state of neurological aspects of humans results have generated using using EMG, EDA, ECG, medical sensors, and CNN. 5 affective states are there which are delighted, dissatisfied, unhappy, and indifferent. An investigation is done to understand the performance of the proposed method.

In [4] the author stated about emotionomics based on neuromarketing, eye tracking, as well as visual attention. The author also says that in retailing analysis the use of Eye Tracking and Neuromarketing is becoming very famous because it gives great results for marketing research, new ideas, product quality development, advertisements, consumer satisfaction, loyalty, and many other topics. Brand value, division, item choices, valuing choices, spot and advancement choices, and social advertising, are things that influence a customer while purchasing a product. Eye-tracking is used here to analyze the process of how a decision is being taken. Since neuromarketing and eye-tracking have grown in popularity, in academia and the business world, there is no doubt that there is more research to be done. Lastly, the author says that we should expect to see this equipment and research line used much more frequently than usual, and It would soon be a standard feature of marketing.

In [6]the author says that emotionomics is based on emotions that are related to both in the marketplace and at work, there are numerous business opportunities. In today's highly competitive world where many products are similar, a product's emotional benefit can make a difference in a customer's perception. The author talks about advertisement's impact on emotions based on consumer choice, behavior, and store atmosphere. The research paper is based on the study of current knowledge in the following areas, consumer's choice(neuromarketing), neuroeconomics, and visual merchandising. At first, music which is a fast type of Slovak was selected. As the type of the music was relatively fast, it could have influenced the customers faster through the stores and therefore it could have increased the sales. In this paper, the author talks about using two different styles of music, which are considered to be the most suitable to create the necessary atmosphere in the wine shop. This paper is aimed to present results based on data gathered from various neuroimaging techniques using an electroencephalograph and a biometric method using an eye camera in a lab setting, based on subconscious reactions of music-influenced consumer behavior. In the retail store, based on the type of merchandise it is necessary to adjust the music (music's genre, style, tempo, and quality) to create a great atmosphere. The highest average levels of emotional engagement were computed for

French music in the middle of the observed time period based on EEG measurements.

In paper [9] based on EEG signal, EMD signal the author has stated about advertisement impacts on consumer's emotions. The examination results indicate that the given strategy acts as a domain, as the fractal dimension. The author has proposed a method in this paper where extracting emotional recognition in the EMD domain is used. EEG signals are likewise acquired as an example at regular intervals all over the test, so it might give another answer for continuous feeling acknowledgment. In view of the quickest changes in EEG signals, the author affirms that the derivation of feeling is more comparative with high recurrence parts. The author again says that a few highlights from IMF1 with 8 channels by the strategy which has been proposed should be acquired, which would be tedious and likewise lessen the weight. Comparing methods in the time domain the author says that EMD provides the benefit of being able to use greater information about frequency. The trial results show that the proposed strategy circulates the technique in a time area. For example, fractal measurement and test entropy in this paper, the author proposed a strategy for inclusion extraction for feeling acknowledgment in EMD space, another part of the view. The proposed technique simply needs to separate highlights with Eight channels from IMF1, which would lessen time and calm calculation load. Compared to strategies, EMD has the upside of using more recurrence data.

In [7] on the basis of neuromarketing, eye tracking, visual attention, consumer behavior, marketing research emotionomics has been described by the author. In this because of a section of the paper analysis, a study was identified in the literature that examined the following topics, segmentation, Targeting, and Positioning. According to the researches, in this paper, the author discussed benefits that can be in many groups such as choices and Customer preferences, decision-making process, buy intent, regional brain activation in response to marketing stimuli, hormones, and unconscious mind. The benefits referring to the aspects of the business, specifically segmentation, targeting, and positioning, include goods and services, brand, advertising, packaging, in-store solutions, or business online. On the basis of the research, companies might be benefitted from the Eye Tracking market research techniques. Promoting is a vital industry in the present computerized world, where commercials feature their things and administrations as significant and advantageous.

In [3] the author has proposed an emotion-enabled algorithm. EEG signals are utilized for determining a user's ongoing emotional state. Since EEG gadgets have grown more portable, affordable, and simple to set up, they can be utilized in advertisements to measure emotion. As a result, clients' true thoughts toward the commercial might be gathered in real-time and used as input to change the advertisement's scenes. It is possible to construct advertising films in which some aspects of the sceneries, such as shapes, sizes, or colors, are altered to elicit the feelings desired by the advertiser. using a combination of these properties, A polynomial kernel SVM (Support Vector Machine) classifier is trained and observed for use in real-time recognition. The EEG data are received via a bandpass filter during the recognition phase. The analytical attributes are then taken out then put into the classifier of SVM that's been built during the training session. This paper it has also discussed that allowing for real-time personalization of advertising videos, EEG feeling identifiers can serve to improve the effectiveness of the advertisement.

This study [12] examines the consumer brain reactions that support passive viewing of expensive compared to regular which means peak emotional value compared to lower emotional value branded items, the authors applied electroencephalogram methods to collect potentials when female participants peacefully watched photos of expensive and regular branded items, following social facilitation theory. They examined the amplitudes of event-related data in three time periods as well as LPP. The amplitude of LPP for was bigger expensive branded items than for regular branded items, but only when both people were present, implying that brand type has a stronger emotional impact when both people are present. The findings suggest that higher attention allocation and motivational significance are represented in LPP frequency during passive watching of suitable marketing portrayals . The ability to respond to stimuli with higher personal significance is strengthened by the presence of another person.

As per the author [1], his paper illustrates how merchandising can be profitable from neuroscience. Here the author examines each product advertisement instance study. Pretests of two copies of this TV or video advertisement demonstrated that, despite being nearly identical, the two versions had significantly various impacts. The authors intended to see if neurophysiological techniques might catch contrasts in customer reactivity on minor variations in marketing stimuli. He performed EEG, EMG, and SC measures for seeing if noticeable differences in frontal cortical activity present there, facial nerve-muscle action, or incentive altitude after observing 2 different copies of the stimulation component. After completing the estimation there would be noticeable statistical variation among pants. The activity of the brain was examined for both the entire ad and specific sequences. The signal was modified utilizing diverse inquiry, various algorithms, wavelet conversion, as well as calculating and analytical methods. On a broad level, the results demonstrated people's beliefs that the brain can detect even minor differences.

In [13] observes an effective audiovisual advertisement dataset that shows accurate emotions in participants. They presented one of the few studies to investigate ad emotions and improve on previous findings. They also discuss advertisement feelings in terms of explicit human perceptions as well as underlying audiovisual and EEG characteristics. They also presented 100 ad affective datasets including related affective ratings that have been thoroughly controlled. According to statistical analysis, the advertisement dataset is able to inspire consistent feelings over a wide range of people. They demonstrated that CNN-based transfer learning can effectively capture audiovisual feelings by fine-tuning the Places205 Alexie. They compared AR attained with content-based CNN attributes which is based on content versus AR attained with CNN attributes which is based on users. Emotional qualities are best encoded using an EEG-based CNN model. Ad AR also benefits greatly from multi-tasking learning to identify commonalities among emotionally similar advertising.

Author [8] describes a unique approach for evaluating video-advertisement ratings on the basis of MMF that combines physiological examination of the user with they have found a global opinion rating on the internet. To better understand the user's preference, they combined the user's electroencephalogram (EEG) waves with the video's worldwide textual remarks. They asked participants to watch a audiovisual advertisement while their EEG signals were recorded in their system. For each video, valence scores were calculated based on self-report. A higher valence refers to the user's inherent attractiveness. Furthermore, the multimedia data, which included Natural Language Process was used to collect and process the comments of overall viewers. Furthermore, the multimedia data was extracted and processed for sentiment analysis using a NLP methodology, which includes comments provided by overall viewers.Furthermore, the multimedia data was extracted and processed for sentiment analysis using a Natural Language Processing (NLP) methodology, which includes comments provided by global viewers. Textual elements from review comments were evaluated to establish a score to identify the sentiment type of the video. Using EEG data, a regression methodology based on Random forest was utilized to predict an advertisement sentiment outcome to improve the overall prediction. This research was conducted using 15 video clips of advertising. In order to assess their suggested system, they attracted the attention of twenty-five people. When compared to prediction utilizing only EEG data, the suggested multimodal method can lower RMSE in ranking prediction.

In the study [2],the creator recommends a multimodal approach in which the mind signal (EEG information) is gathered as a physiological reaction to the substance being watched by the client and utilized as the essential element. In order to progress the client's expectation, the traditional part is combined by researching comments posted by overall individuals. To anticipate three basic methodologies are usually utilized. One of them makes use of the body's physiological response. User, in which the users' EEG signals are examined while they watch the video parts for promotion. EEG signals have been demonstrated to be very useful in the examination of emotions. From the result of Regression analysis techniques, Using Random forest regression, the minimal Root Mean Square Error (RMSE) and Rating were 0.9056 and 95.05 percent, respectively, for 300 trees. The worth of RMSE and Residual difference didn't change a lot after 900 trees, as can be found in the line graph. These RMSE and Rating change pointers for the selected attributes were calculated using the values' significance as described.

Chapter 3

Background Study

Advertising has varied effects on the local target audience. They also have an impact on multinational brands in terms of brand image and equity. It's no surprise that the corona outbreak has drastically altered consumer behavior and ushered in a wave of global digital upheaval. The grocery market appears to have profited the most throughout these times. In fact, 76 percent of brand marketers now acknowledge that "retail media advertising is critical to their business growth" [37] . Different forms of adverts can be seen in different sorts of media. The most well-known marketing methods rely heavily on visual, more specifically, on motion advertisements. Video advertisements are more flexible to engage a viewer and convince them to buy a product as a story can be presented more clearly than any other formats. As it conveys a huge amount of information in a short time, the internet also facilitates video format of commercials [19] . As the TVCs have broadened their range on social media as well, the impact has been huge lately. So analyzing motion adverts and bringing out some important outcomes could help the process to proceed more effectively.

