

Long Island University - MDA 720 Capstone

# Predicting the #1 Song on Spotify

An Analytics Approach for Kalshi Market Insights

Presented by Kenneth Lent

Long Island University - MDA 720 Capstone

# Overview

**01** The Challenge

**02** Project Goal

**03** Fueling the Model

**04** Forecasting #1

**05** Prediction Check

**06** Prediction to Profit

**07** Kalshi Web Page

**08** Conclusion & Path  
Forward

**09** Acknowledging Limits  
& Next Steps

Long Island University - MDA 720 Capstone

# The Challenge

An Analytics Approach for Kalshi Market Insights

- Kalshi offers daily prediction markets: "Which song will be #1 on Spotify USA tomorrow?"
- Market settles at 11 a.m. ET based on official Spotify U.S. Top 50 chart.
- Prices (1¢-99¢) reflect market's collective belief in a song's probability of hitting #1.
- **Opportunity:** External popularity signals (streaming, search, lyric engagement) often arrive before market settlement, creating potential for a data-driven trading edge.

The Kalshi logo is a teal square with the word "Kalshi" in white, bold, sans-serif font.

**Kalshi**

Long Island University - MDA 720 Capstone

# Project Goal

Streaming-Sentiment Signals

- **Primary Objective:** Develop an analytical tool to predict the probability of a song reaching #1 on Spotify USA, providing actionable insights for Kalshi market traders.
- **Core Idea:**
  - Ingest and process timely data signals (Spotify charts, API metadata).
  - Engineer features capturing song momentum and characteristics.
  - Train a predictive model to generate calibrated probabilities.
  - Identify mispricings in the Kalshi market.

Long Island University - MDA 720 Capstone

# Fueling the Model

Data & Features

**01**

**Spotify Charts CSVs**

Daily rank, streams,  
track/artist IDs.

**02**

**Spotify Web API**

Track popularity,  
duration, explicitness,  
release date.

**03**

**Google Trends**

Rate limiting challenges.  
Didn't end up needing  
these features for a well-  
performing model.



Long Island University - MDA 720 Capstone

# Fueling the Model

Data & Features

## Key Engineered Features (Data Mining):

- days\_since\_release
- rank\_change (daily)
- stream\_momentum (vs. 3-day rolling average)





