

Predicción de éxitos musicales usando Spotify

*¿Qué características hacen a un tema
musical un éxito?*



Agenda

01 Motivación y audiencia

02 Preguntas de interés

03 Metadata

04 Análisis Exploratorio



Agenda

05

Insights

06

Modelos de Machine Learning

07

Importancia de variables

08

Conclusiones



Motivación y audiencia

Motivación: Analizar características de canciones utilizando datos de Spotify para predecir su éxito. Se emplean técnicas de visualización de datos para explorar la relación entre características como tempo, energía y popularidad con el éxito.

Audiencia: El proyecto tiene implicaciones para la industria musical y puede ayudar a productores y artistas a seleccionar características que aumenten el éxito en el mercado.

Limitaciones:

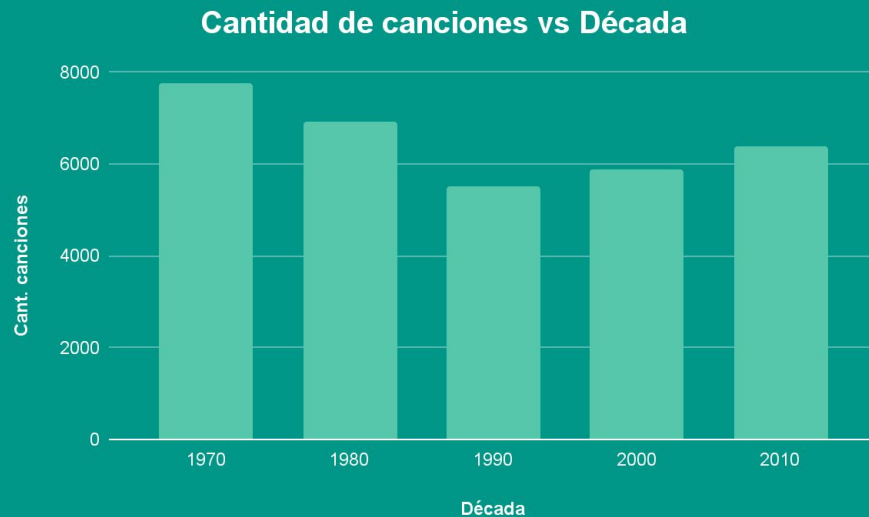
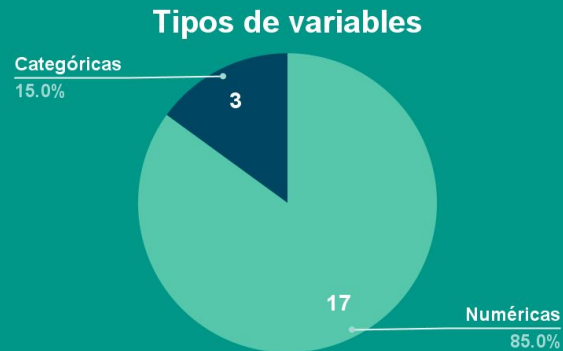
- **Datos insuficientes:** Los datos de Spotify pueden ser limitados para ciertas canciones o artistas.
- **Cambios en la popularidad:** Las preferencias musicales de la audiencia pueden cambiar rápidamente.
- **Ausencia de variables importantes:** El éxito de una canción también puede verse influenciado por el marketing, promoción y contexto cultural, que no se consideraron en este proyecto.

Preguntas de interés

- ¿Existe una **correlación entre las características de una canción**, como el tempo, la energía y la popularidad, **y su éxito**?
- ¿Se pueden utilizar las **características** de una canción **para predecir su éxito**?
- ¿Es útil el **análisis de características** de las canciones utilizando **datos de Spotify** para predecir el éxito de las canciones?
- ¿El uso de **técnicas de aprendizaje automático mejora la precisión de las predicciones** sobre el éxito de las canciones en las listas de éxitos?

