CSE 4839 | Human-Computer Interaction
Final Project Report

Measurement of Movie Popularity using Audiences' Facial Expression Recognition

Submitted By

170042004	Fahim, Arsad Nafis
170042035	Muhtasim Abid
170042049	Maysha Afrin Munia

B.Sc in Software Engineering, Department of Computer Science and Engineering, Islamic University of Technology

Submitted To

Mohammad Ishrak Abedin

Lecturer, Department of Computer Science and Engineering

Islamic University of Technology

Submission Date

16 April 2022

Table of Contents

Abstract

The movie industry is evergreen. It is ever-growing with the global film industry being approximately over $150 billion. As time is passing, it is getting more difficult to receive feedback on movies and calculate a proper rating. A high rated movie may generate more revenue for the company, theatre and other stakeholders. Our problem is collecting authentic reviews from the audience in an effective and efficient way. Our proposed approach is a system that captures the facial expressions of the audience, screening films. The system classifies the data and provides a decision about the general enjoyment of the cinema. According to our findings from the data gatherings, around 35% of the movie audience feel peer pressure while expressing their critics regarding the movie. Around 60% of the movie audience talked with their companion for a mentionable amount of time (>5 min) during the movie run time. Around 90% of movie directors and producers are frustrated with the fake reviews regarding their movies.

Introduction

The success of any movie can be judged based on the opinion of the audience. If the audience likes the movie, the show is supposed to generate revenue for the movie producers, directors and hall managers. Again, if the movie fails to meet the expectation of the audience, it is bound to flop. It is not possible to know the opinion of all the audiences in a manual method. This method is time-consuming, cost-inefficient and biased. **Failure to collect authentic reviews** from the audience is causing loss to the stakeholders. Again, the users are providing fake reviews facing peer pressure.

The methods we have used to collect data are (a) Direct Observation - which allowed us to gain first-hand experience and understand user experience. Natural observation helps reduce bias but is time-consuming. (b) Indirect Observation- which allowed us to gather a lot of data and specific answers. (c) Interview- Helped us gather qualitative data assuming that the interviewees offered honest answers. (d) The study of existing literature - allowed us to gather knowledge about previous experiences as further discussed. We used Affinity Diagrams to organize our gathered information, finding relationships between concepts or ideas.
Our system will evaluate the quality of a film and the audience's general liking of it based on a few variables: (a)Camera distance, (b) Facial features, and (c) face direction. A cognitive walkthrough is done to evaluate the results.

Methodology

The overall work of this project has been divided into many parts. The project work started with gathering the data. It is followed by the need identification, persona & scenario creation, requirements establishment and prototyping.

Data Gathering

We collected data from a group of potential stakeholders at the beginning of the project.

Direct Observation

As the system will be directly implemented in real-time, we have considered the direct observation method as the primary data gathering method. We have visited the 3 different movie theatres at 3 different times to observe the movie audience during movie playtime. During the observation, we have tried to note down the following factors - (1) The attention of the audience watching the movie, (2) The facial expression of the audience watching the movie, (3) The frequency of talking to other audiences during the movie, (4) Frequency of leaving their seat.

Indirect Observation

We have collected relevant data from stakeholders like Theater Intendant, and the Cleaning staff of movie theatres. The data collected from their opinion has worked as an indirect observation for finding out the user requirements. The main focus of our data collection during this method is - (1) An estimate of people leaving their seats during the movie playtime. (2) A pattern of in-theatre chatter during movie playtime, (3) An estimation of how many people watch the full movie.

Interview

We have also interviewed several people such as movie Cinema Hall Managers, Staff, Audiences and Movie Directors. We asked specific questions to each user type. This helped us know about their personal goals and frustrations which helped us design our required personas and scenarios.

Existing Literature Study

This research paper " ***Factorized Variational Autoencoders for Modeling Audience Reactions to Movies"*** formulated a new non-linear variant of tensor factorization using variational autoencoders, which they called Factorized Variational Autoencoders (FVAE).

After observing an audience member for a few minutes, FVAE was able to reliably predict that viewer's facial expressions for the remainder of the movie. Furthermore, FVAEs were able to learn the concepts of smiling and laughing, and that these signals correlate with humorous scenes in a movie.

