**Certified Specialist in Machine Learning and Artificial Intelligence**

**Exit Test Part 2**

**Highlights Extractor**

* **Abstract**

Video analysis is extensively being used for several applications in different domains like sports, entertainment, healthcare, etc. For instance, it helps to perform in-depth analysis of scenarios and human behaviours to learn from or plan strategies in the case of sports.

* **Aim**

An application to extract the identified scenarios – Goal moments, Happy moments

and Loss moments – from the input video. The extracted video segments would be of 30 s to 1 min duration and would be capturing the emotions of the people in the video as well. Extraction should happen in such a way that if “goal” is at “x” min in the video, the extracted segment should start from “x-y” min of the input video and end at “x+y” where “y” is the time in seconds to be considered before and after the shot.

* **Supporting diagram**

Output

* Goal moments
* Happy moments
* Loss moments

Input video

Highlight extractor



* **Approach**

The analysis of Video Highlight Extractor was done using two steps:

**1.Emotion Tracking**

**2.Object tracking**

**1.Emotion Detection**

It was implemented using Opencv deep learning techniques and python

The vocal inflections and facial expressions in the visual data, along with the language appearing in a textual transcript, provide important cues to better identify affective states of opinion holders, creating a more robust emotion recognition model. Though the primary focus of the experiment I will summarize below was to classify sentiment in videos (positive, neutral, or negative), their findings can be incorporated into the more specific task of emotion classification across a wider spectrum of basic emotional categories (anger, happiness, sadness, neutral, excitement, frustration, fear, surprise, and other).

**MobileNet**, a predefined model was used for Emotion detection

MobileNet is a streamlined architecture that uses depth wise separable convolutions to construct lightweight deep convolutional neural networks and provides an efficient model for mobile and embedded vision applications.

**Keras Image Data Generator** is used for getting the input of the original data and further, it makes the transformation of this data on a random basis and gives the output resultant containing only the data that is newly transformed. It does not add the data. Keras image data generator class is also used to carry out data augmentation where we aim to gain the overall increment in the generalization of the model. Operations such as rotations, translations, shearin, scale changes, and horizontal flips are carried out randomly in data augmentation using an image data generator.

**ModelCheckpoint** Callback to save the Keras model or model weights at some frequency. Model Checkpoint callback is used in conjunction with training using model.fit() to save a model or weights (in a checkpoint file) at some interval, so the model or weights can be loaded later to continue the training from the state saved.

**2.Object tracking**

Object tracking is the process of locating a moving object in a video.

Object tracking can be performed using Machine learning as well as deep learning-based approaches. The deep learning approach on the one side provides better results on complex tasks and is pretty generalized, requiring a lot of training data. Whereas ML-based approaches are quite straightforward but are not generalized..

Videos are made of frames. Frames are nothing but one of many still images that together make up the whole moving picture. The next step will be reading those frames using the VideoCapture() function in OpenCV and using the while loop, we can see the frames moving. You can adjust the speed of the video using cv2.waitKey(x) which pauses the screen for x milliseconds.

The UI was built using streamlit app and uploaded video was converted to set of frames.Those frames were stored and then those stored images to be analyzed with the emotion detection model and object tracking model which will set a output whether the frame highlights Happy,Goal or loss moments