

## Introducción

En los últimos años la popularización de los chatbots ha incrementado, esto gracias a la inteligencia artificial que estos bots poseen y al machine learning que permite que estos chatbots puedan reconocer lenguaje humano e interpretarlo, aprendiendo de él y aplicarlo a futuras conversaciones. Gracias a esto cada día vemos la inclusión de estos chatbots o bots de conversación en sitios web que nos ofrecen compra en línea con la finalidad de que encuentres lo que estás buscando y puedas comprarlo sin mucha dificultad. A lo largo del tiempo estos chatbots pueden mantener una conversación con una persona real y que esta crea que quien le está respondiendo es otra persona y no una inteligencia artificial entrenada.

## Objetivos

El objetivo de este proyecto es el que nuestro chatbot se haga pasar por un autor reconocido, que la gente lo crea y que pueda conocer más al autor platicando con el chatbot. Esta idea surge como el objetivo de que las personas puedan ampliar su conocimiento de un autor o conocerlo. De igual forma que la gente les diga a sus conocidos y que este chatbot llegue a ser muy popular. Con esto aportamos un plus a la tecnología y la aplicamos de manera divertida para que la gente interactúe y se involucre. La inteligencia artificial permite el desarrollo de eso y se puede pulir con el uso de machine learning.

## Metodología

El data set usado fue la mitad del libro Ojos de Perro Azul por Gabriel García Márquez, se limpiaron los espacios innecesarios y la parte introductoria. Usamos la mitad del texto para evitar el overfitting y por tiempo de entrenamiento. Haciendo uso de este texto se hace un arreglo bi dimensional en el que la primera dimensión divide el texto en 100 caracteres y la segunda dimensión es el carácter que le sigue a esa cadena de caracteres. Es decir:

X	Y
[0, 1, 2, 3, 4]	[5]
[1, 2, 3, 4, 5]	[6]
[2, 3, 4, 5, 6]	[7]

Después de esto se hace un reshape del arreglo y se divide todo el arreglo después del reshape para normalizar los datos.

Para crear la red neuronal construimos el modelo usando Keras. Se uso un modelo secuencial con una capa LSTM, dos capas escondidas igual LSTM con un dropout de 0.2 para evitar el overfitting. Al final tenemos una capa en la que su función de activación es Softmax. Cada capa de LSTM cuenta con 300 neuronas. El modelo fue entrenado durante 50 épocas con un batch de 200.

Este modelo fue entrenado en la plataforma de google cloud para disminuir los tiempos de entrenamiento. La primera corrida fue con 400 neuronas por capas sobre 4 tarjetas NVIDIA Tesla K80, por problemas con el paralelismo de Keras y overfitting, este modelo tuvo que ser desechado. El segundo, que es el que está en uso, se corrió con 300 neuronas sobre una NVIDIA Tesla K80.

## Conclusiones

La generación de lenguaje natural sigue siendo una gran área de oportunidad en el campo del ML. Para esta investigación, se tomaron en cuenta los caracteres para la creación de los datasets, no las palabras. Esto resulta en la generación de algunas palabras mal escritas o inconsistencias semánticas. Un acercamiento que mapee las palabras en lugar de los caracteres crearía texto con un poco más de sentido, aunque sería un difícil reto realizarlo en la lengua Española por la gran cantidad de conjugaciones que existen.