A Multi-Domain Feature Framework for Predicting Music Streaming Success

Benjamin Morris, M.S. in Data Analytics and Visualization

FACULTY MENTOR: Andrew Catlin, Industry Professor

Katz School of Science and Health

Introduction

Music labels rely on simplified representations of music to model and predict performance. Spotify, for example, characterizes songs using a limited set of audio features such as tempo, energy, and valence. While useful, these features provide only a surface-level view and fail to capture deeper musical, lyrical, and contextual drivers of performance. Key dimensions such as lyrics, structure, production, and context are excluded, limiting predictive accuracy and insight generation.

To address this gap, this project introduces a multi-domain feature framework that expands beyond Spotify's standard feature set to include hundreds of attributes spanning lyrical analysis, sonic characteristics, harmonic structure, production elements, and contextual metadata. Songs are treated as structured objects, enabling a more comprehensive representation of their underlying "musical DNA."

The objective of this study is to evaluate whether this expanded feature set improves the ability to model and predict streaming performance compared to Spotify's existing framework. By integrating these features with historical streaming data, this project aims not only to enhance predictive accuracy, but also to support more informed, data-driven creative decisions and release strategies for music labels.

Approach

An end-to-end pipeline was developed to transform songs into structured data for predictive modeling. Starting from raw audio, lyrics, and metadata, features were extracted across multiple domains, including lyrical content, sonic characteristics, harmonic structure, and contextual signals. This process expands Spotify's feature set by decomposing high-level attributes into more granular components, creating a higher-dimensional feature space. The enriched dataset was then combined with historical streaming data and used to train models that predict early streaming performance. Model outputs were compared against predictions derived from Spotify's standard features to evaluate improvements in accuracy.

