Association for Information Systems
AIS Electronic Library (AISeL)

Research Papers | ECIS 2017 Proceedings

Spring 6-10-2017

UNDERSTANDING MUSIC TRACK POPULARITY IN A SOCIAL NETWORK

Jing Ren
Singapore Management University, Singapore, jing.ren.2012@phdis.smu.edu.sg

Robert Kauffman
Singapore Management University, Singapore, rkauffman@smu.edu.sg

UNDERSTANDING MUSIC TRACK POPULARITY IN A SOCIAL NETWORK

Research paper

Jing Ren, Singapore Management University, Singapore, jing.ren.2012@phdis.smu.edu.sg

Robert J. Kauffman, Singapore Management University, Singapore, rkauffman@smu.edu.sg

Abstract

Thousands of music tracks are uploaded to the Internet every day through websites and social networks that focus on music. While some content has been popular for decades, some tracks that have just been released have been ignored. What makes a music track popular? Can the duration of a music track's popularity be explained and predicted? By analysing data on the performance of a music track on the ranking charts, coupled with the creation of machine-generated music semantics constructs and a variety of other track, artist and market descriptors, this research tests a model to assess how track popularity and duration on the charts are determined. The dataset has 78,000+ track ranking observations from a streaming music service. The importance of music semantics constructs (genre, mood, instrumental, theme) for a track, and other non-musical factors, such as artist reputation and social information, are assessed. These may influence the staying power of music tracks in online social networks. The results show it is possible to explain chart popularity duration and the weekly ranking of music tracks. This research emphasizes the power of data analytics for knowledge discovery and explanation that can be achieved with a combination of machine-based and econometrics-based approaches.

Keywords: Econometrics, Machine Learning, Music Social Networks, Track Popularity

1 Introduction

With contemporary digital entertainment, people can easily access large music collections and stream content via social networks such as Last.fm, Spotify and YouTube. Streaming music and social networks have changed listener behavior dramatically. They can "listen," "like," and "comment" on music tracks, and communicate with and affect other listeners through social communication. In comparison to an album or a radio broadcast, listeners can make much richer selections. They can listen to tracks repetitively or freely shift to other content too.

Music is a *durable information good* that can bring utility to listeners and value to artists. Even one strong and widely appreciated song can lead to the rise of a new music superstar, such as "Rolling in the Deep" for Adele, or "Poker Face" for Lady Gaga. Moreover, a classic track can make people remember the singer, even many years after release. Examples include "Hey Jude" by The Beatles, which *Billboard* named the tenth most popular song of all time in 2012, although it was first released in 1968 (Bronson, 2012). Great business value for music and musicians is the natural outcome. *Forbes* (2016) has reported that Beyoncé's net worth was around US$265 million in June 2016 (Greenburg, 2016), and Adele's around US$80 million. So the music labels have paid attention to how music can be promoted by social networks to maximize its market value (IFPI, 2015).

Researchers and industry pros have been exploring the ingredients for music to achieve sustainable popularity (Chon et al., 2006; Karydis et al., 2016; Nunes et al., 2015). It has become possible to explain how a song became popular, and predict future music superstars. Most have considered music and artist factors, or market and social factors (e.g., Bischoff et al., 2009; Koenigstein and Shavitt, 2009). This work tries to determine how effective music promotion investment activities will be, based on analyzing the popularity performance of a large set of tracks since their release in social media.

This research explores: in a music social network, what factors produce a popular track? (1) Is the music content most important? (2) Can a song's popularity duration be predicted, based on hidden factors? (3) How much does the social context for a track affect the duration of its popularity? (4) And are there discernible popularity patterns for music tracks that are suggested by our research inquiry? This research applies *computational social science* methods that combine machine-based methods for data analytics from computer science (CS) and explanatory methods from information systems (IS) research (Chang et al., 2014). This permits a researcher to capture and analyze different kinds of data that would not be possible using non-machine methods, secondary datasets, or interviews.

2 Related Research

Music popularity analytics have attracted wide attention in multiple research fields, covering IS, CS, Society Science, and Psychology. Track popularity and related research methods are discussed here.

2.1 What Is Music Track Popularity?

There are various ways to define the popularity of a music track. They include: sales volume; the amount of audience listening that occurs via streaming music services; track performance on top-rank charts; and music industry awards received. No single standard to define popularity is recognized.

Most of the research on music popularity has been based on data from public sources (Chon et al., 2006; Herremans and Sörensen, 2014). Some studied *Billboard* rankings (Karydis et al., 2016; Lee and Lee, 2015; Nunes et al., 2015; Singhi and Brown, 2015). Others chose rankings like UKTopChart, or streaming music services, such as Last.fm, Spotify, and Twitter (Dhanaraj and Logan, 2005; Kim et al., 2014; Pachet and Roy, 2008). Some have observed music track performance since they reached top-chart ranking (Frieler et al., 2015; Karydis et al. 2016; Ni et al., 2011). Many used a binary variable to define *popular* or *non-popular* music track, based on chart ranking at a point in time. Such as if a track reached the Top 1, it was labeled as popular; if it never climbed above Top 90, it was non-popular (Nunes and Ordanini, 2014). Lee and Lee (2015) explored various definitions of popularity, for chart performance based on the chart debut position, total weeks on the chart, and so on.

