

Clasificación de géneros musicales a partir de la letra

Javier Mejía

Universidad del Valle de Guatemala

3 de junio de 2021

Contexto

La clasificación de la música en géneros musicales es un tema de discusión en la teoría musical. Existen múltiples puntos de vista para cómo clasificar una canción como un género u otro. Esto lleva a que las clasificaciones sean difusas, y poco confiables.

Por otro lado, la poesía suele tener una clasificación más confiable.

Por lo que es razonable plantear la posibilidad de clasificar los géneros musicales a partir de la letra de las canciones.

Para el trabajo se contó con 128,083 canciones de 6 géneros, extraídas de Kaggle.

Los datos

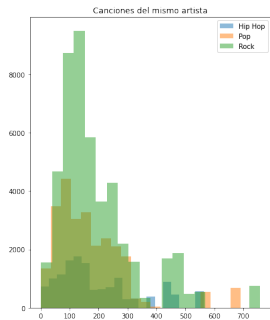
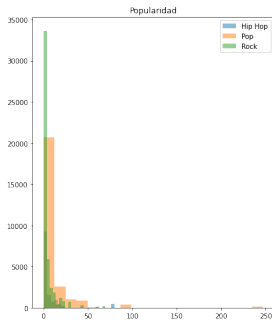
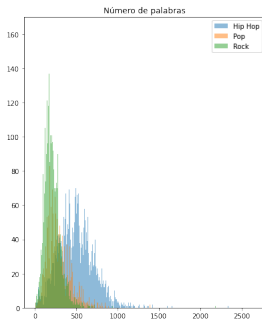
La base de datos inició con 128,083 canciones de 6 géneros. Luego de dejar sólo las canciones en inglés y solo los géneros Hip Hop, Rock, y Pop (puesto que los otros 3 eran atípicos en inglés), y ajustar la base de datos, quedamos con los siguientes datos:

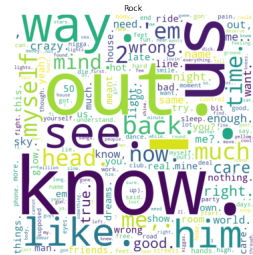
	SName	Lyric	Songs	Popularity	Genre	Number_of_words
0	More Than This	I could feel at the time. There was no way of ...	110	0.3	Rock	145
1	Because The Night	Take me now, baby, here as I am. Hold me close...	110	0.3	Rock	238
2	These Are Days	These are. These are days you'll remember. Nev...	110	0.3	Rock	172
3	A Campfire Song	A lie to say, "O my mountain has coal veins an...	110	0.3	Rock	263
4	Everyday Is Like Sunday	Trudging slowly over wet sand. Back to the ben...	110	0.3	Rock	118
...
87145	Smells Like Teen Spirit	(Chorus). Hello, hello,hello,how low. Hello,he...	71	0.8	Samba	149
87146	Valerie	Well sometimes I go out by myself. And I look ...	71	0.8	Samba	273
87147	Wake Me Up	Feeling my way through the darkness. Guided by...	71	0.8	Samba	251
87148	Don't	Don't, don't, that's what you say. Each time t...	106	4.2	Samba	112
87149	Deixa A Vida Me Levar (translation)	I've been. For almost everything in life. In t...	309	13.2	Samba	378

87150 rows x 6 columns

Distribución de variables cuantitativas

Distribución de variables por género





[illegible]

Preguntas de investigación

- ¿Hasta qué punto es posible reconocer el género musical de una canción solo con su letra?
- ¿Qué información le podemos extraer a la letra de una canción?

Los géneros y supergéneros musicales

Los géneros musicales son categorías en las que se divide la música. Usualmente se clasifican según los aspectos más globales y técnicos:

- Los medios sonoros empleados.
- Su función.
- Su contenido.