Long Island University - MDA 720 Capstone

# Forecasting #1

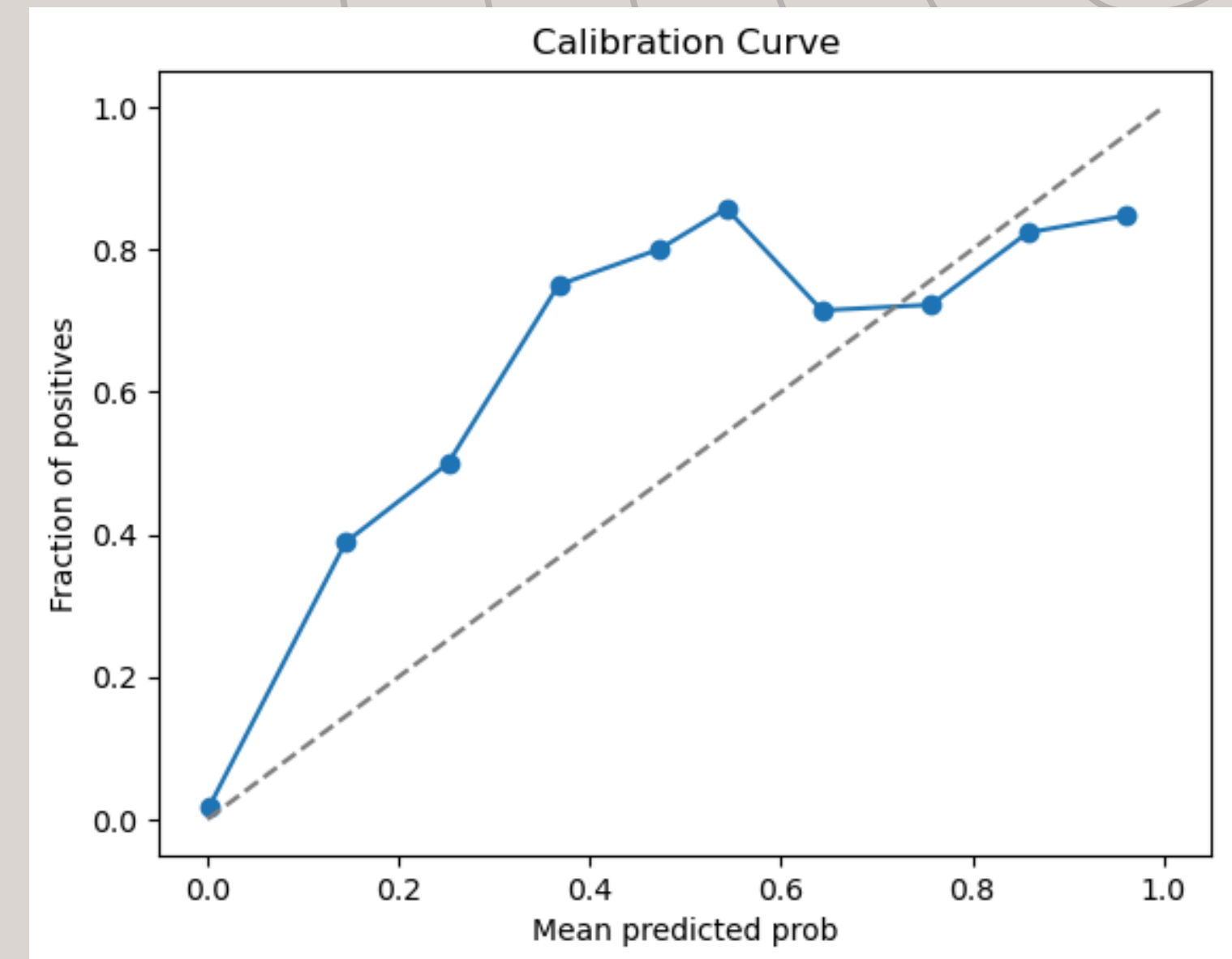
The Model

- **Model:** LightGBM Classifier
  - Chosen for efficiency and performance.
  - `class_weight='balanced'` to handle #1 song rarity.
- **Training & Testing Strategy:**
  - Time-Based Rolling Window:
    - Evaluated daily from Jan 1, 2025, to May 11, 2025.
    - Train on all prior data, predict for the next day.
    - Simulates real-world application.
- **Key Features Used:** rank, streams, rank\_change, stream\_momentum, popularity, duration\_ms, explicit, days\_since\_release.



# How Well Does It Predict?

- **Overall Predictive Power (Aggregated Multi-Day Results):**
  - ROC AUC: 0.97 (Excellent discrimination)
  - Average Precision (PR Curve): 0.81
- **Daily Prediction Averages (for #1 song):**
  - Precision: 0.68
  - Recall: 0.68
- **Calibration:** Model probabilities reasonably well-calibrated.



"Predicted probabilities are generally reliable."