All of them focused on just one stage of a track's developing popularity: after it reached top-chart ranking. This reflects a bias for understanding how a track's popularity developed: it missed the stage of run-up to top-chart ranking. Some of the popular tracks reached a top or even Top 10 ranking immediately, such as "Bad Romance" for Lady Gaga; some others may spend a long time till first appearing in the top-chart ranking. They may become very popular, such as "Little Lion Man" for Mumford & Sons. So how to properly define the *popularity of a music track* is still a challenge.

2.2 Background for Music Track Popularity Analysis and Prediction

CS Researchers have sought to predict music popularity with musical and non-musical features. In contrast, IS researchers have explored what factors explain the observed outcomes for music popularity. The former emphasizes accuracy, while the latter focuses on causation. Some authors have assessed *album popularity*, including those released during Christmas, and the impact of release timing on their success (Bhattacharjee et al., 2007). Other things that promoted *album popularity* are highly correlated with *artist reputation* and *superstardom* (Chung and Cox, 1994; Hamlen, 1991), *label association*, and the debut rank on *Billboard* (Strobl and Tucker, 2000). Nunes and Ordanini (2014) used logit regression to test the relationship between the number of instrumentation combination and the probability of high versus low-ranked tracks. Nunes et al. (2015) used survival analysis for how a song's chorus lyrics affected how fast *Billboard* Hot 100 song reached the Top 1 place.

Various CS authors have tried to find different combinations of musical and non-musical features to increase the accuracy of popular and non-popular track prediction. They used machine-based methods to extract feature sets for prediction, such as acoustic features (Borg and Hokkanen, 2011; Herremans and Sörensen, 2014; Frieler et al., 2015), social information (Koenigstein and Shavitt, 2009; Schedl, 2011; Kim et al., 2014), lyrics plus acoustic features (Dhanaraj and Logan, 2005; Singhi and Brown, 2015) and acoustic features plus early stage popularity (Lee and Lee, 2015). The prediction methods that have been used include *support vector machine* (SVM), *random forest* (RF), Bayesian network analysis, and so on. Most achieved no more than 67% in prediction accuracy. The best performance was achieved by Kim et al. (2014), with 92% accuracy for Top 10 song prediction, but with a limited dataset of 168 tracks over 10 weeks. So no general conclusions were able to be drawn.

It is also hard to compare the performances of different feature sets: there has been no standard dataset. Karydis et al. (2016) is the first work to construct a sharable musical track popularity dataset. This dataset covers 10 years of music ranking data from Last.fm, Spotify and *Billboard*. For each track in the dataset, its artist, album, acoustic features, ranking in the three charts, and similar tracks are included. And yet, other research has broadly shown that track popularity, especially in the social environment, cannot be explained or predicted by these attributes alone.

So the present research aims to: (1) offer new measures for music track popularity that consider the lifespan of a song from first release to top-chart popularity to chart drop-off; (2) construct a relatively complete and very large dataset which covers musical and non-musical constructs that describe the social and market aspects of tracks; and (3) implement machine-based and social science methods to understand and predict track popularity and ranking. This kind of research has become possible in recent years with methods for big data and computational social science research (Chen et al., 2012).

3 A Model for Popularity Duration in a Music Social Network

3.1 Music Popularity Measurement

Music streaming services, such as Last.fm and Spotify, integrate music listening, social network activities, and social recommendation into a single platform. They are appropriate for a study that seeks to understand music popularity development in a "small society" context. Compared to some public music ranking charts, such as *Billboard* and UKTopChart, streaming services record the listening log of each track over time, and rank their weekly listening time. In addition, music streaming services have been shown to be good proxies for a music track's ranking based on their high correlation with *Billboard* (Kim et al., 2014; Koenigstein and Shavitt, 2009; Schedl, 2011).[1]

This research work leverages the record to learn the development of music popularity in Last.fm. It focuses on the full lifespan of a track from its release, to when it reached a top-ranking on the chart, all the way until it dropped off and was no longer popular. Two measures are:

- *Time2TopRank*: Total weeks from a song's release date to the first date it reached a top-chart ranking. It shows how quickly it took for a song to get enough attention to reach a top ranking.
- *Duration*: The total weeks a song appeared in the top-chart ranking for popularity. It suggests how long a song matched people's tastes and was highly rated on Last.fm and *Billboard*.

The measures describe speed for achieving top recognition, and popularity sustainability over time.