Feature Space Expansion

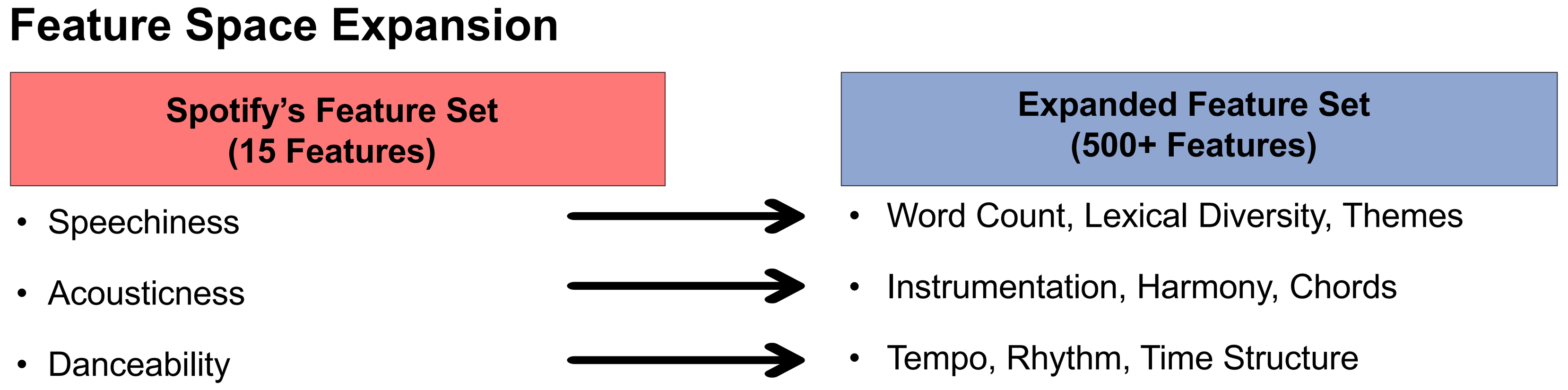


Figure 1: Feature Space Expansion

Modeling Pipeline Architecture

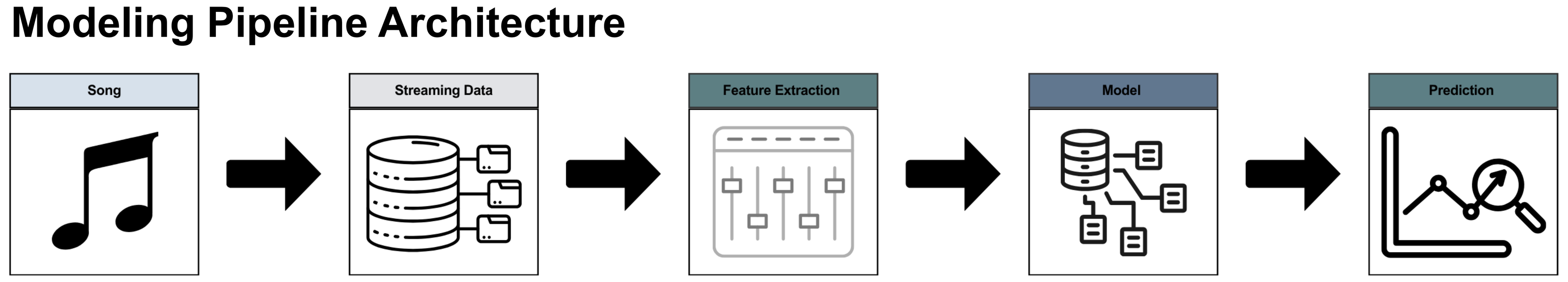


Figure 2: Modeling Pipeline Architecture

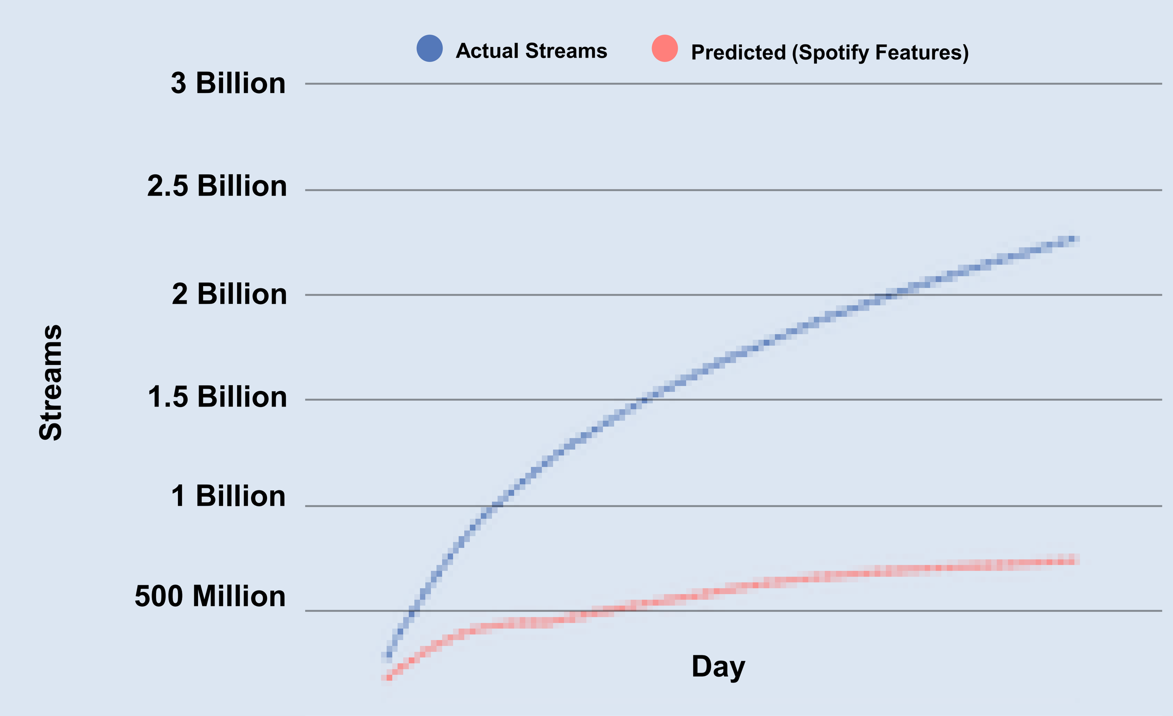
Findings

Using a consistent neural network approach, models were iteratively evaluated (~40 runs) across both baseline and expanded feature spaces, with the expanded dataset consisting of hundreds of engineered features derived from historical streaming data

Across iterative runs, the expanded feature set produced more consistent and stronger predictive performance relative to the baseline features, with predictive error remaining within approximately 4% across multiple time horizons (e.g., 7, 30, and 60 days post-release), showing that **richer feature representations better capture the drivers of performance**.

Overall, the results support the hypothesis that richer, multi-domain feature representations provide a more robust and realistic framework for modeling music performance.