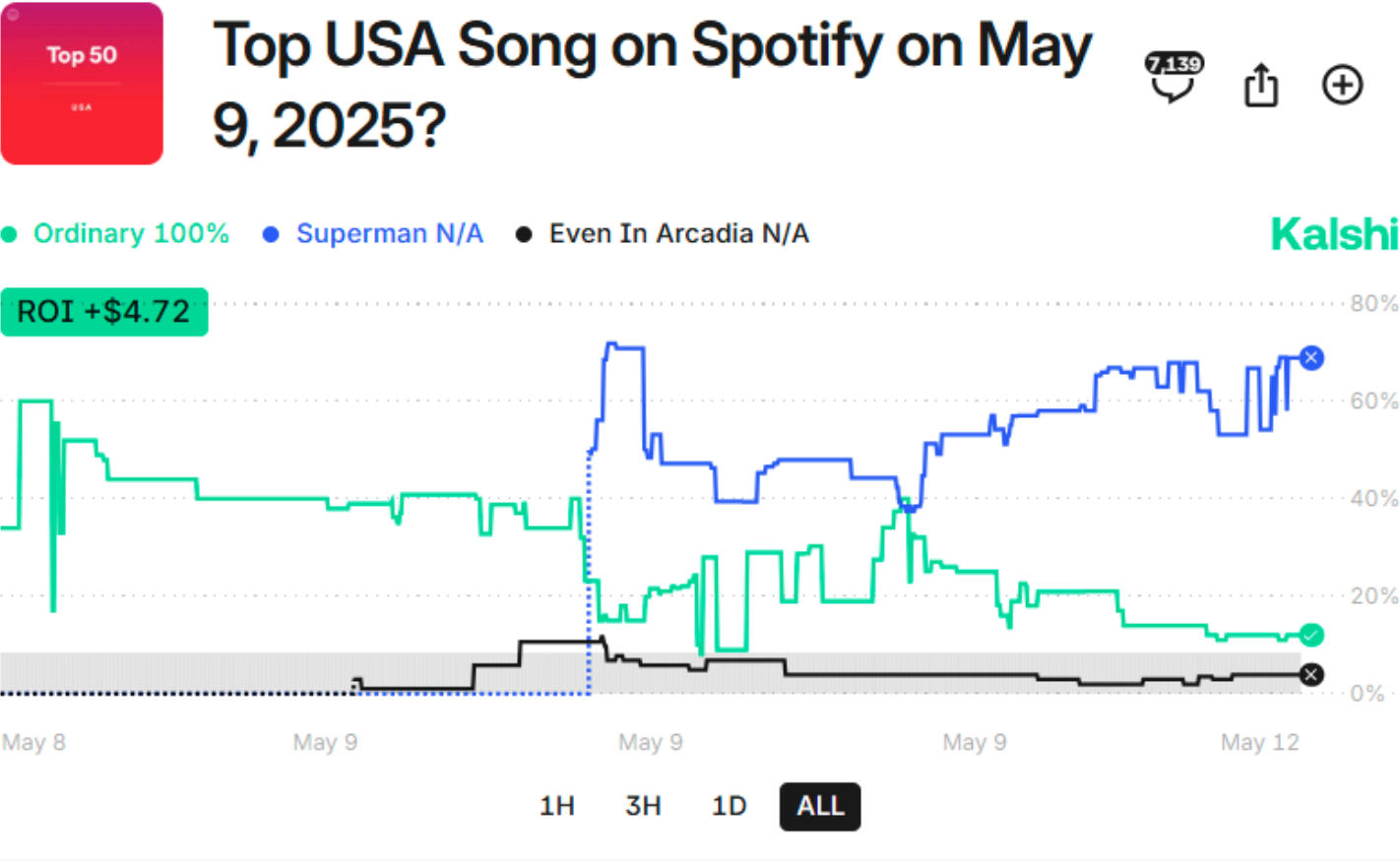
# From Prediction to Potential Profit

## Kalshi Market Edge

- Model generates probabilities before market settlement.
- Allows identification of potential over/undervalued contracts by comparing model probability to market-implied odds.
- **Important Note:** Predicting #1 is the first step. Actual profitability requires backtesting against Kalshi market conditions (liquidity, spreads).

## Key Feature Drivers

- streams and current rank are highly influential.
- stream\_momentum and general Spotify popularity also important.



May 9, 2025 Chart (+2)

\$70,028 open int

	Ordinary	Alex Warren	Yes
+\$11.36 (245%) total return			
	Superman	Morgan Wallen	No
-\$5.28 (-100%) total return			
	Bliss	Tyla	No
	Look To Windward	Sleep Token	No



### Determined

Market feedback

Top 50

Top USA Song on Spotify tomorrow?

7,139

May 13, 2025 Chart (+1)

Chance

\$93 open int

	BIRDS OF A FEATHER	Billie Eilish	<1%	Yes 1¢	No
	Ordinary	Alex Warren	--%	Yes 99¢	No 95¢
	NOKIA	Drake	--%	Yes	No
	luther (with sza)	Kendrick Lamar	--%	Yes	No
	undressed	sombr	--%	Yes	No
	back to friends	sombr	--%	Yes 97¢	No
	Die With A Smile	Lady Gaga	<1%	Yes 1¢	No
	Pink Pony Club	Chappell Roan	<1%	Yes 1¢	No



Top USA Song on Spotify tomorrow?

Buy Yes · Ordinary

Alex Warren

BuySellLimit

Pick a side

Yes 99¢No 95¢

Contracts0

Limit price0¢

Set expiration☐

Submit as resting order only☐

Estimated cost\$0

Payout if Yes wins\$0

Review

# Conclusion & Path Forward

- **Conclusion:**
  - Successfully developed an analytical tool predicting Spotify's #1 song with strong performance (ROC AUC 0.97).
  - Demonstrates value of API usage, feature engineering, and predictive modeling for market insights.
- **Key Recommendations:**
  - **Develop Backtesting Framework:** Crucial for assessing real-world trading viability against Kalshi odds.
  - **Expand Data Sources:** Incorporate web scraping (social media, news) and other APIs (Genius) to bolster the model.
  - **Monetization Potential:** Data pipeline and model can be basis for a subscription insights service.

# Acknowledging Limits & Next Steps

## Limitations

- Google Trends unfeasibility (due to rate limits, but model still performed well).
- Relies on data availability; no live trading implemented.

## Future Work

- Implement financial backtesting.
- Integrate diverse data (social media, news).
- Explore advanced models (ensembles, neural networks).
- Develop a user-friendly dashboard.
- Automate the daily pipeline.

Long Island University - MDA 720 Capstone

# Thank You & Questions



[linkedin.com/in/klent-ai/](https://www.linkedin.com/in/klent-ai/)