Artificial Intelligence-based Business Strategy for Optimized Advertising

Lahiru Kannangara  
*Faculty of Computing*  
*Sri Lanka Institute of Information Technology*Sri Lanka  
lahirukannangara2809@gmail.com

Chandula Siriwardhana  
*Faculty of Computing*  
*Sri Lanka Institute of Information Technology*Sri Lanka  
chandulatw@gmail.com Supun Harsha  
*Faculty of Computing*  
*Sri Lanka Institute of Information Technology* Sri Lanka  
[supunharsha12r@gmail.com](mailto:supunharsha12r@gmail.com)

Thushan Isuru  
*Faculty of Computing*  
*Sri Lanka Institute of Information Technology*Sri Lanka  
thushan966@gmail.com

*Abstract*— Television commercials are a passive type of advertising technique that does not consider consumer demographics who are viewing the television at a specific time. As a result, the user sees irrelevant advertisements, which tends to reduce user engagement and sales conversions.As Sales ,which is the expected target of any advertisement campaign, a user-based advertising approach can be considered as a solution to mitigate the negative aspects. A user-based advertisement suggesting system for television, which is extensively utilized in every other digital media, is expected to be given as the solution. For the suggestion process, user attributes such as age, gender, peer group, and the mood identified in which the advertising is shown were taken into consideration. This will result in more relevant commercials for consumers, making television advertisements more user-friendly, resulting in greater sales conversion for the advertising agency.

Keywords—Advertisements,Age,estimation,Genderestimation, Image processing, Ad Recommendation, Human Emotions, Data Security, Blockchain, Mongo based Blockchain

# Introduction

An advertisement or advertising is the promotional act of selling a product, service, or brand. It builds an attraction in the target audience [1]. Nowadays advertisements are socialized in a variety of ways, such as Display Advertising, Video Advertising, Mobile Advertising, and Native advertising. In any form, the main goal of advertising is to reach the audience most likely to be willing to pay for the products or services and convince them to buy.

Advertising is the most effective means of reaching out to the audience. Audiences are more informed about the brands available in the market and the range of items available to them due to advertising. Compared to society even kids, teenagers, and adults are exposed to advertisements [1].

Currently, advertising is used far and wide such as in newspapers, television, social media, and other e-commerce platforms. Although rapid development has happened in digital media like YouTube and other social media, television advertising still follows a traditional method which is based on the location and schedule of the tv programs, but most of the other digital media used a user centered targeted mechanism.In this research, the intention is that Tv advertising should also shift from traditional to user-centric contextual advertising.

In this research, the main goal is to propose a system that recommends TV advertising by evaluating demographics, facial features, and emotions. This research project is mainly focused on three major factors to deliver the final which is the advertisement. Recommendation system.

The main requirement is that anyone who uses our system must have a Smart TV and a camera because with the help of the camera and the AFFEDEC software that have in the training model is used in gathering emotions, facial features, and demographics (age and gender) are captured and processed in the front end because as mentioned above user privacy is the major thing, and those snaps will never be saved or stored in this system. The main purpose of using the AFFEDEX tool is that to gather relevant data (emotions, facial features, and demographics) in real-time in a way that does not harm the users. Much research has validated AFFEDEC and its ability to classify facial features and emotions.[2][3] At that point, no need to save captured snapshots in the encrypted database and just save & store the features as JSON files. in the blockchain encrypted database.

# Literature Survey

[Xiaohua Huang](https://www.researchgate.net/profile/Xiaohua-Huang-12) have proposed a method to recognize arbitrary-view facial expressions by using discriminative neighborhood preserving embedding and multi-view concepts. Initially, it captures the discriminative property and afterward, it investigates the proximity of intra-class samples in a low-dimensional subspace from any viewpoint. this method produces promising results for recognizing face emotions with arbitrary perspectives, according to experimental findings on the BU-3DFE and Multi-PIE databases [4].

