Aim: accomplish text to image process

Generate image and provide CLIP score

The corresponding agent will enhance it

4 types:

Animal, Landscape,

Portrait, others

Trained LSTM Model to recognize the type of sentence

Input the sentence