

Rohan Nain

2020102023

1. Literature:

I have used references, help of batchmates to discuss the problem and come up with a submission.

2. Approaches Studied:

Wishper was fine tuned on the given dataset domain. Further we can add a LLM block to even make the ASR give the prompt also not only just the transcription.

Similarly we can tune the wav2vec2 with the same approach.

3. Findings and results:

```
[ ]: trainer.train()
```

[1804/4000 1:48:51 < 2:12:40, 0.28 it/s, Epoch 100.17/223]

Step	Training Loss	Validation Loss	Wer
1000	0.000000	0.004393	34.390486