This conducted research describes a computerized technique for presenting interactive advertisements based on one embodiment of the invention [5]. The research has shown how to enable interactive ads to be loaded on mobile devices and how to interact with such interactive adverts on their mobile devices while offline [6]. An interactive advertising board is another invention based on environmental factors [7].

An improved content activator incorporated in interactive advertising has been described by Huber, Lemmons, Zenoni, and Hensgen [8]. In [9], the influence of mobile advertising content on purchase intention was shown using a framework developed from previous models and theories of technology adoption. Park and Choi [10] developed a method for broadcasting interactive advertising with extensive information that uses secondary information tags. In the meanwhile, style is available in an off-line mode.

# Methodology

## Overal System

During the initial step the user will be captured using the camera and the extracted features will be sent to database as a JSON format.The captured data is trained against age , gender ,emotional and facial features in order to predict the most relevant advertisement for the user.A blockchain db is used in between to verify the security of the system and users privacy. **Fig.1.** depicts the overall sytems diagram which depicts how all the components are connected to each other and how the advertiser works.

.

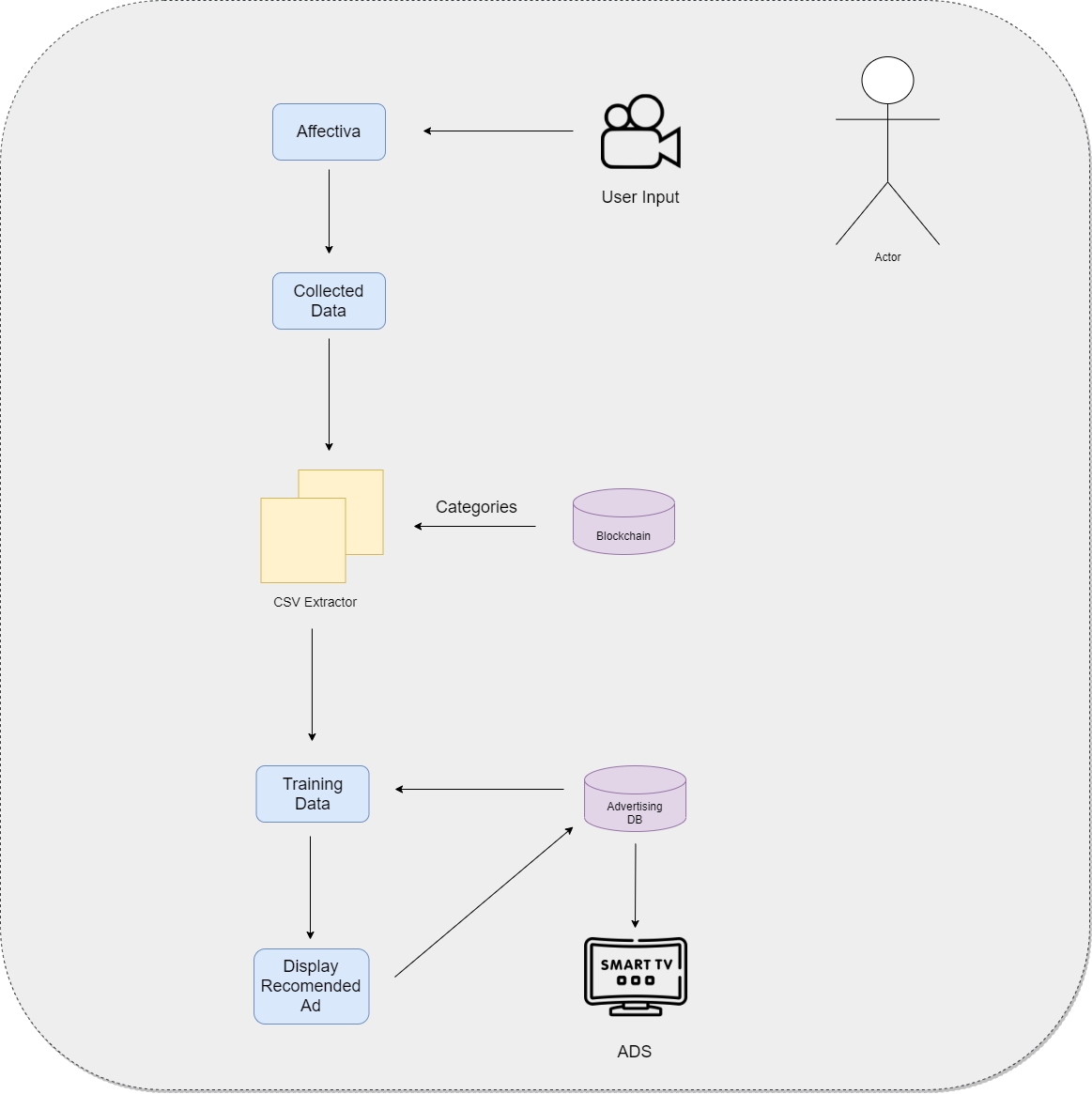


Fig.1.Overall System Diagram

## Data Gathering

A web application was implemented that lets the users watch a small movie clip with advertisements. In addition, Emotional and facial features were captured while the user is watching the movie. At last, A little survey was given and asked for the most remembered advertisement out of the four ads.

Extracting facial features and emotions from an image is a well-researched area. [5]. And our system needs to inspect the user/users constantly while they are watching the movie. And the images are not saved or send to the backend but to classify and extract facial features and emotions from the client-side. AFFDEX was used to extract these features from the user’s face. Finally, the saved data is sent to the backend. And the backend encrypts the received data with blockchain encryption, save to a mongo instance.

## Data Preprocessing

This study is mainly focused to understand the seven basic emotions proposed by Ekman: two positive (joy and surprise) and five negatives (anger, disgust, sadness, contempt, and fear)[11]. In addition, since valence describes the intensity of the emotions, valence was coupled with the feelings. So, in total, eight emotional features for each timestamp per user are captured. Along with the expressive features, AFFDEX extracts twenty facial features per user for each timestamp. In addition to dynamic and facial features, the users were asked for their most memorable ad at the end of the session. And that data was recorded.

