B20AI016_SU_ASSIGNMENT_03

Audio Deepfake Detection Evaluation

Task 1: Results

For Task 1, the SSL_W2V model, pre-trained on LA and DF tracks of the ASVSpoof dataset, was evaluated on a custom dataset. The AUC and EER metrics obtained are as follows:

Metric Name	Value
AUC	0.5229
EER	0.4918

Task 2: Performance Analysis

The performance of the SSL_W2V model on the custom dataset indicates challenges in effectively distinguishing between real and fake audio samples. The model exhibits limited discriminative power with an AUC of 0.5229, only marginally better than random guessing (0.5). The EER of 0.4918 also suggests an imprecise balance between false positives and false negatives, indicating unreliable classifications.

The observed subpar performance could stem from a mismatch between the training and custom datasets. While the model was trained on the ASVSpoof dataset, the custom dataset may introduce different audio manipulations or conditions not encountered during training. This discrepancy hampers the model's ability to generalize well, resulting in decreased performance.

To address this, fine-tuning the model with the custom dataset is recommended. This process allows adaptation to the dataset's specific characteristics, enhancing classification accuracy. Exploring alternative model architectures or additional training data could further improve performance.