They used their approach on an audience facial expression dataset collected from an instrumented 400 seat theatre that hosted multiple viewings of multiple movies over a twelve-month period. This paper is a reference for our own system that will perform in a similar manner with some varying functionalities and applicability in mind.

Identifying Needs

After collecting the required data, we have tried to sum up all the data. As the collected data was not structured, we have considered analyzing the data manually. Brainstorming has appeared as the most effective strategy to identify the user's needs. We have used Miro to create an affinity diagram of the overall system.

Affinity Diagram

The affinity diagram helps to organize the individual interpretation or idea into a hierarchical diagram grouping the data into key issues under the label. It reflects the stakeholder needs and requirements. The affinity diagram echoes the voice of the users. We have chosen to identify the need of the users by using an affinity diagram because it is one of the easiest and most time-efficient methods of brainstorming.

Firstly, we have analyzed all the unstructured data that have been collected during the data gathering process. The key points have been noted down in digital square-shaped sticky notes. After completing this process, we have grouped the sticky notes according to their contents. In this case, we have grouped them into 4 subcategories -
(1) Workflow
(2) Prospectyed Stakeholders
(3) Metrics, and
(4) Goal of the project.

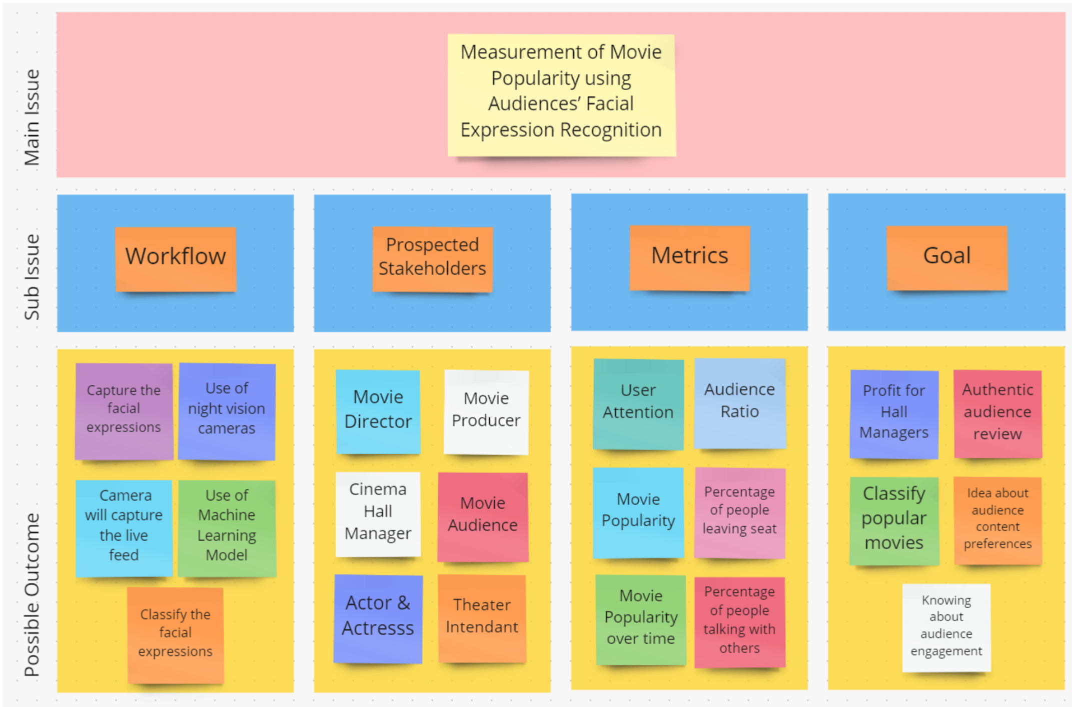


Figure 1: Affinity Diagram

Persona & Scenario Creation

Based on our brainstorming, we have figured out the major stakeholders of our project. I order to capture the characteristics of the stakeholders, personas can be an effective solution. Though personas are not real-life people they are synthesised from real-life users with a focused goal related to the system.

