

Generador de versos musicales en el idioma inglés por medio de procesamiento de lenguaje natural y redes neuronales

Espinosa de los Monteros Lechuga Jaime Daniel, Nava Romo Edgar Adrián, Salgado Gómez Alfredo Emilio
Kolesnikova Olga, López Rojas Ariel
Escuela Superior de Cómputo I.P.N. México D.F.
Tel. 57-29-6000 ext. 52000 y 52021.

E-mail: jamesdlechu@gmail.com, edgar.adrian97@gmail.com, alfredoe.sgomez@gmail.com

Abstract—The terminal work presented consists of the development of a web application, which allows users who access it, to generate lyrics of the pop genre through a model made by us, this was created using bidirectional LSTM neural networks.

Index Terms—Bidirectional, LSTM, Music, Neural Networks.

Resumen—El trabajo terminal presentado consiste en el desarrollo de una aplicación web, la cual permite a los usuarios que acceden a ella, generar letras musicales del género pop a través de un modelo generado por nosotros, este fue creado usando redes neuronales del tipo LSTM bidireccionales.

Palabras Clave—Bidireccional, LSTM, Música, Redes neuronales.

I. INTRODUCCIÓN

En la actualidad la industria musical obtiene ganancias a través de la creación y divulgación de la música de manera física y digital [1], dejando que aficionados y emprendedores de la música no tengan oportunidad de avanzar en su carrera, ya sea por la falta de creatividad, tiempo y/o recursos, haciendo que la creación de nuevas letras para sus canciones sea un gran obstáculo. Teniendo esto en cuenta, vimos una oportunidad y haciendo uso de las nuevas tecnologías, decidimos desarrollar una aplicación web la cual permite la generación de letras, las cuales pueden ser usadas posteriormente para integrarlas con sus temas musicales, creando canciones.

La generación de texto es una de las tareas más populares y desafiantes en el área de procesamiento del lenguaje natural. Recientemente, se han desarrollado una gran cantidad de trabajos (Generating Text with Recurrent Neural Networks [2], Convolutional Neural Networks for Sentence Classification [3]) los cuales propusieron generar texto utilizando redes neuronales recurrentes y redes neuronales convolucionales. Sin embargo, la mayoría de los trabajos actuales solo se enfocan en generar una o varias oraciones, ni siquiera un párrafo largo y mucho menos una canción completa.

Las letras de canciones, como un tipo de texto, tienen algunas características propias, estas se constituyen de coros, estribillos, rimas, versos y en algunos casos, patrones de repetición. Se entiende por coro que se trata de una estrofa la cual se repite varias veces dentro de una composición, acentuando la idea más importante de la canción, el estribillo

es un conjunto de palabras o frases con la que se inicia la composición y la cual se repite al final de una estrofa de esta. Mientras que el verso es la parte encargada de comenzar a desarrollar la idea a transmitir, trata de contarnos el tema de la canción. Estas características hacen que generar letras musicales sea mucho más difícil que generar textos normales.

La generación de letras por medio de inteligencia artificial dado un estilo y un tema es un asunto poco trabajado. Por lo tanto, nos enfocamos en trabajar este problema. Estamos interesados en ver si el modelo propuesto puede aprender diferentes características en un género musical y generar letras que sean acorde a este.

II. METODOLOGÍA

Se entiende por Inteligencia Artificial a una simulación de procesos computacionales cognitivos para que simulen el comportamiento de una mente humana, estos comportamientos abarcan diferentes áreas de investigación, como el aprendizaje, comunicación, percepción, razonamiento y la capacidad de desarrollarse en entornos más complejos.

A. Red neuronal

Una red neuronal es una herramienta derivada de la inteligencia artificial que utiliza modelos matemáticos para ser utilizada como un mecanismo de predicción del texto. Son una forma de hacer que las computadoras aprendan, donde la computadora aprende a realizar alguna tarea analizando ejemplos de entrenamiento. Por lo general, estos ejemplos han sido etiquetados previamente.

Las redes neuronales son modeladas vagamente asemejando el cerebro humano, esta consiste en miles de millones de nodos que realizan el procesamiento de la información los cuales están densamente interconectados.