3.2 Music Track Popularity Duration Model

A duration model is used to estimate when top-rank drop-off occurs. A *hazard function* specifies the duration until time t when this event happens. A *proportional hazard* (PH) *model* for this setting is: $\lambda\ (t \mid \boldsymbol{X_i}) = \lambda_0(t)\ exp\ (X_1\,\beta_1^{PH} + X_2\,\beta_2^{PH} + \ldots) = \lambda_0(t)\ exp\ (\boldsymbol{X_i} \cdot \boldsymbol{B})$. Here, $\lambda_0(t)$ is the *baseline hazard*, which represents the likelihood that a track drops off the top-rank chart. $\boldsymbol{X_i}$ are explanatory variables

[1] Koenigstein and Shavitt (2009) and Kim et al. (2014) reported on the strong correlation between song popularity on *Billboard*'s list and the extent of social media activity related to it in P2P networks and instant listening on Twitter. Schedl (2011) also offered evidence for high correlation between an artist's popularity on Twitter and the artist's ranking in Last.fm.

for a track i (i = 1, 2, …), and βi^{PH} are parameters to be estimated for all the data to gauge if there are modifications to the hazard rate of top-chart drop-off due to their values (Kleinbaum and Klein, 2006).

A Weibull hazard function $\lambda(t)$ for duration follows a monotonic curve, $\lambda(t) = \lambda z t^{z-1}$. In this model, λ is a scale parameter, and z is a shape parameter. The z value makes it so the hazard function can be constant, or steeply declining or increasing at an accelerating rate. This also fits situations in healthcare, finance, marketing and e-commerce. Other distributions are non-monotonic, such as the log-logistic hazard, with $\lambda(t) = \lambda z t^{z-1}/(1 + \lambda t^{z})$: it decreases after peaking. It captures the dynamics of situations that involve an initially increasing and later decreasing hazard rate, as with the diagnosis and treatment of leukemia and cancer. In contrast, a log-normal distribution follows a normal distribution, with a positive, skewed distribution with a lower mean time to event and a higher variance.

The present work considers these three hazard function models for the analysis of track popularity *Duration*. A linear model estimated with *ordinary least squares* (OLS) (Bhattacharjee et al., 2007) is also considered. With a log-transformation, this approximates the more refined hazard models, and can act as a baseline for estimation. Time-invariant musical constructs, such as genre and mood, are used. Non-musical constructs, such as the artist reputation and social context, are time-varying. By including fixed and time-varying covariates for each track, the general vector form of this model is: $\lambda(t) = f(\lambda, z, t; \boldsymbol{X}_{Music}\boldsymbol{B}_{Music}{}^{OLS}, \boldsymbol{X}_{Artist}\boldsymbol{B}_{Artist}{}^{OLS}, \boldsymbol{X}_{Social}\boldsymbol{B}_{Social}{}^{OLS})$.

4 Setting, Dataset and Machine-Based Data Extraction

4.1 Research Setting

Spotify and Last.fm are the two widely-adopted music streaming services. Both of them offer PC and mobile phone access. And both have a *scrobble function*, which connects a user's listening profile to other music-streaming services. This function links several music-streaming services to Last.fm and Spotify, such as Pandora Radio, iTunes, Windows Media Player, and Deezer, and supports the tracking of complete listening trends over time. Spotify has listening limit for free-access use, while Last.fm has essentially unlimited listening.[2] In addition, since this research studies a setting in which social sharing, comments and interaction are unconstrained, Last.fm is a better choice for this research.

Last.fm puts out a Weekly Listening Chart based on its users' activities. It reports on the top-150 music tracks each week. For the 10 years of data, track popularity *Duration* in Last.fm was 44.2% correlated with song popularity duration for the *Billboard Hot 100*, as well as and 34.3% correlated with *Billboard's Streaming Songs*, based on Spotify's data. This helps to verify Last.fm as a representative source of track popularity data, though some data were omitted due to imprecise song names that were hard to match across the services. A ranking dataset from February 2005 to May 2015 was collected from Last.fm. This yielded 532 weeks and 12+ million streaming music tracks. Relatively few made it to the top-150 chart ranking though: only 4,410 tracks or 0.04% of the total.

Two popularity measures – *Time2TopRank* and *Duration* – were obtained for tracks. The dataset had a long-tail distribution: 80%+ of tracks were ranked for less than 18 weeks. Tracks released earlier had a greater top-rank chance, and some achieved this after as much as 10 years. Others only lasted in the top-rank chart for 1 week. For the skewed data, the log number of weeks for a track's popularity duration for its release date was taken: ln(*NormalizedDuration*) = ln (*Duration*(*Weeks*) /*Weeks - SinceTrack Release*). ln(*NormalizedDuration*) is Gaussian, with skewness (0.89) and kurtosis (2.97).