100 users were interviewed, and approximately there were 1.3 million data points in total. After preprocessing and removing outliers and undefined values, and finally, got 77 users. Our movie is 10 minutes long and consists of 4 movie parts and four advertisements. Aggregated each emotional and facial feature per movie part and advertisement part. So, in the end, for a user, and got eight values per feature.

The set of features were filtered that belong to the favorite advertisement of the user. Then, as per record, the most memorable ad per user, A supervised network was decided to implement based on emotions, facial features, and demographic data to assess the feasibility of implementing a recommender system based on the features above. For all the trained models, and partitioned 50%,20%, and 30% as the datasets for training, validations, and test data.

## Approach

The advertiser has designed the ads to target the age groups: children, young, young adults, and elders. Therefore, A correlation analysis of age groups was done against the most memorable ads.

**Chart

Description automatically generated**

Fig.2.Corealtion staus with most frequently viewed advertisement

As mentioned above, separate machine learning models were implemented with emotional, facial features, and demographic data to recommend advertisements. For each approach, Random Forest and Support vector machine was used as our baseline. Further, the Deep learning model was implemented for each feature set, afterward compared the accuracies with the baseline models. and ensure to keep the architecture of the deep learning model the same to avoid any conflicts and get the maximum accuracy [12][13].

In view of fact that the results obtained by the SVM and RF models are not accurate enough, and implemented a deep learning Sequential model [14][15] inspired by previous work as depicted in Fig.3. Each set of features is processed by a stack of fully connected dense layers with the Leakyrelu activation function.

Diagram

Description automatically generated

Fig. 3. Model Architecture

The deep learning model consists of three fully connected hidden dense layers. The first layer is always dense, but the input size of the first layer is different from each model according to the size of the feature set. And the last layer is always a dense layer with SoftMax activation function, which classifies into four classes.

The model was trained in batches of 32 using categorical cross-entropy as the loss function. The Adam was used optimizer with the learning rate η = 0. 0001 and decay rates β1 = 0.9, β2 = 0.999.The maximum epochs were set to 500, but also set early stopping criteria of 50 epochs. So, the training will be stopped if the validation loss did not improve after 50 consecutive epochs, and the best model weights will be retained.