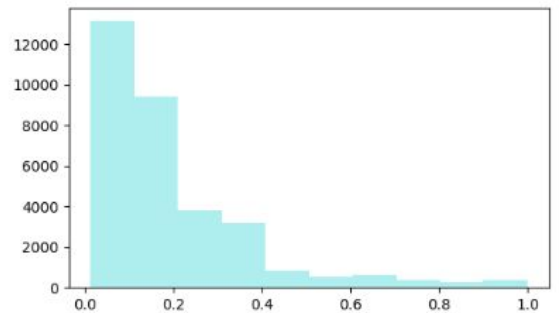
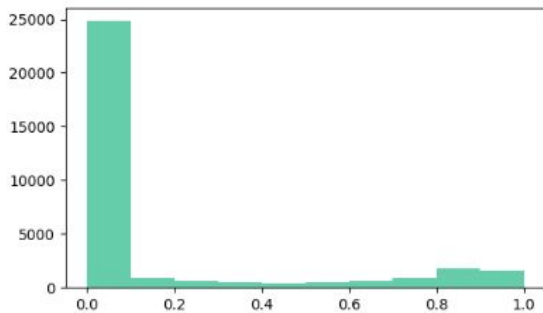
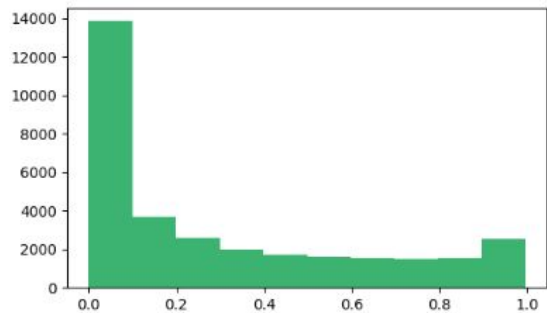
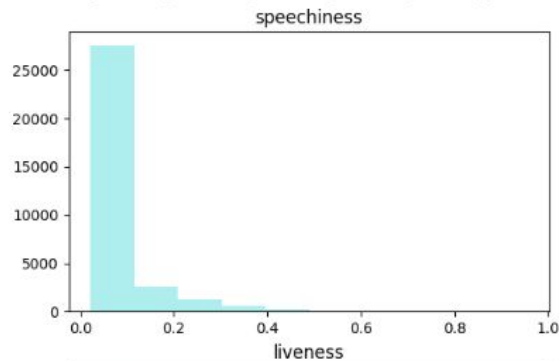
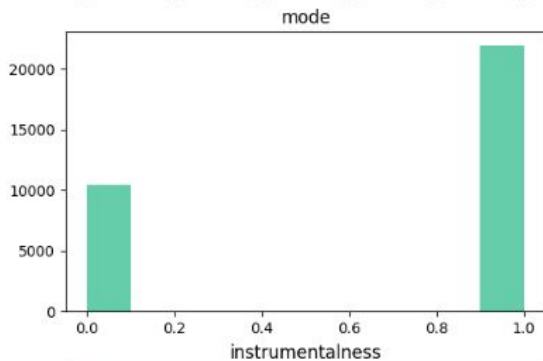
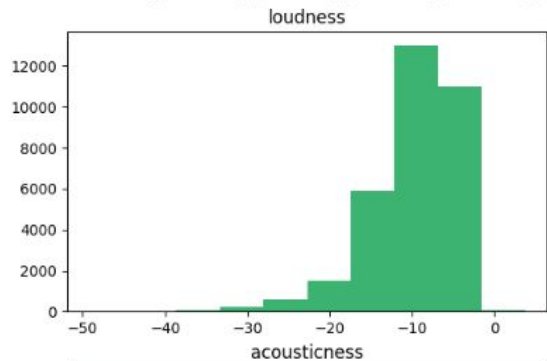
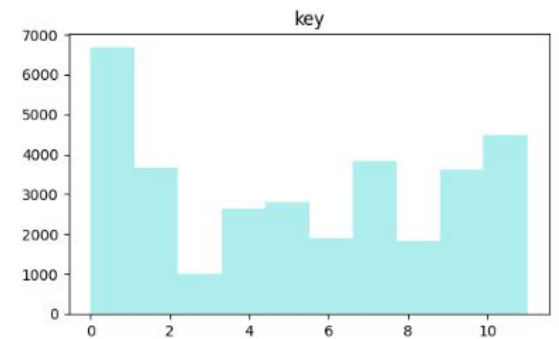
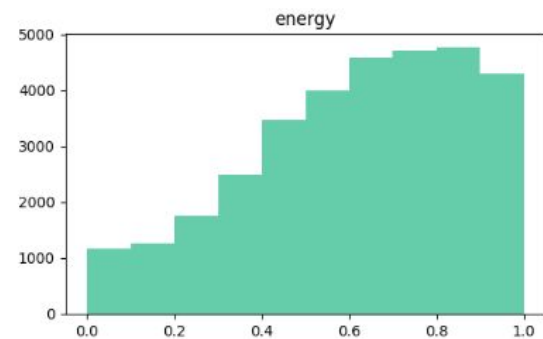
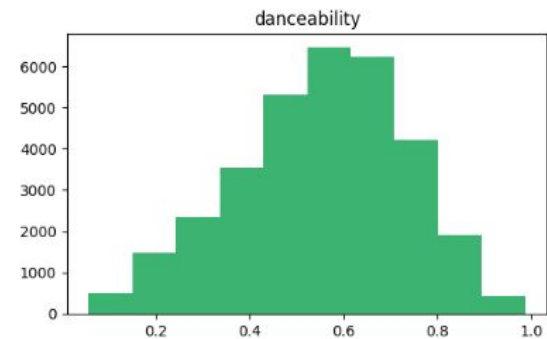
Metadata