Some tracks had left-censored observations of their top-chart popularity durations. Bob Dylan's "The Times They Are a Changin" was released in 1964, but only reached Last.fm top-chart ranking in

[2] Flacy (2012) quoted Spotify's new terms of service in January 2012, when the free trials started to expire. It could be "accessed as an ad-supported free-to-the-user service having no monthly cap on listening hours or a cap on number of plays of a unique track during the first 6 months following creation of your Spotify account, but thereafter a cap of 10 listening hours per month and a cap of 5 plays per unique track." Last.fm users were less limited: to 1 million songs, and ~3,000 songs a day.

March 2009, for instance. Right-censored observations include tracks that were popular in the study period, such as Coldplay's "The Scientist," released in 2002. For analysis of popularity duration, all censored data, including 421 left-censored and 108 right-censored tracks, were removed. Overall, this left 3,881 tracks by 477 artists for subsequent use. Without censored data, a smoother Gaussian distribution was observed, with more modest skewness (0.42), and kurtosis (2.45). (See Table 1.)

DATASET (ALL OBS.)	MIN	MAX	MEAN (SD)	1ST QUARTILE VALUE	MEDIAN	3RD QUARTILE VALUE
Duration	1	532	17.9 (47.2)	1	3	11
Time2TopRank	1	473	20.1 (53.7)	2	4	10
DATASET (WITHOUT CENSORED OBS.)						
Duration	1	504	13.1 (31.6)	1	3	9
Time2TopRank	1	395	11.4 (27.7)	2	3	9
Notes. Dataset with all obs.: 4,410 tracks, 550 music artists; dataset without censored obs.: 3,881 tracks, 477 music artists; values in weeks. Study period: February 2005 to May 2015, 532 weeks. In addition to the minimum, maximum and mean values of dependent variables' long-tail distributions, also included are the quartile values of the distributions.						

Table 1. Duration, Time2TopRank in weeks for all observations and without censored observations.

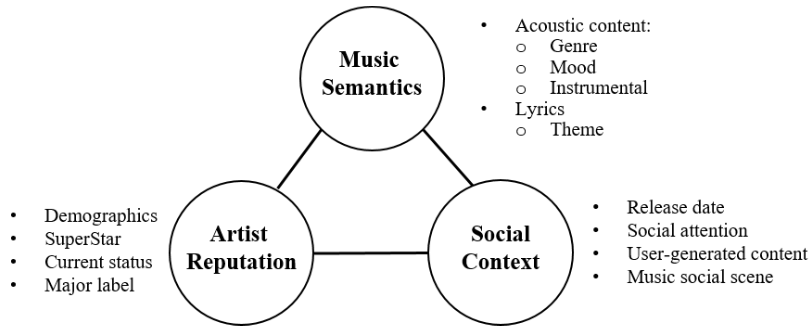

Fig. 1. Drivers of track popularity in a music social network.

Raw *Duration* and *Normalized Duration* in weeks are used to measure the duration of a track's popularity in Last.fm. In the social network environment, three kinds of constructs are relevant: (1) *track semantics*, acoustics and lyrics; (2) *artist reputation* and profile; and (3) *social context* data. They were extracted from multiple sources. Through the measures associated with this *musical construct vector* (MCV), it is possible to assess how they affect popularity duration. (See Figure 1.)

4.2 Musical Construct Vector (MCV)

Music semantics. A music track has two components: *acoustic content* and *lyrics*. The content can be characterized as a *musical construct vector* (MCV), with the Theme, Mood, Instrumental, and Genre, reflecting how acoustic content is perceived. High-level semantics can be extracted from lower-level musical features, such as timbre, rhythm, and tempo (Kim et al., 2010). Machine-based methods were used in this research to extract the music semantics.

*Acoustic content***.** For each track, a 30-second sample was collected from 7Digital or YouTube.[3] A four-step method was used to learn the constructs and implement filtering (Cheng and Shen, 2016):

- **Step 1.** *Segment music tracks* into clips of 1 to 5 seconds in length.

[3] A music track usually has an *Introduction, Verse, Chorus, Bridge,* and *Conclusion*. The *Chorus* is the key element of a track, and its music and lyrics are repeated. It is almost always of greater musical and emotional intensity than other structures in the track. 7Digital supplies 30-second samples for listeners to decide whether they would like to pay for an entire track. By 7Digital's design, most of these samples include the Chorus, while some offer Verse content in their 30-second clips. Our approach with 30-second samples of tracks is similar for downloads from YouTube. We manually chose the Chorus of tracks.

- **Step 2.** *Extract acoustic features* to identify a multi-dimensional low-level acoustic feature vector for all clips (Janani et al., 2012), via: *spectral features* (70 dim.); *timbral features* (23 dim.); *rhythmic features* (12 dim.); and *temporal feature* (62 dim.).
- **Step 3.** *Estimate musical construct probabilities*, based on track tags statistics for 18 genres on Last.fm,[4] 12 types of instrumentation (Zhang et al., 2009), and 5 moods that were selected from the MIREX mood classification (Napiorkowski, 2015) for learning in the musical construct models. (See Table 2.)

