# Hackathon Report: Gricean Maxmemes

## IIIT-H Confessions Generator based on Neural methods

### Introduction

This is a web-app for generating confessions, using a LSTM Recurrent Neural Network with a dataset of 20k confessions (pre-cleaning) scraped from Facebook Confessions Pages of various universities. The language model used the pre-trained GPT-2 Model, by OpenAI. For better speed and usability, the entire web-app was based on vanilla JavaScript, and uses the pre-generated dataset.

Go to the website https://prajneya.github.io/Confessions-Generator/ where you can add parameters such as name and length of confession and it will generate one.

#### Dataset

We scraped ~21k Confessions from Facebook pages of various US Colleges, using a Selenium Bot. This was done in reference to this project on mental health. This process took time due to the anti-bot features of Facebook which we circumvented using timeouts and random behavior. This data had to be majorly cleaned as it was not in the UTF-8 Encoding standard we were basing our model on. Thus, we used Pandas to read the csv and encode it in UTF-8, and then only use ASCII characters from the data left. We ran multiple regular expressions on this model for cleaning, and getting read of residual encoding errors, URLs, and so on. Further on, to avoid dealing with the complex challenge of code-mixing, we used a modified version of the NLTK English Language Wordlist, adding text-speech to properly illustrate college language, and then created a Non-English Rate, and removed all sentences with N-E rate above 30%. This was chosen as several words like "sooooo" or such were detected as N-E, and we wished to preserve them. This brought our dataset down to 13k sentences. This data was stored as quoted strings due to errors from automatic delimiters in CSV renderers such as Libre. We used regular expressions to "nativize" our data for consistent content words (such as IIIT instead of all colleges). This was done by checking the data and regularly running it through various such scripts.

### The Neural Network && The Database Generation

We have used a LSTM model with embeddings to capture character level dependencies. The embeddings layer converts the words into 256 dimensional tensors, which are then fed into the LSTM layer. The LSTM layer takes into account all the previous characters it has seen till now. The speciality of the LSTM cell is that it is able to capture long ranged dependencies as well, as is seen in the case of gender inflections. The short term dependencies helps it get the spellings correct. At the end, we put the returned sequences into a Dense layer, which now predicts the character to be output next. After the model has been trained for a sufficient number of epochs, we start our database generation. We now give the model an empty string to work upon. We render a string with a character-generation model, which incrementally adds character as generated by our recurrent neural network. We used a pre-trained language model by OpenAI, specifically the GPT-2 model for a transfer learning approach on our finetuned dataset for better adaptability and accurate word embeddings. Following this procedure, we can proceed with the previously mentioned LTSM approach.

### Website

To display our model, we designed a Vanilla web application, which can be accessed at https://prajneya.github.io/Confessions-Generator/ . This web-app serves as a single page application, where the user can input a name on which they desire the confession to be generated. We do not run the model each time the user wants to generate a confession, since it would lead to a huge consumption on the backend, making the application slower, especially on devices with low-bandwidth internet connections. Recognizing this problem, the model generates a large number of confessions, which are stored in a .json file. The website fetches confessions from this file and uses local browser cookies to remember which confession has already been displayed, thus displaying a new confession each time.