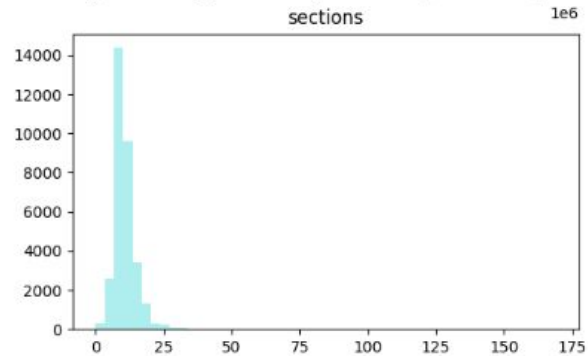
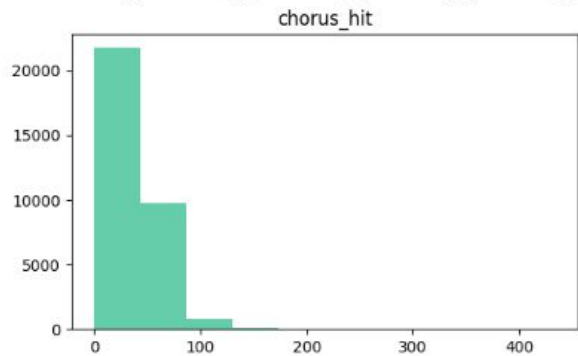
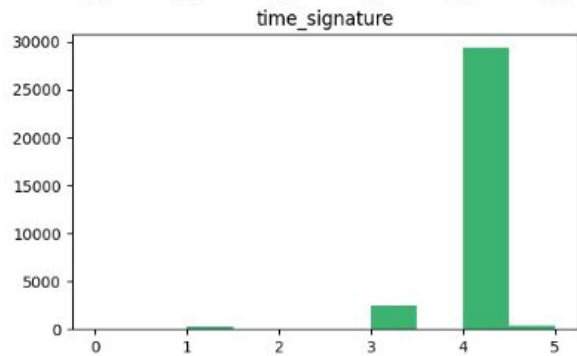
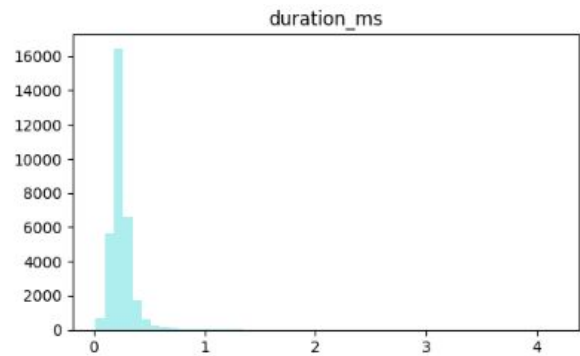
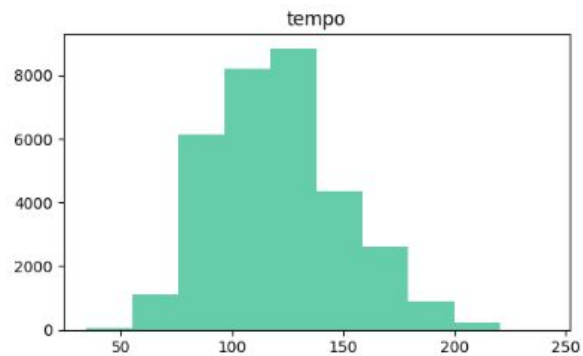
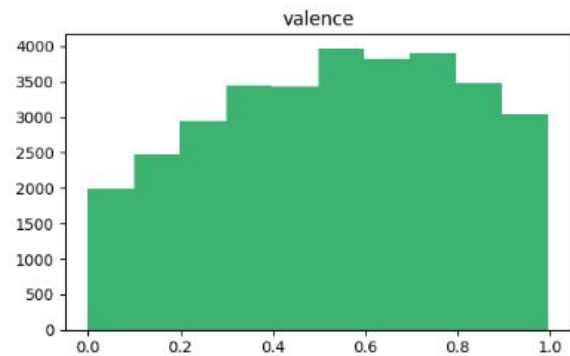
Dataset final	
Filas	Columnas
32464	20