3.1 Algorithm

3.1.1 Random Forest

RF [29] is an adjustable, convenient, and trouble-free machine learning algorithm that outputs, indeed without hyper-parameter tuning, an incredible outcome a number of times. RF is also amongst the most used computations since it is trouble-free and different. RF [28] can be pursued for both categorization and regression assignments.Random forest could be an administered learning calculation. Here, "forest" helps to create a gathering of step by step decision trees. One massive purpose is that it might be equalized for the both categorization and regression factors, which frame the high number share of current ML frameworks. By utilizing this equation, we can discover the value from random forest algorithms.

$$RFfi_i = \frac{\sum_j normfi_{ij}}{\sum_{j \in all\, features, k \in all\, trees} normfi_{jk}}$$

Figure 3.1: Random Forest

3.1.2 Logistic Regression(LR)

Logistic Regression [21] is one of the most commonly used Machine Learning algorithms. It is a statistical model that works together with calculated capacity to introduce a paired ward variable in its most straightforward form, though there are similar more complex extensions available. It is a technique for assessing the boundaries of a calculated model (a type of paired relapse) in regression examination. A binary logistic model has an established variable with two feasible values, such as pass/fail, that is expressed by an indicator variable, and the two values are called "0""1" in mathematics. It provides probabilistic values ranging from 0 to 1. Instead of fitting a curve, we have a tendency to match associate "S" fashioned supply operations in supplying regression, that predicts 2 most values (0 or 1). Using continuous and distinct datasets, it would give chances and classify new knowledge. Logistic regression may be wont to classify perceptions supporting varied styles of knowledge and can easily identify the most useful variables for classification. The Logistic Regression equation is given below:

Y = e^(b0+ b1*x) / (1 + e^(b0 + b1*x))

Figure 3.2: Logistic Regression (LR)

3.1.3 SVM

SVM [24] could be a directed machine learning computation which was used for each categorization or regression challenge. In any case, it's typically used in categorization issues. In the support vector machine algorithm, we lay respective informative data like a point in the n-proportional region (here n indicates the amount of attributes), with the value of an individual attribute which is the value of a specific interrelated, by performing categorization by detecting the hyper-plane that separates the two categories exceptionally fine. We can evaluated the SVM value from the given equation:

$$h(\boldsymbol{x}_i) = sign\left(\sum_{j=1}^{s} \alpha_j \; y_j \; K(\boldsymbol{x}_j, \boldsymbol{x}_i) + b \right)$$

$$K(\boldsymbol{v}, \boldsymbol{v}') = \exp\left(\frac{||\boldsymbol{v} - \boldsymbol{v}'||^2}{2\gamma^2}\right)$$

Figure 3.3: SVM

3.1.4 K Nearest Neighbour(KNN)

K nearest neighbor (KNN) [42] can be utilized for both classification and regression prescient issues. In any case, it is more broadly utilized in classification issues within the industry. To assess any method, we, for the most part, see at three vital angles, which are, facility to explicate output, calculation of time, prescient Power. KNN is a passive learning, distribution-free algorithm. KNN calculation fairs overall parameters of contemplations. Nevertheless, generally, it is utilized due to its ease of translation and less calculation time. KNN calculates the space between data ends. We can apply this easy Euclidean Distance equation to calculate the distance.

$$d(\mathbf{p},\mathbf{q}) = d(\mathbf{q},\mathbf{p}) = \sqrt{(q_1-p_1)^2+(q_2-p_2)^2+\cdots+(q_n-p_n)^2}$$
$$= \sqrt{\sum_{i=1}^{n}(q_i-p_i)^2}.$$

Figure 3.4: K Nearest Neighbour(KNN)

3.1.5 Naive Bayes

NB [26] is a categorization procedure based on Bayes's Theorem with a supposition of self-determination amid predictors. In the middle of more accessible phrases, a NB classifier expects that the extent of an explicit attribute in a class is disconnected from the extent of any other attributes. This algorithm is straightforward to construct and particularly valuable for exceptionally expansive data sets. Along with clarity, NB is known for beating indeed profoundly 4 advanced classification strategies. Bayes theorem gives a way of calculating the posterior probability. By applying the equation stated below, we can find out probability:

$$\underbrace{P(c\,|\,x)}_{\text{Posterior Probability}} = \frac{\overbrace{P(x\,|\,c)}^{\text{Likelihood}}\,\overbrace{P(c)}^{\text{Class Prior Probability}}}{\underbrace{P(x)}_{\text{Predictor Prior Probability}}}$$

$$P(c\,|\,X) = P(x_1\,|\,c)\times P(x_2\,|\,c)\times\cdots\times P(x_n\,|\,c)\times P(c)$$

Figure 3.5: Naive Bayes

3.1.6 ANN

ANN [14] may be a part of a process system that mimics however the human brain analyzes and processes information.Artificial neurons are a series of connected units or nodes in an exceedingly ANN that loosely type the neurons in a biological brain.

Each affiliation can send a touch to different neurons, a lot like synapses during a biological brain.An artificial somatic cell that receives a symptom, processes it and might send signals to different neurons. The activity of every nerve cell is computed by some non-linear function's input's total. Edges are the terms for the links. The strain of edges and neurons is usually obtained as learning improves. The signal intensity with the link is improved or unimproved by the weight. The neurons that have a threshold that permits them to submit a sign given that the mixture signal exceeds that. Neurons square measure typically sorted into layers. On their inputs, totally different—completely different layers would apply different transformations. It functions in the same way as the human brain does. ANN is made up of a huge number of interconnected processing units that work together to process data. As a result, useful outcomes are often produced by them. We may use neural networks for more than just classification. It can also be used for continuous goal attribute regression. By applying the following equation, we can calculate the ANN.

$$z = f(b + x \cdot w) = f\left(b + \sum_{i=1}^{n} x_i w_i\right)$$

$$x \in d_{1\times n},\ w \in d_{n\times 1},\ b \in d_{1\times 1},\ z \in d_{1\times 1}$$

Figure 3.6: ANN

3.1.7 XGBoost

XGBoost [34] is a decision-tree-based Machine Learning algorithm that makes use of an angle boosting technique. Artificial neural systems tend to outperform all other calculations or algorithms in forecasting challenges involving unstructured input images, text, etc. Despite the fact that it was built for speed and performance. Inauguration to XGB, Code of XGBoost Algorithm, Advanced functionality Algorithm of XGB. The goal is to obtain an optimal output value for the leaf in order to reduce the overall equation to a minimum. Because that starts with a value of y0, the next prediction is always equal to the i-1th forecast plus the output value from the ith tree. By applying the equation stated below, we can find out XGBoost:

Real value (label) known
from the training data-set

$$\mathcal{L}^{(t)} = \sum_{i=1}^{n} l(y_i, \hat{y}_i^{(t-1)} + f_t(\mathbf{x}_i)) + \Omega(f_t)$$

Can be seen as f(x + Δx) where x = $\hat{y}_i^{(t-1)}$

Figure 3.7: XGBoost

3.1.8 Chi Square

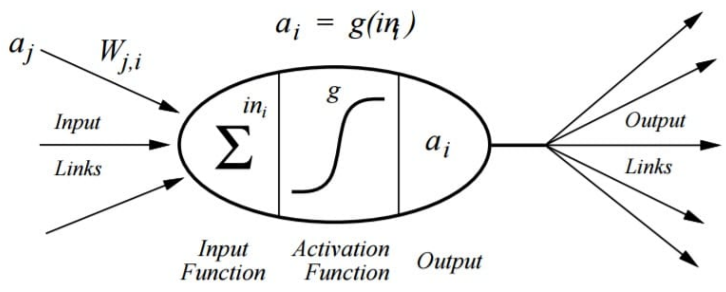
The chi-squared [33] experiment is an analytical hypothesis experiment that is also known as the 2 test that is reliable at the point when the test measurement is chi-squared circulated under the invalid speculation. It includes Pearson's CS experiment and alternatives. In one or more categories of a possibility table, a CS experiment is performed to see on the off chance that there is a really critical distinction between the normal and noticed frequencies. The observations are partitioned into commonly exclusive sorts in normal employments of this test. There are no differentiations between the classes in the population is valid and 2 frequency distribution is observed in the test measurement got from the perceptions. The test's goal is to discover how likely the noticed frequencies are if the invalid speculation is right. When the variables are not dependent on anything and normally disseminated, test statistics follow 2 distribution, which is a general idea and it is supported by the central limit theorem. There are additional two criteria for determining if two random variables are independent based on the pairs' observations. Feature selection is a critical topic in machine learning since we had multiple features in line and must choose the best ones to build the model. By examining the relationship between the features, the chi-square test assists us in solving the problem of feature selection.

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$$

Figure 3.8: Chi Square

3.1.9 Multi Layer Perceptron (MLP)

MLP [39] is a sort of counterfeit neural organization that uses feedforward learning (ANN). The name MLP is confusing, referring to any feedforward ANN at times, and networks built of many layers of perceptrons at other times (with threshold activation). MLP, especially those with a solitary secret layer, are sometimes indicated as "vanilla" neural networks. There are three degrees of nodes in a MLP. Without the data centers, every node is a neuron with a nonlinear initiation work. Backpropagation which is a method of supervised learning technique used by MLP during training.The different layers and nonlinear linear enactment recognize MLP from a direct perceptron. It can show the difference between data that is linearly inseparable and data that is. This algorithm can address issues that aren't straightly divisible and are intended to inexact any consistent capacity. By using the following equation, we can calculate the MLP(Multi Layer Perceptron):



$$a_i = g(\sum_j W_{j,i} a_j)$$

Figure 3.9: Multi Layer Perceptron (MLP)

3.2 Emotion Analysis

3.2.1 Valance

The positive or negative consequences of any circumstance are referred to as valance [41]. The valance is defined as the degree to which the news or visual representation is favorable. It can also be negative. Some questions in our questionnaire are designed to assess valence. After seeing the advertising, the participants are asked if they would like to buy or purchase the product. They were given five possible answers.Firmly Differ,Differ,Impartial, Consent, and firmly consent. The lower mark is marked equal to 1 and higher is marked equal to five. Neutral is 3. The impression of the advertisement is also part of valance. The lower mark is equal to 1, and the higher mark is equal to 5. The number 3 stands for neutral. Valance also includes the advertisement's general impression.

3.2.2 Dominance

Dominance [38] is the portrayal of the prevailing idea of the emotional tion or the controlling of the feeling. Emotion can be grouped into two sorts. Sub note and Dominant. For instance, outrage is prevailing while dread is known as an accommodating feeling. In our questionnaire the members are found out if the message of the commercials are justifiable or not and whether the message passed on in the notice is clear or not. These inquiries assist us with estimating the predominance of the data.

3.2.3 Arousal

Arousal [35] is a measurement of the information's excitement and calmness. Participants are questioned whether the advertising is relevant or not, and whether the message of the advertising is believable or not, in our questionnaire. There are five sections to the answers. Firmly Differ,Differ,Impartial, Consent, and firmly consent these are the five sections.

3.2.4 SAM Scale Survey

The SAM scale study is to quantify the excitement, pleasure, and the dependency level of the data.

3.2.5 Emotionomics

Emotionomics [20] appears to be a motivator in terms of consumer and business prospects. In today's extremely competitive marketplace, a product's emotional worth will make a difference. A company with a highly emotional workforce usually have a competitive advantage. The book "Emotion: Leveraging Emotion for Business Success" by Dan Hill is based on facial coding observations. It demonstrates how branding, product design, advertisement, revenue, customer satisfaction, and workforce management can all be used to control passion for company success. This book will not only help to get closer to consumers and employees, but it also helps to move ahead of the competition.