According to the affinity diagram created in the previous steps, we figured out the prospective stakeholders of the systems. Xtensio has been used to design the personas. We have considered 4 personas as our primary stakeholders of the project - movie director, hall manager, hall intendant and movie audience. For creating each persona, we have considered details like personal information, a short biography of the persona, motivation, goals, frustrations, preferred channels and brands & personality traits. Considering the mentioned information, we have assumed scenarios that will suit our personas and their requirements.

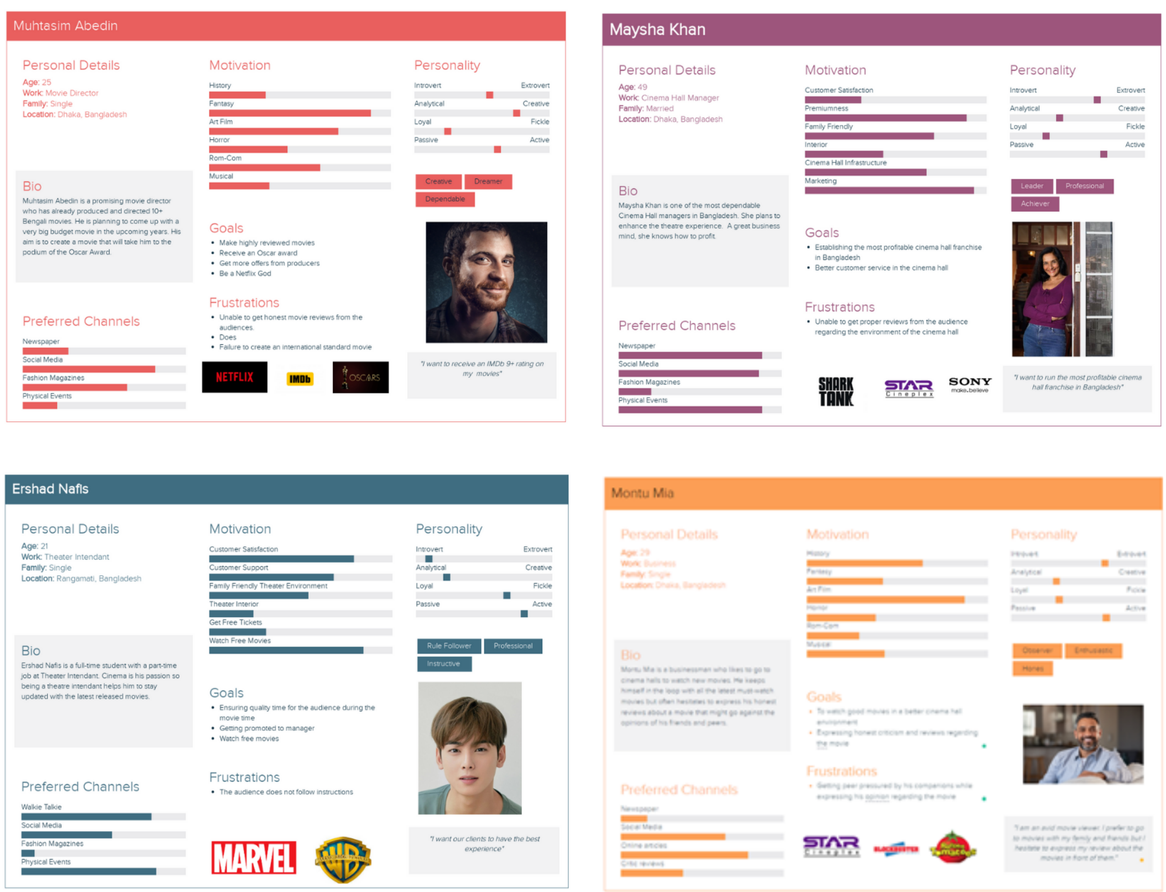


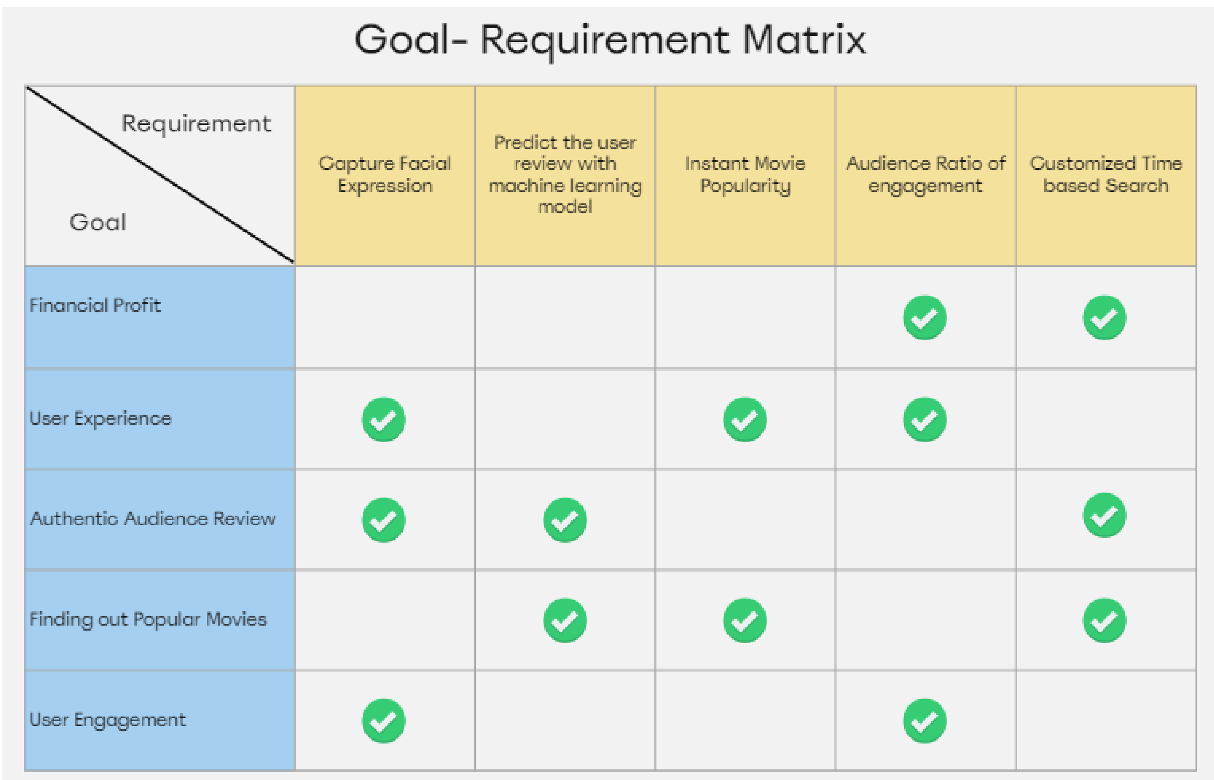
Figure 2: Persona of primary stakeholders

Establishing Requirements

Identifying the needs is not enough for an experimental project. We need to establish the requirements of the user keeping in mind the overall goal of the system. The goal-Requirement matrix is an effective tool to show the relationship between the goal and the requirements of the system. We have used [Miro]() to create a goal-requirement matrix of the overall system.

Goal-Requirement matrix

In order to create the Goal-Requirement matrix, we have listed down the goals of the system column-wise and the requirements of the system row-wise. If we go through the cells of the table, each cell will represent the correlation between the goal and the requirement. The green tick represents the strong correlation between that bisected goal and requirement.



Goal- Requirement Matrix

Requirement / Goal	Capture Facial Expression	Predict the user review with machine learning model	Instant Movie Popularity	Audience Ratio of engagement	Customized Time based Search
Financial Profit				✓	✓
User Experience	✓		✓	✓	
Authentic Audience Review	✓	✓			✓
Finding out Popular Movies		✓	✓		✓
User Engagement	✓			✓	

Figure 3: Goal-Requirement matrix

Prototyping

Considering evaluation and feedback as the central motive of the interactive design, prototyping is very important to explore its suitability. It offers users a limited representation of the system with that users can interact. It helps to test ideas and encourages their reflection on them.

We have prepared a wireframe of our system to clarify the functional requirement of the system. It gives us an idea about the design compatibility of the users as well. The wireframe is considered a low fidelity prototype because of its quick, cheap and easily changeable features. Though it does not look exactly similar to the final product, it comes up with alternative designs and ideas for the system.

