

# CS 583 - Deep Learning Final Project

## Automatic Speech Recognition

SHIVANI MOGILI- 10473465

Dataset: <https://data.keithito.com/data/speech/LJSpeech-1.1.tar.bz2>

READ ME:

### TRANSFORMER MODEL

In the transformer, multi-head self-attention is used for representing the input sequence and the output sequence, though the decoder has to preserve the auto-regressive property via a masked version. Both the residual connections and the layer normalization in the transformer are important for training a very deep model. The position wise feed-forward network in the transformer model transforms the representation at all the sequence positions using the same MLP.

### STEPS TO RUN THE PROJECT -

#### Step 1-

Google collab has been used to run the model ,I used transformer model and its taking more time to run on the local machine.()

#### Step 2- Download all the packages

Pip Install TensorFlow and all the other packages are mentioned in source file.

#### Step 3- Download the Data-Set

Keras.utils.get\_file()

Link : <https://data.keithito.com/data/speech/LJSpeech-1.1.tar.bz2>

#### Step 4- Train the model

batch = next(iter(val\_ds))

#### Step 5- For inference

class Transformer(keras.Model)