3.2.6 EEG Sensor

The electroencephalogram (EEG) [25] records the electrical activity of the brain. The waveforms were gathered with the intention of simulating the activity of the brain's superficial layer. They move through the synapse, which connects one brain cell to the next. Neurotransmitters are chemicals that help bridge the gap between the dendrites of one brain cell and the dendrites of another. To achieve a balance among various neurotransmitters in the brain, the brain must go through a difficult process. Simple terms, the brain has to go through a lot of trouble to achieve a balance between each of these neurotransmitters. The EEG [25] test measurement is quite low since it is written in microVolts (V) and has a high scale frequency of roughly 30 Hertz (Hz).

Chapter 4

Our Proposed Approach

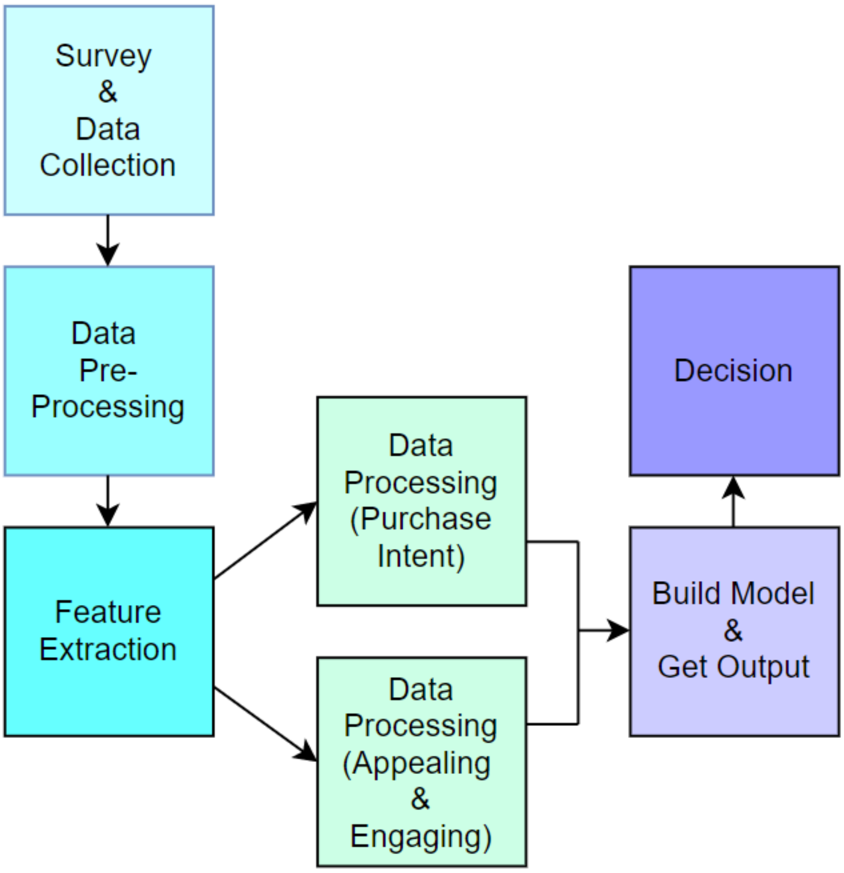


Figure 4.1: Proposed model for prediction of sucess factors of TV commercials

4.1 Survey and Data Collection

For the text dataset, we divided the questionnaire into three parts. Basic information, advertising details and SAM-scale survey of our survey subjects. All this data is protected. All data is collected from specific questions. The textual data set is mainly used for our main investigation. We got people's reactions to certain advertisements. Although we did not get any emotional response from the textual data.

4.1.1 Basic Information of the people

The age, gender, salary, preferred medium for watching advertisements, and total time spent watching advertisements in the preferred mediums are the basic data we want for our research. A total of 170 persons participated in the survey. They are of all genders, ages, and use various methods for watching commercials, among other things. Everyone does not have the same amount of time to watch commercials. As a result, everything must be studied in order to categorize them. These basic details enable us to classify them into categories. These data were gathered using a Google form.

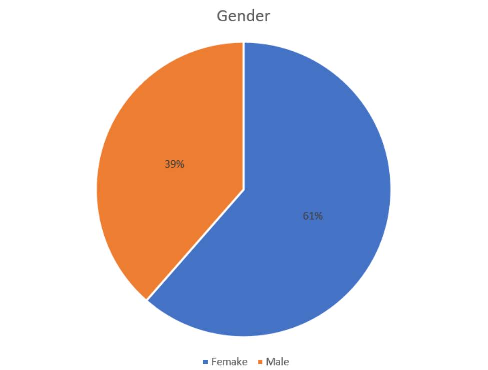


Figure 4.2: Pie Chart of Gender

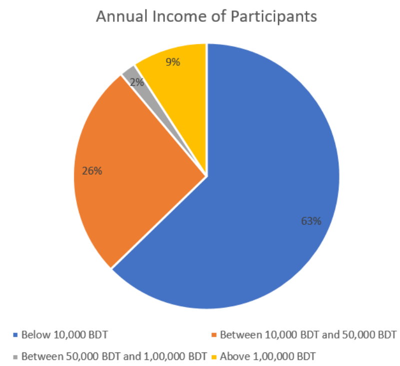


Figure 4.3: Annual income of Participants

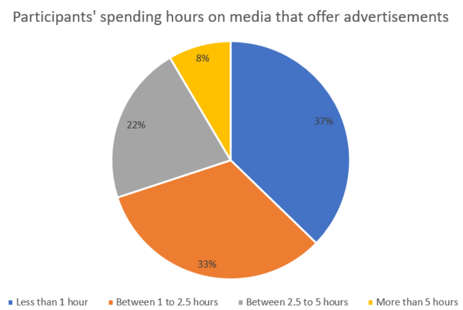


Figure 4.4: Participant's spending hours on media

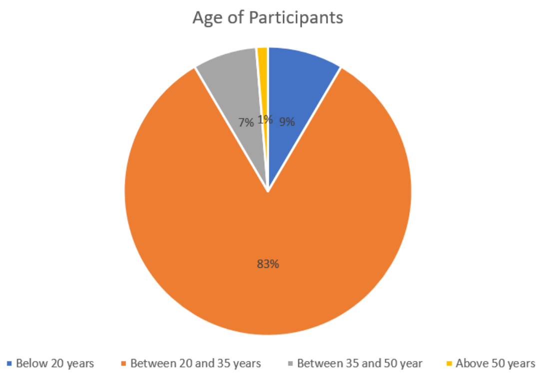


Figure 4.5: Age of Participants

The participants of various classes who engage in the survey are shown in the figures above. The survey included people of all ages as well as persons from various economic backgrounds. We can also see that the number of male and female participants is nearly equal. The duration of time individuals spend watching advertising is also shown in the pie charts and bar diagrams.

For example, the first pie chart shows the gender ratio of male-to-female, which is 4:6. The second one represents the participant's annual income, which shows that the majority of the participants have an annual income of less than 10,000 BDT and

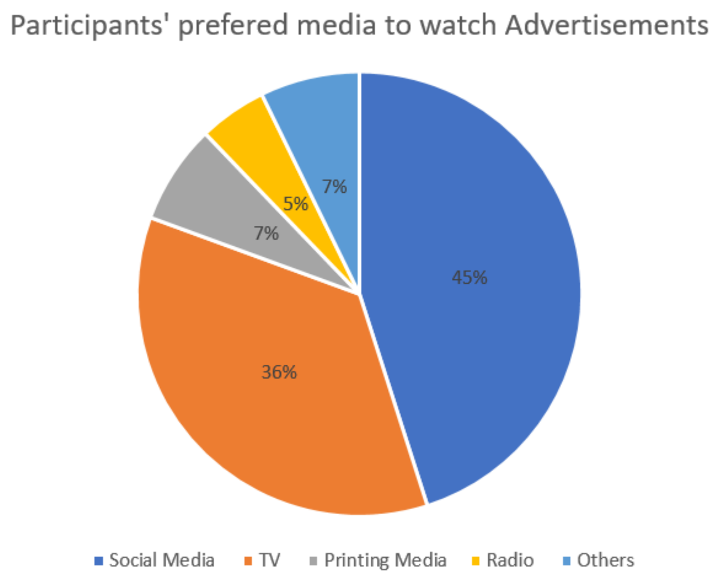


Figure 4.6: Participant's prefered media

they are between 20 and 35. The fourth pie chart illustrates participants' spending hours on media that shows adverts as well as the amount of time spent on them (more than 5hours). Lastly, we figure out that the majority of participants spend the majority of their time on social media and watching television.

4.1.2 Advertisement Information

The advertisement information is used to basically identify the characteristics necessary for the success of the advertisement. Here, the members needed to watch a commercial and they needed to respond to the inquiries. We select ads based on people's comments. For investigating the advertisement part, we group the inquiries into three scales. Dominance, Valance and Arousal. These are the pieces of Liker Scale. Each question bears any of the scales. Every one of the inquiries are replied in five unique sorts. Firmly Differ, Differ,Impartial, Consent, and firmly consent these are the five sections.. We investigate the valance, excitement and dominance as per these inquiries.

4.1.3 Texual Dataset Description

[15] There was no data set accessible for our research as it one of the absolute first studies in participants' reaction while watching FMCG[22]product's commercial area. So we needed to make this data set without any preparation. So we made a few fields to accumulate the data we look for. We thought about the inquiries that may emerge against what general individuals may feel while watching a commercial. FMCGs[22] are items that sell rapidly for their tolerably minimal price. FMCGs[22] have a shortage of time span of since of its high usability shopper interest (e.g., cold drinks, sweets) on the other hand on the grounds that they are transitory (e.g. meat, dairy items, prepared merchandise). These products are bought on a regular basis, consumed fast, and reviewed very low and are sold in large amounts. Essentially we

passed on an overview on the inquiries we made and made the dataset dependent on that.We took the name, sex, and pay reach to distinguish which standards of individuals are keen on which sort of promoting and items. We kept specific fields to remove the specific sentiments watchers have while watching a commercial. Questions were like "if the ad was visually appealing and engaging? "Have you seen this advertisement before?" "Based on the advertisement, how likely would you be the next time you need a product of this category" "if the duration of the ad was too long or too short?" "If the ad's message was understandable?" "if the ad's message was relevant?" are the issues to check the straightforwardness of the genuine message it was wanting to give.Also, the pleasure dependency, and excitement level on the SAM scale decides how the watchers felt sincerely while the advertisements played. The general objective of this was to discover the connection between watchers or shoppers and promotions and concentrate on what we need for our examination.

