**ARTIFICIAL INTELIGENCE BASED BUSINESS STRATEGY FOR OPTIMIZED ADVERTISING**

**2021-152**

**A picture containing text, clipart, vector graphics

Description automatically generated**

**Individual Final Project Thesis**

**Kannangara K K A L**

**B.Sc. (Hons) Degree in Information Technology**

**Specialization: Information Technology**

**Department of Information Technology**

**Sri Lanka Institute of Information Technology**

**Submitted on the 13th of October 2021**

October 2021

I certify that I have read this thesis and that it is entirely acceptable, in content and quality, as a Bachelor of Science thesis.

Ms . Dinuka Wijendra

|  |  |  |
| --- | --- | --- |
| **IT Number** | **Name** | **Signature** |
| IT18507652 | Kannangara K K A L | Diagram  Description automatically generated with low confidence |

Approved for Research Project:

Research Project Coordinator, SLIIT

Ms. Jenny Kishara

# DECLARATION OF ORIGINALITY

I declare that this is our work, and that this proposal does not contain any material previously submitted for a degree or diploma at any other university or institute of higher learning without acknowledgment, and that it does not contain any material previously published or written by another person, except where acknowledgement is made in the text. In addition, I hereby give to Sri Lanka Institute of Information Technology the nonexclusive right to reproduce and distribute my dissertation in whole or in part in print, electronic, or other media, in whole or in part. I maintain the right to incorporate this information into future works in whole or in part (such as articles or books).

Signed

Date

# ABSTRACT

Advertising is the mode of communication between the consumers and the products or services. With the rapid development of technology advertising campaigns carried out through digital media like social media , YouTube and ecommerce have been improved in a way which caters the user according to their preference. Television advertisements which also takes a prominent role in that domain sticks around with outdated commercialization mechanisms which does not consider the user’s preferences. It may be lead the users special younger generations to unrelated advertisements. As a solution for the above mentioned facts , a user based advertising approach is suggested. The system compromises machine learning and deep learning techniques to analyze demographics of a person which are important in classifying advertisements. The concept prototype is evaluated against a sample video consisting of 4 advertisements ,One of the main concern is that the study is completely based on Sri Lankan context. This will result in more relevant commercials for customers, making television ads more user-friendly and increasing the advertising agency's sales conversion.

# ACKNOWLEDGEMENT

The history of all great achievements shows that no great attempt has ever been achieved without the active or passive support of one's immediate surroundings. As a consequence, I'd want to express my gratitude to the project supervisor, Ms. Dinuka Wijendra, and the co-supervisor, Ms. Jenny Kishara for the  direction and support offered throughout the study period, as well as the happy attitude, resulted in the research being completed in a pleasant and stress-free working atmosphere. As a result, without his essential support, expert guidance, and encouragement, the effort would not have been a success.

I'd like to express my gratitude to our course coordinator, Mr. Jagath Wickramarathne, for his invaluable assistance in providing course materials and for sharing his vast knowledge of the module. Finally, I'd want to express my gratitude to my parents and group members for guiding me and supporting me in solving problems I faced when developing the system, as well as my SLIIT colleagues who supported me often when working on the assignments and other aspects of the project. Of course, my deepest appreciation goes out to everyone who helped in any way to the success of this study, even if their names aren't included above.

# DEDICATION

I would want to dedicate this research to my University, Parents, and friends, who have been an everlasting source of inspiration to me. They have given us courage, spirit, and direction.

My heartfelt gratitude also goes to everyone who assisted in various ways; without their love and support, this project would not have been possible.

# Table of Contents

[DECLARATION OF ORIGINALITY 2](#_Toc85061278)

[ABSTRACT 3](#_Toc85061279)

[ACKNOWLEDGEMENT 4](#_Toc85061280)

[DEDICATION 5](#_Toc85061281)

[Table of Contents 6](#_Toc85061282)

[LIST OF FIGURES 7](#_Toc85061283)

[LIST OF TABLES 7](#_Toc85061284)

[1 INTRODUCTION 8](#_Toc85061285)

[1.1 Background and Literature Survey 10](#_Toc85061286)

[1.2 Research Gap 14](#_Toc85061287)

[1.3 Research Problem 15](#_Toc85061288)

[1.4 Research Objectives 16](#_Toc85061289)

[1.4.1 Main objective 16](#_Toc85061290)

[1.4.2 Specific objectives 16](#_Toc85061291)

[2 METHODOLOGY 17](#_Toc85061292)

[2.1 Methodology 17](#_Toc85061293)

[2.1.1 Requirement analysis process 17](#_Toc85061294)

[2.1.2 Approach of the study 19](#_Toc85061295)

[Training phase for age and gender prediction 25](#_Toc85061296)

[Testing phase age and gender prediction 26](#_Toc85061297)

[2.1.4 Implementation of models 27](#_Toc85061298)

[2.1 Commercialization aspects of the product 34](#_Toc85061299)

[2.2.1 Introduction to the market 34](#_Toc85061300)

[2.2 Testing of the product 36](#_Toc85061301)

[3 RESULTS AND DISCUSSION 37](#_Toc85061302)

[3.1 Results 37](#_Toc85061303)

[3.1.1 Accuracy of models 38](#_Toc85061304)

[3.1.2 Research findings 42](#_Toc85061305)

[3.2 Discussion 47](#_Toc85061306)

[4 CONCLUTION 48](#_Toc85061307)

[References 49](#_Toc85061308)

[APENDIX 50](#_Toc85061309)

[Similarity Index 50](#_Toc85061310)

# LIST OF FIGURES

[Fig . 1 User response 9](https://mysliit.sharepoint.com/sites/CDAPSubmissionCloud/2021RegCloud/2021-152-Students/6.%20Final%20Report%20%26%20Presentation/Final%20Reports/IT18507652_Final_Report.docx#_Toc85061311)

[Fig . 2 Rate for user preferred advertisements 10](https://mysliit.sharepoint.com/sites/CDAPSubmissionCloud/2021RegCloud/2021-152-Students/6.%20Final%20Report%20%26%20Presentation/Final%20Reports/IT18507652_Final_Report.docx#_Toc85061312)