## Age and Gender Models

Based on the two features, age and gender, the accuracy of the RF and SVM algorithms is the same as 54%. **Fig.4** shows a detailed report.

**Table

Description automatically generated**

Fig. 4. Age and Gender results of SVM and RF

The deep learning model also failed to increase the accuracy in classifying the dataset. Moreover, as this model only has two input features and the input features are categorical values, the performance worsened. Fig.5 shows a detailed report.

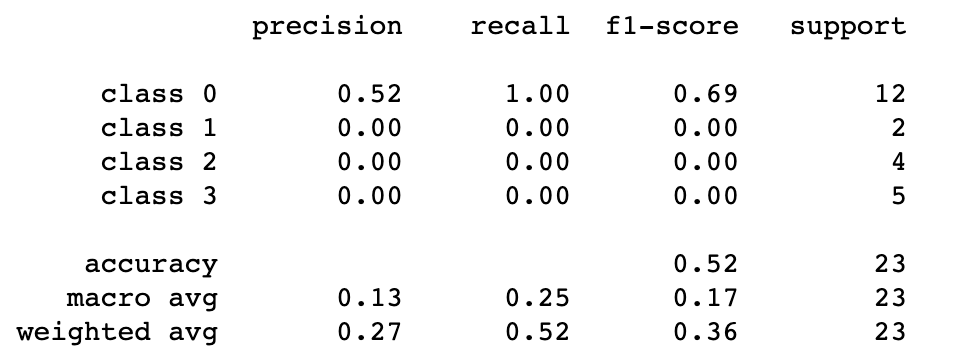
****

Fig.5. Age and gender Deep Learning Results

## Emotional Models

We wanted to evaluate how good is emotional features alone in recommending advertisements to the users. As mentioned earlier, and have eight emotional features. RF and SVM classified data based on emotional features, and both algorithms acquired the same accuracy of 55%, as shown in Fig.6.

**Table

Description automatically generated**

**Table

Description automatically generated**

Fig. 6. Emotional Data with SVM & RF models

The accuracy of the trained model is 65%, AUC is 80% for the deep learning emotional classifier. **Fig.7** shows the confusion matrix of the classifier.

**Chart

Description automatically generated**

Fig 7. Confusion Matrix

The techniques that used in this research have recovered the best model before overfitting with the help of early stopping middleware. **Fig.8** shows the accuracy and loss curves for the training and validation dataset.

**Chart, line chart

Description automatically generated**

Fig. 8. Accuracy and Loss curve

## Facial Feature Model

The maximum accuracy of SVM and RF algorithms for facial features is 57%. **Fig.9** shows the classification report.

Fig. 9. Facial feature data with SVM and RF models

However, the accuracy of the deep learning model increased significantly for the facial features dataset.

**Fig.10 & Fig.11** shows the classification report and the confusion matrix obtained using the deep learning model.

And the accuracy is 86%, with AUC 94.99%.