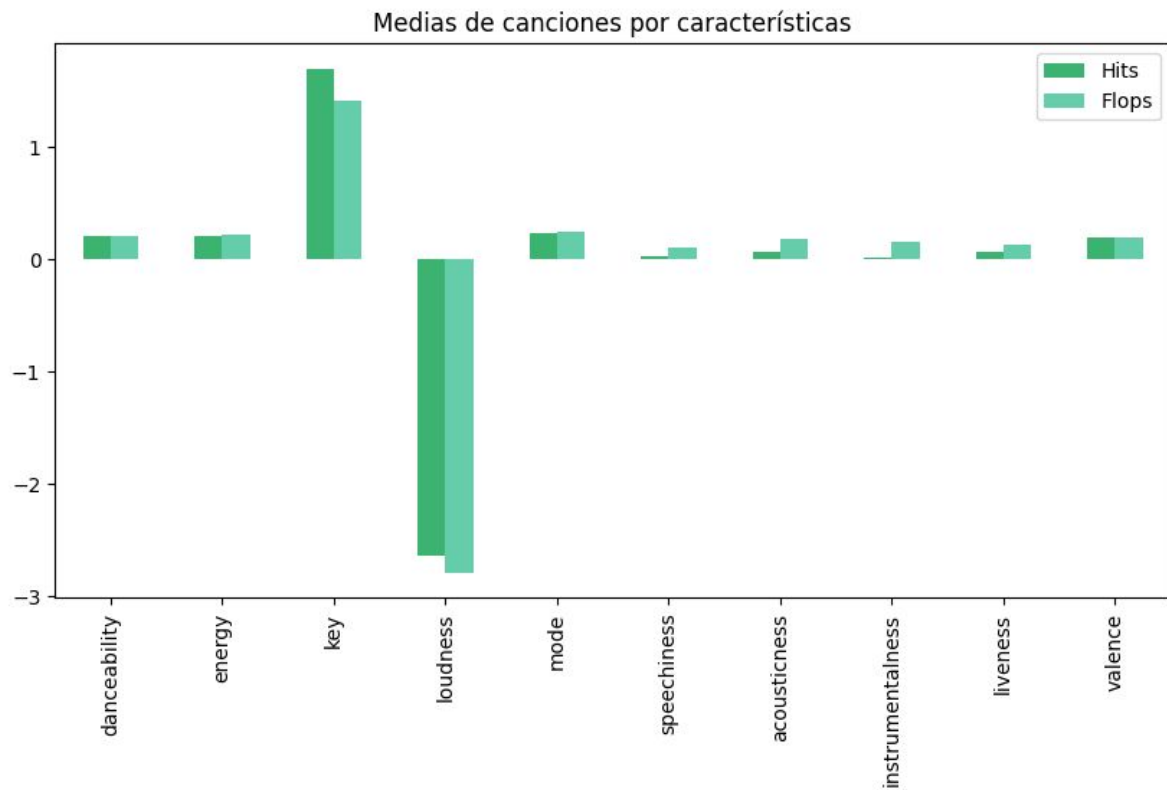
Análisis exploratorio

Distribución de las variables



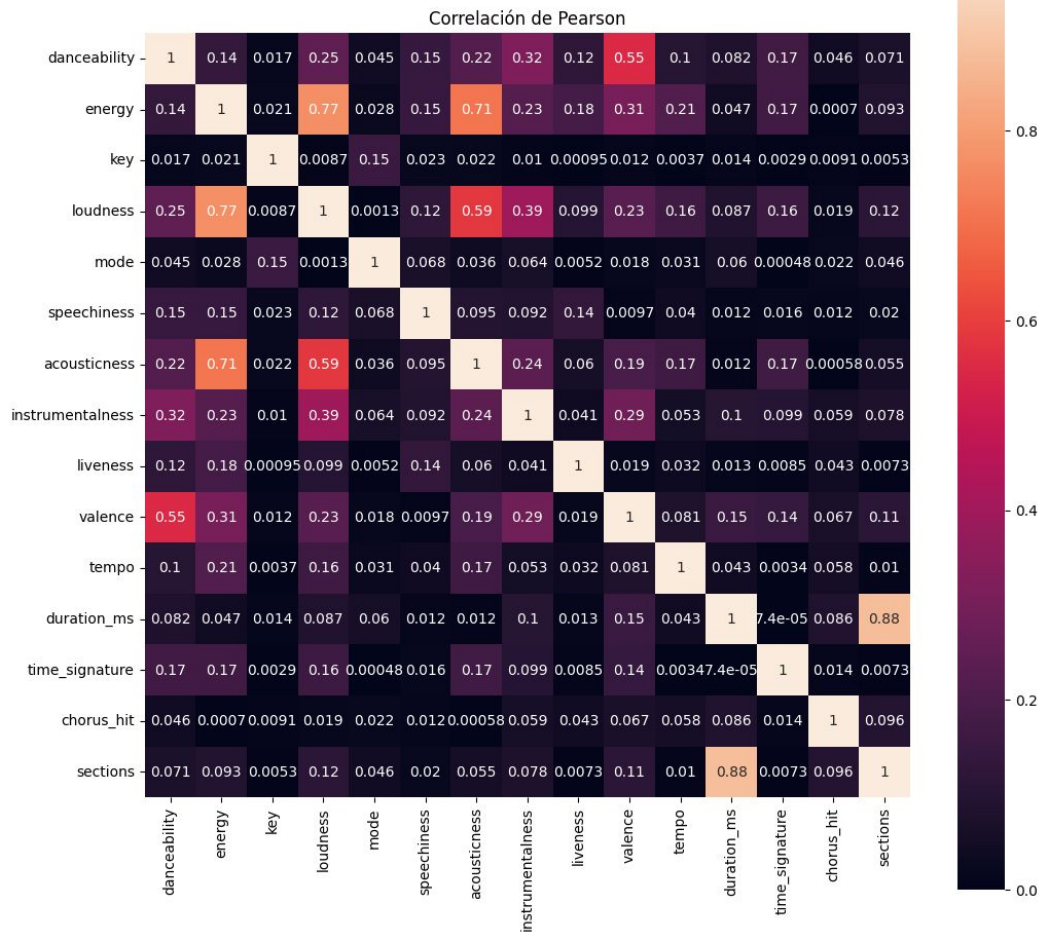


Comparación de medias



Diferencias en las medias entre éxitos y fracasos indican una posible distinción entre ambos tipos de canciones.