[Fig . 3 Data Collection Interface 19](#_Toc85061313)

[Fig . 4 Hyperplane Plot 20](https://mysliit.sharepoint.com/sites/CDAPSubmissionCloud/2021RegCloud/2021-152-Students/6.%20Final%20Report%20%26%20Presentation/Final%20Reports/IT18507652_Final_Report.docx#_Toc85061314)

[Fig . 5 Prediction process of advertisement - Random Forest 21](https://mysliit.sharepoint.com/sites/CDAPSubmissionCloud/2021RegCloud/2021-152-Students/6.%20Final%20Report%20%26%20Presentation/Final%20Reports/IT18507652_Final_Report.docx#_Toc85061315)

[Fig . 6 Deep learning model architecture 22](#_Toc85061316)

[Fig . 7 Flow of implementation 23](https://mysliit.sharepoint.com/sites/CDAPSubmissionCloud/2021RegCloud/2021-152-Students/6.%20Final%20Report%20%26%20Presentation/Final%20Reports/IT18507652_Final_Report.docx#_Toc85061317)

[Fig . 8 System Overview Diagram 24](#_Toc85061318)

[Fig . 9 Data set training (code base) 25](#_Toc85061319)

[Fig . 10 Preparing the data 28](#_Toc85061320)

[Fig . 11 Iteration through the dataset 28](#_Toc85061321)

[Fig . 12 Locating the favourable ad 29](#_Toc85061322)

[Fig . 13 Random Forest Classifier 30](#_Toc85061323)

[Fig . 14 Support Vector Machine Classifier 30](#_Toc85061324)

[Fig . 15 Initializing train ,test and validation data 31](#_Toc85061325)

[Fig . 16 Deep Learning Model 32](#_Toc85061326)

[Fig . 17 Implementation of result plotting 33](#_Toc85061327)

[Fig . 18 Facebook Page Overview 35](#_Toc85061328)

[Fig . 19 Concept Impressions 36](#_Toc85061329)

[Fig . 20 Correlation between demographic features and most favorable advertisement 38](#_Toc85061330)

[Fig . 21 Detailed report of SVM Classification 39](#_Toc85061331)

[Fig . 22 Deep learning model training 40](#_Toc85061332)

[Fig . 23 Results of Deep learning classifier 41](#_Toc85061333)

[Fig . 24 Accuracy curve 41](#_Toc85061334)

[Fig . 25 Loss curve 42](#_Toc85061335)

[Fig . 26 Advertisement and movie part analysis for age group 5 to 19 42](#_Toc85061336)

[Fig . 27 Advertisement and movie part analysis for age group 19 to 35 43](#_Toc85061337)

[Fig . 28 Advertisement and movie part analysis for age group 35 to 55 44](#_Toc85061338)

[Fig . 29 Mean score of the preferences of age groups 45](#_Toc85061339)

[Fig . 30 Mean score for attention toward the video of age groups 46](#_Toc85061340)

[Fig . 31 Similarity Index of the report 50](#_Toc85061341)

# LIST OF TABLES

[Table 1 :Research gap 14](#_Toc85026964)

[Table 2: Test Cases 33](#_Toc85026965)

# INTRODUCTION

In today's digital age, the majority of consumers around the world choose a service or a product based on advertisements. This is the same as for Sri Lanka also. Advertising is one of the most important things for a company since it tends to form the first impression of the consumer towards any product or service. It is also a method of building the brand in order to enhance sales or improve awareness of products or services among the target audience. There are various ways of advertising digitally such as Social Media ,E-commerce platforms and Television.

Social media which one of the most used advertising medias , is vastly spread over the globe. It drives in a lot of digital engagement of the mid age consumers toward the product and services advertised. The number of likes, shares , comments , views , follows and click are some of the factors considered when measuring the engagement of social media advertisements. E-commerce platforms also follows the same mechanism. Even though social media and E-commerce platforms have improved to advertise more effectively , television advertising which reaches a wide audience is still following tradition mechanisms to cater the audience. The television advertising is based on a timely basis which displays advertisements to all the people watching television from an advertisement database (classification of advertisements based on the time of the day).

Unlike social media and E-commerce platforms , television cannot consider likes, shares , comments for advertisement classification. But user demographics such as age and gender is very important to classify advertisements. The decisions/selections made by a person to consume a specific brand or product is due to the likes and dislikes of the person considering the factors age and gender . With the help of artificial intelligence mechanisms , if evaluation of the user preferred advertisements can be displayed in the television , it will be advantageous for the consumers as well as the product owner in generating sales. There is no such system which caters the Sri Lankan context to classify and evaluate the user preferred advertisements considering user demographics.

The survey we collected clearly states that most of the users even prefer watching advertisements if the advertisement really maches their likes and dislikes.

Chart, pie chart

Description automatically generated

Fig . User response

Teams

Description automatically generated

Fig . Rate for user preferred advertisements

In this research I have used existing age and gender prediction modal to the gather the relevant data for the analysis. Followed by a supervised learning Machine Learning and Deep Learning algorithms to train the gathered demographics against advertisements. Furthermore as most people watch television in groups, the analysis is taken forward to not only individual users but also peer group classification. Deep learning algorithms are also used to build this Proof of Concept(POC) to analyze the data collected .

## Background and Literature Survey

In the modern world there are several research papers and some implementations done to detect human objects detection and age gender prediction. Although there are many researches in that domain , there I no such research or analysis conducted to evaluate how affective age ,gender and peer groups are to classify user preferred advertisements. Below I have described some of researches conducted on television advertising , human object detection and age gender prediction, emotional analysis and facial features. Pragya Keshari and Sangeeth Jain have been conducted a research as “Effect of Age and Gender on Consumer Response to Advertising Appeals”[1].Although many researches have been conducted to evaluate the effectiveness of advertising in number of factors like presentation style, attractiveness of the endorsers ,creative appeal, etc. These two researcher’s have examined the effect of demographic factors on consumer response to advertising appeals. Their attempt was to study the responses given by the consumers to rational and emotional advertising appeals on the factors age and gender. They have considered a sample size of 348 consumers of the cities Indore and Ujjain. The data has been collected through a self-designed instrument and the data was analyzed using the methods paired-sample t-test and Multivariate Analysis of Variance (MANOVA).

