



# MUSIC SUCCESS PREDICTION



Kush Ashesh  
Shah



Dhvanil  
Sanjay Shah



Hardik  
Prakashbhai  
Shah



Adélie  
Plumasseau

## Problem

## Statement

# CLASSIFYING HIT SONGS

- ❖ Can we predict song popularity based on lyrics ?
- ❖ Can lyrical patterns optimize songwriting, marketing, and music recommendations?
- ❖ Can songwriters, producers, and streaming platforms benefit from AI-driven insights?



# Business Objectives

WHO  
**Benefits**  
FROM THIS  
PROJECT?



## **Music Producers & Labels**

They can use AI insights to spot potential hits before investing in production and marketing.



## **Marketing & Promotion Teams**

They can tailor campaigns based on lyrical trends that resonate with audiences.



## **Streaming Platforms (Spotify, Apple Music)**

Personalized playlists and AI-driven recommendations can improve based on lyrical patterns.



## **Independent Artists**

Up-and-coming musicians can analyze trends to craft lyrics that increase their chances of success.



# Overview



**Can we predict song popularity based on lyrics using NLP & ML?**

To answer this, we structured our project into 6 sections:

- Data Collection & Preprocessing 🚒
- Feature Engineering 🎵
- Model Selection and Training
- Result & Performance Analysis 📊
- Challenges & Limitations ⚠
- Demo!! 🚀🎵



# Data Collection & Preprocessing

 **Dataset:** [Billboard Hot weekly charts](#)

 **Goal:** Ensure high-quality, structured data for accurate NLP modeling.

## **Methodology**

- **Deduplicated Records:** Grouped by **Song & Performer**, using mode for categorical & mean for numeric features.
- **Feature Selection:** Removed irrelevant columns and merged tables on SongID.
- **Lyrics Collection:** Built a **multi-threaded scraper** to fetch lyrics.
- Applied lower casing, treated non-alphabetic characters , filtered lyrical tags, handaled contractions, and remove stop words.



# Feature Engineering

- **Word2Vec:** Captured word relationships through context.
  - **TF-IDF:** Ranked word importance based on frequency.
  - **Sentiment Analysis & Word Patterns:** Analyzed emotions and lyrical complexity.