These are some broad information found from a review. We made an overview on 200 unique individuals. After the review we began to make a dataset. Considering our examination and our fundamental object of the exploration we construct the dataset. It contains fourteen significant highlights which are considered as the principal highlights of the ads. Mostly we attempted to tackle two issues. One is discovering the buy rate examining the components. Another is discovering the highlights of making a positive impression of notices among individuals. Despite the fact that every one of the notices should not have these components which lead them to turn into a fruitful notice. Impression is a factor that shows the positive and pessimistic perspectives on individuals to a promotion. In the event that individuals acquire an uplifting perspective from the promotion of the item, individuals will have a positive impression. This factor is vital while buying an item in the wake of reviewing the promotion. We may think about whether individuals will purchase a similar result of another organization by watching another notice or not. Promotion status implies the distinction among the ads of the same item. The correlation among the notices of the same item.

Positive allure of the items may assume a significant part to buy the item and furthermore may make a decent impression. Engaging a lot is the impression of the amount it makes in the human psyche for purchasing the item. Drawing in demonstrates the visual portrayal of the notice is equipped for getting public feeling or not.

Then again, in short narrative or social message type commercials at times make extensive promotions. In this way, the term of the notices is a significant factor in anticipating the achievement pace of notice. This sort of notice really is the synopsis of a huge however viable social message. Conveyed Message is the message which is conveyed by the promotion about the item. On the off chance that the item is addressed appropriately with a legitimate message, individuals will have their consideration towards the item. On the off chance that the message of the item isn't cleared to individuals to the clients, they won't be motivated to buy the item. Above all among every one of the variables, a conveyed Message is the most significant piece of a promotion.Persuaded implies individuals are prepared to purchase this item or not. On the off chance that they are not persuaded they won't accept the item. So the commercial ought to be convincing. In the event that individuals are persuaded, they will be persuaded to purchase the items. Along these lines, an ad should have to have a convincing force.

Preferred means individuals might want to watch such ads in the future or not. An ad as of now exists and it has numerous elements. On the off chance that individuals like the variables of the ad and in the event that they need to watch such sort of ad in future, the elements present in the commercial will be considered as significant variables. item. So the ad ought to be convincing. In the event that individuals are persuaded, they will be persuaded to purchase the items. In this way, a commercial should have to have a convincing force.Understandable means the message conveyed by the ad is great or not. So individuals won't get legitimate data from the commercial. The message ought to be clear and justifiable is the critical factor of an ad. The fundamental saying of a commercial is to mindful individuals about the notice. On the off chance that individuals won't comprehend the message, they won't accept the item.

Significant is how much relatable the commercial with the item. The publicized ment ought to be pertinent to the item. For instance, you are watching a promote of a portable yet it is seeing different things as opposed to versatile it won't be important. We plan to purchase any item just when we get a pertinent and legitimate message from the promotion of the item. men promote the messages to apply To the notice ought to be applicable to the item.Believable benefits mean the message conveyed in the ad is authentic or not. In the event that a commercial is seeing the benefit of utilizing any item more than the capacity of the item, it won't be a conceivable promotion. Appropriate ad has consistently shown the genuine advantages of an item. We are searching for that ad consistently where we track down the appropriate message with great visual portrayal. So this factor is additionally assisting with discovering to foresee the achievement pace of a promotion.Pleasure is the estimation of the notice is lovely or upsetting. It is a significant factor of a commercial. On the off chance that individuals may not track down any charming message from the promotion, the notice won't stand out enough to be noticed. Dependency Is the buying of the item is how much ward on the advertisement. Pretty much we purchase the items by watching the notice. Since this is the most widely recognized method of spreading the information of any item to individuals. Excitement is what amount energized individuals feel while watching the ad. Actually the fervor level relies upon the ad's visual show and the internal significance. More noteworthy shows will acquire wonderful individuals' consideration.As we referenced in our poll depiction we utilized liker scale for addressing our inquiries identified with promotion and SAM scale study. Firmly Differ,Differ,Impartial, Consent, and firmly consent these are the five sections. Thinking about the assessment, we make a data set where every one of the qualities were 1 to 5. Our preparation set for each data set were Purchased and Appealing. In each data set we dealt with both of these sections. This is one of our data set. Our data set is the place where every one of the qualities are assessed 1 to 5 without the Purchased section. Here the section Purchased is either 0 or 1 and the neutral value 3 is considered as 1. Here 1 is keen on buying and 0 is not interested. Once more, there is another data set we chipped away at the impartial worth of the Purchased section was considered as 0. The equivalent goes for the Appealing segment for once the nonpartisan appealing worth was 0 and some other time the unbiased worth was 1. We additionally work on a data set where every one of the qualities are 0 and 1. Where, 0 is a negative reaction and 1 is a positive reaction. For this situation we additionally thought to first the nonpartisan worth is 0 and again the unbiased worth 1. For this situation 0 is additionally considered as a bad introduction or negative buy aim and 1 is positive impression or buy expectation. We partition the dataset among these segments since we need to break down every single condition in regards to this dataset. That may assist us with getting amazing precision or best factors of notices.

4.1.4 Dataset Visualization

Stacked column Chart

A stacked section graph [30] is a standard Excel diagram that permits us to contrast leaves behind wholes over the long haul or across classes. In a stacked segment diagram, information arrangements are stacked in vertical sections one on top of the other.
Because total column lengths are easy to compare, SCS can demonstrate change after some time. However, with the exception of the first data series (close to the x-hub) and overall bar period. However, comparing the relative sizes of the components that make up each bar is difficult, with the exception of the initial sequence of data (close to the x-hub) and total bar period. SCS immediately becomes complex when more classifications or sequences of data are added. Here, We used color to illustrate data-value interactions that would be much more difficult to understand if displayed numerically in a graph.

Line Chart

Line charts [31] are useful for displaying patterns over time. A simple illustration would be how the stock value of a certain company increases over time on the stock market. It does not, however, have to be time on the x-axis. A function can be plotted using any data that is consistent with the variable on the X-axis. Time flow and shift rate are more important than quantity in a line chart. We want continuous data if we have it. A line diagram is a fantastic alternative if we have continuous data that we wish to convey through a diagram. This graph is very suitable for analyzing a trend or pattern in your data, such as seasonal effects or significant changes over time.

Scatter Plot

A scatter plot [16] represents values for two distinctive mathematical factors by using dots. Individual dot's location on the flat and vertical end addresses a worth for a unique piece of information The SP is used for taking a snap at the connection between factors. Scatter plots are mainly used to display and track relationships between two numeric values. In a scatter plot, the dots not only show where the data is coming from, but they also show where the data is coming from. When the data is evaluated as a whole, the dots in a scatter plot indicate not only the values of the data points, but also trends. We can classify data points into classes based on how close points cluster sets together. Scatter plots can also identify any unanticipated data gaps or outliers. The dots in a scatter plot reflect not only the values of data points, but also trends when the data is seen as a whole used.

Textual dataset where all the values of the features are classified 1 to 5:

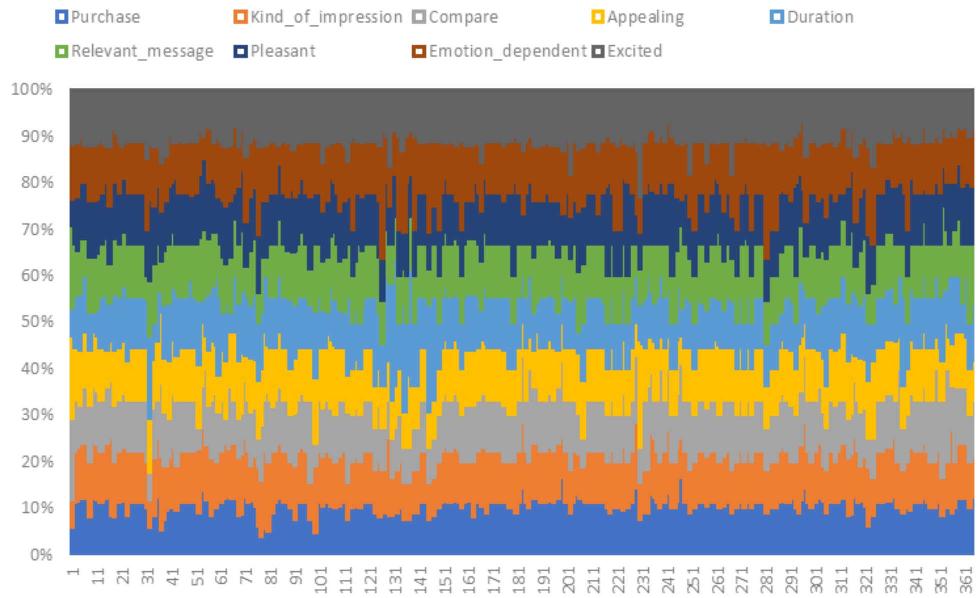


Figure 4.7: Stacked Column Chart for the features in Dataset

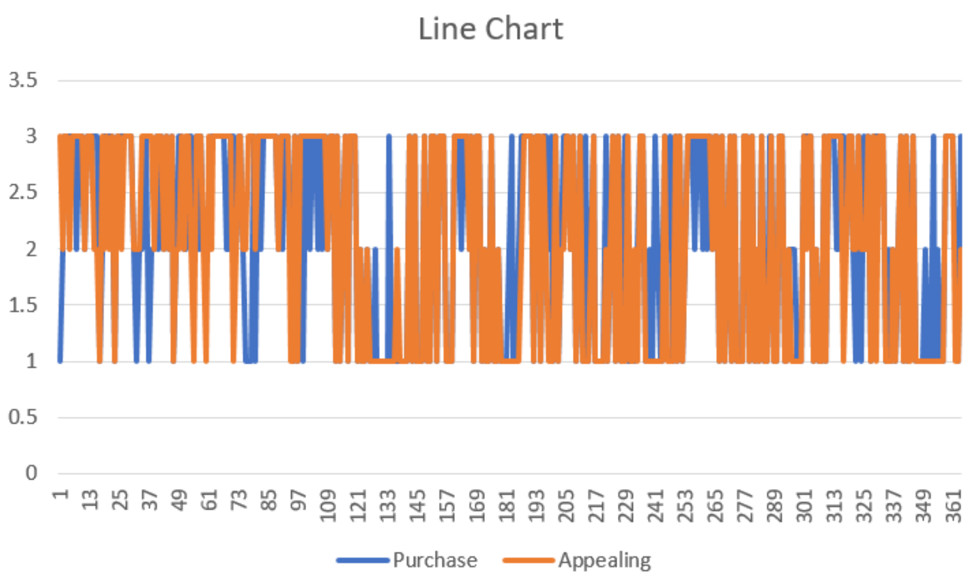


Figure 4.8: Line Chart for the fetures in Dataset

https://www.overleaf.com/project/60b6b96f167feaea86a086ac

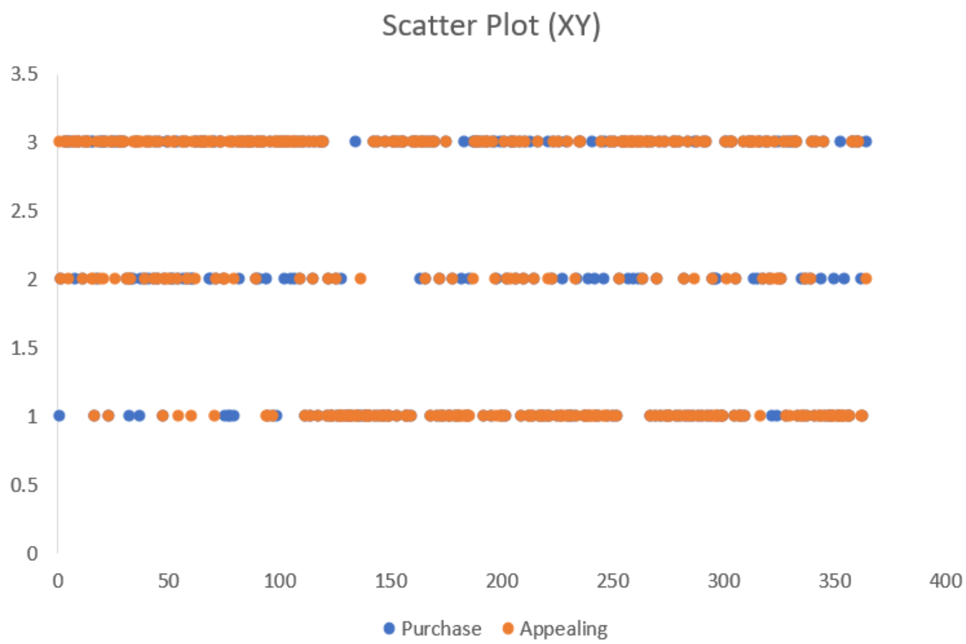


Figure 4.9: Scatter Plot (XY) for the features in Dataset

4.1.5 EEG Sensor Dataset

[18] The EEG[25] sensor Dataset Descriptionaset consists of sensor data values and participant feedback, such as impressions of advertisements and purchase rates based on advertisements.
As previously stated, participants were required to watch 10 videos and provide feedback. Evaluated from the above. There were 30 people that took part in this survey. Due to technical difficulties and internet disruptions, two respondents were unable to finish the survey. There were 4 females and 24 males among the 28 contestants.
We determine the mean, median, and entropy of each participant's advertising. We also work on the sensor's 28 channels. As a result, each channel provided 3 values: mean, median, and entropy. The "Appealing" and "Purchase" are then linked to the sensor channel values.

4.2 Data Preprocessing

At the beginning, for textual data, we ran a survey and collected data from people like consumers and advertising experts. We catagorized the users by their age, sex, annual income,their preferred media for advertising and so on. The experts are from different FMCG based companies and advertising agencies. We used 2 different dataset,primarily textual and secondarily brain signal data collected by EEG (Electroencephalogram) sensor. After collecting data, we cleaned the dataset and kept the key information from the survey reports by discarding all the errors, missing and irrelavent data. After that, we checked if the dataset is normalized by implementing Shapiro Wilk and found out that the dataset is normalized. The secondary dataset that we used for EEG signals was clean itself and ready to use.

4.3 Feature Extraction

For reducing the number of features, we applied Feature Extraction in our self reported dataset of advertisements by creating features from that already exist. Next we discarded the less important original features. We used Chi Square technique for feature extraction which illustrated the important factors that have an impact on advertisement success. The majority of the information contained in the previous set of features was summarized in this new reduced set of features.

4.4 Data Processing

We separated the data to predict Purchase Intent and Appealing or Engaging. In the questionnaire for each specific factor, the variability of the answers from the participants were 5, which is for example, very bad, bad, neutral, good, very good. Then we resampled (down-sampled) this data and reduced the variances into 3 in order to get better predictive performance. Then we scaled both the Textual and EEG dataset using standard scaler. We split the data for training and testing in 8:2 ratio.

4.5 Build more model and get output

We implemented machine learning and deep learning classifier models. To measure the impact of the factors we applied Logistic Regression, KNN, SVM, Naïve Bayes, Random Forest etc. Random forest performed comparatively a bit poor from others. So, we implemented another tree based algorithm which is XGBoost. It is a boosting algorithm which not only reduces variances but also biases. For deep learning we approached with Artificial Neural Network (ANN) but it did not work well on our dataset. The outputs were fluctuating too much. So, for better performance we used Multilayer Perceptron (MLP) which is a specific feed-forward neural network architecture. This is also known as the simplest model of deep neural network. As the outputs, we compared all the algorithms that we implemented based on accuracy and ROC curve.

4.6 Decission

In this way we attempted to determine the impact of different human emotions responsible for the purchase decision of any FMCG commodity while watching TV commercials. We also found out the impact of the same factors to make a TVC appealing and engaging. We also sorted out the important factors responsible for Purchase Intent and Engaging or Appealing in ascending order.

Chapter 5

Result Analysis and Discussion

Data that has been collected through google form to create the dataset there, values indicate 1 =Not at all likely, 2 = Not likely, 3 = Neutral, 4 = Likely, 5 = Very likely. Then these 1 to 5 values have been compressed to 3 values which indicate 0 = Not likely, 1= Neutral and 2 = Likely. After that, we again evaluated the values to 0,1 and then we chose value 1 for Neutral and did run the dataset again we chose 0 for Neutral and ran the dataset. Then we found out the accuracy without doing scaling and to increase the accuracy we eventually found out the values with scaling as well.

5.1 Textual Dataset Result

5.1.1 Logistic Regression(LR)

Logistic Regression processes [21] the likelihood of an occasional event. It's another case of direct relapse where the objective variable is a natural occurrence. It uses huge chances as the reliant value. Using a logic work, calculated regression predicts the likelihood of an event of a paired occasion.

Table 5.1: Accuracy from Logistic Regression (On the basis of Purchase)

Dataset	Without scaling	Standard scaling
Scale(0-2)	86%	80%
Neutral as 0	83%	83%
Neutral as 1	93%	96%

Using the data we got from Google form, we investigated the impact of people's purchases from the responses we received. The criteria for the question are kind of impression, compare, appealing , duration, relevant message, pleasant , emotion dependent and excited. Most of the responses received based on these questions are that people are likely to buy the product. According to the algorithm that has been applied to the data set, the algorithm also shows the highest accuracy (without and standard scaling) 96% (From table 5.1).
The information we got from Google Forms, we checked this information, and people were very interested in the response we got. The categories of inquiries are about purchase , kind of impression, compare, duration, relevant message, pleasant , emotion dependent and excited. Most of the responses received based on these questions

Table 5.2: Accuracy from Logistic Regression (On the basis of Appealing)

Dataset	Without scaling	Standard scaling
Scale(0-2)	83%	85%
Neutral as 0	86%	87%
Neutral as 1	91%	92%

are that people might buy the product. Judging from the algorithms that have been applied to the data set, the data set also shows the highest accuracy (without and standard scaling) 92% (from Table 5.2)

5.1.2 SVM(Support Vector Machine)

SVM [24] is equipped for order just as relapsed. SVM which is Non-linear indicates that the algorithm's determined border does not must be a linear line. The advantage is that we can catch undeniably more perplexing connections between our information focuses without doing any complicated modifications ourselves.

Table 5.3: Accuracy from SVM Algorithm (On the basis of Purchase)

Dataset	Without scaling	Standard scaling
Scale(0-2)	78%	83%
Neutral as 0	91%	87%
Neutral as 1	83%	91%

According to the algorithm that has been applied to the data set, from the algorithmwe get the highest accuracy 91% (From table 5.3)

Table 5.4: Accuracy from SVM Algorithm (On the basis of Appealing)

Dataset	Without scaling	Standard scaling
Scale(0-2)	82%	83%
Neutral as 0	83%	86%
Neutral as 1	92%	94%

Judging from the algorithms that have been applied to the dataset, the dataset also shows the highest accuracy 94% (From table 5.4)

5.1.3 K Nearest Neighbour(KNN)

The output of KNN [42] is a class membership, which can be utilized for categorization. A majority of an object's neighbors vote to categorize it, and with the object being classified in the most popular class among its nearest neighbours. It can also be used for regression, with the output being the object's value. This value represents the average (or median) of its nearest neighbors' values.
Based on the algorithms that have been applied to the dataset, the dataset shows the highest accuracy 93% (From table 5.5).
This dataset also has the maximum accuracy of 96%,(From table 5.4) according to the algorithms that have been applied to it.

Table 5.5: Accuracy from KNN Algorithm (On the basis of Purchase)

Dataset	Without scaling	Standard scaling
Scale(0-2)	81%	87%
Neutral as 0	90%	87%
Neutral as 1	93%	91%

Table 5.6: Accuracy from KNN Algorithm (On the basis of Appealing)

Dataset	Without scaling	Standard scaling
Scale(0-2)	87%	92%
Neutral as 0	88%	93%
Neutral as 1	96%	96%

5.1.4 Naive Bayes

For multi-class expectation issues, Naive Bayes [26] is a decent decision. In the event that the case of highlight autonomy stays valid, it can perform different models when utilizing undeniably less preparing information. All out input factors are more fit to Naive Bayes than mathematical info factors.

Table 5.7: Accuracy from NB Algorithm (On the basis of Purchase)

Dataset	Without scaling	Standard scaling
Scale(0-2)	84%	85%
Neutral as 0	89%	89%
Neutral as 1	90%	94%

According to the algorithms that have been applied to the dataset, the dataset shows the highest accuracy 94% (From table 5.7).

Table 5.8: Accuracy from NB Algorithm (On the basis of Appealing)

Dataset	Without scaling	Standard scaling
Scale(0-2)	81%	84%
Neutral as 0	91%	89%
Neutral as 1	89%	92%

Based on the algorithms that have been applied to the dataset, the dataset shows the highest accuracy 92% (From table 5.8).

5.1.5 Chi Square

To find the impact of the individual factors for 'purchase intent' and 'appealing and engaging' , Chi Square [33] was implemented. It helps to find out the significance of the factors and find the most important ones in an order.