CONSTRUCT	SUBCONSTRUCTS (VARIABLES)
Genre (18)	*Rock, Alternative, Indie, Pop, HipPop, Rap, Electronic, Metal, Folk, Soul, Blues, Country, R&B, Punk, Classic, Jazz, Experimental, Reggae*
Instrumental (12)	*Cello, Guitar, Drumkit, Violin, Piano, Tuba, Flute, Clarinet, Saxophone, Trombone, Trumpet, Snare*
Mood (5)	*Passionate, Lively, Brooding, Humorous, Intense*

Table 2. Musical constructs used for the machine-based content analytics.

100 labeled tracks were selected per subconstruct to train a multi-state vector model for each construct. An SVM with a Gaussian *radial basis function kernel* was trained on 80% of randomly-selected labeled clips of tracks, and tested on the remaining 20% with 10 repetitions. 5 segmentation sets with lengths of 1 to 5 seconds were explored. 2-second clips were most effective for 53,296 clips, with prediction accuracies for: Genre – 70.5%; Instrumental – 85.6%; and Mood – 57.5%. The trained models were used to label each clip for tracks, using a 15×35 *acoustic MCV probability matrix*.

- **Step 4.** *Filtering of the learned constructs* resulted in only the useful ones being retained, while the noisy ones were cut. A 35-dimension acoustic MCV was produced for each track.

Lyrics. They complement the acoustic content, and give the artist's meaning behind the music (Hu et al., 2014). *Latent Dirichlet allocation* (LDA, Blei et al., 2003) was used to build a topic model to learn the semantic themes from the dataset of 4,410 tracks. This resulted in 5 topics, with LDA hyperparameters of $\alpha = 2.0$, $\beta = 0.1$, and 3,000 iterations. Table 3 shows the themes that emerged with representative words. About 65% of the tracks were about "love" and "life" (Themes 1, 2, and 4).

Artist reputation. The popularity of a track depends on who performs it to some extent, although other considerations may arise for some tracks and artists. Famous artists attract larger audiences. How to best measure reputation is open to debate though. The present research measures artist reputation, and leverages information on news on the Grammy, American and *Billboard* awards. Also relevant are their labels. Major labels have more resources to produce and promote high-quality tracks. This study covers 10 years, and 20 sub-labels associated with the 3 major labels were considered.

LABEL	MUSIC SEMANTICS THEMES	REPRESENTATIVE WORDS	# TRACKS (# IN SUBSET)
1	*Life, Dance, Passion*	We, like, dance, young, live, good, sweet, dream	589 (514)
2	*In Love, Relationships*	You, love, like, baby, wanna, need, girl, feel	967 (880)
3	*Soul*	Eyes, heart, soul, fall, cold, dark, blue, blood, left	1,041 (918)
4	*Sad Life, Love*	Back, alone, long, over, wrong, lost, leave, remember	1,290 (1,105)
5	*Anger, Hostility*	Like, fuck, shit, rock, bitch, fucking, hit, damn	523 (446)

Notes. The right-most column is the number of track with the labeled themes as their first-ranked theme. The numbers in parenthesis correspond to the number of tracks in the track popularity duration analysis dataset.

Table 3. Music themes and the representative words for each.

[4] Last.fm offers well-defined categories for user tagging tracks, and most are genre related, but the tags may be noisy. There are spelling errors and incorrectly applied labels. And, tagging in Last.fm tends to lack appropriate balance. So popular tracks tend to have more tags, less popular tracks less so, and niche tracks may have none. To make sure each track had proper genre tags, machined-based methods were used to label them by training genre models based on the top 18 genres of Last.fm.

Data were collected from Wikipedia and *Billboard* charts, and 8 dimensions were extracted and built:

- *Vocal*. Solo male, solo female, and group (3 dim.)
- *Major label*. Whether artist belonged to a major record label (1 dim.).
- *Pre-2005 reputation.* Times nominated or won award pre-2005, before a new track (2 dim.).
- *Post-2004 reputation*. Times nominated or won award post-2005, before a new track (2 dim).

Social context. Last.fm had 59.2 million users in July 2015 when user growth plateaued. Its social environment is different from YouTube, Pinterest and Twitter. Users can "tag," "like," and "comment" on tracks and artists. Social comments offer a way to figure out what people are interested in and replace survey methods. Artists attract a group of followers as time passes, even when they are not famous. The social context subconstructs are as follows:

- *EarlyStageComments*. Cumulative comments since track release, time t, supporting diffusion.
- *Top-rank before release*. If in Top 50/100/150 before new track released (sensitivity analysis).
- *Holiday debut*. If released in December holidays in North America and Europe.
- *First top-chart rank.* The ranking when a track first reached a top-chart ranking on Last.fm.