Los géneros y supergéneros musicales

Utilizando la clasificación usual, se puede clasificar toda la música en inglés en 6 grandes supergéneros:

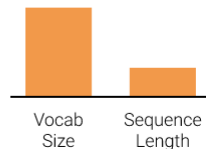
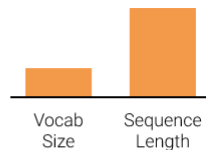
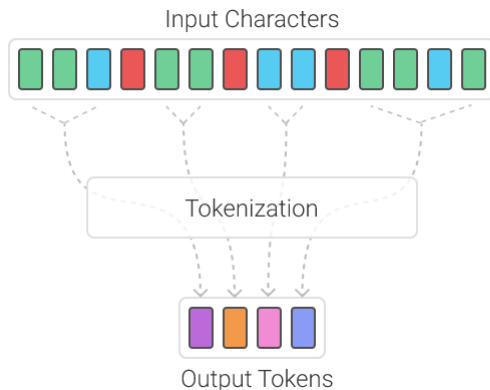
- Pop-like (Dividido en regiones)
- Rock-like (Originado del Blues)
- Hip Hop (O música urbana)
- Música Clásica
- Música Electrónica
- Jazz

Procesamiento de lenguaje natural (NLP)

El problema que se presenta es un problema relacionado al procesamiento de lenguaje natural. Para atacarlo es necesario convertir la información que está en palabras a información numérica, es decir, hacer una codificación de las palabras. Existen dos métodos populares: Tokenizar, y Vectorizar.

Tokenizar

Tokenizar consiste en convertir las oraciones a una cadena de Tokens (números enteros), usualmente 1 por palabra. Al tokenizar por palabra se crea un diccionario en el que cada palabra reconocida tiene un número asignado.



Vectorizar

Vectorizar es similar a tokenizar con la diferencia de que a cada oración (o elemento) se le asigna un vector de varias dimensiones. Una forma común de vectorización (o embedding) es utilizar una Bag Of Words. Los BOW cuentan el número de ocurrencias de cada palabra del vocabulario en las oraciones.

Document	the	cat	sat	in	hat	with
<i>the cat sat</i>	1	1	1	0	0	0
<i>the cat sat in the hat</i>	2	1	1	1	1	0
<i>the cat with the hat</i>	2	1	0	0	1	1

Tratamiento de las letras

Se utilizaron ambas formas mencionadas anteriormente, para probar cuál es más efectiva en cuál método.

- BOW de vocabulario de las 4500 palabras más importantes en el corpus del problema. Con scikit-learn CountVectorizer.
- Tokenizador con el vocabulario de las 10,000 palabras más importantes en el corpus. Con keras Tokenizer.

Nota: El tratamiento de Tokenizar fue casi 1000 veces más ligero que el tratamiento de BOW.

Modelos utilizados

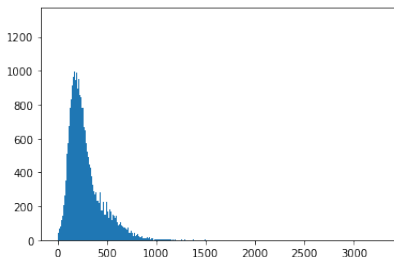
Se plantearon dos modelos sencillos como base comparativa:

- Regresión logística.
- Descenso gradiente simple.

Luego, se buscó superar los resultados de estos modelos con una red neuronal Convolutional, en secuencia.

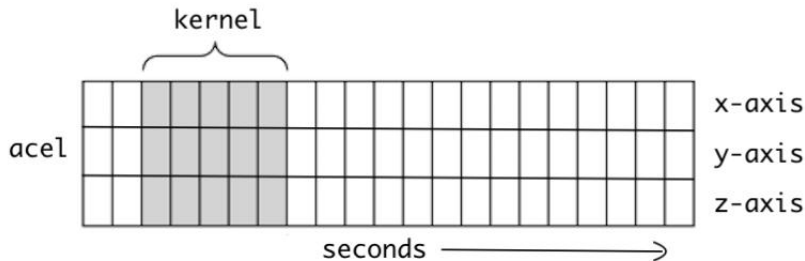
Nota

Para la CNN se trabajó solo con las primeras 1024 palabras de las canciones, pues el 99.7 por ciento de las canciones tienen menos de 1024 palabras.



