
Prediciendo *Despacito*

Universidad Nacional de Colombia, Sede Bogotá

Introducción a la Inteligencia Artificial

2022-II

Juan Pablo Ortega Bermudez

jortegab@unal.edu.co

Ana Paulina Castillo Velásquez

ancastillov@unal.edu.co

Sebastian Leonardo Molina Diaz

smolinad@unal.edu.co

Resumen

La industria musical, como muchas otras industrias del entretenimiento, buscan captar la atención del público dentro de un mar de contenido audiovisual, por lo que conocer la receta del nuevo *hit* mundial es de vital importancia para mantener el negocio rentable. Como aplicación de redes neuronales y *random forests*, los agentes presentados pretenden predecir la popularidad de una canción dada, siguiendo los parámetros y variables definidos por una de las plataformas musicales más importantes de la actualidad.

1 Introducción

En la industria musical, la cual mueve alrededor de 25000 millones de dólares anualmente, es necesario distinguir producciones exitosas de las que no lo son. Esto crea una oportunidad de incursionar con la inteligencia artificial, en particular con el aprendizaje de máquina, ya que muchos métodos de esta área son comúnmente usados con el fin de categorizar la música.

En particular, en este proyecto se intentará predecir el posible éxito de una pieza musical a partir de sus características, como lo son el tempo, el instrumental o la vocalidad. Plataformas como Spotify, Youtube o Apple Music categorizan la música según una cantidad de parámetros determinada, en los cuales basaremos nuestros análisis y modelos de predicción.

Dentro de la baraja de opciones que se tienen para abordar este problema resaltamos dos métodos; *neural networks* y *random forest regression*. Estos sistemas son un conjunto de herramientas y técnicas que permiten construir predicciones y categorizaciones respecto a una base de datos de entrada, los cuales se encuentran dentro del campo del aprendizaje de máquina, y en particular, del aprendizaje supervisado.

Problemas relacionados a la categorización de la música han sido tratados con dichos métodos en varias oportunida-

des, por ejemplo en [3], [4], [5], [7], [8], entre otros, por lo que abordaremos el problema de la predicción de la probabilidad usando estos métodos.

2 Materiales y Métodos

2.1 Neural Network

Una red neuronal —en inglés *neural network*— es un sistema de tecnologías inspiradas en la estructura del cerebro humano. Las redes neuronales artificiales pueden describirse como modelos de dos capas como mínimo; una de entrada y una de salida. En la mayoría de casos también existen capas intermedias, las cuales almacenan las denominadas neuronas, que son contenedores de información que se comunican con la siguiente capa. Entre más complejo el problema, más capaz son requeridas.

Mediante estos enlaces y ponderaciones se modela un algoritmo que genera un resultado para cada información de entrada. Este algoritmo puede ser modificado gracias al entrenamiento de la red, que consiste en dotar a la máquina de entradas y salidas específicas a modo de que el algoritmo encuentre los patrones necesarios para futuros casos.

La pertinencia de este modelo para la solución del proble-

ma de predicción se debe a su uso extendido en la clasificación de géneros de canciones.

2.2 Random Forest Regression

Random Forest Regression es un algoritmo de aprendizaje supervisado que utiliza un método de aprendizaje conjunto para la regresión.

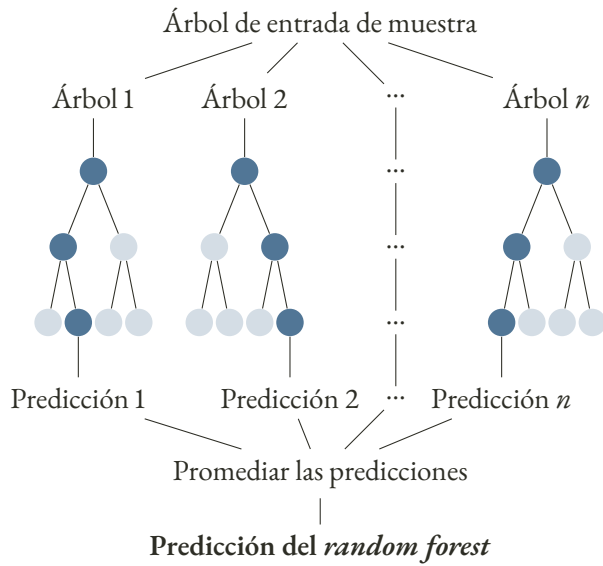


Figura 1: Diagrama de funcionamiento del algoritmo de *random forest*.

Aquí, varios árboles corren independientemente. Su operación consiste en construir varios árboles de decisión durante el tiempo de entrenamiento y generar la media de las clases como la predicción de todos los árboles. El algoritmo de *random forest regression*, obtenido de [2] se encuentra especificado en el Algoritmo 1.

