Satisfaction Scouts

Erdos Summer 2024

Concert Lineup Planning with Data Science

Vishal Bhatoy, Peter Graziano, Anish Joseph, Rachel Lopez, Eric Malitz, and Obada Nairat.

Introduction

Concert lineup planning is traditionally a manual process, requiring significant time and resources from concert organizers. This method is not only inefficient but also prone to subjective decision-making, which can result in less-than-optimal lineups. In an effort to streamline this process and improve lineup quality, we have developed a machine learning model designed to rank selected artist and predict the three nearest neighbor artists for each part of the lineup. This approach ensures a cohesive and engaging experience for the audience while significantly reducing costs and effort for organizers. The market size, measured by revenue, of the Concert & Event Promotion industry was $35.7 billion in 2023, highlighting the substantial financial impact of optimizing concert lineups.

Problem Statement

Concert organizers face numerous challenges in manually selecting and ranking artist for events. These challenges include extensive research to identify suitable artists, coordination with booking agents, and ensuring that the selected lineup appeals to the target audience. Our solution automates this process using data-driven recommendations, allowing for more efficient and accurate lineup planning.

Dataset

Our model utilizes data sourced from the Spotify API, encompassing a wide range of audio features and metadata. Key features include danceability, energy, loudness, speechiness, acousticness, instrumentalness, valence, tempo, genre, popularity, and explicit content flags. This comprehensive dataset provides a robust foundation for our machine learning model to analyze and predict artist compatibility.

Exploratory Data Analysis

To understand the dataset and identify key trends and relationships, we performed extensive exploratory data analysis (EDA). This involved visualizing the distribution of various audio features, examining correlations between features, and identifying patterns that could inform our modeling approach. The insights gained from EDA were crucial in guiding our data preprocessing and feature selection steps.

Methodology

The methodology for our concert lineup planning model consists of several key steps:

1. **Data Collection**: Using Spotify's Web API, we gathered detailed audio features and metadata for a large set of tracks.
2. **Data Preprocessing**: We cleaned the data, and scaled numerical features to ensure consistency and readiness for modeling.
3. **Modeling**: We employed the K-Nearest Neighbors (KNN) algorithm to identify the nearest neighbors for a given track based on its audio features. This choice was driven by KNN's effectiveness in handling similarity-based recommendations.

Model Training

The model training process involved several iterative steps:

- **Initial Training**: We first trained the KNN model using the preprocessed features.
- **Feature Enhancement**: Predictions from initial models were used as additional features to refine the KNN model. This iterative enhancement process improved the accuracy and relevance of our recommendations.

Application

Our user interface is designed to be intuitive and user-friendly:

- **Input**: Users input the artist and track they are interested in.
- **Output**: The system provides the nearest neighbors and suggests a complete lineup.

Benefits:

- **Time Savings**: Automating the lineup creation process significantly reduces the time required for research and decision-making.
- **Enhanced Audience Experience**: Data-driven recommendations ensure more cohesive and engaging lineups, improving the overall concert experience.

Cost Savings

Our machine learning model offers substantial financial benefits:

- **Research Cost Savings**: Automating research reduces costs
- **Marketing Efficiency**: Improved targeting increases ticket sales
- **Administrative Cost Reduction**: Reduces coordination time and costs
- **Increased Ticket Sales**: Optimized lineups attract more attendees

Future Work

To further enhance our model, we plan to:

- **Expand the Dataset**: Include more tracks and features to improve accuracy.
- **Incorporate Real-Time Data**: Use live data to adapt recommendations based on current trends.
- **Add Features**: Implement automated scheduling, budgeting tools, and lead capture mechanisms.

Conclusion

Our data science-driven approach to concert lineup planning offers a significant improvement over traditional methods. By leveraging machine learning, we can create more cohesive and engaging lineups, save time and resources, and enhance the overall concert experience. This innovative solution promises better lineup curation and greater audience satisfaction, demonstrating the effective application of data science in the entertainment industry.