Some researches of Sri Lanka Institute of Information Technology has conducted a research “Ads-In Site: Location based Advertising Framework with Social Network Analyzer”[2]. This research provides an application which enables user preferred advertisements by considering their social profile. They have build a framework which also considers the user demographics when providing the suitable advertisement for the user by tracking the location of the shops located in a 5km radius of the home location of the user. The social profiles are used to categorize the advertisement according to the likes and dislikes of the user. They have concluded that there system is able to deliver the most likely purchasing product advertisement to the potential customer who steps into the shop.

Sen Zhang , Qiang Fu and Wendong Xiao have been conducted a research on “ Advertisement Click-Through Rate Prediction Based on the Weighted-ELM and Adaboost Algorithm”[3].This study is based on advertisement Click Through Rate (CTR) prediction. They have implemented an online CTR system which consider the historical data and user profile information through a real time binding advertising mechanism. They have clearly proved that by using the system not only the product awareness but also the revenue of the company can be improved. They have also mentioned two unresolved problems of their study which have been planed to resolve in future works as low prediction accuracy and lack of real time advertisement binding.

Some researches of University of Rome and IULM University , Italy has conducted a research study “ Gender and Age Related Effects While Watching Tv Advertisements: An EEG Study”[4]. This study shows how EEG frontal cortical is correlated to the average satisfaction experienced when watching TV commercials. They have considered the demographic factor age and gender as parameters. An experiment have been performed for gender using a car advertisement and for the age , a chewing gum advertisement was experimented. As for the results obtained by the experiment 1 mostly males liked the advertisement while most of the females did not enjoy it. As the experiment 2 was performed considering the age , the results were as most of the older adults did not like that advertisement. As stated in this paper in all investigations, there was no measurable difference in the scene compared to the product offering between the experimental and control groups, indicating that there was no bias in the assessed populations. Finally as the conclusion they have stated that by designing creative advertisement can attract the targeted audience.

Two researchers Vaclav Stritesky and Marek Stritesky of the Prague University of Economics and Business has conducted a study “ Significance of demographic Variables for Targeting of Internet Advertisements”[5]. Internet is one of the mostly used mediums of advertising a product or service so in this research they have taken a sample population of internet users in Czech Republic. There are so many ways of delivering targeted advertisements and in this study they have considered the demographic variable of internet users. The purpose of this article is to discover gender and age-based variations in Czech internet users' internet content consumption patterns. The study attempts to assess from the Netmonitor project's comprehensive research, which was supplied to the Association for Internet Development for use in this study (SPIR). The findings of the study demonstrate that on the Czech internet, the conventional affinity-based approach of gender and age targeting is still viable. On either side, the standard method of affinity-based targeted advertisements might result in wasted advertising revenue that miss the specified target audience in some situations.

Human age and gender detection is one of the most popular research topics, most of these researches have used the active appearance model (AAM),which is a statistical model that captures shape and texture variations. Ali Maina Bukar, Hassan Ugail, [David Connah](https://www.spiedigitallibrary.org/profile/David.Connah-99530) has conducted a research on “Automatic age and gender classification using supervised appearance model”[6].In this research they have considered about some of the drawbacks in AAM when used for age and gender classification. As the solution for the drawbacks found in the AAM , they have proposed a supervised learning model. By substituting PCA with partial least-squares regression, the supervised appearance model (sAM) improves on the AAM. The gender and age categorization issues are then solved using this segmentation algorithm. Our tests reveal that sAM outperforms traditional AAM in terms of predictive capability. This demonstrates the sAM's strong predictability and image intensity ability. As future works they have planed to examine the possibility to reconstruct the human face using sAM , with the goal of doing automatic facial age synthesis.

Alina , who is a researcher has published an article on the topic “Age Friendly Advertising: A Qualitative Research on the Romanian Silver Consumers”[7] .She has mainly focused consumer attraction based on age groups. As people of age 55 and above are not much active on social media , they have different expectations when consuming products , at the same there are no much researches carried out to observe expectations of older adults of age 55 or higher. This research analyzes how retired seniors of Romania who lives in rural areas perceive ageing in relation to how they are depicted in contemporary advertising, as well as the specific characteristics of offensive and non offensive promotional activities aimed at them.A sample size of 23 people are taken into consideration for this analysis. Participants who lost their status during aging preferred to utilize chronological rather than cognitive pictures of their age, according to the findings. Furthermore, both the young and older people in this study has regarded the use of natural environments in advertising to be empowering.

There are many research papers conducted to predict the user preferred advertisement for social media and web applications using machine learning and deep learning algorithms. The system proposed in this report , unlike every other study has novelties like using Sri Lankan context for the analysis and predicting advertisements for television based on demographic factors.

## Research Gap

|  |  |  |
| --- | --- | --- |
| **Research** | **Affect of age and gender for television advertising(Sri Lankan Context)** | **Age groups effectiveness against emotional features**  **(Sri Lankan Context)** |
| Ads-In Site: Location based Advertising Framework with Social Network Analyzer | NO | NO |
| Gender and Age Related Effects While Watching YouTube Advertisements | NO | NO |
| Effect of Age and Gender on Consumer Response to Advertising Appeals | NO | NO |
| Artificial Intelligence based Business Strategy for Optimized Advertising | YES | YES |

Table 1 :Research gap

The primary need of the project is to bridge in between and optimize advertising on television in Sri Lanka to the next level. Moreover, considering the user’s demographics when placing advertisements on television is to be extended. Advertising is the attempt to impact the buying pattern of customers with the ability to deliver the message of trading goods and services. Among various modes of advertising, television commercials take a prominent position. The reason for television advertisements to be popular is that it reaches a vast number of audiences which will directly impact their buying patterns.