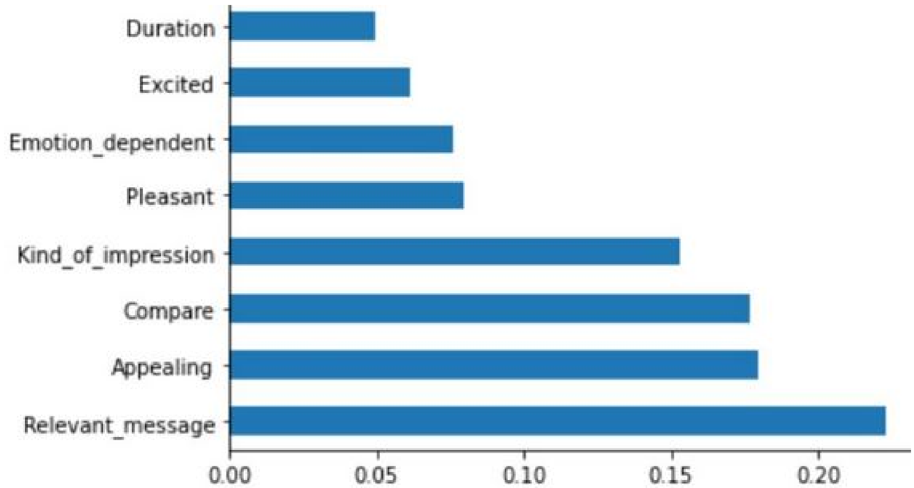


Figure 5.1: Accuracy from Chi Square Algorithm (On the basis of Purchase)

Here the bar chart illustrates that the most important feature to determine whether the audience would like to purchase a product or not is the 'Relevant message' that the advertisement provided. So relevance of the message that an ad delivers is the most important factor to convince a person to buy a product according to the research. After 'relevant message', other important factors are consecutively, 'Appealing', 'Compare', 'Kind of Impression', 'Pleasance', 'Emotion', 'Dependence',' Excited', 'Duration'.
In the same way, to make an ad 'Appealing and Engaging' the most important factor according to chi square are consecutively 'Relevant Message', 'Kind of Impression', 'Pleasant', 'Purchased', 'Compare', 'Excited', 'Duration', 'Emotion Dependent'.

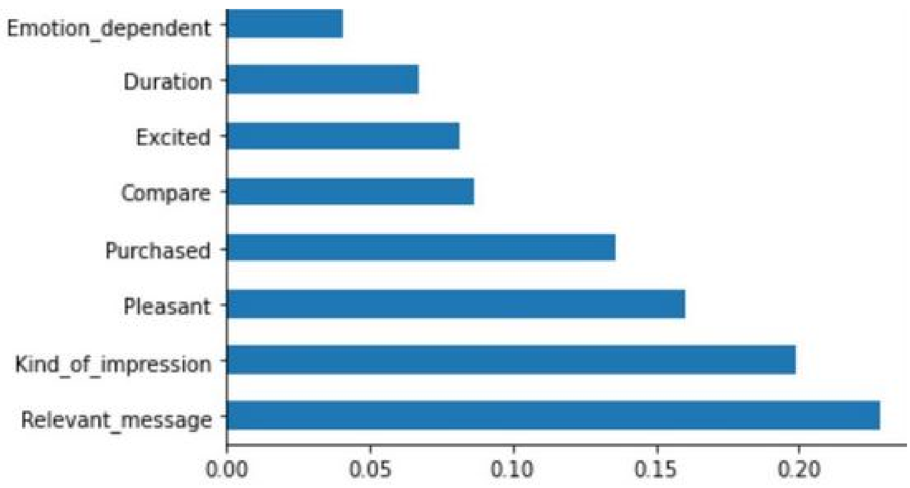


Figure 5.2: Accuracy from Chi Square Algorithm (On the basis of Appealing)

5.1.6 Random Forest(RF)

RF [29] adds an extra anomaly in the model, while fostering a tree. Instead of searching for the fundamental segment while separating a hub, it searches for the best element between an unpredictable subset of attributes. This results in a wide assortment that generally achieves a better model.
From the corresponding algorithms that have been applied to the dataset, the dataset shows the highest accuracy 80% (From table 5.9).

Table 5.9: Accuracy from RF Algorithm (On the basis of Purchase)

Dataset	Without scaling
Scale(0-2)	84%
Neutral as 0	87%
Neutral as 1	80%

Table 5.10: Accuracy from RF Algorithm (On the basis of Appealing)

Dataset	Without scaling
Scale(0-2)	82%
Neutral as 0	79%
Neutral as 1	84%

From the calculation that has been applied in the dataset that likewise shows the highest accuracy 84% (From table 5.10).

5.1.7 XGBoost

Since it was generated and constructed for the main objective of computational speed and model performance, For boosted trees algorithms, XGBoost [34] has proven to test the envelope of computer power. It was specifically designed for tree boosting algorithms, make the most of every bit of memory and hardware resource available. According to the algorithms that have been applied to the dataset, the dataset shows the highest accuracy 94% for Purchased (From table 5.11).

Table 5.11: Accuracy from XGBoost Algorithm(On the basis of Purchase)

Dataset	Without Scaling	F1 Score(Weighted Avg)
Scale(0-2)	89%	89%
Neutral as 0	90%	90%
Neutral as 1	94%	94%

Table 5.12: Accuracy from XGBoost Algorithm(On the basis of Appealing)

Dataset	Without Scaling	F1 Score(Weighted Avg)
Scale(0-2)	0.87	0.87
Neutral as 0	0.87	0.87
Neutral as 1	0.93	0.93

From the algorithm that has been applied to the dataset, the dataset shows the highest accuracy 93% (From table 5.12) for Appealing.

5.1.8 ANN

Neural networks usually require much more data than traditional machine learning algorithms, as in at least thousands if not millions of labeled samples. This isn't an easy problem to deal with and many machine learning problems can be solved well with less data if you use other algorithms. Although there are some cases where neural networks do well with little data, most of the time they don't.

In our case ANN [14] is not performing so well. The outputs are not stable as well as the accuracy of the model is fluctuating up and down. Our dataset consists of less than 400 tuples of data to feed the model. Because of the less amount of data ANN is not performing upto the mark. For that reason, we performed another deep learning technique, Multi layer perceptron (MLP) [39] which is a supplement of feed forward neural network.

5.1.9 MLP

MLP [39] deep learning method is commonly used to solve problems that require supervised learning. MLP uses a subjective number of covered up layers that are the genuine computational motor of the MLP. MLPs with one secret layer are fit for approximating any persistent capacity. With a small dataset, it works comparatively better than conventional ANN.

Table 5.13: Accuracy from MLP Algorithm(On the basis of Purchase)

Dataset	Without Scaling	With Scaling
Scale(0-2)	80%	81%
Neutral as 0	77%	80%
Neutral as 1	91%	91%

According to the algorithms that have been applied to the dataset, the dataset shows the highest accuracy 91% for Purchased (From table 5.13).

Table 5.14: Accuracy from MLP Algorithm(On the basis of Appealing)

Dataset	Without Scaling	With Scaling
Scale(0-2)	81%	84%
Neutral as 0	78%	87%
Neutral as 1	88%	93%

According to the algorithms that have been used to the dataset, the dataset shows the highest accuracy 93% for Appealing (From table 5.14).

5.1.10 ROC Curve

To measure the comparative performance among machine learning and deep learning methods on a particular dataset ROC (Receiver Operating Characteristics) curve and AUC (Area under the Curve) are term that we rely on. AUC-ROC, also known as AUROC (the receiver area's operating characteristics), are used to test or simulate the quality of a multi-classification task. The AUC-ROC curve represents an estimate of the measurement's efficiency at various thresholds.

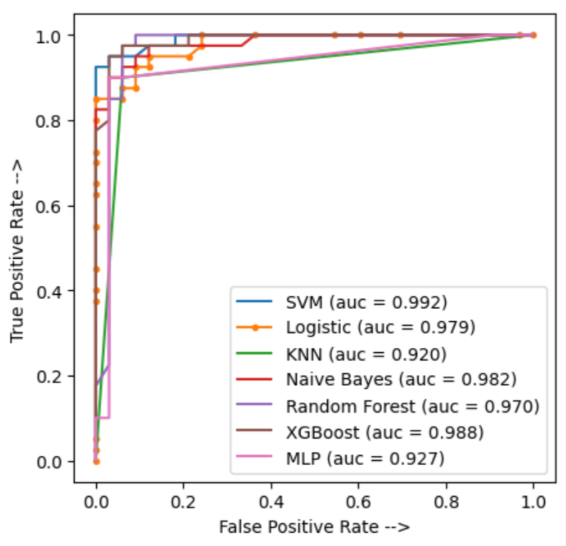


Figure 5.3: ROC curve of Purchase Intent

All the algorithms that we applied performed more or less upto the mark on our data. To justify our claim AUC-ROC is applied. In the ROC Curve above, all of the machine learning and deep learning algorithms that used are showing a comparison between their performances. The ROC curve that we presented, means true positive rate by it's y-axis and false positive rate by it's x-axis. All of the models that we ran for this research shows good AUC area, for purchase intent ranging from 92.7% (MLP) to 99.2% (SVM) and for advertise appeal and engagement ranging from 88.3% (MLP) to 94.4% (SVM), which proves that the features that we predicted as important for advertisements success are valid and gives good percentage of true positive results for both of purchase intent and advertisement engagement.

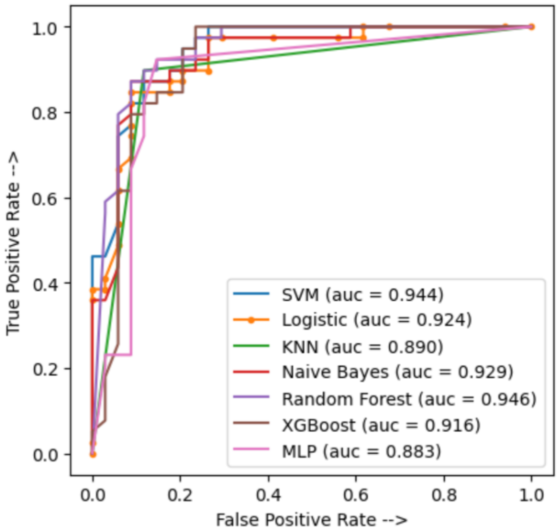


Figure 5.4: ROC curve of Appeal and Engagement of TVC

5.2 EEG Signal

Table 5.15: Accuracy of EEG Signal(On the basis of Purchase)

Dataset	Accuracy for Purchase
SVM	51%
KNN	50%
XG Boost	54%
MLP	46%

The inquiries' standards[25] were centered around kind of impression, compare, appealing, duration, relevant message, pleasant, emotion dependent and excited. Individuals are well on the way to purchase the item, contingent upon the most widely recognized reaction dependent on the notices they have watched. From SVM, KNN, XG Boost, MLP algorithms that have been used to apply in the dataset that shows the accuracy is from 51% to 46% (from table 5.15).