Correlación de variables



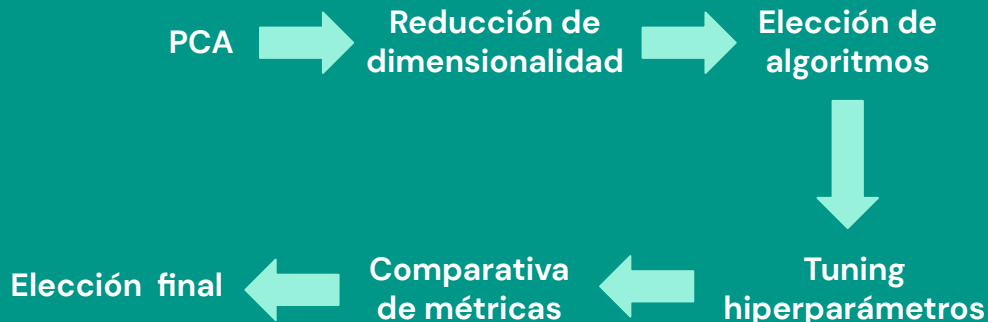
Los valores de correlación entre dos variables cualesquiera no superan el 90%.

Las únicas variables que podrían llegar a estar correlacionadas, son las de **duration_ms** y **sections**.

Insights

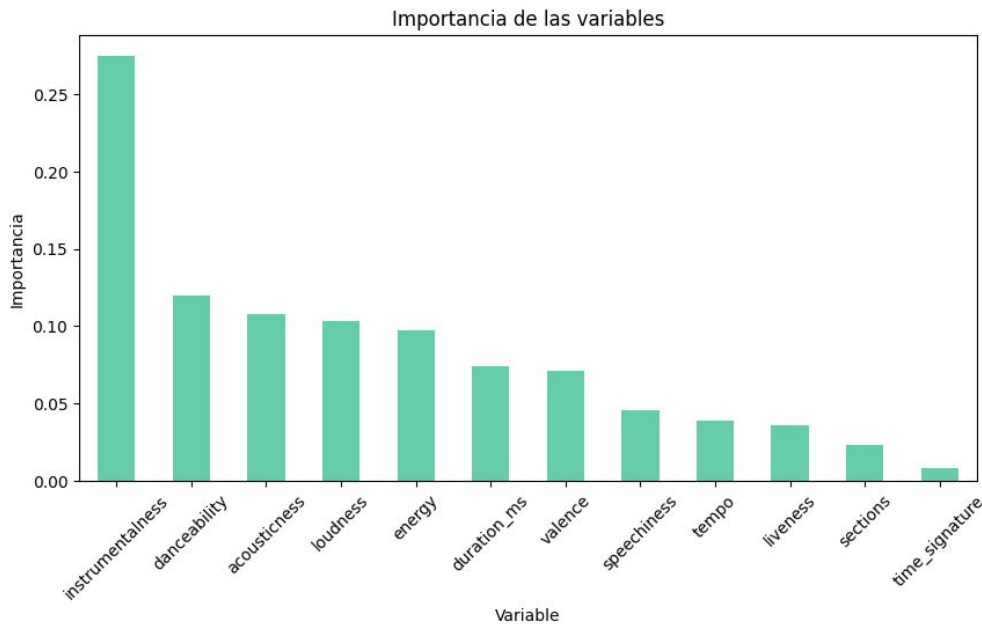
- Las canciones tienden a ser **más bailables** que menos bailables.
- Las canciones tienden a tener **más energía** que menos energía.
- La **tonalidad de Do** (C) es la más popular.
- Las canciones tienden a tener **menos de 10 dB**.
- La mayoría de las canciones están en **escalas mayores**.
- La mayoría de las canciones contienen más **música que voz**.
- La mayoría de las canciones **no son en vivo**.
- La mayoría de las canciones **no son acústicas**.
- La mayoría de las canciones **contienen música**.
- Hay una buena **mezcla de canciones alegres y tristes**.
- La mayoría de las canciones tienen entre **80 y 90 bpm**.
- La mayoría de las canciones están en **4/4**.
- Existen claras **diferencias entre las características** de canciones **exitosas vs los fracasos**.