For tracks that reached top-chart ranking, the median time t was 3 weeks, with a skewed distribution. So it was appropriate to use several different periods to build an effective observation window. If the first few weeks of comments were sufficient to predict a track's popularity duration, then the number of weeks was set to the appropriate value of t. Overall, 54 MCV dimensions emerged for explanation.

4.3 Time-Wise Musical Construct Vector (TMCV)

To learn more about the development of a track's popularity over time, a *time-wise musical construct vector* (TMCV) was built and applied. Some dimensions of a track do not vary over time, and are non-musical constructs: the artist's voice; whether the artist had a major label; and so on. Others vary: pre- and post-new track release awards, top-rank in the past month or year, and social comments:

- *Artist awards, past month.* Times nominated or won an award one month before current week.
- *Artist awards, past 3 years*. Times nominated or won in past 3 years before the current week.
- *Track comments, past month*. Comments at $t, \ldots, t-3$ before the current week.
- *Track comments, past year*. Comments in the past year based on the top-chart ranking list.
- *Top-rank, past year*. Track rank for top-chart rank across the Top 50/100/150.
- *Prior rank change.* Rank change at week $t-1$ from $t-2$: positive if rank was ascending (worsening), and negative if rank was descending (improving).
- *Holiday debut*. Whether current week is in month of December (North America and Europe).
- *Similar tracks, past month.* Similar tracks reaching top-chart rank in Weeks $t, \ldots, t-3$.
- *Similar tracks, past year.* Similar tracks reaching top-chart rank in the past year.

The first 5 constructs are similar to those in MCV, but were calculated across different times in the 10-year dataset. The last two constructs describe the Last.fm effect. It offers a recommendation service for similar tracks to users in its network, so a user's listening choices may be affected. The similarity of two tracks is gauged via the *conceptual Euclidean distance* between their 167-dimension low-level acoustic features. For a track in a week, a 54-dimension TMCV was produced.

Overall, 78,697 observations for TMCV were used to explain a track's ranking in a week, so right-censored data (when a track dropped off the chart) were not a problem. But the observations started in February 2005, 3 years after Last.fm was launched. For artists who obtained early social attention, no data were available. So tracks before February 2006, and for artists who were active before February 2005 were removed. This yielded: 67,508 observations on 2,989 tracks, and 450 music artists.

5 Estimation Results: Top-Chart Ranking Popularity Duration

5.1 Estimation Results for Music Track Popularity Duration

Table 4 presents the music track popularity duration modeling estimation results.

CONSTRUCTS AND VARIABLES	LINEAR		WEIBULL HAZARD		LOG-LOGISTIC HAZARD		LOG-NORMAL HAZARD	
Genre								
Pop	0.75***	(0.07)	0.47***	(0.04)	0.45***	(0.04)	0.41***	(0.04)
Indie	0.42***	(0.05)	0.25***	(0.03)	0.24***	(0.03)	0.23***	(0.03)
Alternative	0.21**	(0.07)	0.10**	(0.04)	0.09**	(0.03)	0.09**	(0.03)
Soul	0.53***	(0.12)	0.25***	(0.06)	0.31***	(0.07)	0.27***	(0.06)
Folk	0.39***	(0.10)	0.25***	(0.05)	0.25***	(0.06)	0.24***	(0.05)
Electronic	0.13*	(0.06)	0.16***	(0.04)	0.09**	(0.03)	0.09**	(0.03)
Experimental	-0.55***	(0.11)	-0.45***	(0.06)	-0.22***	(0.05)	-0.26***	(0.06)
Country	-0.40	(0.24)	-0.37**	(0.12)	-0.38**	(0.12)	-0.32*	(0.13)
Punk	-0.26	(0.23)	-0.27*	(0.12)	-0.14	(0.11)	-0.13	(0.12)
Instrumental								
Piano	-0.43*	(0.07)	-0.21*	(0.11)	-0.26*	(0.11)	-0.23*	(0.11)
Guitar	-0.05	(0.07)	-0.05	(0.03)	-0.04	(0.04)	-0.05	(0.03)
Trombone	1.55	(0.92)	0.82	(0.47)	0.83	(0.51)	0.85	(0.48)
Theme								
Life	0.35*	(0.14)	0.13·	(0.07)	0.21**	(0.08)	0.20**	(0.07)
LoveRelations	0.50***	(0.13)	0.27***	(0.07)	0.32***	(0.07)	0.30***	(0.07)
Soul	0.20	(0.14)	0.06	(0.07)	0.17*	(0.07)	0.14	(0.07)
SadLifeLove	0.16	(0.13)	0.06	(0.07)	0.14	(0.07)	0.11	(0.07)
Hostility	0.30	(0.16)	0.14	(0.08)	0.22**	(0.08)	0.18*	(0.08)
Reputation								
MajorLabel	0.02	(0.04)	0.02	(0.02)	0.01	(0.02)	0.02	(0.02)
Post-2004Awards	0.02	(0.02)	0.02	(0.01)	0.02*	(0.01)	0.02	(0.01)
Post-2004Nominations	0.07***	(0.02)	0.04***	(0.01)	0.04***	(0.01)	0.04***	(0.01)
Pre-2005Awards	0.005	(0.010)	-0.004	(0.005)	0.003	(0.005)	0.002	(0.005)
Pre-2005Nominations	-0.04***	(0.011)	-0.03***	(0.006)	-0.02**	(0.006)	-0.02**	(0.006)
SocialContext								
HolidayDebut	0.06	(0.070)	0.003	(0.036)	0.03	(0.04)	0.01	(0.04)
FirstTop-Rank#	-0.004***	(0.000)	-0.002***	(0.000)	-0.003***	(0.000)	-0.003***	(0.000)
EarlyStageComments	0.003***	(0.001)	0.002**	(0.001)	0.001*	(0.001)	0.001**	(0.000)
Top-Rank 51-100	0.11***	(0.034)	0.045*	(0.019)	0.06**	(0.018)	0.06***	(0.017)
Top-Rank 100-150	-0.10***	(0.027)	-0.035*	(0.015)	-0.04*	(0.014)	-0.05**	(0.014)
Model fit: Adj. R^2 or *LL*	Adj. $R^2 = 0.269$		*LL* = -4,592.4		*LL* = -4,327.8		*LL* = -4,268.8	
Shape Parameter	—		2.016		0.291		0.505	
Notes. The *Intercept* for the models was never significant. *LL* = log-likelihood. Mood-related variables (*Passionate, Lively, Brooding, Humorous, Intense*) also were not, and were omitted. Signif: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. A similar analysis was run for *NormalizedDuration* with similar results, so they are not shown.								