**Table, calendar

Description automatically generated**

Fig. 10. facial feature deep learning model

**Chart

Description automatically generated**

Fig.11. Confusion matrix of facial features

**Fig.12** shows the accuracy and loss curves for the training and validation dataset. for Facial Features Models

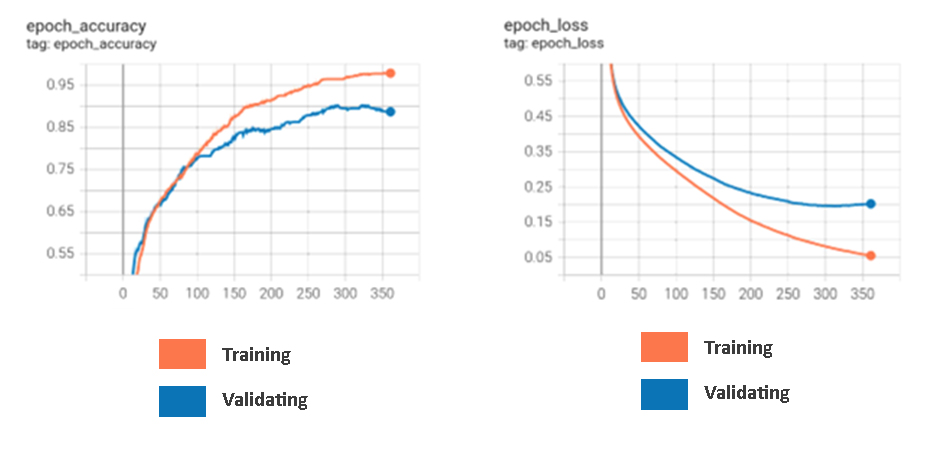


Fig.12. Accuracy & Loss curs for facial features

## Ensemble Models

All the above models were somewhat good at recommending user-preferred advertisement. All the deep learning models performed better than standard machine learning models. Enough studies have been performed on recommending ads based on emotions, demographics, and facial features individually. To the best of our knowledge,Tthe first were to analyses facial features, emotions, age, and gender to recommend user-preferred advertisements on television.

The accuracy of the SVM and RF were the same as facial features for the ensemble model. However, the accuracy sof the deep learning model increased significantly. According to the classification report in **Fig.13**, the accuracy is 90%, and the AUC is 97.38%.

**Table

Description automatically generated**

Fig. 13: Ensemble Model Results

The ensemble model shows that using demographics, emotions and facial features improve the quality of recommended systems.**Fig.14** and **Fig.15** respectively shows the confusion matrix and the loss and accuracy curves for the ensemble model.

**Chart

Description automatically generated**

Fig.14. Confusion Matrix

Chart

Description automatically generated

Fig. 15:Loss and Accuracy Curves of Ensemble model

## Comparison of Each Model

In this study three factors have been used to train the data set against advertisements in order to select the most prefered advertisement.tha **Table.1** bellow depicts the comparison of the results obtained.As the first model , Age and Gender data which are the demographic data was trained using three algorithms and among those three the deep learning model has obtained the highest test accuracy which is 70% .The same method have been followed regarding the facial and emotional data using the deep learning model  and  has obtained a test accuracy of 80% and 95% respectively.Then finally all the factors together was trained in the ensemble model and was able to obtain the highest accuracy of all models which is 97%.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Model** | **Test Accuracy %** | **AUC %** | **Precision** | **Recall** | **F1** |
| Age & Gender | 52 | 70 | 0.27 | 0.52 | 0.36 |
| Emotion | 65.05 | 80.15 | 0.66 | 0.65 | 0.64 |
| Facial | 86.55 | 94.99 | 0.86 | 0.86 | 0.86 |
| Ensemble | 90.32 | 97.38 | 0.90 | 0.90 | 0.90 |

Table :Comparison of Models

# Discussion

A sample of 100 users were studied for their emotional, facial and demographic feature changes by monitoring and logging the changes closely. 77 data points were retained after preprocessing and filtering. Joy, sad, … were belong to emotional features while eye lifting, laughing, .. were belong to facial features. Age and gender features were belonged to demographic data. Even though the previous studies have suggested that the recommending ads based on demographic data are a great success, analysis of this study shows otherwise. Nevertheless, this study proposes a new and affectional method for recommending ads while preserving the privacy of the users.

Recent studies suggest that a considerable percentage of the Sri Lankans watch live tv across all age groups [16].The same study points out that the advertisers cost the advertisements according to the time of broadcasting. But this study suggested a more robust and affective way for the advertisers to broadcast ads and more importantly they can evaluate the performance of the suggested advertisements in realtime.

Chart, bar chart

Description automatically generated

# Future Works

The user studies were performed during the pandemic and tried best to minimize the environmental variables for each user because environmental factors matter the most for the emotional features. But it was a challenging task for us to interview each user in a specific place at a specific time due to the covid outbreak

In the future, the performance of the recommender system is being planned to evaluate. Furthermore, the team is interested in building an unsupervised recommender system that recommends new advertisements to the users based on their emotional, facial, and demographic features. Nevertheless, the advertisement database is being planned to expand as well.

