TrackTracer: Analyzing Audio Features for Social Media Track Popularity Estimation

Presented by Team 59

Haider Ali | Dhruv Arora | Roshni Balasubramanian | Camilla Kang

Short video advertising & BGM

- Short video advertising on social media platforms, such as TikTok and Reels, has emerged as a crucial marketing strategy for **building brand awareness**.

- When creating in-feed video ads, the **choice of BGM** is vital in engaging our audience.

- Smooth and trendy BGM **captivates user's attention in the first few seconds**, which is crucial for short videos as that is the time users scroll on.
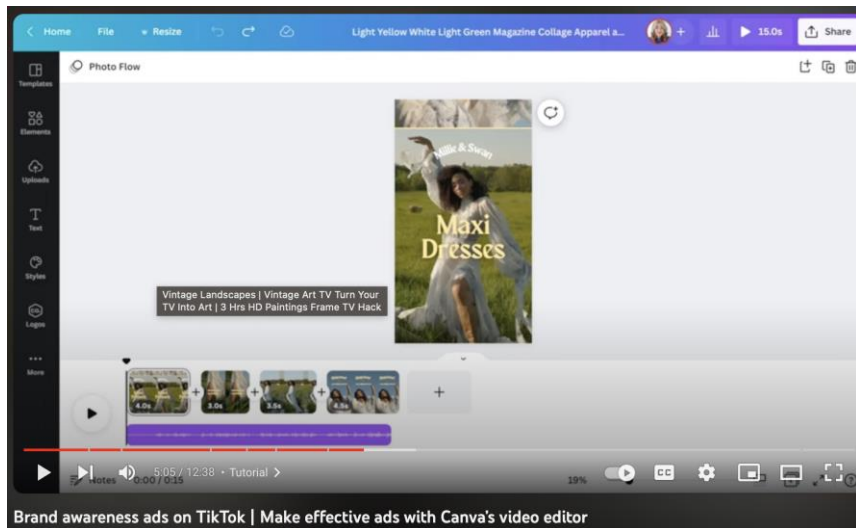

99% of the time, TikTok users are fixated on their screens while an ad is playing, compared to **76%** of the time on other platforms.

"Well-engaged BGM have the potential to reach a massive audience within a short span of time and helps establish a sonic identity for the brand."

-Tiktok for Business

How is BGM a business problem?

Ad Templates

Video Editor

Edit your existing assets into TikTok-ready ad creatives with our free online Video Editor.

Add royalty-free commercial music to any of your videos.

Create new

Therefore, in order to **capitalize short video advertising** and be ahead of the trend, **hit songs can have a template made for them ahead of time**.

Tiktok has a long partnership with Canva to **utilize existing TikTok templates** on Canva, explore best practices for creating TikTok ads, and hear from a TikTok expert as they analyze best-performing designs.

Creative Center

Explore the latest trends and tools to help you create content that gets results.

Learn more

WE AIM TO INVESTIGATE WHAT SONGS ARE BEST USED FOR COMMERCIALIZATION THROUGH DATA MINING ON SPOTIFY.

▪ Successful BGM contributes much to an advertisement, and alternatively, some commercials have even lead the trend of hit music so much that they become part of brand identity.

▪ Thus, analyzing and predicting potential hit soundtracks that are best used for reels and tiktoks commercialization is crucial in short video marketing.

Data Collection & Data Attributes

Data collection

- **Choosing Spotify**: Given its global popularity and extensive music database, Spotify was our preferred choice for sourcing data.
- **Data Extraction**: Leveraged the Spotify API to fetch crucial audio features and attributes of songs.
- **Depth of Data Collection**: 17 diverse playlists on Spotify, 1,091 unique songs
- Attributes: Song Attributes (Track name, Track Popularity, etc.) & Audio Features (Danceability, Acousticness, Key, etc.)

Track Popularity

Track Popularity Value	Count of Songs
>=50 (Hit songs)	595
<50 (flop songs)	496

Track Spotify Popularity Scores
Understand and build your Spotify growth with day by day tracking of songs popularity scores on Spotify.
See some examples