La complejidad de las neuronas reales se abstrae mucho cuando se realiza el modelado de neuronas artificiales. Estas consisten básicamente en entradas, que se multiplican por pesos (fuerza de las respectivas señales), y luego calculada por una función matemática que determina la activación de la neurona. Otra función (que puede ser la de identidad) calcula

la salida de la neurona artificial (a veces en dependencia de un cierto umbral). Las redes neuronales combinan neuronas artificiales para procesar información.[**neuronart**]

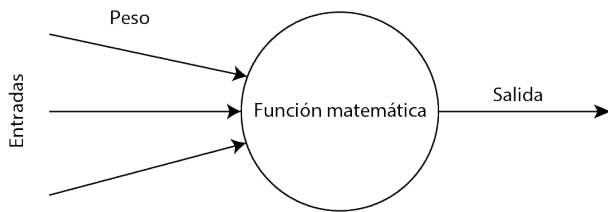


Fig. 1. Neurona artificial.

B. Recurrent Neural Network (RNN) – Long Short Term Memory (LSTM)

Una red neuronal recurrente es un tipo de red neuronal artificial donde la salida de alguna capa en particular es salvada y sirve para retroalimentar la entrada de esta capa, lo cual ayuda a predecir futuras salidas de esta. La primera

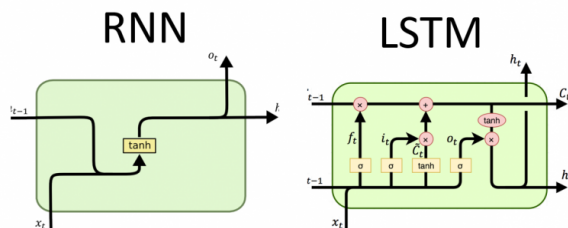


Fig. 2. Diagrama de las redes neuronales RNN y LSTM.

capa está formada de la misma manera que la Feed-forward Neural Network, es decir, solo pasa la información que entra a la siguiente capa inmediata, posteriormente la siguiente capa con el paso del tiempo comenzará a retroalimentarse, pero manteniendo la propagación frontal. Haciendo uso de esta retroalimentación la capa en futuras operaciones puede realizar predicciones, si estas predicciones no son los resultados esperados, el sistema aprende y trabaja para corregir sus futuras predicciones.[**redneurnn**]

Este tipo de redes neuronales se distinguen por su “memoria”, ya que toman información de entradas anteriores para influir en la entrada y salida actuales.

C. Generación del modelo

c.1) Obtención del conjunto de datos

El primer paso para poder empezar a trabajar con nuestro modelo fue obtener un conjunto de datos que cumpliera con las características que se necesitaban, por eso nos dimos a la tarea de buscar uno. En Kaggle [**kaggle1**] se encontró un dataset llamado “Song lyrics for 6 musical genres” [**kaggleDataset**] el cual contenía los datos necesarios y contaba con un total de 160,790 letras, las cuales pertenecían a 6 tipos de géneros musicales:

- Rock
- Pop
- Hip Hop
- Samba

- Sertanejo
- Funk Carioca

Debido a que estas letras de canciones pertenecían a distintos géneros e idiomas se procedió a realizar una limpieza a la base de datos para quedarnos únicamente con las letras de canciones pertenecientes al genero pop y que estuvieran en el idioma inglés, además haciendo uso de la librería de “Pandas” y siguiendo el estándar de limpieza de datos: [**data`cleaning**]

- 1) Remover caracteres innecesarios
- 2) Eliminar Duplicados
- 3) Evitar errores ortográficos de similitud
- 4) Convertir los tipos de dato
- 5) Tratar los valores nulos o faltantes

Se obtuvo como resultado un conjunto de datos útil, el cual solo cuenta con los datos necesarios para después poder tratarlos y utilizarlos para empezar el entrenamiento del modelo de la red neuronal, cabe destacar que este proceso es necesario hacerlo solo una vez para cada género musical ingresado, en caso de que se lleguen a necesitar más datos, será necesario repetir el proceso.

c.2) Modelo

c.3) Entrenamiento

III. RESULTADOS

Si hubo resultados

IV. CONCLUSIONES

Pues estuvo chido el proyecto