Improved Predictive Fit with Expanded Feature Space
Spotify Model Predictions vs. Actual Streams



Expanded Model Predictions vs Actual Streams

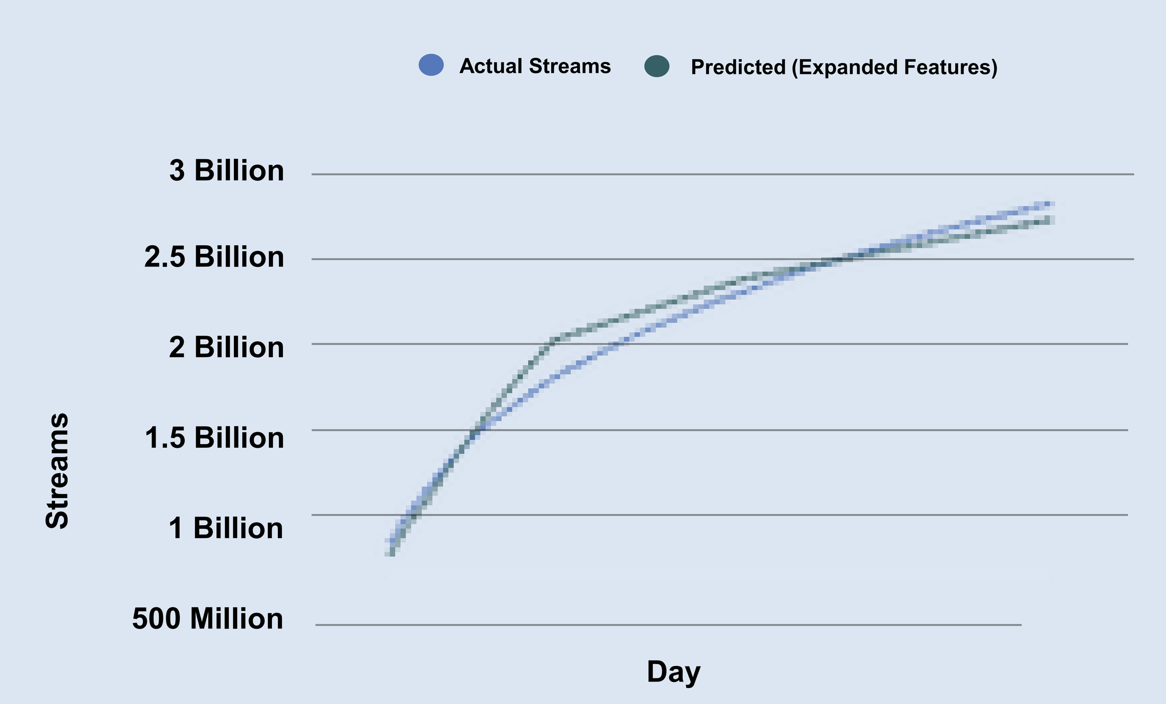


Figure 3: Improved Predictive Fit with Expanded Feature Space

Conclusions

This work challenges the adequacy of simplified feature representations commonly used in music analytics. While platforms such as Spotify provide useful baseline descriptors, they do not fully capture the complexity of how music is created, experienced, and consumed.

By reframing songs as structured, multi-dimensional objects, this project demonstrates a more effective approach to modeling streaming performance, one that aligns more closely with both artistic composition and listener behavior.

The findings suggest that future predictive systems in music should prioritize feature depth and cross-domain integration rather than relying on limited, surface-level metrics.

Beyond prediction, this approach has broader implications for feature-based systems, including recommender systems, similarity modeling, and content discovery frameworks, where richer representations may lead to more meaningful and accurate results.

A key limitation of this study is that the analysis was conducted on a single album; future work should validate the framework across a broader, more diverse set of artists, genres, and release context

From an applied perspective, this framework may inform release strategy, track sequencing, and creative decision-making by providing a data-driven understanding of factors associated with sustained engagement.

Key Insights

- Predictive performance is driven by the quality and depth of feature representation, with expanded feature spaces better capturing the underlying drivers of streaming performance
- Accurate modeling requires integrating lyrical, sonic, and contextual features to reflect real-world listening behavior
- Once predictive performance is established, these models can support data-driven creative decisions aimed at improving streaming outcomes

Acknowledgments

I sincerely thank Professor Andrew Catlin for his time, guidance, and continued support throughout this project.

References