### Screenshots from the website
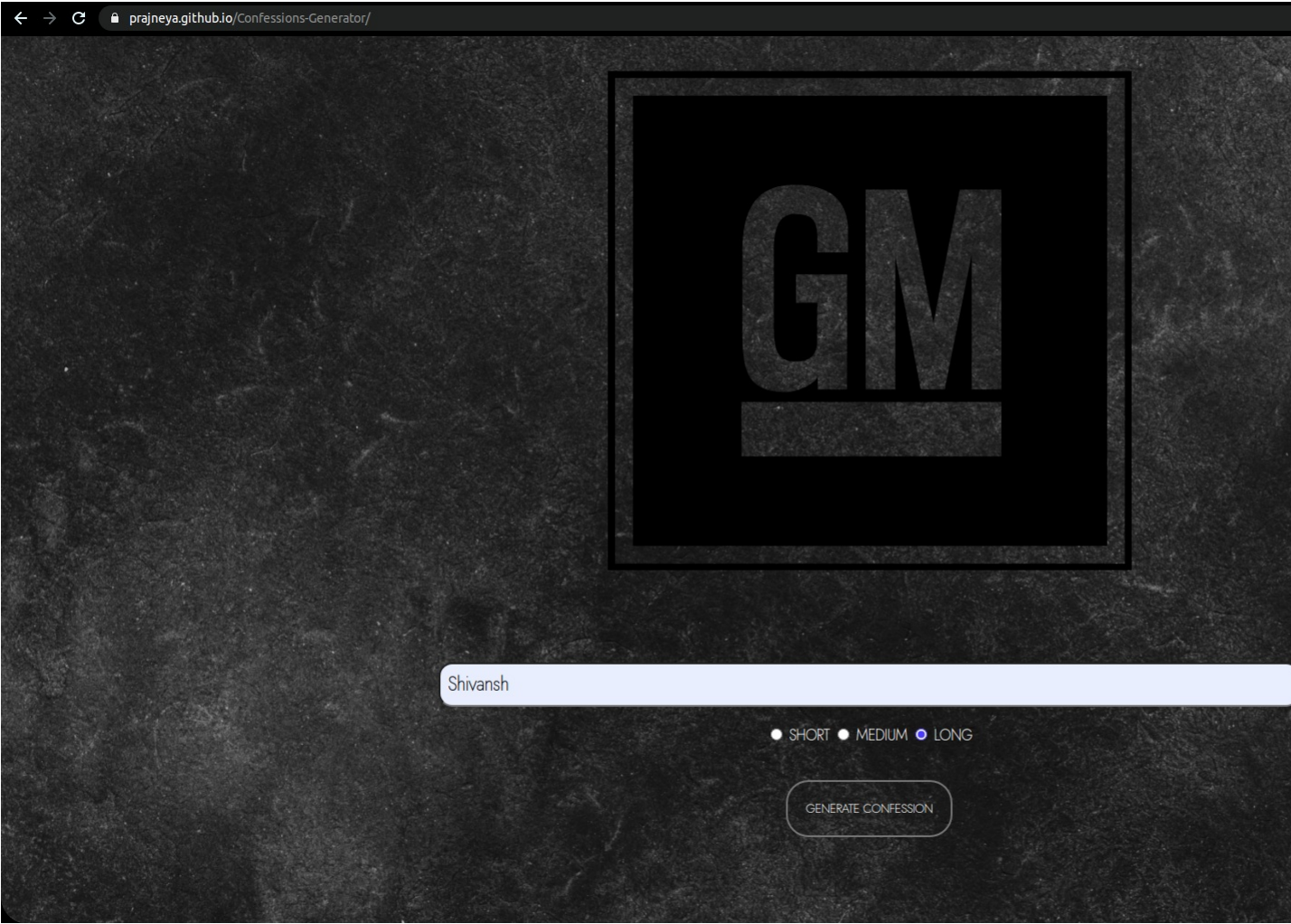
Dear prajneya, It was the most disappointing experience of my life . I want to be a feminist but I feel like I am too stupid to think about it . I feel like I am too smart to think about it . I am a bit of a guy , but still very feminine . I wear bright colors and I am so stupid . I am a virgin , but I am a virgin . I am a huge admirer of the girl that I love at iit ! I am the only one in the group that is having a hard time with my GPA . I am a freshman in an Ivy league college . At iit , I have the highest GPA and I am the only one who is having problems with getting into college . I'm having trouble getting into my classes , but I really want to get into them . I actually wish I found a mentor of my own . I am a freshman in an Ivy league college . I have the highest GPA and I am the only one who is having problems with getting into college . I am a freshman in an Ivy league college . I have the highest GPA and I am the only one who is having problems with getting into college . I am the only one who is having trouble at iit . I am a freshman in an Ivy league college . I have the highest GPA and I am the only one who is having problems with getting into college . I am in a relationship . I am a virgin and I am having some problems with getting into college . I am amazing in my classes . I have been in relationships for a long time and I feel very fortunate to have had so many wonderful relationships . Is this the only reason I am a virgin ? Maybe not and I don't know . I am an Ivy League college freshman . I am currently enrolled in A.S.M.U . I'm having a bad day today . I don't know what to do . I am tired . I felt like I was in a lot of pain and has been getting really bad grades all semester . I'm now in a depression . I don't know what to do . I think I am having suicidal ideation , but I don't know . I don't know what to do . I think that there are so many people who are not on this page . I am still feeling depressed . I am so sorry . I don't know what to do . I feel like I am in a lot of pain and I'm scared to go somewhere because I am just so sick . I'm addicted to DEET . I dont even know how to put it together . I am still addicted . I'm a freshman in an Ivy league college . I have the highest GPA and I am the only one who is having problems with getting into college . I am a male . I am a virgin . I am trying to get into college but I feel like my GPA is getting in the way of my life . I am in a relationship . I am a virgin . I am having some problems with getting into college . I am a freshman in an Ivy league college . I have the highest GPA and I am the only one who is having problems with getting into college . I am a virgin . I am a virgin . I am having some problems with getting into college . I feel like I am in a lot of pain . I am tired . I feel like I am in a lot of pain and I'm afraid to go somewhere because I am so sick . I don't know what to do . I feel as though I am in a lot of pain . I'm a virgin . I'm having some problems with getting into college . I am a virgin . I am a virgin . I am having some problems with getting into college . I am in a relationship . I am a virgin . I am having some problems with getting into college . I am a virgin . I am a virgin . I am having some problems with getting into college . I'm having some problems with getting into college . I am in a relationship . I am a virgin . I am having some problems with getting into college . I am in a relationship . I am a virgin . I am having some problems with getting into college .

SUBMIT CONFESSION

GENERATE CONFESSION