Table 5.16: Accuracy of EEG Signal(On the basis of Impression)

Dataset	Accuracy for Impression
SVM	64%
KNN	66%
XG Boost	71%
MLP	67%

Purchase, compare, appealing, duration, relevant message, pleasant, emotion dependent, and exciting were some of the factors used in the surveys. People are well on their way to purchasing the item, based on the most frequently known reaction based on the advertisements they've seen. SVM, KNN, XGBoost, and MLP algorithms were implemented in the dataset that shows the accuracy from 64% to 67% (from table 5.16).

Chapter 6

Conclusion

6.1 Conclusion

The major goal of our research was to predict a TVC's success rate by studying its factors and gather information about human emotional response on an advertisement. For detecting advertising variables, we used machine learning and deep learning algorithms. In our research, firstly on textual data, we have worked on FMCG [40] based TVCs and predicted success factors of adverts for a customer to get convinced to purchase a product [36]. Also, we predicted the factors of a TVC to make it appealing and engaging. On EEG data we analyzed the same factors to find out purchase intent, appealing and engaging based on our interviewees brain response. In [25], the research target was quite similar to ours where different machine learning algorithms were applied and the dataset was based on the food and beverages industry. In our approach we applied deep learning methods successfully alongside ML algorithms [17] and found better results in terms of confusion matrix and overall accuracy. Improvements on some subjective measurements such as the amount of data, quality of textual responses and some objective measurements like improving noise factors while collecting EEG data [25] would allow us to go further with this work. The goal of our study was to use the factors to predict the purchasing behavior and measure the engagement power of an advertisement. We believe that this will be useful to advertising agencies and relevant organizations to be more specific when creating commercials and give them proper ideas about where to emphasize and vice versa. We implemented our method to observe the most crucial factors for a TVC and, as a result, showed the relationship between important factors of advertisements and the viewers' thoughts on purchasing that product. Our research aimed to find out the crucial factors of a successful TV commercial that makes a product appealing and engaging to the viewers and influence their purchase intent. Successful outcome from our research will aid the advertisement industry to be more successful than any time before.

6.2 Future Work

After this thesis we would like to take our work a lot further. Because of Covid 19 situation we could not visit the university and take EEG data [25] by ourselves. We have plans to collect accurate EEG data by ourselves on FMCG [40] industry TVCs. Removing the drawbacks that we had in our subjective and objective measurements will make the research outcome better. So we aim to use a bigger dataset in the future to get better results while applying deep learning techniques. Our thesis focused on human emotion-related brain activity. However, we'd like to work on a human's facial expression while watching a commercial. For the purpose of analysis, we shall compare the facial expression to the message supplied in commercials. We will apply the CNN algorithm [27] to improve accuracy in this case. These will aid us in creating more resource efficient adverts in the future.

Bibliography

[1] O. Rafal, R. Dorota, W. Dawid, and C. Anna, "Analysis of neurophysiological reactions to advertising stimuli by means of eeg and galvanic skin response measures," 2009.

[2] D. Hill, "Emotionomics: Leveraging emotions for business success," 2010.

[3] O. Sourina and M. R. Hafiyyandi, "Eeg-based emotion-adaptive advertising," 2013.

[4] R. D. O. J. dos Santos, J. H. C. Oliveira, J. B. Rocha, and J. M. E. Giraldi, "Eye tracking in neuromarketing: A research agenda for marketing studies," vol., no., 2015, ISSN:

[5] M. G. R. Alam, A. K. M. Masum, L.-S. Beh, and C. S. Hong, "Critical factors influencing decision to adopt human resource information system (hris) in hospitals," Aug. 2016.

[6] J. Berčík, E. Horska, J. Gálová, and E. S. Margianti, "Consumer neuroscience in practice: The impact of store atmosphere on consumer behavior," Jun. 2016.

[7] H. Kumar, "Neuromarketing: A new agenda for marketing researches with particular reference to eye tracking," May 2016.

[8] H. Gauba, P. Kumara, P. P. Roy, P. Singh, D. P. Dogra, Balasubramanian, and Raman, "Prediction of advertisement preference by fusing eeg response and sentiment analysis," Aug. 2017.

[9] N. Zhuang, Y. Zeng, L. Tong, C. Zhang, H. Zhang, and B. Yan, "Emotion recognition from eeg signals using multidimensional information in emd domain," Aug. 2017.

[10] R. Govindasamy, S. Arumugam, J. Zhuang, K. M. Kelley, and I. Vellangany, "Cluster analysis of wine market segmentation-a consumer based study in the mid-atlantic usa, economic affairs," 2018.

[11] M. G. R. Alam, S. F. Abedin, S. I. Moon, A. Talukder, and C. S. Hong, "Healthcare iot-based affective state mining using a deep convolutional neural network," vol., pp. 75 189–75 202, May 2019, ISSN: 18749212. DOI: 10.1109/ACCESS.2019.2919995. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/8726298.

[12] R. Pozharliev, W. J. Verbeke, and R. P. Bagozzi, "Merely being with you increases my attention to luxury products: Using eeg to understand consumers' emotional experience with luxury branded products," Mar. 2019.

[13] A. S. andShruti Shriya Gullapuram, H. Katti, M. Kankanhalli, S. Winkler, and R. Subramanian, "Recognition of advertisement emotions with application to computational advertising," Jan. 2020.

[14] S. Agatonovic-Kustrin, “Basic concepts of artificial neural network (ann) modeling and its application in pharmaceutical research,”

[15] A. Alshamsi, R. Bayari, and S. Salloum, “Sentiment analysis in english text.”

[16] R. A. Becker and W. S. Cleveland, “Brushing scatterplots,”

[17] G. Bonaccorso, “Machine learning algorithms,”

[18] H. Cho1, M. Ahn2, S. Ahn1, M. Kwon1, and S. C. Jun1, “Eeg datasets for motor imagery brain computer interface.”

[19] S. Day, “The 9 key advantages to video advertising,”

[20] D. Hill, “Emotionomics (2nd ed.).”

[21] D. W. Hosmer and J. S. L. N. R. X. Sturdivant, “Applied logistic regression,”

[22] M. I. Hossain1, M. R. H. Polas2, M. M. Rahman3, T. Islam4, and Y. Jamadar5, “An exploration of covid-19 pandemic and its consequences on fmcg industry in bangladesh.”

[23] S. Jangid, “Global digital ad spending 2019 – reports, statistics market trends,”

[24] T. Joachims, “Svmlight$_{s}upport_{v}ector_{m}achine$,”

[25] R. Kannana, S. S. A. Alia, A. Faraha, S. H. Adilb, and A. Khan, “Smart wearable eeg sensor,” pp. 17–20.

[26] M. Karim and R. M. Rahman, “Decision tree and naïve bayes algorithm for classification and generation of actionable knowledge for direct marketing,”

[27] T. Kozek, T. Roska, and L. Chua, “Genetic algorithm for cnn template learning,”

[28] A. D. Kulkarni and B. Lowe, “Random forest algorithm for land cover classification,”

[29] W. Lin, Z. Wu, L. Lin, A. Wen, and J. Li, “An ensemble random forest algorithm for insurance big data analysis,”

[30] C. Long, “Stacked column chart,”

[31] R. M., K. Jr.1, D. E. Kostewicz2, K. M. Brennan2, and S. A. King, “Critical review of line graphs in behavior analytic journals.”

[32] Y. Ma and G. Guo, “Support vector machines applications,”

[33] M. L. McHugh, “The chi-square test of independence,”

[34] R. Mitchell and E. Frank, “Accelerating the xgboost algorithm using gpu computing,”

[35] A. Moors, J. D. Houwer, D. Hermans, S. Wanmaker, K. van Schie, A.-L. V. Harmelen, M. D. Schryver, J. D. Winne, and M. Brysbaert, “Norms of valence, arousal, dominance, and age of acquisition for 4,300 dutch words,”

[36] D. G. Morrison, “Purchase intentions and purchase behavior,”

[37] M. Prasad, “A look at the full revenue impact of retail media ad platforms,”

[38] B. Reuderink, C. Mühl, and M. Poel, “Valence, arousal and dominance in the eeg during game play,”

[39] H. Taud and J. Mas, “Multilayer perceptron (mlp),”

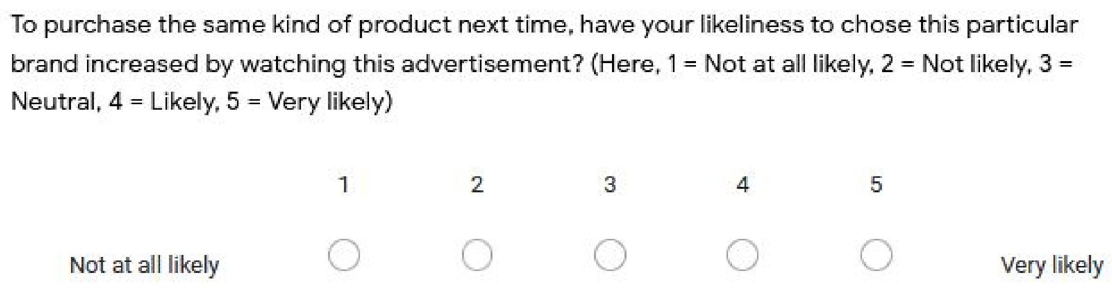
[40] D. S. S. Vernekar and P. Wadhwa, “Green consumption an empirical study of consumers attitudes and perception regarding eco-friendly fmcg products, with special reference to delhi and ncr region abstract,”

[41] A. B. Warriner, V. Kuperman, and M. Brysbaert, “Norms of valence, arousal, and dominance for 13,915 english lemmas,”

[42] Z. Zhang, “Introduction to machine learning: K-nearest neighbors,”

Appendix A

We tried to conduct a survey in order to create our dataset. We have a tendency to gather responses from a wide range of people. They must answer questions on two distinct datasets. From an EEG sensor [25], a textual dataset and an emotion dataset were generated. These responses enable us in evaluating the exact components of advertisements and their correlations between advertisement impression and product. purchase rate after seeing them.