Flowers by Miley Cyrus   Real Spring by Skrillex   WARNING   As It Was by Harry Styles
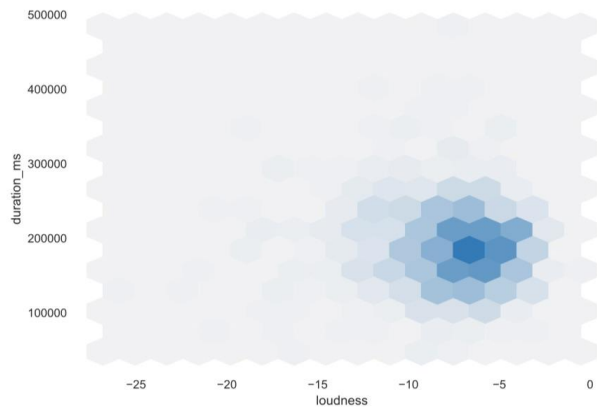
Paint The Town Red
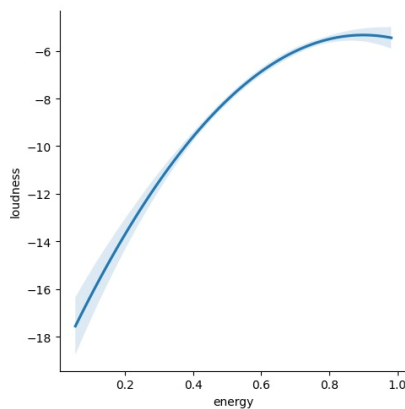Doja Cat

100% current popularity on Spotify

Regardless
Naomi Sharon
obsidian

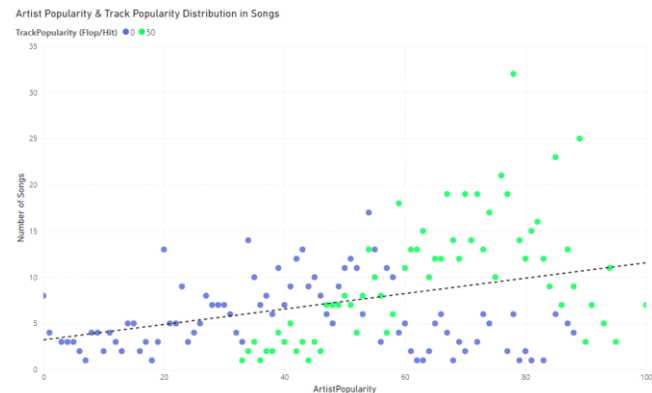**35%** current popularity on Spotify

Q1: Are louder songs shorter or longer in duration?

Q2: Are more energetic songs louder too?

Q3: Does an artist's popularity affect the track popularity?

1. Predicting the Track Popularity for a new song

Model	Interaction	MSE	R^2 Score
Linear Regression	No	347.162	0.404
Random Forest Regressor	No	332.036	0.522
Linear Regression	Yes	359.364	0.483
Random Forest Regressor	**Yes**	**328.162**	**0.528**

3. Scoring Based on Similarity of Songs' AudioFeatures (MSE out of all sample records: **0.1693**)

4. Deep Learning (MSE of **388.91**)

2. Classification of Hit/Flop Song

Model	Interaction	Validation Method	Accuracy
Logistic Regression	No	Train/Test Split	0.7716
AdaBoost Classifier	No	Train/Test Split	0.8082
Random Forest Classifier	**No**	**Train/Test Split**	**0.8173**
Logistic Regression	Yes	Train/Test Split	0.8264
AdaBoost Classifier	**Yes**	**Train/Test Split**	**0.8447**
Random Forest Classifier	Yes	Train/Test Split	0.8219
Logistic Regression	Yes	10-Fold Cross Validation	0.8103
AdaBoost Classifier	Yes	10-Fold Cross Validation	0.7735
Random Forest Classifier	**Yes**	**10-Fold Cross Validation**	**0.8038**

Evaluation of Best Model:

We ultimately got 3 best models:

- For train/test split without interactions: **Random Forest Classifier**
- For train/test split with interactions: **AdaBoost Classifier**
- For 10 fold CV with interactions: **Random Forest Classifier**

We have decided to go with the 3rd option:

Random Forest Classifier with interactions and 10 Fold CV

Validation for the following reasons:

- Our dataset was limited. We are more likely to encounter outliers and other noise in the real world. This model will be more robust in such a scenario.
- This model will not tend to overfit with the training data.

If our team were a strategy consulting agency offering insights to TikTok…

- **Curating a selection of songs** that we anticipate will become popular among the recently added tracks on Spotify
- **TikTok** could elevate its platform by either **introducing templates** that align with these potential hits or **actively promoting** these songs. Also, identifying hits early offers TikTok a chance to **secure favorable licensing deals**.
- **Content creators**, armed with advance knowledge of these predicted hit songs, could seamlessly incorporate them into their content, potentially **expanding their audience reach**.

SAY MY GRACE: Offset ft. Travis Scott

Hope You Know: Kodak Black

My Simple Jeep: Eyedress ft. Mac DeMarco

The above songs are going to become hits based off of our consultants' analysis. TikTok will benefit greatly from allocating resources on these songs

COST OF DECISION

- **5%** increase in engagement, assuming each % point translates to an additional $1 million in ad revenue, could boost TikTok's coffers by an estimated **$5 million**
- If negotiations reduce licensing costs by 10%, with the average hit song's license costing around $100,000, TikTok could save approximately **$10,000 per song**.

THANK YOU