## Word Cloud for Hit Songs

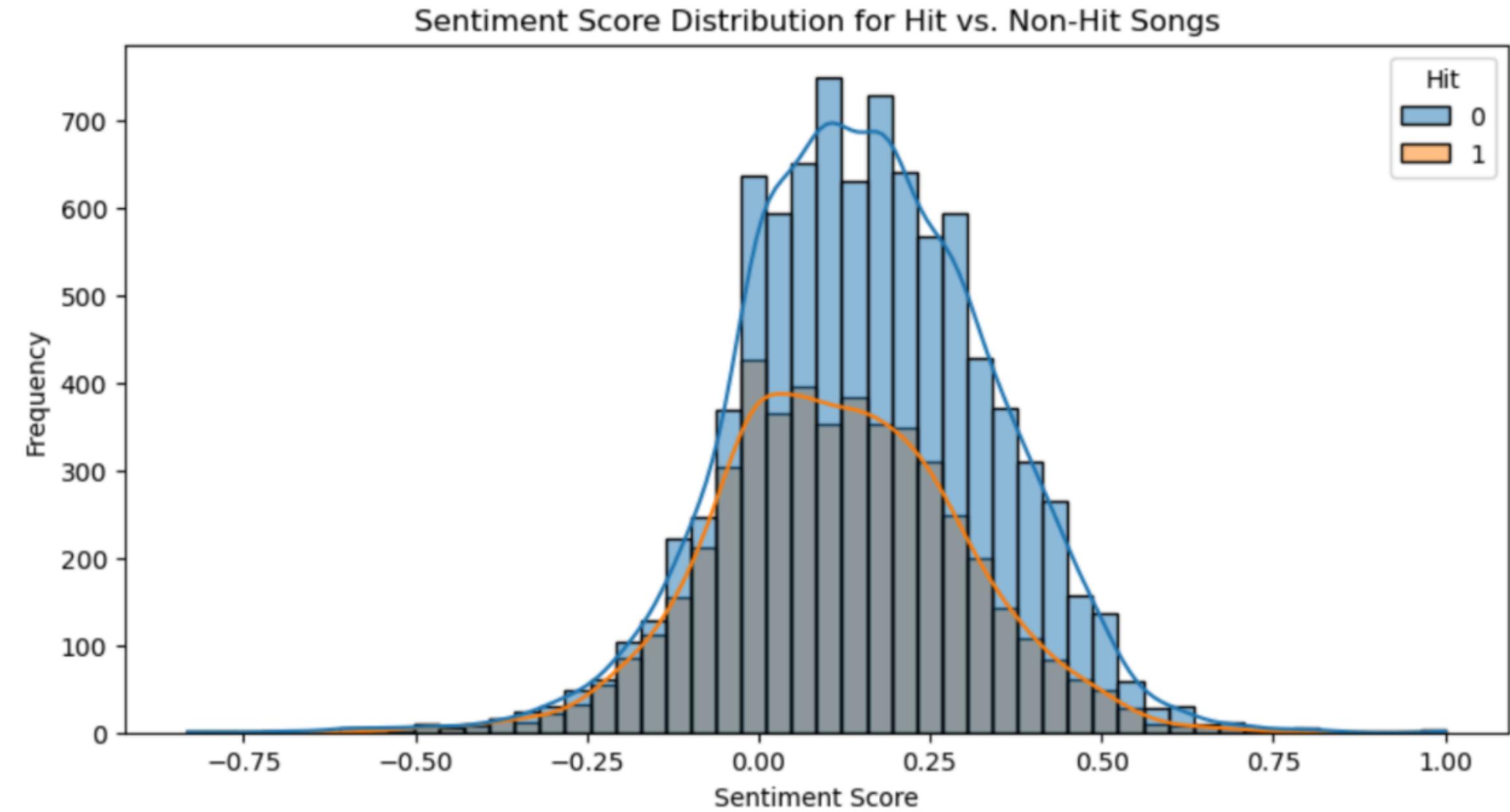


Words like "know," "got," "one," "love" appear prominently in **both hit and non-hit songs**, indicating their widespread usage in song lyrics.



# Feature Engineering♪

- **Word2Vec:** Captured word relationships through context.
- **TF-IDF:** Ranked word importance based on frequency.
- **Sentiment Analysis & Word Patterns:** Analyzed emotions and lyrical complexity.



**Balanced Sentiment:** Most songs, whether hits or non-hits, have neutral sentiment scores

**Lack of Extreme Sentiments:** Very negative ( $< -0.5$ ) or highly positive ( $> 0.5$ ) lyrics are uncommon in both categories



# Model Selection & Training

- **LSTM (Best for sequences, captures context).**
- **Random Forest (Structured feature analysis).**
- **AutoGluon (Best accuracy, automates optimization).**

<u>Model</u>	<u>Accuracy</u>	<u>Strength</u> 💡	<u>Limitation</u> ⚠️
LSTM 🏆	64%	<b>Best for sequential text &amp; context learning.</b>	Needs large dataset, struggles with short lyrics.
Random Forest	60.5%	Works well with structured data and ensures an ensemble approach.	Doesn't understand word order.
SVM	62.3%	Good for text classification (TF-IDF, Word2Vec).	Struggles with complex, long sequences.
XGBoost	61%	Strong on structured features.	Not ideal for lyrics-based NLP.
AutoGluon 🏆	72.2%	Best performer—automates model tuning & combines multiple approaches.	No major downside!

## Results & Performance Analysis

**AutoGluon**  
**72.2% Accuracy**

**LSTM**  
**64% Accuracy**

- **Random Forest:** 60.5% Accuracy
- **SVM:** 62.3% Accuracy
- **XGBoost:** 61% Accuracy



**LSTM captures lyrical sequence but with slightly lower accuracy.**

### **Song Structure Insights:**

**Highly repetitive lyrics** → more likely to be a hit.

**Emotionally positive songs** → slightly more likely to be a hit.

### **Unexpected Findings:**

Word count doesn't significantly impact success.

Sentiment alone isn't a strong indicator of a hit song.



## Challenges & Limitations !

**Music Success is NOT Just About Lyrics!**



### **Challenge:**

Our model only analyzes lyrics, but **song success depends on many factors**—production, marketing, artist reputation, and trends.

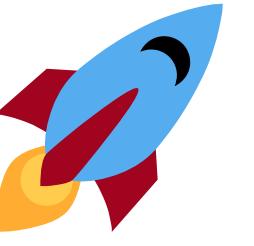


### **Limitation:**

Success is **subjective**, so we need to **consider cultural, emotional, and industry influences** beyond text analysis.



# Demo!!



# Song Popularity Predictor

Predict song's popularity based on its lyrics

Singer/Artist Name

Ellie Goulding

Song Name

Enter song name

- Ed Sheeran
- Ellie Goulding X
- asd
- hey
- 25Eminem
- Eminem

Find Lyrics

# Song Popularity Predictor

Predict song's popularity based on its lyrics

Singer/Artist Name	Song Name
<input type="text" value="Ellie Goulding"/>	<input type="text" value="Love Me Like You Do"/>
<input type="button" value="Find Lyrics"/>	

## Lyrics

You're the light, you're the night  
You're the colour of my blood  
You're the cure, you're the pain  
You're the only thing I wanna touch  
Never knew that it could mean so much, so much

You're the fear, I don't care  
Cause I've never been so high  
Follow me through the dark  
Let me take you past the satellites  
You can see the world you brought to life, to life

So love me like you do, lo-lo-love me like you do  
Love me like you do, lo-lo-love me like you do  
Touch me like you do, to-to-touch me like you do  
What are you waiting for?