# Conclusion

This novel study was done to find out how affectively tv advertisements can be recommended by considering the demographic features (age, gender and peer groups), facial features and emotional features. Therefore a sample of 100 users was let to watch a short movie clip with four advertisement and was able to gather their emotional, facial and demographic features. Then at the end of the movie, the user was asked a questioner “what was the most memorable advertisement played?” as if to verify the collected data when training. All the three factors was trained against the advertisement data using a deep learning model and the results obtained were that there is no strong correlation between the three factors against the user-preferred advertisement. In the end ensemble model outperformed all with 97% accuracy. Therefore it is concluded that an ensemble model with demographic, emotional and facial features is best at recommending preferred advertisements.

##### References

1. Advertising *|* Definition, History, Objectives, Techniques, Examples, & Facts. (n.d.). Encyclopedia Britannica. Retrieved August 26, 2021, from https://www.britannica.com/topic/advertising
2. Lewinski, Peter & Fransen, Marieke & Tan, Ed. (2014). Predicting Advertising Effectiveness by Facial Expressions in Response to
3. Stöckli, S., Schulte-Mecklenbeck, M., Borer, S. *et al.* Facial expression analysis with AFFDEX and FACET: A validation study. *Behav Res***50,**1446–1460 (2018). https://doi.org/10.3758/s13428-017-0996-1
4. Huang, Xiaohua & Zhao, Guoying & Pietikäinen, Matti. (2013). Emotion recognition from facial images with arbitrary views. BMVC 2013 - Electronic Proceedings of the British Machine Vision Conference 2013. 10.5244/C.27.76.
5. Moshe, E.: Method and system for interactive advertisement. US Patent Application No. 10/700,476 (2005) 17. Aufricht, M., Weinstein, R.,
6. Broadwell, G.: Interactive advertisement mechanism on a mobile device. US Patent Application 09/864,293, 2 May 2002 18. Yen, D.,
7. Huang, T.H.: Interactive advertisement board. US Patent Application 11/429,986, 1 Feb 2007
8. Huber, T., Lemmons, T., Zenoni, I., Hansgen, D.: System and method for providing interactive advertisement. US Patent 9,865,017, 9 Jan 2018 14.
9. Hashim, N.H., Kassim, N.M., George, R.A.: Revisiting the effect of mobile advertising content on buying intention. Gading J. Soc. Sci. 20(01), 1–8 (2017) 15.
10. Park, W., Choi, D.: Method of realizing interactive advertisement under digital braodcasting environment by extending program associated data-broadcasting to internet area. US Patent Application 11/575,975, 13 Mar 2008
11. “What Are 7 Universal Facial Expressions? An Introduction to the Science.” The Emotional Intelligence Academy, 16 Apr. 2018, https://www.eiagroup.com/facial-expressions-explored/.
12. [1] Keras Team, “The sequential model,” Keras.io. [Online]. Available: <https://keras.io/guides/sequential_model/>. [Accessed: 16-Aug-2021].
13. Amusing Persuasive Stimuli. Journal of Neuroscience, Psychology, and Economics. 7. 1. 10.1037/npe0000012. J. Clerk Maxwell, A
14. Timme S, Brand R (2020) Affect and exertion during incremental physical exercise: Examining changes using automated facial action analysis and experiential self-report. PLOS ONE 15(2): e0228739. <https://doi.org/10.1371/journal.pone.0228739>
15. Taggart, R. W., Dressler, M., Kumar, P., Khan, S., and Coppola, J. F. (2016).  Determining emotions via facial expression analysis software. Available at:  http://csis.pace.edu/~ctappert/srd2016/2016PDF/c2.pdf (Accessed July 20, 2020).
16. Media, R. S. G. “TV Audience Measurement.” *RSG Media*, https://www.rsgmedia.com/tv-audience-measurement/. Accessed 28 Aug. 2021.