AIGES RIDER CAPSTONE PROJECT

APPROACH TO USER STORIES SIX AND SEVEN

1.0 Problem Statement

1.0.1 Building a neural network that summarizes ride videos focussing on curves.

As a user I would like to see a summary of my ride video, Therefore I need a neural network that summarized my ride video. I enjoy seeing curves in my ride video, hence I wish to see more curves and turns in my summarized video

1.0.2 Build a neural network that summarizes ride videos focussing on speed.

As a user I would like to see a summary of my ride video, Therefor I need a network that summarizes my ride video. I enjoy seeing top speeds in my ride video; hence I wish to see portions of top speed in my ride summarized video.

1.1 Method

This assignment comes with two modes of datasets. The first dataset is portions of videos recorded from a moving motorcycle. The second video a tabular data from sensor reading of the moving motorcycle. The tabular data contains the recordings of the speed and curvature of the motorcycle per second at every point in the video.

To get the dataset ready for the model, there must be some pre-processing steps on both the tabular and video data. It was observed that to build a model for video summarization, the data must contain certain attributes and features. The videos in the data were all 4 minutes and 9 seconds long. There were ten sets of videos in the folder to the model training. I observed that the best possible way to build a summarization model from a sensor reading will require a supervised machine learning approach that will label from the dataset. I choose the speed and curvature column in the dataset as the labels for the dataset. The speed column corresponds to the speed associated with a single frame in the video at a specific time. The main idea of this project is to build an A.I model that can select keyframes of a video based on the speed and curves from the sensor readings. Below is an outline of the steps to going through the thought process in building the model.

1. To build a robust dataset for the model. Five of the videos from the Appenzell sensor were concatenated to form a single video. The corresponding tabular data were also joined together.
2. The video was then sectioned into frames by the OpenCV library.

3. By using a threshold of the seventy-fifth percentile of the dataset, a manual summary was generated from all the segments of the videos.
4. The Google net pretrained model was utilized to learn the features of the dataset.
5. The various components of the video include:
    a. the feature which is the features of the frames in the video data.
    b. Label which includes the scores of the speed recorded by the sensor in the motorcycle.
    c. Length: this is the length of the entire video data.
    d. Change points: This component is the number of frames per segment in the video dataset.
    e. The number of frames per segment is also another component that records the number of frames in each segment of the dataset.

These components were saved into Hierarchical Data Format version 5 (HDF5). The HDF5 data format is an open-source data format that allows saving large and complex data into a structured format.

The approach to building this summarization model was to adopt an image segmentation model. This is to give scores to the various frames in the video dataset. The scores here were taken to be the curvature and speed readings from the sensor. The scores were then mapped to the various frames. The model is trained using the frames and their corresponding scores. The appropriate model for this video summarization task was the Fully Convolutional Sequence Networks.

Steps in building the model

1. Initialize the weights
2. Convert a video to a feature
3. Make the AIGES summary dataset
4. Calculate the Knapsack distance to a given value
5. Convert a video to feature format.
6. Compute the aggregation summary of each user.
7. Make a dataset from the video dataset
8. Get the training data
9. Calculate oracle score for the given user summary
10. Select keyframes from the video