The most traditional process of advertising which is carried out on television is in the form of passive advertising which is delivered based on a schedule, which does not take into consideration of the user's demographic data such as age, gender, peer groups, etc. In advertising, user demographics are very important. The cost of advertising is very high as the advertisements are displayed to all users who are watching television. This will directly affect the sales conversions of the ads. The proposed system consist of an image recognition system that identifies the user, predicts the age and gender to produce an accurate analysis of demographic based advertising.

## Research Problem

Profitability is one of the main goals of an organizations achievements. Advertising helps an organization to achieve that goal. It has been proven that companies tend to spend a higher percentage of their revenue on marketing and sales. Some examples of these are: Google, Microsoft, and Amazon[9]. It is the way of connecting products and services with customers. In this process, the customer should be the main focus. Platforms such as social media, e-commerce, and video sharing platforms are strictly following user-centric strategies for advertising which considers the user's likes, dislikes, search history, and social profile, etc.

Television a mode of advertising which caters a large number of consumers stills follows the tradition method of screening advertisements according the location and time schedule. As an improvement for this problem this study provides several analysis on How effective demographic features are to select the most relevant and user preferred advertisement. Another specialty of this research which is that the analysis is based of Sri Lankan context. This will directly affect the sales conversions of any product or service .

## Research Objectives

### 1.4.1 Main objective

The main aim of this project is to increase user preferences by optimizing user-preferred and user-centric advertisements. This means the audience is capable to view advertisements related to their interests. Here factors like age, gender, and peer groups are taken into consideration when determining the most suitable advertisement on the TV during a commercial break. This project will be beneficial for both the audience and the advertisement firm by letting the audience make an interest in interested goods which will be advertised and the ability to advertising firms to create advertisement content based on age, gender, and peer group which will directly affect the users in purchasing goods and services more effectively and efficiently.

### 1.4.2 Specific objectives

* Gathering a Sri Lankan data sets and training the model.
* Detect human objects from the captured snapshots.
* Extract features to predict gender.
* Extract features to predict the age.
* Detect the number of human objects in the captured snapshot, and classifying them into peer groups.
* Analyze age and gender factors against advertisements
* Analyze age and gender factors against advertisements considering emotional features joy and sadness.

# METHODOLOGY

## Methodology

### Requirement analysis process

Due to the pandemic situation requirement gathering was one of the challenges I had to face . To identify age and gender of the user, a large data set was to be collected as the idea was to train my own dataset. However I was able to collect a dataset of 100 users. The data set was collected from my colleagues of Sri Lanka Institute of Information Technology and relations as I needed a dataset with people of different ages and gender.

As there were travel restrictions imposed I implemented a front end application to gather the required data. The implemented application was hosted in a server to give full control of the application to the user. I used WhatsApp to share the instruction sheet on how to navigate through the web application. The image fig .2.1 represents the front end application which was build to gather the age and gender data set. As the initial step , the user should enter a user id in the text box provided in the format “Name\_age\_sex”. As the next step the user should give access to the camera which capture the user .Then the camera preview will appear on the screen. In the next step , the user will need to click on the start button which initializes the capturing. The instructions are given regarding the position the user should remail for capturing as ;

* Remain 5 seconds front facing the camera
* Tilt the face to the left and remain for 5 seconds
* Tilt the face to the right and remain for 5 seconds.

After the above steps , the user is instructed to click on the Stop button which will terminate the capturing process. The figure Fig.3 bellow illustrates the interface used to collect data from the user.

Graphical user interface, text, application

Description automatically generated

Fig . Data Collection Interface

After a few seconds a notification will be popped up which confirms the data has been submitted to the cloud.

### 2.1.2 Approach of the study

After carful observation and conducting the feasibility study for the component , the technologies and hardware specification were selected. Since this study is conducted to evaluate the demographic factors against the advertisements I choose supervised learning machine learning models. It is a method used to build Artificial Intelligence applications by training the computer on the data input .Then the data is labeled to a certain extend until the algorithm understands the patterns and the correlation between the input and output data. The purpose of using test data is to check how accurately the algorithm performs on unlabeled data.

The algorithms used for the study was Support Vector Machines and Random Forest. SVM is an algorithm which is verified by professional to have a very high accuracy with a low computational power[10].The aim of this technique is to find a hyperplane by considering the given factors age group and the emotional factors. The diagram Fig .4 illustrate the plot diagram of the hyperplane of SVM algorithm.

Chart

Description automatically generated

Fig . Hyperplane Plot

Parallelly to the SVM algorithm , an analysis was conducted with Random Forest model too to detect the best algorithm which suits the analysis. RF is also one of the best supervised learning classification techniques [11] . it a decision tree based classifier .The factors age , gender and emotional features will be selected randomly when slitting the node. The diagram Fig.5 bellow illustrates how the algorithm classifies the final output of the scenario.

Diagram, engineering drawing

Description automatically generated

Fig . Prediction process of advertisement - Random Forest

Due to lack of data points SVM and RF models failed get a better accuracy of the prediction. As a solution for it , I used deep learning model to the analysis . For deep learning a sequential model was used with a batch size of 32 . Each set of features are processed by a stack of fully connected dense layers with the LeakyReLU activation function. The following diagram Fig.6 demonstrates the architecture of the deep learning model used in this analysis

Diagram

Description automatically generated

Fig . Deep learning model architecture

The flow chart Fig .7 below represents the implementation of the component discussed in this report.