Random forest agrega aleatoriedad adicional al modelo mientras crecen los árboles. Al dividir un nodo, busca la mejor función entre un subconjunto aleatorio de funciones en lugar de buscar la función más importante. Por lo tanto, reduce el problema de sobreajuste en los árboles de decisión y disminuye la varianza, mejorando la precisión. Esto es bastante útil a la hora de hacer predicciones ya que puede manejar fácilmente datos faltantes.

Para este proyecto, los datos fueron extraídos de Kaggle, en particular de [6], los cuales fueron generados con la API de Spotify para desarrolladores. Nuestra base de datos cuenta con 18454 canciones con 25 atributos, 9 de los cuales fueron descartados porque no se consideraron útiles para el

Algoritmo 1: Random Forest Regression

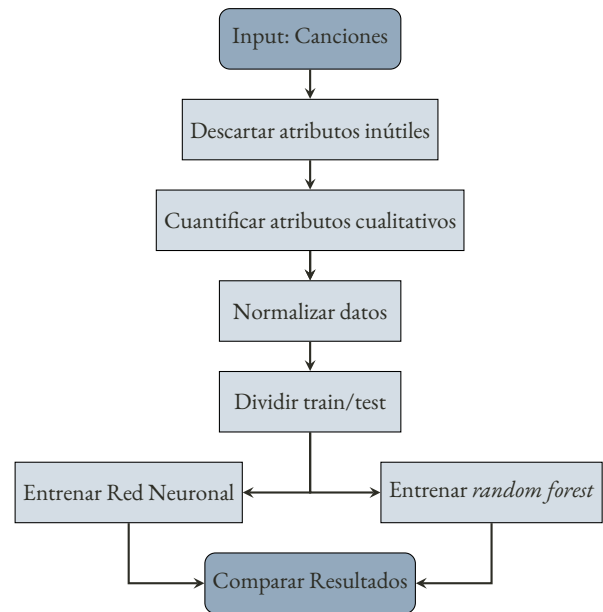
para $b = 1$ *hasta* N **hacer**

1. Obtener una muestra de los datos por medio de *bootstrapping*.
2. Crear un árbol aleatorio a partir de la muestra obtenida.
3. Repetir los siguientes pasos recursivamente mientras que el tamaño de los nodos es mayor que n_{\min} :
 - (a) Elegir aleatoriamente m parámetros que categorizan a los datos.
 - (b) Encontrar la variable más importante (*split point*).
 - (c) Dividir el nodo en dos hijos por medio del *split point*.

fin

devolver Predicción: $\frac{1}{N} \sum_{b=1}^N T_b(x)$

modelo. El diagrama de bloques del proyecto es el siguiente:



Los atributos son cualitativos son convertidos en numéricos utilizando la técnica de *one-hot encoding*. Además, se usó el análisis de componentes principales para reducir la matriz obtenida al combinar los atributos numéricos resultantes. Este proceso se aplicó a los atributos de género, idioma, modo y clave. Después de realizarlo, conservando los 11 componentes principales, obtenemos una razón de varianza ex-

plicada de 79% para estos atributos.

La distribución de la popularidad de las canciones, el cual es el atributo que se quiere predecir, se puede ver en la Figura 2.

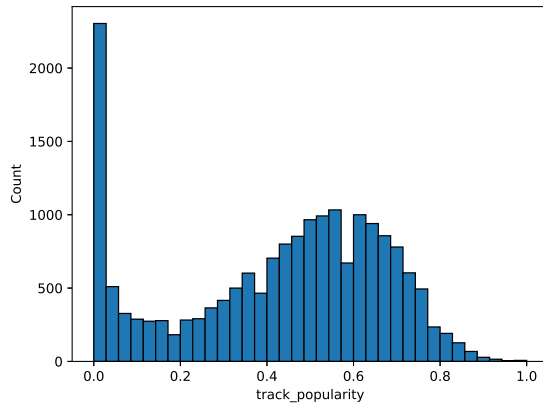


Figura 2: Distribución de la popularidad de las canciones.

Decidimos excluir del experimento a las canciones con popularidad debido a su gran cantidad, pues el claro desbalance podría ocasionar sesgo. Estos datos fueron distribuidos de manera aleatoria y uniforme en entrenamiento y test, con una proporción de 80% y 20% respectivamente.