Table 4. Estimation results for music track popularity duration

Music semantics. In the past 10 years (2005-2015), the *Pop* music Genre seems to have most easily achieved longer popularity – 59.9% ($p < 0.01$) longer, a positive impact. This result follows since the dependent variable is in log form while the explanatory variable is not. Comparing *Pop* and non-*Pop* music, the difference is $1 - e^{0.47} = 59.9\%$ (Bhattacharjee et al. 2007). The *Indie*, *Soul*, *Alternative* and *Electronic* music Genre had longer popularity too, while *Experimental* music had 36.0% ($p < 0.01$) shorter popularity. *Country* and *Punk* also had less sustainable popularity, and Instrumental music tracks with *Piano* or *Guitar* were less successful in maintaining high listener appeal. For Theme, tracks representing *Life*, *Love,* and *Relationships* were popular longer. For example, music related to *LoveRelations* had 31.0% ($p < 0.01$) more sustainable popularity duration.

Artist reputation. The ArtistReputation construct-related *Vocal* variable was not significant. Many tracks were vocal works, which may suggest regression to the mean for their popularity. *Major Label* was not significant, which suggests a different impact than for album sales, for which *Major Label* was positive and significant. But tracks in the same album are likely to be cointegrated, and exhibited correlation over time in their popularity, even if they are not identical.

People attend to recent tracks of famous artists more than older, less active ones. Nobel Prize winner, Bob Dylan, has over 60 music award nominations. His pre-2005 reputation was high, but not so post-2004. He had 2 tracks in our study that charted, were popular for 1 week, and dropped off. In contrast, Adele had no pre-2005 reputation, but her album "21" won music awards. She shot to stardom, and her tracks are top-ranked for a year now. *Primacy and recency effects* are at work it seems.

Also, *Post-2004Nominations* had a positive impact on the popularity for tracks released later, while

Post-2004Awards did not. There were few awards and many nominations, so the econometric estimation may not have been able to use information beyond what was present in the nominations. Still, if a musician was nominated, it had a reputation effect for the track's popularity.

Social context. *HolidayDebut* racks released at Christmas in North America and Europe had longer popularity on average, but not significant. When a track first rose to top-chart ranking is important, and the higher its first rank, the longer should be its popularity duration. The number of *EarlyStageComments* in each of the first 8 weeks were tested too. Those in the first 4 weeks had some explanatory power for popularity. We adopted $t = 4$ to maximize the likelihood of discovering an effect. Prior top-ranked tracks before a new track appeared demonstrate an artist's social network power. Locally, each artist has followers, and they will tend to adopt the artist's next album. The top-chart ranks from 51 to 100 had a positive impact, while those from 101 to 150 had a negative impact on popularity sustainability.

5.2 Music Track Popularity Duration Prediction

Music popularity is predicted using the MCV subconstructs (variables) that were significant in the empirical results. Different combinations of construct and subconstructs to predict popularity duration are used with multiple classification methods: *support vector regression* (SVR), *bagging*, and *RF*. These were the most used in methods in previous research on popularity classification. Hierarchical prediction tests with *10-fold cross validation* (10-fold CV) are shown in Table 5.

