



Affective Movie Evaluator

PROJECT EVALUATION

Ibrahim Mohamed Shaatha

Submitted to the

**FACULTY OF COMPUTING AND DIGITAL TECHNOLOGY
(SCHOOL OF INFORMATION AND COMMUNICATION TECHNOLOGY)**

**In partial fulfilment of the requirements
for the degree of**

HUBSC2

HELP UNIVERSITY

JANUARY - APRIL 2019

Project Evaluation

1 Introduction

I am working on a system that is going to analyse an audience's reaction from a video and try to predict whether they are enjoying a movie they are watching. We spent the first half of our semester trying to analyse the problem, and design the system and we started working on the implementation on 11th of March.

So far, we were able to successfully complete most vital tasks we set out and planned for this iteration, though the goals and objectives for the iteration were kept as simple as possible to give us a timeframe to learn. The completion of this portion of the project means that, we are halfway towards achieving our primary goal, and also that we have completed roughly half of the project.

2 Evaluation of Objectives

In our project proposal we set out 5 different objectives. Our main goal has always been to create an application that can analyse a video file and give it a score that represents audiences sentiment for the movie. In order to achieve these goals we set out objectives in our proposal, which we still believe are simple, achievable, though in retrospect we should have re-written the objectives in a more clear and concise terms for clarity.

2.1. To create metrics by which the system gives score.

In order for our idea to work, we need to define how the system is going to rate the audience reaction. In other words we need to establish what exactly is our system going to output, once the analysis is complete.

We considered categorical rating and numerical ratings, while we were creating this objectives. What we didn't give much thought was about the fact, that this metrics is going to affect how the machine learning system was going to work. For example, if the metrics is categorical like A, B, and C the most suitable machine learning algorithm we choose for analysis is a "Classification" type of algorithm. If it is a continuous score like a percentage, we can treat the whole thing as a regression problem.

Building upon the existing works, we have considered to use a continuous scale between 0-1, but we have not finalised the metrics we will use and this objective is still not considered achieved.

2.2. To research on the requirements of the dataset and methodologies before March.

Before we even started requirement analysis, we actually thought researchin on existing projects, datasets and methodologies will give us a good idea of how we are going to proceed. While it was beneficial for the project, in the end it took us back by one week, because we were getting used to how we are going to manage the project.

This object was achieved altogether, once we completed our literature review.

2.3. To Create and Collect Datasets for testing and developement

To clarify all datasets with the exception of “Video Recordings” datasets, are supposed to be used in all iteration for testing. Information extracted from our “Video Recordings” will be used to train and develop our movie evaluation ML model.

Unfortunately, this objective is considered partially complete. It was really supposed to be completed by this iteration but we had several issue collecting the dataset, because most of the datasets we collected did not meet the testing requirements . Here is the status:

1. Facial Dataset with labelled Emotion: Pending
2. Pose Dataset with body pose labelled: Pending
3. Video Recordings of audience reacting to Movie: Completed

We are able to complete the objectives we set for dataset “video recordings” of audience reacting, because our criterea for this iteration was to record atleast 3 subjects watching more than two movies. We were able to record more than that.

One of the reason for failing to achieve this objective completely is that we actually put a really low priority on all data collection and during our planning stage we let the system development tasks take precedance over them.

2.4 To determine accuracy of the project by comparing its output to existing scoring systems.

The purpose of this objective is to let us know that our system is working as we planned. A better accuracy score means the project is more successful than we anticipated, mediocre score means that we may still have to work on it until we get better results.

We did not set out to complete this objective for FYP 1. It is possible the project might meet this requirement be completed after the planned third iteration.

2.5 To produce accuracy reports for the analysis subsystems

We have to produce the accuracy report for the two pre-trained models, FER and BEGR. We need to do this in order to evaluate the accuracy of the overall pipeline and choose better trained models if possible.

Accuracy report for FER was initially planned, but we decided to move it to the iteration two, because of lack of space and priority over the system development task. Our priority was to build a “Demo” app for the presentation, and we allocated most of our resources on that task.

3 Further Improvement

We mixed SDLC phases and SEMMA terminologies, for iteration 1. Iteration 1 is mostly system development, while we expect iteration 2 to be more balanced and the final iteration to be mostly focused on the machine learning aspect. Therefore, we will likely not use SDLC names for phases later on during our project.

During analysis, we had difficulties measuring the time it will take for us to work on tasks and we actually ended up over-estimating time, except for three tasks, which gave us a lot of slack time. I believe a better approach would have been to consider the work items and measure the size and time relatively instead of by numbers. Then spend more time figuring out how long we actually take for the simplest task. Then try to estimate by comparing from there on.

4 Future Work

In the next few iteration, I plan to spend more time on the Machine Learning algorithm itself. There are two main parts of the system, I plan to improve. One is the BEGR system needs to be trained to possibly give an emotion on a valence arousal scale. The other is that a movie evaluation system must be developed and trained by us. The system already have undocumented and extra features such as the ability to download movies from YouTube, I do also plan to add more interesting features as such so that the application may be a tool that holds the user within the systems as much as possible.

5 Conclusion

Working on and developing this project gave me knowledge and experience on two different fronts. One is the python based software development and the other is the use of ML and Computer Vision technologies, such as pose estimation models. I learned more details on how deep learning models can be integrated with software designed with object oriented tools.