We have used balsamiq to create a rapid low fidelity prototype of the system. Our prototype consists of 3 screens of the system. The Home Screen (Figure 4) represent the summary of the audience's attention to the movie. Here users can select the camera from a drop-down menu to see the instant live feed of the theatre. If the 'See Details' button of the Home Page is clicked, the user will be redirected to the Analytics Page (Figure 5). The 'Analytics Page' displays important analytical data like movie popularity over ti,e, audience ratio etc. If the user wants to get access to more customized information, they can click on the 'Customized Search' button which will lead them to the 'Data Filter Page' (Figure 6). The user can set his preferred date and time to perform a customized search.

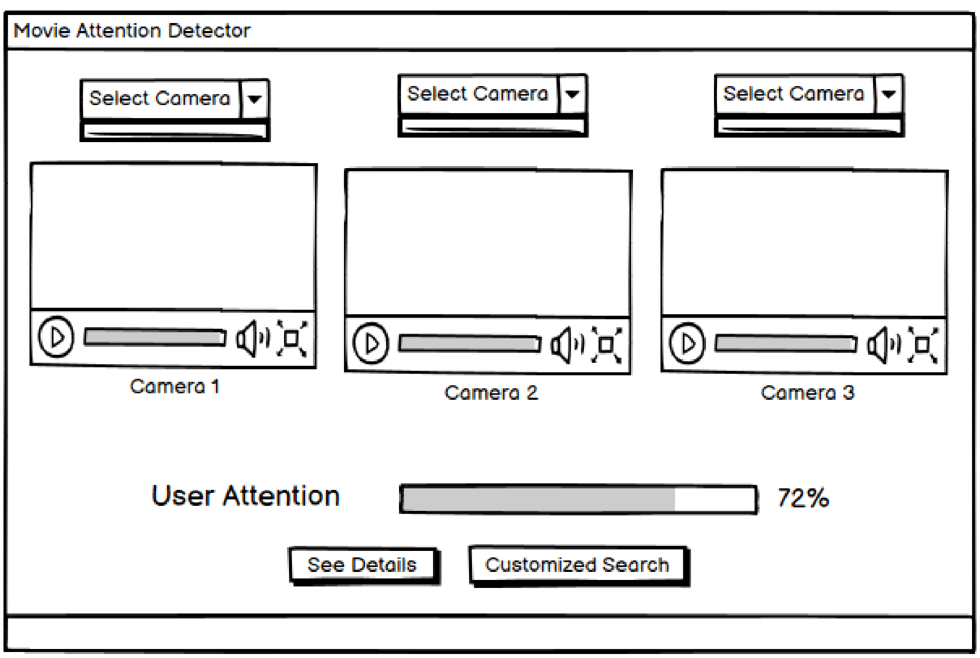


Figure 4: Wireframe of the 'Home Page'