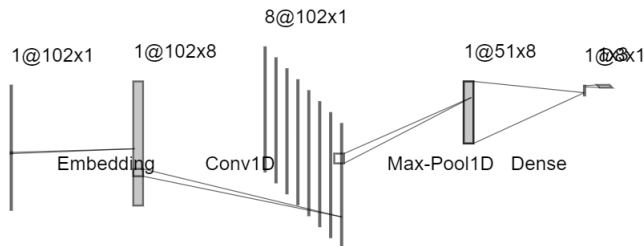
Arquitectura de la red

Se utilizó una red neuronal convolucional de 1 dimensión, también conocida como una convolución temporal.



Arquitectura de la red

La arquitectura de la red se resume en el siguiente diagrama:



Arquitectura de la red

Model: "CNN1D"

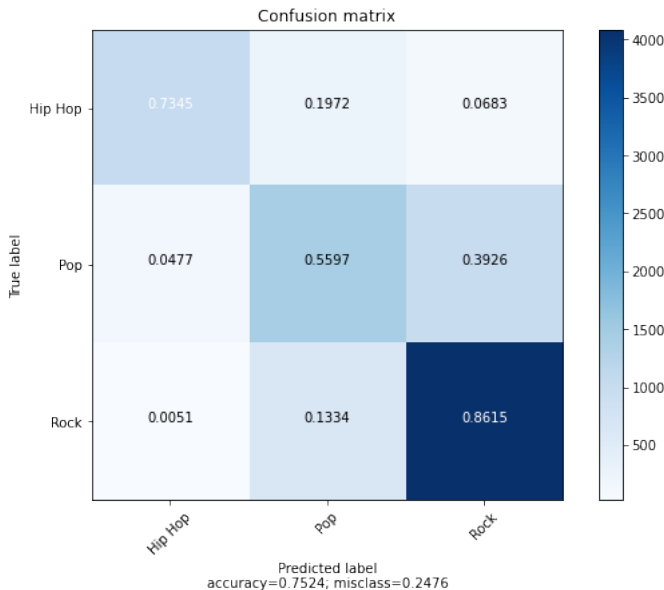
Layer (type)	Output Shape	Param #
Embedding (ModuleWrapper)	(None, 1024, 8)	80000
Conv1d (ModuleWrapper)	(None, 1024, 8)	4104
SpatialDropout (ModuleWrapper)	(None, 1024, 8)	0
Maxpooling1d (ModuleWrapper)	(None, 512, 8)	0
Flatten (ModuleWrapper)	(None, 4096)	0
Dense (ModuleWrapper)	(None, 8)	32776
Dropout (ModuleWrapper)	(None, 8)	0
Output (ModuleWrapper)	(None, 3)	27
Total params: 116,907		
Trainable params: 116,907		
Non-trainable params: 0		

Comparación de métricas

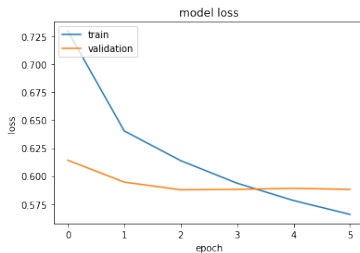
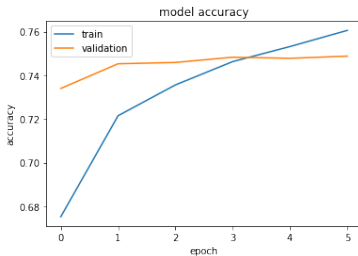
Resultados			
Modelo	Precision	Recall	F1
Regresion logística V	0.70	0.69	0.70
SGD V	0.71	0.67	0.68
Regresion logística T	0.57	0.54	0.53
SGD T	0.48	0.51	0.49
CNN1D T y V	0.76	0.72	0.74

En mi investigación no encontré un clasificador similar que superase el 0.66 de precisión.