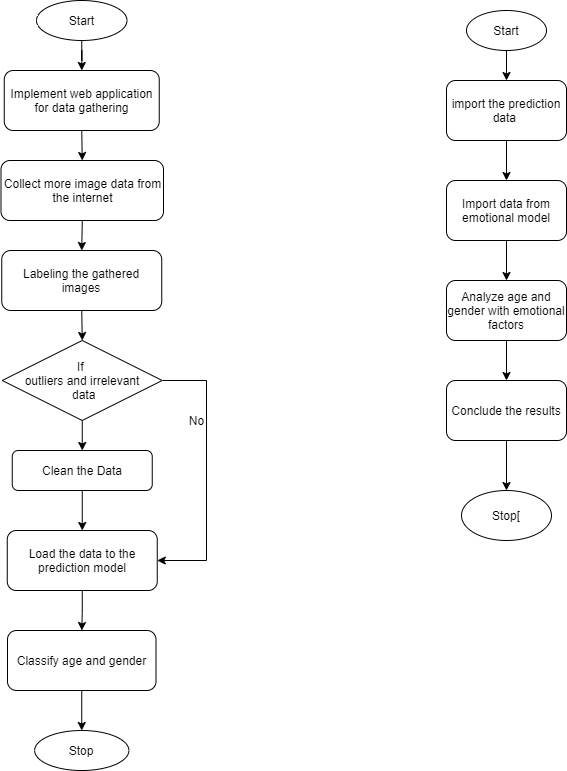


Fig . Flow of implementation

The proposed system is implemented to classify advertisements considering the demographic factors age , gender and the peer group. The below diagram Fig .8 depicts the system overview diagram of this components.

Diagram

Description automatically generated

Fig . System Overview Diagram

### Training phase for age and gender prediction

Age and gender classification is implemented using an existing Convolution Neural Network (CNN). As the initial step for classification , collected data was subjected for the data cleaning process . First the duplicate and irrelevant data was to be removed. This is usually done to speedup the analysis and reduce distractions for the main purpose of the data set . It also increases the performance of the dataset . The next step is to filter unwanted outliers so that a statistically significant result can be received. As the pandemic made a huge influence in data collection , had to collect an online data set for the age and gender prediction. An existing Machine Learning model [Name of the model] was used to train the user images in order to classify age and anger of the user. The Screen shot Fig.9 bellow illustrates the code base of the training model.

Text

Description automatically generated

Fig . Data set training (code base)

As the amount of data which was able to gather was only 100 user , an online data set containing facial images of people of all ages were user to train the model. The model which was used to train the data set was Rude Carnie which is a well known deep learning model with TensorFlow. The importance of this model is that it supports multiple face detection and in the training process all the facial images are trained individually. Open CV library is used to detect the faces .The facial images from the online dataset are imported in a csv for which organized and labeled so that the training will be better than success

This model comprises of maxpooling , convolution and flatten layers. At the first maxpooling layer the important feature of the image are extracted .The autoencoder is used to copy the input image to its output. Then the images are passed through the convolution layers. Following the convolution layers , a maxpooling layer and s flatten layer is used to convert the image into 1 dimensional array. Finally the 1d array is loaded to the SoftMax layer .The next phase is to test using the gathered images to validate the prediction.

### Testing phase age and gender prediction

In the testing phase , the user images I gathered were used to test the age and gender classification . Although the dataset consist only 100 user image , the machine learning model was able to obtain a better test accuracy.

### 2.1.4 Implementation of models

The analysis was carried out in order to achieve stated objectives of the research. Data gathered from the web application was taken into account through out the analysis with the help of emotional data gathered for the emotion prediction component. To detect the emotions of the user , a short video was used with 4 advertisements embedded in between the video clip so that the change of emotions can be clearly identified. The users who watched the video clip was the same users who’s demographic data was collected. Then the gathered data was processed and exported to a csv files such as demographics , emotions facial feature , advertisements and movies. The ages of the users were groups as 5 – 19 , 19 – 35 , 35 – 55 and 55 – 100 which will mean kids and teenagers , young adults , older adults and elders respectively. This is done because there are emotions that varies for different age groups. Evaluating the effectiveness of demographic factors in line with emotions for advertisement classification is experimented using three methods states bellow.

Illustrated below in Fig.10 are the code segments used to implement the classification model using Deep learning and Machine learning. Graphical user interface, text, application, email

Description automatically generated

Fig . Preparing the data

Text

Description automatically generated

Fig . Iteration through the dataset

In the above code segment first the data sets have been loaded such as movie part , advertisement part demographics and the emotions. To embed the age groups with the advertisement the pointer iterates through the movie part to find the user id which matches the id in the demography data set. Then a new column is added to the loaded data as age group. Age groups are distributed and defined as 5 – 19 ,19 – 35 , 35 – 55 and 55 – 75. A get method is used to return the defined value of age groups. Then data is saved into a csv named moviePartsWithAge. A for loop is used to iterate through the data set to identify the emotions regarding the movie part and advertisements in relation with the age group. This is conducted for all four age groups one at a time for all the users representing the age group considering the 4 movie parts and 4 advertisement parts respectively.

The bellow code Fig .12 is implemented to locate the favorite advertisement which is labeled in the data set. It also verified by the user at the end of data gathering process as the most remembered advertisement. The response of the user was assumed to be the most favorable advertisement.

Graphical user interface, text, application

Description automatically generated

Fig . Locating the favourable ad

**SVM and RF Classifiers**

With the use of the above implemented code segments first the analysis was conducted using SVM and RF classifiers. The parameters age and gender are used as x and y coordinate of the train and test split of the model. The code segment Fig .13 bellow demonstrates the implementation of RF classifier .

Graphical user interface, text, application, email

Description automatically generated

Fig . Random Forest Classifier

Train and test data is used to train the model and do the actual prediction. Then the results are illustrated using a confusion metrics. The same method was also experimented with SVM model to ensure and receive a better accuracy, but both the models was able to get and equal number of accuracy .The following code segment Fig . 14 illustrated the implementation of the SVM classifier.

Graphical user interface, text, application, Word

Description automatically generated

Fig . Support Vector Machine Classifier

**Deep Learning Classifier**

Since the machine learning classifiers were not able receive a better accuracy the analysis was conducted with use of deep learning. The following code Fig .15 segments depicts the implementation of the classifier.

Graphical user interface, text, application

Description automatically generated

Fig . Initializing train ,test and validation data

Prior to starting the analysis , the data set is split into three as train , test and validation data.