Fading in, fading out

Uniqueness

30.32%

Repetition

76.98%

Lexical Diversity

25.69%

Rhythm Density

75.06%

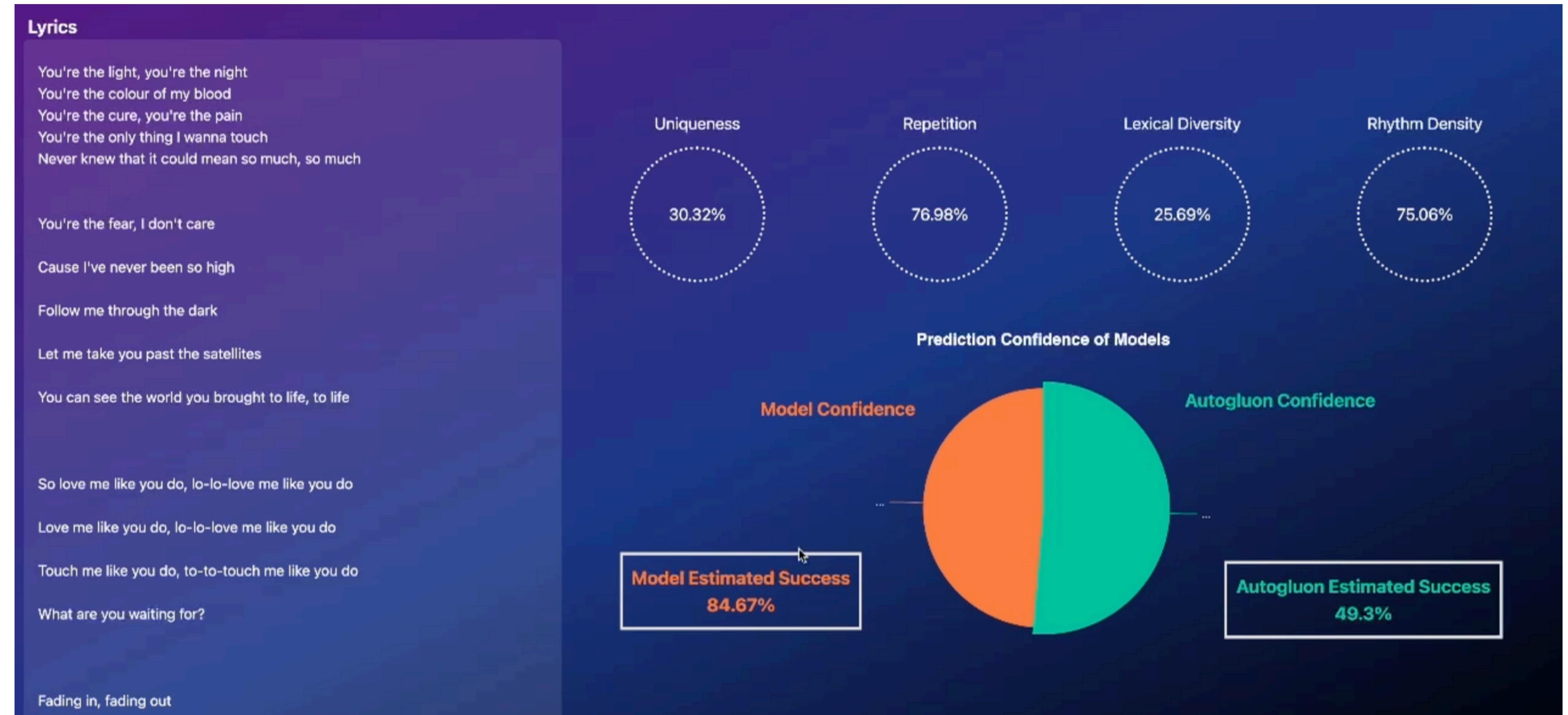
Prediction Confidence of Models

Model Confidence

Model Estimated Success  
84.67%

Autogluon Confidence

Autogluon Estimated Success  
49.3%



## Submit Lyrics

### Lyrics

...  
That I know you can't afford  
Bite that tattoo on your shoulder  
Pull the sheets right off the corner  
Of that mattress that you stole  
From your roommate back in Boulder  
We ain't ever getting older  
We ain't ever getting older  
We ain't ever getting older

Submit Lyrics

## Lyrics

Hey, I was doing just fine before I met you  
I drink too much and that's an issue, but I'm okay  
Hey, you tell your friends it was nice to meet them  
But I hope I never see them again  
I know it breaks your heart  
Moved to the city in a broke-down car, and  
Four years, no calls  
Now you're looking pretty in a hotel bar  
And I, I, I, I, I can't stop  
No, I, I, I, I, I can't stop  
So, baby, pull me closer  
In the back seat of your Rover  
That I know you can't afford  
Bite that tattoo on your shoulder  
Pull the sheets right off the corner  
Of that mattress that you stole  
From your roommate back in Boulder  
We ain't ever getting older  
We ain't ever getting older  
We ain't ever getting older  
You look as good as the day I met you  
I forget just why I left you, I was insane  
Stay and play that Blink-182 song  
That we beat to death in Tucson, okay  
I know it breaks your heart  
Moved to the city in a broke-down car, and  
Four years, no call  
Now I'm looking pretty in a hotel bar  
And I, I, I, I, I can't stop  
No, I, I, I, I, I can't stop



Prediction Confidence of Models



# Reference



**Can Song Lyrics Predict Hits**