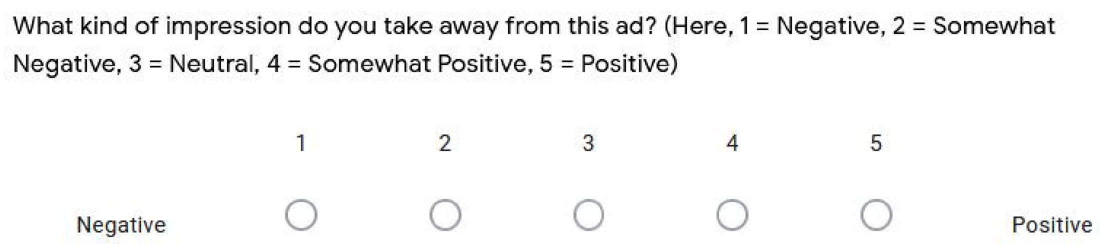

To purchase the same kind of product next time, have your likeliness to chose this particular brand increased by watching this advertisement? (Here, 1 = Not at all likely, 2 = Not likely, 3 = Neutral, 4 = Likely, 5 = Very likely)

1 2 3 4 5

Not at all likely ○ ○ ○ ○ ○ Very likely

Figure 6.1: Convinced by Ad

This figure shows a question where participants were asked whether their likeness increased by watching advertisements. This is the primary question because if the participant had a positive impression about the advertisement, they thought about purchasing the product again. They also have a positive impression about the product also.


What kind of impression do you take away from this ad? (Here, 1 = Negative, 2 = Somewhat Negative, 3 = Neutral, 4 = Somewhat Positive, 5 = Positive)

1 2 3 4 5

Negative ○ ○ ○ ○ ○ Positive

Figure 6.2: Question Regarding Impression

This figure shows In this question, participants were asked about the type of impression they had on the ad. This is the main problem, because if the participants have a positive impression of the ad, they considered buying the product. They also have a positive impression of the product.

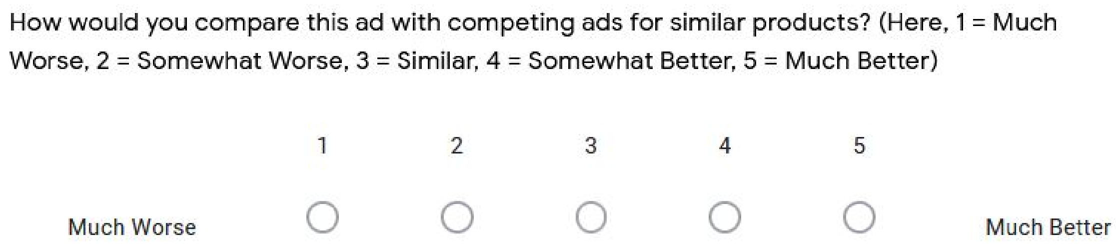


Figure 6.3: Comparison With Similar Ads

This figure shows a comparison of similar product advertisements. This helped us understand the extent to which people feel positive about the product. This article discusses the advertising differences of similar products.

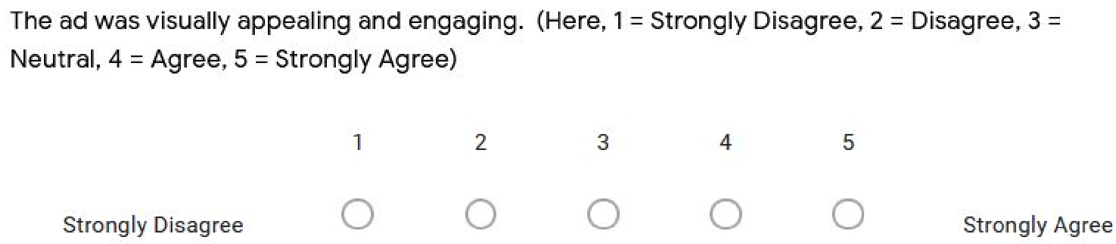


Figure 6.4: Appealing and Engaging

This figure depicts the question about whether the advertisement is convincing or not. The advertisement's inner meaning is important in this case. The advertisement's message about whether or not certain variables are important to participants. We posed this question to see if they were more concerned with the message or the visual representation.

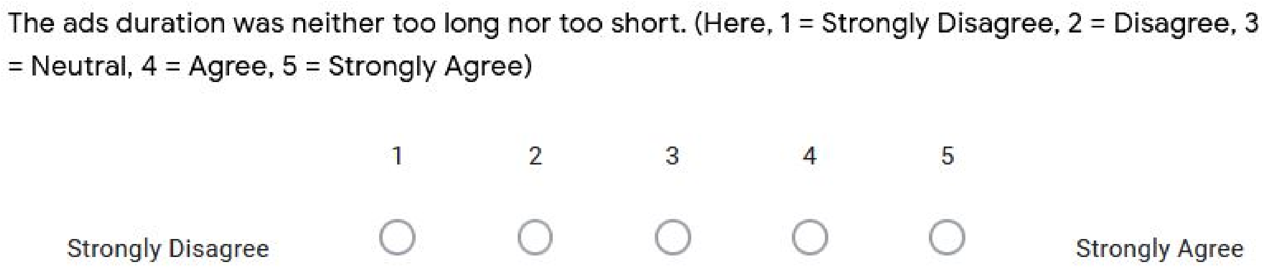


Figure 6.5: Duration of Ad

This figure shows the question about what was the pleasure level of the participants after watching the advertisement. Pleasure level helps to understand where the participant is convinced to purchase the product or not.
This figure depicts the question about whether the advertisement is convincing or not.The advertisement's inner meaning is important in this case. The advertisement's message about whether or not certain variables are important to participants. We posed this question to see if they were more concerned with the message or the visual representation.

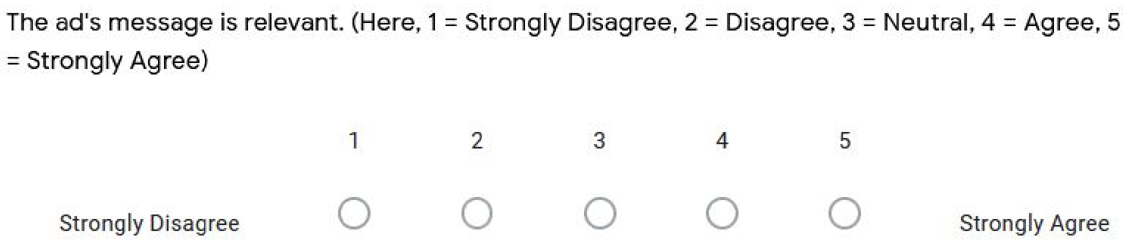


Figure 6.6: Conveyed Message

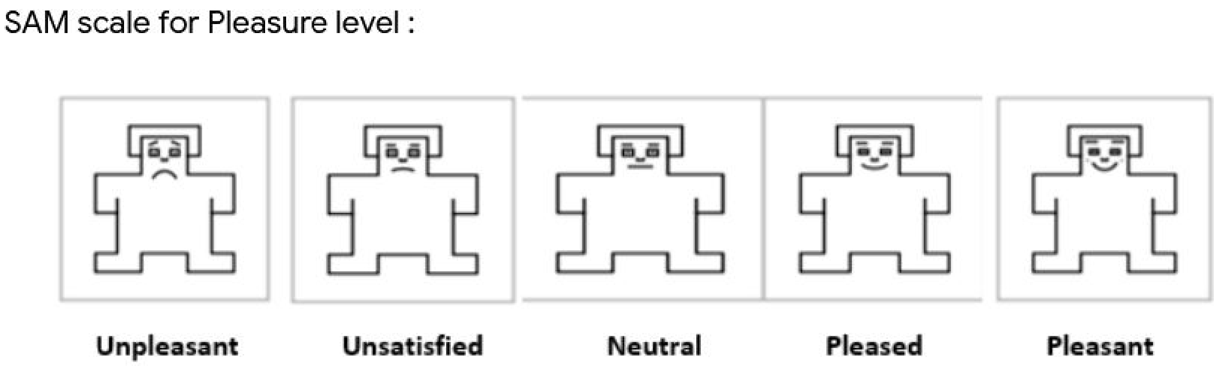


Figure 6.7: Pleasure Level

Figure shows the question about what was the pleasure level of the participants after watching the advertisement. Pleasure level helps to understand where the participant is convinced to purchase the product or not.

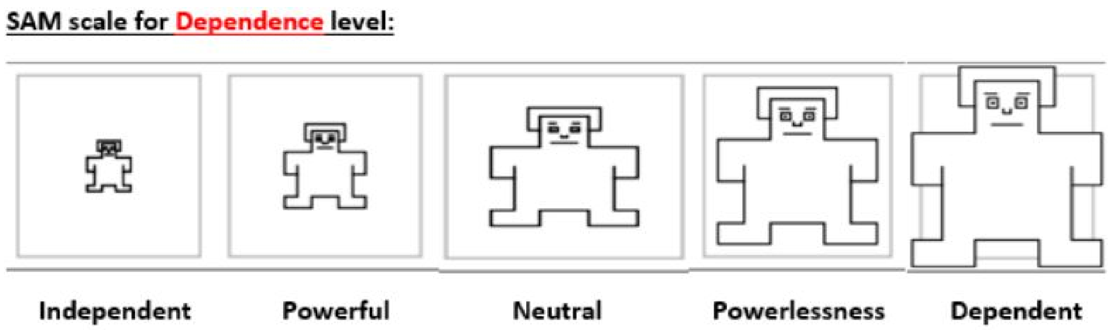


Figure 6.8: Dependence Level

This figure shows the question about what was the dependence level of the participants after watching the advertisement. Dependence level helps to understand where the participant is dependent on their emotion or not.

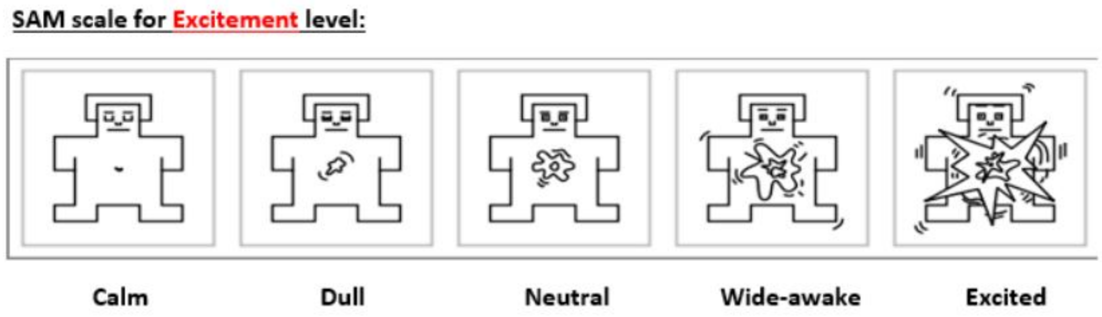


Figure 6.9: Excitement Level

This figure shows the question about whether participants were excited or not after watching the advertisement. Excitement level helps to understand where the participant is excited enough to buy the product or not.