Graphical user interface, text, application, email

Description automatically generated

Fig . Deep Learning Model

For the deep learning algorithm , the same data used for machine learning classifier is considered. The parameters are age , gender and the favorite advertisement of the user .The favorite advertisement is represented as a numerical value. Since categorical method is used and the output layer should consist of 4 neurons , a one hot encoder is build to store the y values of the data set. The patience ,learning rate , batch size and epochs was set to 50 , 0.0001 , 32 and 500 respectively. This sequential model consist of two dimensions as the inputs are age and gender together with fully connected dense layers followed with the activation function LekyReLu .The final layer consist of 4 neurons representing the 4 advertisement and the activation function used is SoftMax. When training the model two middleware’s , visualizer and early stopper are been used. Visualizer is used to save the validation and test accuracies of the model . With the use of early stopper the change of validation accuracy after every epoch can be checked and if there is no significant change for a certain number of epoch’s the early stop criteria with stop the training process and restore the accuracy to the highest point received. This will prevent the model from overfitting. Loss and the optimization functions used in this model are Binary Cross Entropy and Adam optimizer .

Next the results are plotted in separate graphs for each age group and each user. As there are too many graph lines which make the analysis more complex , the mean value was concluded to plot in one graph representing all the users of the 4 age groups. Each age group was analyzed against joy and sadness emotions received from the user in response for the video clip. I have used a chart to plot the results obtained by the analysis to confirm people always does not refuse watching advertisements. The code segment Fig .17 demonstrated the implementation to plot the results of the analysis.

Text

Description automatically generated with medium confidence

Fig . Implementation of result plotting

## Commercialization aspects of the product

Traditional media is still used by the majority of Sri Lankan media users. Traditional media commercials are mostly targeted based on location and time. Because the advertising technique is inadequate, a significant amount of money is being misplaced. This loss affects not only the marketer, but also the content consumer. The consumer is forced to view what he or she does not want to watch.

The goal of advertising is to enlighten the audience about the product and the firm, as well as persuade them that the advertisers' items are the best on the market. The number of impressions and click through rate are used to determine the success rate of digital adverts. However, there is no robust, rapid, or reliable way for quantifying these in traditional advertising. We recommend incorporating digital ad strategies into traditional ad campaigns. Because the majority of Sri Lankans still rely on traditional advertising, there is a significant market for our goods. In addition, we demonstrated in this study that our prototype is capable of recommending user-favored commercials.

Apart from these benefits, with the help of the gadget we are plugging into the television, we can assess the number of impressions and whether the audience appreciated the ad broadcast to them in real-time.

### 2.2.1 Introduction to the market

To introduce the concept to the market , initially we have chosen social media as our marketing mechanism. Facebook is the ideal market place since most people and companies surf through it daily. We have also created a Facebook page since most of the sponsors needed to be aware of the concept they are giving the sponsorship. We have been working to present this prototype concept to some advertising firms and broadcasting companies. At the introduction stage the concept able to test with the real audience free of charge . The below diagrams Fig .18 & Fig .19 demonstrates the Facebook page used to commercialize the concept.

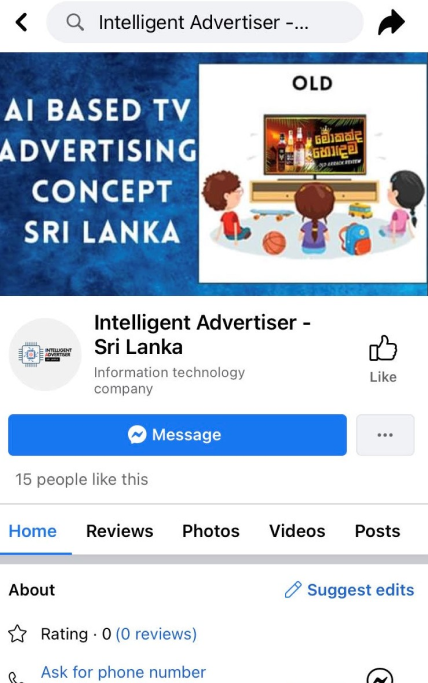


Fig . Facebook Page Overview

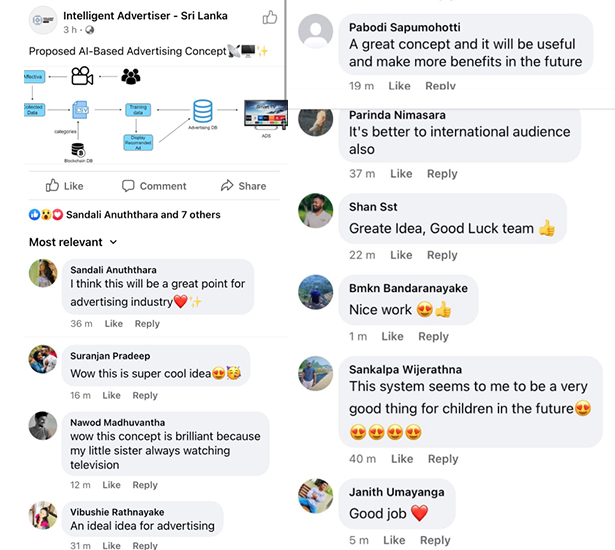


Fig . Concept Impressions

## Testing of the product

For testing purposes a sample number of users are let to watch the same movie clip used to gather data, but unlike the previous movie clip this does not have any advertisements embedded to it. Once the user is detected machine learning model will check for the age group and the gender of the user .Along with the factors and the results obtained from the analysis which was conducted, the most preferred advertisement will be overlayed by pausing the video clip. After the video clip is over the prediction will be further evaluated by taking a survey. Table 2 below demonstrates some of the test cases used for the testing purposes.

|  |  |  |
| --- | --- | --- |
| **Test Case** | **Test Case** | **Expected Result** |
| Age prediction | Locate the user in front of the screen in order to predict the age | Predict the correct age group |
| Gender prediction | Locate the user in front of the screen in order to predict the gender | Predict the correct gender |
| Advertisement prediction | Predict the mot suitable advertisement by considering the results obtained from the analysis | Predict the most suitable advertisement |
| Validate the prediction | Use users response to validate the preferred advertisement prediction | Response matching the prediction |

Table 2: Test Cases

# RESULTS AND DISCUSSION

## Results

Three models have been used to experiment the effectiveness of age and gender when classifying advertisements .To brief the relation between a persons attraction towards an advertisement and the demographic features, a graph is plotted. The figure Fig.20 demonstrates the correlation of demographic factors and advertisements. The advertisement which corelates with advertisement was verified by the response given by the user , which was the most memorable advertisement at the stage of data gathering.

