**PROPOSED TECHNIQUE FOR LGSI PROJECT**

**AI-based audio/video quality checking - solution to measure the quality of audio/video**

Animesh Raj Zayed Haque Ponnuri Aniruddha

RA2112704010005 RA2112704010010 RA2112704010015

CSE DATA SCIENCE CSE DATA SCIENCE CSE DATA SCIENCE

**Objective For the Project**

The intention is to create a solution that can automatically analyze audio and video files and provide a quality assessment.

**Steps for implementing BiLSTM for Audio Quality Assessment**

1. Extract features from the audio signal. The first step is to extract features from the audio signal. This can be done using a variety of techniques, such as MFCCs or spectrograms. MFCCs are a popular choice for audio quality assessment because they are able to capture the pitch, loudness, and timbre of the audio signal. Spectrograms are another popular choice for audio quality assessment because they can capture the frequency content of the audio signal.
2. Feed the features to a BiLSTM. Once the features have been extracted, they can be fed to a BiLSTM. The BiLSTM will create hidden states for each frame of audio. The hidden states for a frame of audio represent the meaning of that frame of audio in the context of the surrounding frames of audio.
3. Apply an Attention Mechanism to the hidden states. The next step is to apply an Attention Mechanism to the hidden states. The Attention Mechanism will create a weighted sum of the hidden states, with the weights representing the importance of each frame of audio. This allows the model to focus on the most important parts of the audio signal when predicting the audio quality.
4. Use the weighted sum of the hidden states to predict the audio quality. The final step is to use the weighted sum of the hidden states to predict the audio quality. This can be done using a variety of techniques, such as a linear regression model or a neural network. We test both the models and depending performance and deployment requirements we choose the best suitable one.

**Steps for implementing BiLSTM for Video Quality Assessment**

1. Collect a dataset of video recordings with known quality labels. The dataset should be large enough to train the BiLSTM model. The quality labels can be obtained from human raters or from objective quality measures.
2. Extract features from the video recordings. The features should be representative of the video quality. Some common features for VQA include DCT coefficients, edge features, and motion features.
3. Choose a BiLSTM architecture. There are many different BiLSTM architectures such as Multi-layer BiLSTM, BiLSTM with pre-trained weights, BiLSTM with attention mechanism, BiLSTM with transfer learningthat can be used for VQA. The best architecture will depend on the size of the dataset, the complexity of the features, and the desired accuracy.
4. Train the BiLSTM model. The BiLSTM model can be trained using a variety of machine learning frameworks, such as TensorFlow or PyTorch. The training process can take several hours or days, depending on the size of the dataset and the complexity of the BiLSTM model.
5. Evaluate the BiLSTM model. The BiLSTM model's performance can be evaluated using a variety of metrics, such as the mean squared error (MSE) or the root mean squared error (RMSE). The evaluation process can take several minutes or hours, depending on the size of the dataset.
6. Deploy the BiLSTM model. The BiLSTM model can be deployed in a variety of ways, such as on a web server or on a mobile device. The deployment process will depend on the specific application.