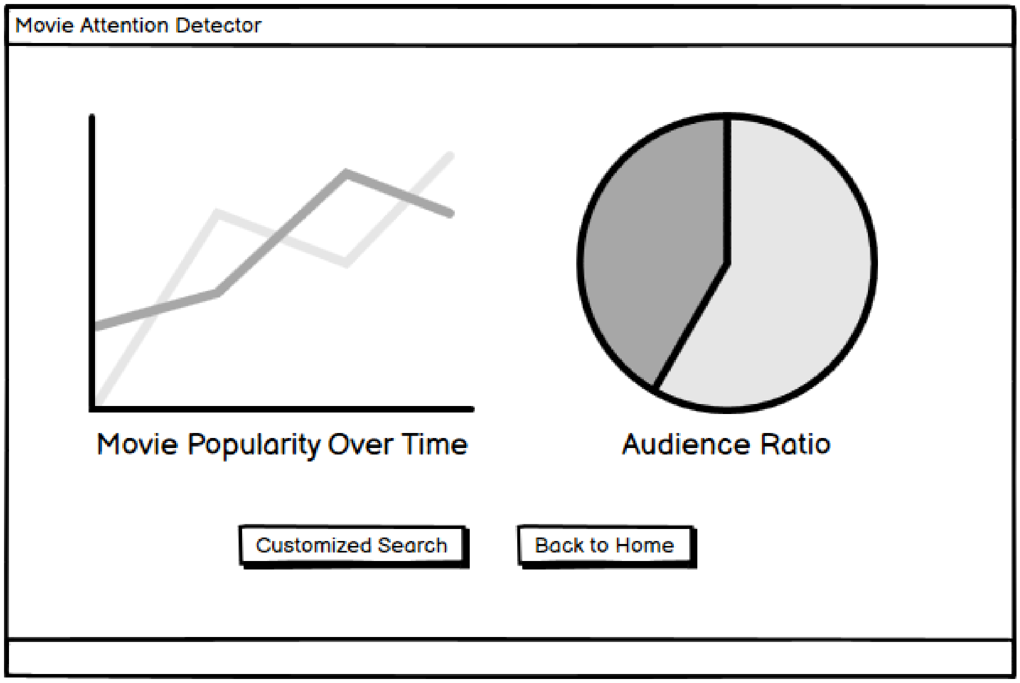


Figure 5: Wireframe of the ‘Analytics Page’

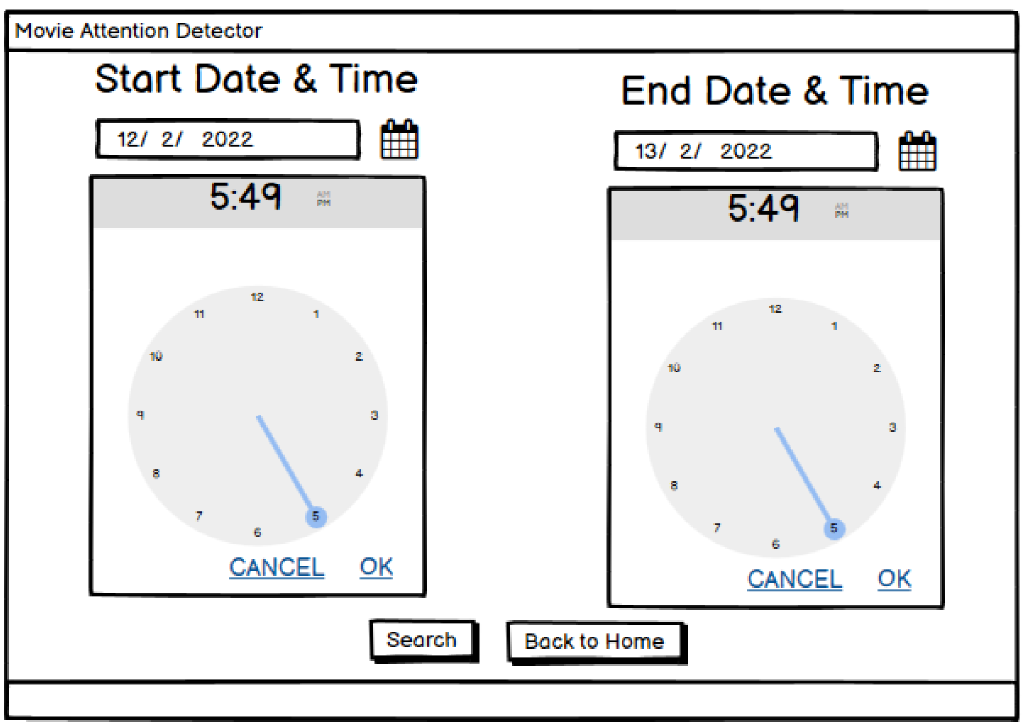


Figure 6: Wireframe of the ‘Data Filter Page’

Mechanism

Our proposed system is completely dependent on the video feed received from the night vision camera set up in the theatre. A night vision camera is set up in the theatre in different positions to get the continuous and live video feed of the audience. The captured live video feed is passed to a trained and advanced machine learning model which can recognize facial expressions. The machine learning model measures audience engagement using various metrics. Finally, the system shows the calculated result of the movie's popularity on the user interface.