**Chart

Description automatically generated**

Fig . Correlation between demographic features and most favorable advertisement

### 3.1.1 Accuracy of models

There were complications due to lack of data points since the pandemic situation was the biggest challenge faced during the analysis. However the Random fores model was able to obtain a validation accuracy of 54%. A detailed report of the classification model is shown by the Fig .21.

Table

Description automatically generated

Fig . Detailed report of SVM Classification

When the prediction was done Support Vector machine the same accuracy could be obtained which is 54%. Then the approach was to use a deep learning method to conduct the analysis since deep learning algorithms are verified in obtaining higher accuracies compared to machine learning models. The Fig.22 shows the results of the deep learning model trained against advertisements to find the favorable advertisement of each age group.

Graphical user interface, text

Description automatically generated with medium confidence

Fig . Deep learning model training

Since there was no significant change in the accuracy for 50 epoch’s starting from 242nd epoch , the early stopping criteria function has stopped the training and restored the models from the end of the best epoch. And the accuracy was as 52%.A detailed report of the classification model is illustrated in Fig .23 .

Table

Description automatically generated

Fig . Results of Deep learning classifier

The following graphs Fig.24 and Fig.25 illustrates the accuracy and loss curves of the deep learning model.

Chart

Description automatically generated

Fig . Accuracy curve

Chart, line chart

Description automatically generated

Fig . Loss curve

### 3.1.2 Research findings

For further clarification, a results obtained from the analysis which was conducted to see how effective demographic features are for advertisement classification I represented in a line graph. The effectiveness was observed for each age group separately. Liked advertisement was classified with the use of joy and sadness factors gathered for the emotional analysis component. Each Fig .26 , Fig .27 , Fig .28 and Fig .29 are the graphs obtained for the four age groups.

Chart, line chart

Description automatically generated

Fig . Advertisement and movie part analysis for age group 5 to 19

Chart, line chart

Description automatically generated

Fig . Advertisement and movie part analysis for age group 19 to 35

Chart, line chart

Description automatically generated

Fig . Advertisement and movie part analysis for age group 35 to 55

Chart, line chart

Description automatically generated

Fig .27: Advertisement and movie part analysis for age group 55 to 100

To get a clear understanding about the preference of the user the above results are summarized and the mean value of each age group is plotted in a graph which represents the level of joy and sadness . Joy and sadness represents how the user responded to the advertisement and the video clip. . The sample video clip we used for the overall study was a Sinhala comedy drama name “Yes Boss” and the 4 advertisements were as follows;

* Gold Marie advertisement
* Smart Cream Cracker Advertisement
* BOC Pension Scheme Advertisement
* Kottu Me Advertisement

**Chart, line chart

Description automatically generated**

Fig . Mean score of the preferences of age groups

Referring to the above graph Fig.29 following observations are made;

* All the age groups has a similar distribution.
* All the age groups have a better joy rate at the beginning of the video clip.
* Joy rate of all age groups have dropped at the start of the first advertisement.
* Graph of age group 55 to 75 shows that they are the most satisfied user group through out the video.
* According to the graph almost every age group has enjoyed the 4th advertisement since it is an entertaining advertisement
* According to the graph all age groups other than 19 to 35 enjoys advertisement compared to the movie parts.

In the later part of the analysis a graph representing the attention to the video of age groups was also predicted .

Chart, line chart

Description automatically generated

Fig . Mean score for attention toward the video of age groups

According to the Fig .30 the age group 5 to 19 shows a higher score of attention towards advertisement as well as the overall video.

## Discussion

This study is mainly conducted to evaluate how affective the demographic factors cause the preferences of advertisements. Television advertisement make a huge impact in consumers selection and also it is mode which caters for a huge audience. Recent studies proves that Almost every personal in Sri Lanka watches live television even though they are not aware of other modern technology[12]. According to the same report, it states advertisers charge for commercials based on when they are broadcast. To take advantage of this problem the analysis done for this component is very important as it has novel finding to contribute to the television advertising domain.

The results prove that the best time for an advertisement which all age groups and genders play attention is at the beginning of any tv program and also the ability to detect the person or the group of personals in front of the tv will cause in significant changes of the sales conversions for that particular item. Also it proved that the age group 5 – 19 which is considered as the youngest age group pays more attention other than the rest of the age groups. This will validate that this system make a great effect to the society since youngers learn from watching so if unwanted ads are played during time they are in front of the television , they will pay attention for that.

# CONCLUTION

Considering the results obtained by the three model , although the accuracy of them did not reach more than 60% , the finding obtained by the analysis confirmed that the advertising strategy proposed was a success. The first observation was that the youngest age group was attending the video clip and the advertisements without any distraction .Considering that we can conclude this concept should be taken forward to implement with more enhancements since younger generations learn things by watching things they see so unwanted and unrelated advertisements may ruin their thoughts. The next observation was that all the age groups did not like when playing the advertisement in the first 5 minutes but in the next five minutes some enjoyed the advertisement other than the movie part. The analysis conducted considering age group reactions towards joy and sadness clearly stated how the users interacted throughout the video. But recommending advertisements considering only the demographic factors will effect of making good impressions from the user.

# References

[1] Keshari, Pragya, and Sangeeta Jain. “Effect of Age and Gender on Consumer Response to Advertising Appeals.” *Paradigm*, vol. 20, no. 1, June 2016, pp. 69–82. *DOI.org (Crossref)*, <https://doi.org/10.1177/0971890716637702>.