Confusiones



Entrenamiento



El alcance de la letra

Observamos que implementar una CNN1D es una buena estrategia para atacar problemas de clasificación de textos, cuando los textos tienen una secuencia bien marcada. Funcionó mejor que modelos de RNN o modelos de convoluciones en varias dimensiones con Embeddings pre entrenados como BERT.

Notamos que se pudo alcanzar una precisión de 0.75 al clasificar las canciones en géneros musicales a partir de su letra. Este resultado lo podemos interpretar como que la letra puede contener hasta el 75 % de la información sobre el género de una canción, pero no la tiene toda.

Además, de forma consistente no fue posible cruzar la barrera del 0.76, por lo que existen razones para creer que el error de Bayes para este problema no está lejos de eso.

El pop como supergénero

Comparación de letras	
Rock	Pop
Sometimes I feel like I don't have a partner Sometimes I feel like my only friend Is the city I live in, the city of angels Lonely as I am, together we cry	Sometimes giving up is the strong thing Sometimes to run is the brave thing Sometimes walking out is the one thing That will find you the right thing

Conclusiones

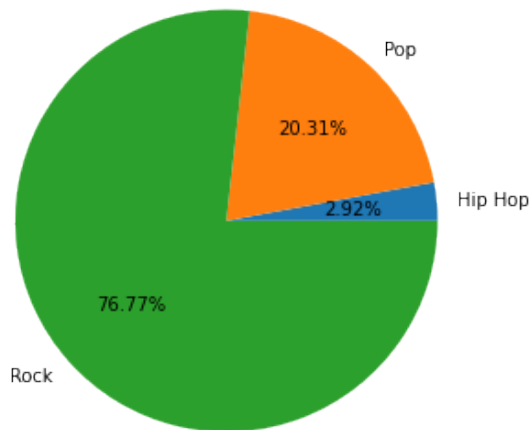
- El mejor procedimiento de codificación de texto depende de los datos y del modelo que se va a implementar.
- Por lo general, 1000 palabras son suficientes para obtener la información de la letra de una canción.
- El vocabulario del Hip Hop es mucho más amplio que el de los otros géneros y lo hace más fácil de clasificar.
- Se puede obtener una clasificación consistente de al menos 0.75 de precisión utilizando únicamente la letra de una canción.
- Con la información disponible, no es posible superar la barrera del 0.76 de precisión utilizando únicamente la letra de una canción.
- La letra de rock es muy similar a la de pop, lo cual hace que estos géneros sean fáciles de confundir.

Recomendaciones

- ➊ Dado el positivo resultado de implementar la arquitectura de red CNN1D, se recomienda su uso para problemas de clasificación de texto, principalmente en casos en los que las RNN o FCNN no dan buenos resultados.
- ➋ Se recomienda trabajar los modelos iniciando por las ideas más sencillas (la primera arquitectura que intenté tardó varias horas en entrenarse).
- ➌ Para clasificar canciones en géneros musicales es mejor contar con más información que solo la letra, pero en caso de que sólo la letra está disponible, es recomendable sólo clasificar géneros que no se parecen mucho entre sí.

Resultados interesantes

Clasificación de poemas de: WILLIAM SHAKESPEARE



Resultados interesantes

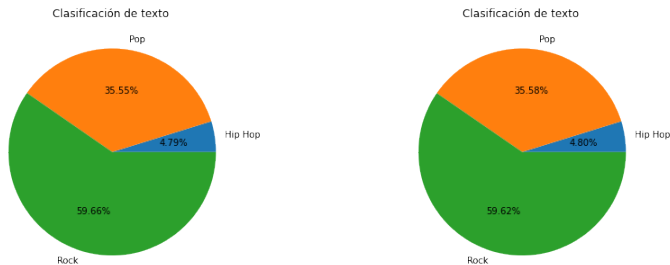


Figura 1: Comparación de la clasificación del manifiesto comunista (izquierda) con el libro MEASURE AND INTEGRATION (derecha)