Utilizamos dos modelos de inteligencia artificial para aproximar la función objetivo, El primero, basado en redes neuronales; el segundo, basado en bosques aleatorios de regresión. La red neuronal cuenta con una capa de entrada, 6 capas ocultas, todas con función de activación *Leaky ReLU*, mientras que la capa de salida posee una sola neurona con función de activación sigmoideal. *Leaky ReLU* es una función de activación definida por

$$f(x) = \begin{cases} x, & x > 0 \\ 0.01x, & \text{en otro caso} \end{cases},$$

donde 0.01 es el gradiente que se asigna a una neurona cuando no es activa. El bosque aleatorio cuenta con 1000 árboles estimadores con profundidad máxima igual a 1000.

3 Resultados

Los siguientes resultados son la finalidad de puntuar el valor real de la popularidad al predicho por los algoritmos mencionados al 20% de los datos restantes. Primero hablaremos a

los resultados obtenidos a la red neuronal descrita, para luego seguir con los resultados del *random forest* y dar una comparativa entre los métodos. Como se puede ver en la Figura 3, usando el coeficiente de correlación de Pearson podemos ver que el índice de correlación es de 0.26113875, lo cual indica una correlación bastante moderada.

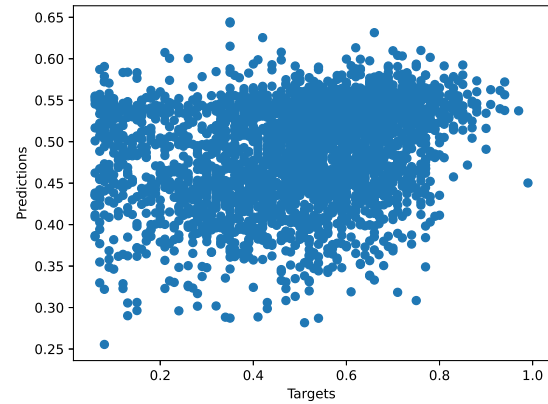


Figura 3: Correlación popularidad real y predicha de las canciones según el algoritmo de red neuronal

Al realizar el conteo del número de canciones por cada escalón de popularidad, podemos ver que hay una distribución bimodal centrada en 0.45 y 0.54, teniendo en cuenta que la media de los valores reales se encuentra en 0.6. Así, podemos decir que con respecto al modelo de redes neuronales hay una tendencia a pronosticar la popularidad de una canción dada con un valor menor al original.

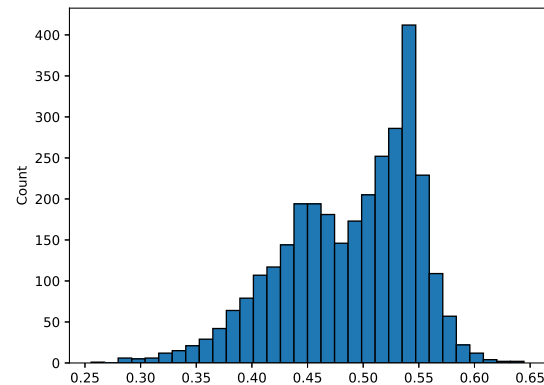


Figura 4: Distribución de la popularidad de las canciones según el algoritmo de red neuronal.

Respecto al modelo de *random forest*, usando el coeficien-

te de correlación de Pearson podemos ver que el índice de correlación es de 0.32601011, un resultado levemente mejor que con respecto al modelo de redes neuronales.

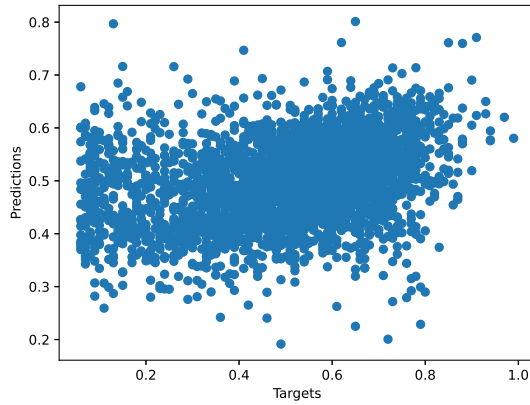


Figura 5: Correlación popularidad real y predecida de las canciones según el algoritmo de *random forest*.

En la distribución de popularidad de las predicciones del modelo de *random forest* de la Figura 6, se puede observar que dicha distribución es de tipo normal, centrada en aproximadamente 0.5. En este caso, es clara una mejoría con respecto a la distribución obtenida con el modelo anterior.

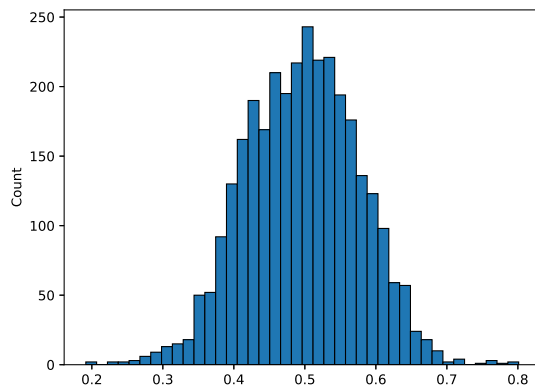


Figura 6: Distribución de la popularidad de las canciones según el algoritmo de *random forest*.