Modelos de Machine Learning



Modelo	Accuracy (%)	Precision (%)	Recall (%)
Regresión logística	77.62	71.66	93.03
Árboles de decisión	81.12	78.17	87.53
Bosques aleatorios	84.73	82.29	89.73

Importancia de las variables



Las tres características más importantes resultaron ser:

Instrumentalness

Pistas con mayor instrumentalidad, menor contenido vocal.

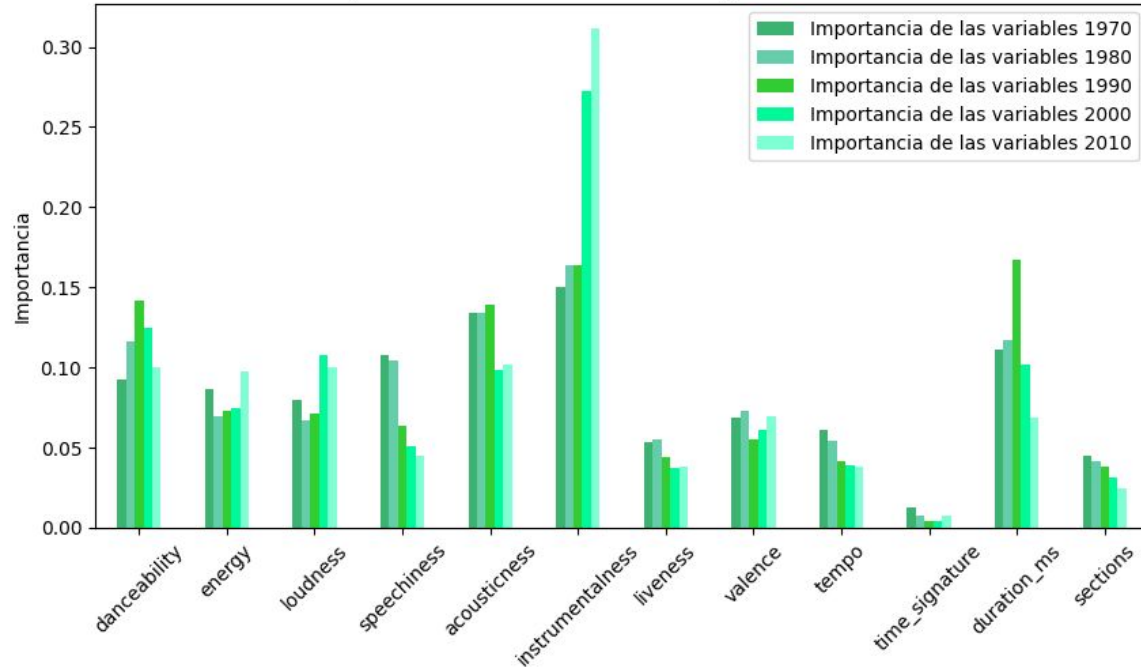
Danceability

Pistas más “bailables” en función de una combinación de elementos musicales que incluyen tempo, estabilidad del ritmo, fuerza del ritmo y regularidad general.

Acousticness

Pistas con más contenido acústico.

Importancia de las variables a lo largo de las décadas



La importancia de las características **varía respecto de las décadas**, sin embargo, el comportamiento es bastante parecido entre ellas.

En las décadas más actuales, la característica de **instrumentality** tomó aún más relevancia.

Conclusiones

- Se analizaron **características** como el tempo, la energía y la popularidad de las canciones para determinar su **correlación con el éxito**.
- Se utilizaron modelos de aprendizaje automático como **regresión logística, árboles de decisión y bosques aleatorios**.
- Los resultados muestran que **las características de las canciones pueden predecir con precisión si serán éxitos**.
- **Bosques aleatorios** fue el modelo con mejor rendimiento.
- Las características más importantes son **instrumentalness, danceability y acousticness**.
- **Limitaciones:** datos insuficientes y factores externos como cambios en la popularidad y marketing.

El análisis de características de canciones en Spotify y el uso de técnicas de aprendizaje automático pueden predecir el éxito, siendo útil para la industria musical y los artistas.