[2] Perera, A. A. G. A. K., et al. “Ads-In Site: Location Based Advertising Framework with Social Network Analyzer.” *2014 14th International Conference on Advances in ICT for Emerging Regions (ICTer)*, IEEE, 2014, pp. 116–23. *DOI.org (Crossref)*, <https://doi.org/10.1109/ICTER.2014.7083889>.

[3] Zhang, Sen, et al. “Advertisement Click-Through Rate Prediction Based on the Weighted-ELM and Adaboost Algorithm.” *Scientific Programming*, vol. 2017, 2017, pp. 1–8. *DOI.org (Crossref)*, <https://doi.org/10.1155/2017/2938369>.

[4] Cartocci, Giulia, et al. “Gender and Age Related Effects While Watching TV Advertisements: An EEG Study.” *Computational Intelligence and Neuroscience*, vol. 2016, 2016, pp. 1–10. *DOI.org (Crossref)*, <https://doi.org/10.1155/2016/3795325>.

[5] Stříteský, Václav, et al. “SIGNIFICANCE OF DEMOGRAPHIC VARIABLES FOR TARGETING OF INTERNET ADVERTISEMENTS.” *Business, Management and Education*, vol. 14, no. 1, June 2016, pp. 1–18. *DOI.org (Crossref)*, https://doi.org/10.3846/bme.2016.277.

[6] Bukar, Ali Maina, et al. “Automatic Age and Gender Classification Using Supervised Appearance Model.” *Journal of Electronic Imaging*, vol. 25, no. 6, Aug. 2016, p. 061605. *DOI.org (Crossref)*, <https://doi.org/10.1117/1.JEI.25.6.061605>.

[7] Duduciuc, Alina. “Age-Friendly Advertising: A Qualitative Research on the Romanian Silver Consumers.” *Romanian Journal of Communication and Public Relations*, vol. 19, no. 2, July 2017, p. 29. *DOI.org (Crossref)*, <https://doi.org/10.21018/rjcpr.2017.2.237>.

[9] Bennett, Coleman & Co. Ltd,"The Economic Times".[Online].

Available: <https://economictimes.indiatimes.com/definition/advertising>

[10] Gandhi, Rohith. “Support Vector Machine — Introduction to Machine Learning Algorithms.” Medium, 5 July 2018, <https://towardsdatascience.com/support-vector-machine-introduction-to-machine-learning-algorithms-934a444fca47>.

[11] “Introduction to Random Forest in Machine Learning.” Engineering Education (EngEd) Program | Section, https://www.section.io/engineering-education/introduction-to-random-forest-in-machine-learning/. Accessed 7 Oct. 2021.

[12] Media, R. S. G. “TV Audience Measurement.” RSG Media, https://www.rsgmedia.com/tv-audience-measurement/. Accessed 28 Aug. 2021.

# APENDIX

## Similarity Index

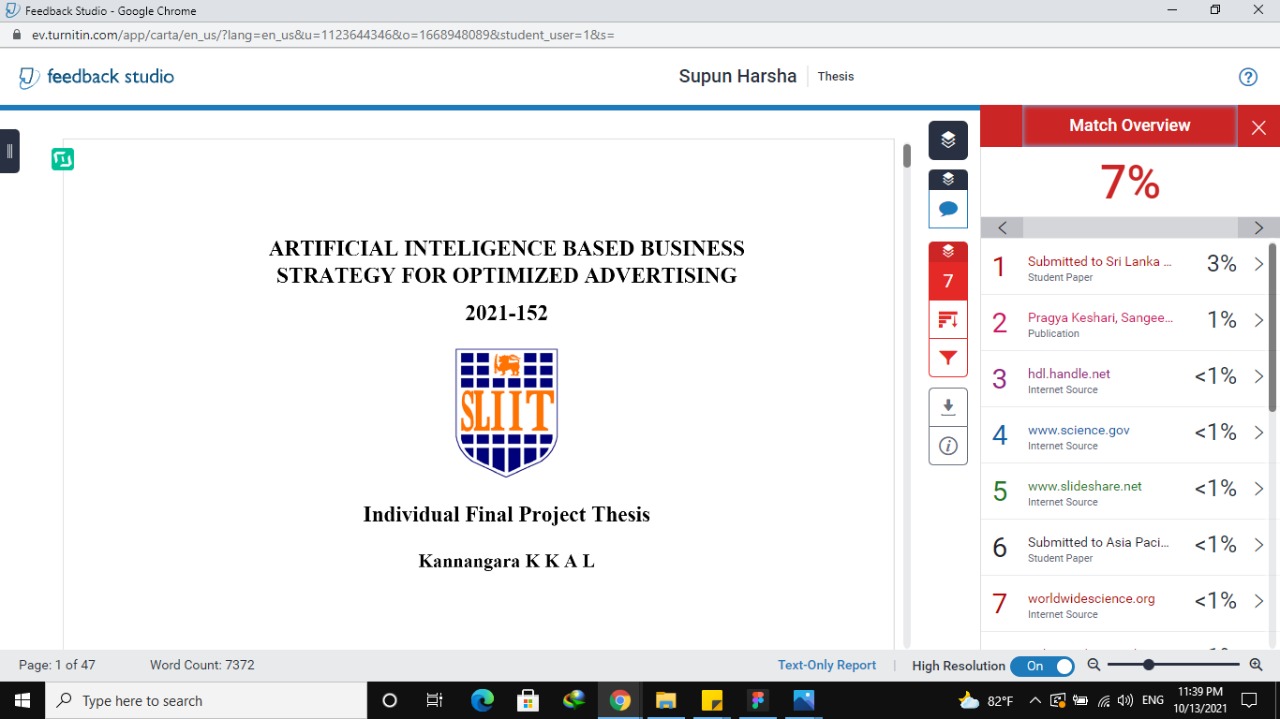


Fig . Similarity Index of the report