Como ejemplo de las predicciones tomamos dos canciones de muestra, una con menor tiempo desde su lanzamiento y la segunda con más tiempo desde su lanzamiento. Los resultados de las predicciones se pueden ver en el Cuadro 1.

4 Discusión y conclusiones

Previo al procesamiento con los agentes ya mencionados, se intentó clasificar las canciones con la técnica de DBSCAN. El proceso de DBSCAN —en inglés *Density-based Clustering Based on Connected Regions with High Density*— consiste en agrupar un conjunto de puntos en una vecindad de radio ϵ y una cantidad mínima de puntos, *MinPts*. Este algoritmo es un estándar en el procesamiento de datos debido a su eficiencia. Sin embargo, al aplicar este método al conjunto de canciones, no se pudo encontrar *clusters* de puntos, es decir, no se encontraron características particulares que permitieran clasificar las canciones populares según un listado de parámetros comunes.

De los resultados obtenidos podemos observar que aunque los modelos no devolvieron los resultados esperados (como mayor correlación y mejor distribución), existe una mejora al utilizar el modelo de *random forest*. Es claro que hay mucho por mejorar, por ejemplo, los modelos implementados fallaron al predecir la desviación en los datos de entrada, enfocándose en valores cercanos a la media.

Cabe resaltar que la capacidad de procesamiento para este proyecto estuvo limitada a computadores comerciales, lo que no permitió, entre otras cosas, usar diferentes funciones de activación y de costo disponibles de manera iterada, lo cual permitiría un análisis más profundo sobre los modelos. Sin embargo, debido a la reproducibilidad del experimento se podría implementar en un ambiente con mayor capacidad.

Proponemos como trabajo futuro investigar nuevas técnicas que permitan obtener resultados con mayor precisión, ya que como se menciona en [4] y como se pudo observar en la presente implementación, los métodos de redes neuronales y *random forests* han tenido éxito limitado y resultados variables.

Además, el uso de los modelos propuestos en nuevas canciones necesitan de los parámetros generados internamente por Spotify, lo cual requiere que la canción sea distribuida en la plataforma. Por lo anterior, puede ser interesante predecir la popularidad de una canción incluso antes de su lanzamiento, lo cual proveería retroalimentación a los artistas, quienes tendrían tiempo de modificar aspectos de su producción e incrementar la posibilidad de tener un tema exitoso.

Canción	Popularidad original	Predicción
<i>Tusa</i> - Karol G	98%	87%
<i>Every Breath You Take</i> - The Police	39%	42%

Cuadro 1: Predicción de popularidad para canciones eleccionadas usando el modelo de *random forests*.

Referencias

- [1] J. Han, M. Kamber y J. Pei. *Data Mining: Concepts and Techniques*. Morgan Kaufmann, 2012.
- [2] T. Hastie, J. Friedman y R. Tibshirani. *The elements of Statistical Learning: Data Mining, Inference, and prediction*. Springer, 2017.
- [3] C. P. Martin et al. «Understanding Musical Predictions With an Embodied Interface for Musical Machine Learning». En: *Frontiers in Artificial Intelligence* 3 (2020). ISSN: 2624-8212. DOI: 10.3389/frai.2020.00006. URL: <https://www.frontiersin.org/articles/10.3389/frai.2020.00006>.
- [4] N. Pelchat y C. M. Gelowitz. «Neural Network Music Genre Classification». En: *Canadian Journal of Electrical and Computer Engineering* 43.3 (2020), págs. 170-173. DOI: 10.1109/CJECE.2020.2970144.
- [5] J. Pham, E. Kyauk y E. Park. «Predicting song popularity». En: *Dept. Comput. Sci., Stanford Univ., Stanford, CA, USA, Tech. Rep* 26 (2016).
- [6] H. S. Saragih. *Exploratory Data Analysis 18k Spotify Songs*. Nov. de 2020. URL: <https://www.kaggle.com/code/harrimansaragih/exploratory-data-analysis-18k-spotify-songs>.
- [7] B. Xu et al. «Classifying very high-dimensional data with random forests built from small subspaces». En: *International Journal of Data Warehousing and Mining* 8.2 (2012), págs. 44-63. DOI: 10.4018/jdwm.2012040103.
- [8] L.-C. Yang et al. «Revisiting the problem of audio-based hit song prediction using convolutional neural networks». En: *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2017, págs. 621-625.