

# CSC-8637(Deep Learning)

Rohit Kumar(C2052473)

Semester 2 - 2023

---

## Text Generation

**Data Processing:** The data is extracted from the given text file and is passed through a series of operations where the text is cleaned of punctuations and special characters. The data was then split into spaces and stored in a list named `text_corpus` and a dictionary of unique words was made. The list of words- `text_corpus` was converted into a list of sentences by having 31 words in each sentence. Each training pattern of the network comprises 30 steps of one character (X) followed by one character output (y). When creating these sequences, you slide this window along the whole text, allowing each character a chance to be learned. In the final step, tokenized the sequences and then converted the tokens into NumPy arrays.

**Modeling:** Long Short term memory was used to generate the text. In this, the embedding layer is used as the input layer which takes the tokenized data followed by two LSTM layers of 64 units each and one dense layer of 128 nodes followed by an output layer.

**Hyperparameters:** The Adam optimizer was employed with a learning rate of 0.001 due to its versatility. For the activation function, Rectified Linear Unit (ReLU) was chosen because it adds non-linearity to the network without consuming excessive computational resources. Also, we have used `sparse_categorical_crossentropy` instead of `categorical_crossentropy` since it does not require one-hot encoding of the target variable which was crashing the memory. To allow for the learning of sufficient features and their storage in memory, a batch size of 32 was selected.

**Input:** *"He moved slowly and with a kind of painful dignity, as a man moves on his way to the firing squad. A rumpled shock of black hair pointed up the extreme pallor of a gaunt face"*

**Output:** *'days on that instantly on the really was item the couldnt or who youll the else you out did he you warily of singhalut the true huge almost left shrugged answer walked the knocked chairman to which against to with had possibility one miles after then from laugh and thousand'*