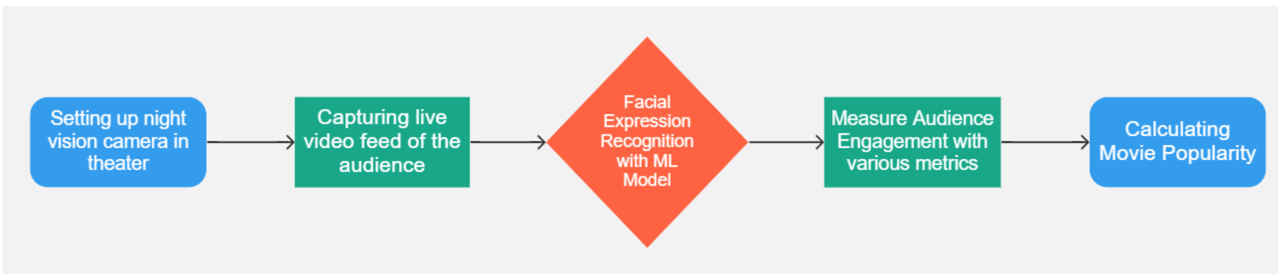


Figure 7: Flowchart of the working mechanism of the system

Experimental Evaluation

Our proposed system does not have an implemented form or a high-fidelity prototype as of yet, but the experimental design and method of evaluation that can be followed to evaluate the implemented form in the future are as follows.

Experimental variables

- Independent variables: Camera distance, facial features, face direction
    - Camera distance
        - Test conditions: Distance from front rows, distance from middle rows and distance from back rows etc.
    - Facial features
        - Test conditions: Facial expression, facial hair, presence of spectacles etc.
    - Face direction
        - Test conditions: Looking directly at the camera or at an angle
- Dependent variables: Reaction prediction
- Control variables: Low-light environment (since cinema halls have low-light environment during movie playtime)

- Random variables: Sleeping audience, distracted audience resulting in facial expressions not related to the movie being shown etc.
- Confounding variables: Different types of cameras are used for different distances of audience seats in the hall, and different face detection models because of different resolutions of cameras for front and back row seats.

Evaluation method to be followed

Cognitive walkthrough method could be used for evaluation as it would be a better fit for our system. The expert will be given assumptions about the user population (which will mainly be movie directors, producers and hall managers), the context of use (to determine movie popularity using the analysis of audiences' facial expressions) and related task details. The experts will then determine if the correct action is sufficiently evident to the user, if the user notices that the correct action is available, and if the user can associate and interpret the response from the action correctly. Using this method, the usability of our system will be sufficiently evaluated.

Conclusion

Our proposed system captures the facial expressions of the audience throughout the movie playtime in the cinema hall. The cameras take the live feed of the facial expressions of the audience and then the system classifies the facial expressions and provides an overall decision on how much the audience is enjoying the movie. The system shows the percentage of the audience who liked the movie and vice versa and also informs which part of the movie was enjoyed most by the audience.

During our data-gathering phase, we found out that around 90% of movie directors and producers are frustrated with the fake reviews regarding their movies and around 35% of the movie audience feel peer pressure while expressing their honest opinion regarding the movie. Our system has the scope to tackle these problems and help the movie directors, producers and cinema hall managers to know about the acceptance of any movie directly from the audience.

As for future work, we could incorporate a few more features to our system that would help in increasing the accuracy of the system's prediction and we can build up a high-quality dataset that would assist researchers in the future to do further work in this field and also for developing AI systems for various relevant purposes because understanding human behaviour is fundamental to developing such systems.

References

- Z. Deng et al., "**Factorized Variational Autoencoders for Modeling Audience Reactions to Movies**," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 6014-6023, doi: 10.1109/CVPR.2017.637.

- O. M. Rajpurkar, S. S. Kamble, J. P. Nandagiri and P. J. Bide, **"A Survey on Engagement and Emotion Analysis in Theatre using Thermal Imaging,"** 2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA), 2020, pp. 905-911, doi: 10.1109/ICECA49313.2020.9297656.

- Almeida, João, et al. **"Emotion Identification in Movies through Facial Expression Recognition.**" Applied Sciences, vol. 11, no. 15, 25 July 2021, p. 6827, 10.3390/app11156827. Accessed 25 Oct. 2021.

- Mezzofiore, Gianluca. "Disney Is Using Facial Recognition to Predict How You'll React to Movies."
  Mashable, 27 July2017,mashable.com/article/disney-facial-recognition-prediction-movie