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

Presented by Juliana Silva

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.

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

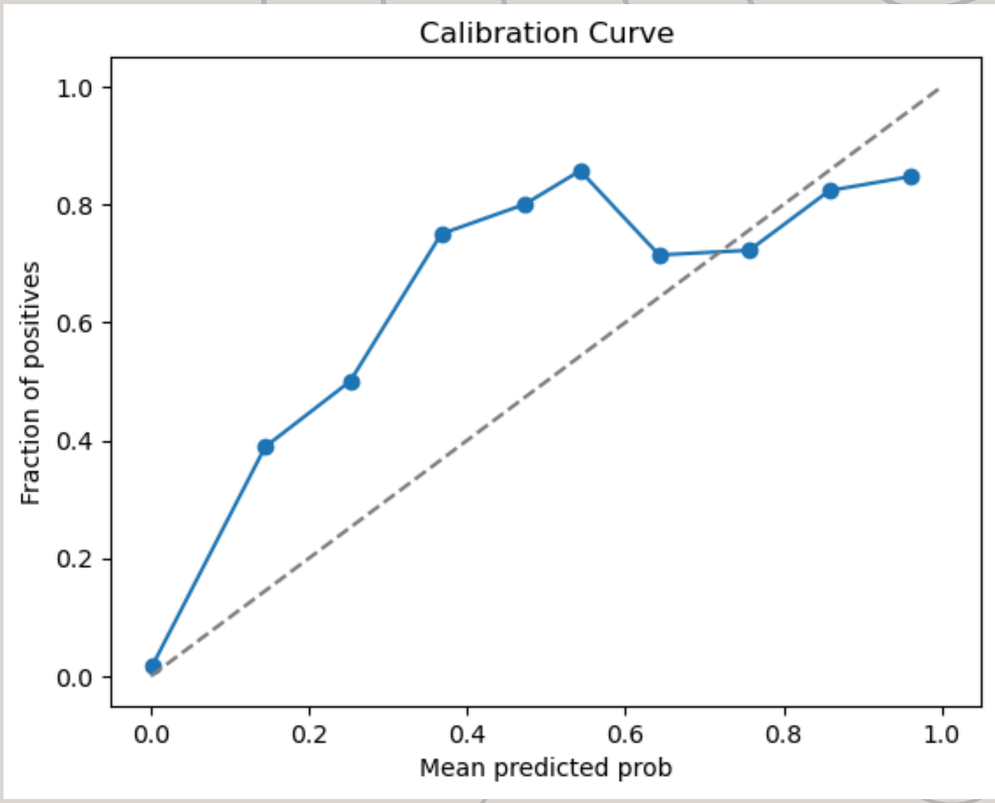
The Model

- **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.

Long Island University - MDA 720 Capstone

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."

Long Island University - MDA 720 Capstone

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.

Kalshi Markets Live Ideas

Search markets or profiles /

$****** Cash

$****** Portfolio

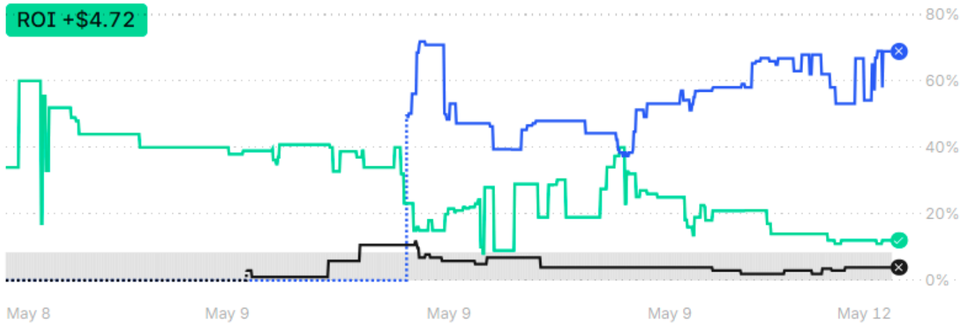
Top USA Song on Spotify on May 9, 2025?

7139

Ordinary 100% · Superman N/A · Even In Arcadia N/A

Kalshi

ROI +$4.72



1H 3H 1D ALL

May 9, 2025 Chart (+2)

$70,028 open int

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

Determined

Market feedback

Kalshi Markets Live Ideas

Search markets or profiles /

$****** Cash

$****** Portfolio

Top USA Song on Spotify tomorrow?

7139

May 13, 2025 Chart (+1)

	Chance		
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

$93 open int

Top USA Song on Spotify tomorrow?

Buy Yes · Ordinary

Alex Warren

Buy Sell Limit

Pick a side

Yes 99¢ No 95¢

Contracts 0

Limit price 0¢

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/