



Universidad Nacional Autónoma de México  
Facultad de Ciencias

Proyecto Final

Reconocimiento de Patrones y Aprendizaje Automatizado  
Semestre 2023-2

Sergio Hernández López

José Alejandro Rojas López, Miguel Andrés Guevara Castro

---

Alberto Olvera Trejo, Ricardo Bernabé Nicolás. Miguel Angel Romero Gonzalez

---

## 1. Introducción y Objetivo

El objetivo de éste proyecto es mostrar el desempeño de diferentes algoritmos de machine learning para resolver la misma tarea, la cual es la clasificación de géneros musicales. Los algoritmos usados son redes neuronales, árboles de decisión, random forest, K-Neighbors, regresión logística, GradientBoost y XGB.

El objetivo es encontrar el algoritmo que tengo un mejor desempeño

## 2. Datos

Se trabajó con un dataset el cual contine diferentes características de una señal de audio. El nombre de algunas columans es: *chroma\_stft\_mean*, *chroma\_stft\_var*, *rms\_mean*, *rms\_var*, *spectral\_centroid\_mean*, *spectral\_centroid\_var*. Como su nombre lo indica, algunas de las variables son obtenidas del espectrograma resultante de aplicar la transformada corta de Fourier.

En la figura 1 se muestra cómo se ve el dataset.

	filename	length	chroma_stft_mean	chroma_stft_var	rms_mean	rms_var	spectral_centroid_mean	spectral_centroid_var
0	blues.00000.wav	661794	0.350088	0.088757	0.130228	0.002827	1784.165850	129774.064525
1	blues.00001.wav	661794	0.340914	0.094980	0.095948	0.002373	1530.176679	375850.073649
2	blues.00002.wav	661794	0.363637	0.085275	0.175570	0.002746	1552.811865	156467.643368
3	blues.00003.wav	661794	0.404785	0.093999	0.141093	0.006346	1070.106615	184355.942417
4	blues.00004.wav	661794	0.308526	0.087841	0.091529	0.002303	1835.004266	343399.939274

Figura 1: Datos con los que se entrenaron los algoritmos

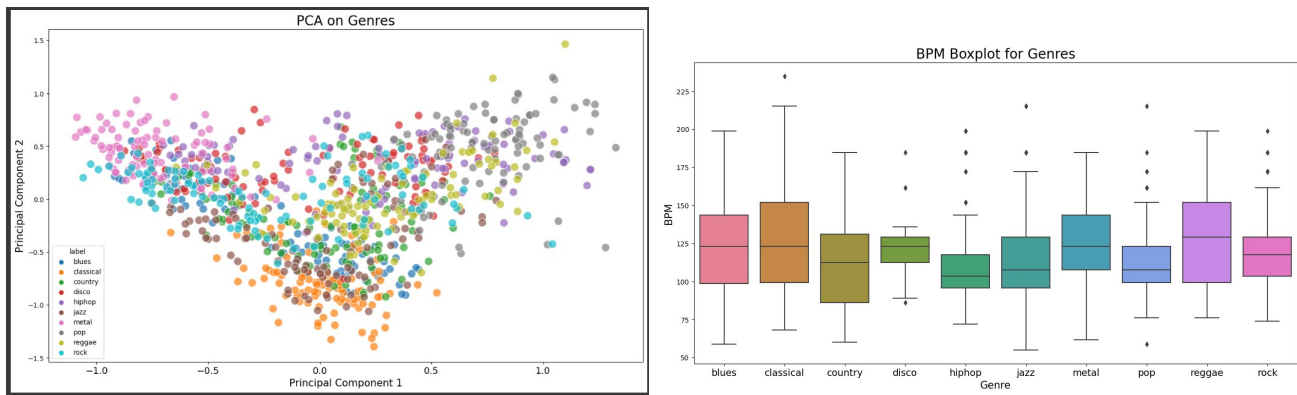


Figura 2: PCA de los géneros musicales y boxplot del rango de los beats por minuto de cada género

### 3. Visualización

Para ver gráficamente cómo se distribuyen los diferentes géneros musicales decidimos realizar un PCA para obtener dos componentes principales y poder graficar los géneros. El resultado obtenido se muestra en la figura

### 4. Entrenamiento y arquitectura de la red neuronal

Para poder entrenar a los algoritmos ya mencionados se utilizó Grid Search, el cual es un método para poder encontrar los mejores hiperparámetros, como por ejemplo para poder encontrar cuál es la mejor profundidad en un random forest, o la  $k$  óptima en K-Neighbors, etc. Durante el uso de Grid Search se empleó cross validation para mejorar los resultados

Por otro lado, en la red multicapa se usó la siguiente arquitectura

- Una capa completamente conectada de 1024 neuronas
- Una capa completamente conectada de 512 neuronas
- Una capa completamente conectada de 512 neuronas
- Una capa completamente conectada de 256 neuronas
- Una capa completamente conectada de 128 neuronas
- Una capa completamente conectada de 64 neuronas
- Una capa softmax para obtener las probabilidades de cada categoría

La función de activación que se empleó fue la sigmoide. Además, entre cada capa se utilizó un dropout que varía entre 0.25 y 0.1.

## 5. Resultados

En la tabla 1 podemos ver el accuracy de cada algoritmo.

Modelo	Accuracy
Decision Tree	0.43
Regresión Logística	0.545
K-Neighbors	0.235
Random Forest	0.69
GradientBoost	0.56
XGB	0.655
Red Multicapa	0.9

Tabla 1: Desempeño de cada algoritmo

Los algoritmos con mejor desempeño fueron las redes neuronales y XGB. En la figura 3 se muestra la matriz de confusión de la red multicapa.

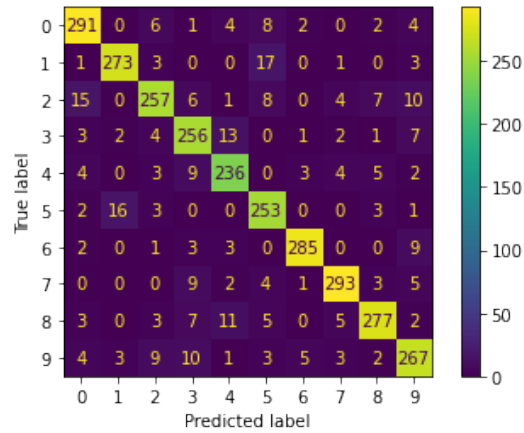


Figura 3: Matriz de confusión de la red neuronal multicapa

## 6. Conclusiones

Podemos afirmar que el mejor algoritmo fue la red neuronal multicapa, pero dicho desempeño tiene su costo debido a que fue el modelo que más tiempo tomar en entrenarse, aproximadamente un día.

Por último, podemos decir que el objetivo del proyecto fue completado debido a que encontramos un modelo (red multicapa) que pudo clasificar los 10 géneros musicales de manera exitosa.