Constructs	Subconstructs (Variables)	SVR Coef.	(SE)	Bagging Coef.	(SE)	RF Coef.	(SE)
Music	See the notes below (Singhi and Brown, 2015)	0.26	(0.03)	0.41	(0.03)	0.38	(0.03)
Non-Music	*ArtistReputation*	0.22	(0.04)	0.42	(0.03)	0.43*	(0.03)
	SocialContext (Schedl, 2011; Kim et al., 2014)	0.37***	(0.03)	0.62***	(0.02)	0.62***	(0.02)
Combined	*Music + ArtistReputation*	0.30***	(0.03)	0.56***	(0.03)	0.58***	(0.02)
	Music + SocialContext (Lee and Lee, 2015)	0.43***	(0.03)	0.69***	(0.02)	0.72***	(0.01)
	ArtistReputation + SocialContext	0.40***	(0.03)	0.68***	(0.02)	0.69***	(0.02)
	Music + ArtistReputation + SocialContext	0.45***	(0.03)	0.70***	(0.02)	0.73***	(0.01)

Note: Music includes ***Genre, Instrumental, Mood***, and ***Theme.*** Related citations shown in table. Correlations between variables with top-rank chart popularity *Duration* are given by: * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Table 5. Prediction results for music content, artist profile and social context.

Prediction performance overall is based on the correlations between the observed and estimated values of *Duration*. Other constructs that are known to have been important in music popularity prediction research are labeled, and prediction were done for the in pairs and with all 3 combined. Based on the results, RF yielded the best performance among the methods. This is consistent with the frequent observation that it is the best algorithm to classify a large real-world dataset (Fernández-Delgado et al., 2014). The evidence is in the right-most column, which has the correlations and significance levels.

The correlation between *Music* and top-chart rank *Duration* suggests a positive but not significant correlation ($\rho = 0.38$, $p > 0.10$). *ArtistReputation* was a more reliable indicator compared to *Music*, which captures previous top-chart ranking performance ($\rho = 0.43$, $p = 0.10$). *SocialContext*'s correlation with duration was highly positive ($\rho = 0.62$, $p < 0.01$). This is consistent with the prior duration analysis because Last.fm's ranking is related to the listening behavior of its social network members. *SocialContext* subconstructs cover the multiple social effects that were operating in Last.fm, and so the prediction should be the close to the true value for a track's popularity duration.

The assessment of combinations of constructs and their subconstructs – in pairs (*Music + ArtistReputation*, $\rho = 0.58$, $p < 0.02$; *ArtistReputation + SocialContext*, $\rho = 0.69$, $p < 0.02$; *Music + SocialContent*, $\rho = 0.72$, $p < 0.01$) – suggests that if listeners do not have knowledge of what a track is about, the artist and social context still will be useful to predict its future top-rank chart list popularity duration. If listeners do not know who an artist is, especially for new artists, the music content and social promotion are likely to be effective generate new interest in the market. When *Music*, *ArtistReputation* and *SocialContext* were all considered, the highest correlation between predicted and actual duration was achieved ($\rho = 0.73$, $p < 0.01$), but only marginal in terms of the new information it offered.

6 Popularity Patterns and Top-Chart Ranking Prediction

Track *Duration* is an *event-based performance measure* (track drops off chart); and *Time2TopRank* is a *speed-based performance measure*. Both supply a possible basis for predicting the popularity patterns of tracks, and how their ranks can be estimated.

6.1 Popularity Patterns

The dataset in this work demonstrates different patterns for track popularity performance with *Duration* and *Time2TopRank*. Some tracks may attract the attention of a large audience immediately –as they are released – and keep satisfying their audience. In contrast, it may take a long time for an artist's track to attract attention – and then may lose audience interest fast. There seem to be different popularity patterns that are at work. Such patterns may create the impetus to forecast the potential value of a track, even if the artist has yet to achieve popularity. If one can predict a new track's future chart performance based on historical data, this will be a key advance toward the loftier goal of predicting emerging superstars. To address this challenge, in this research popularity performance was mined by leveraging the *synthetic minority over-sampling technique* (SMOTE).[5]

This machine-based approach classification results produced 6 popularity patterns for music tracks based on the calibrated 54-dimensional MCV approach. (See Table 6.) The lower and upper bounds on *Duration* and *Time2TopRank* are in weeks, and the names given to the patterns are included. Descriptions of the patterns will enable the reader to interpret how they differ, as follows:

- **Flash in the Pan, Short Popularity (Pattern 1).** The artist's tracks stay in top-chart rank for less than 3 weeks before dropping off. Many tracks have this pattern, but it does not persist. h.
- **Overnight Sensation, Lengthy Popularity (Pattern 2).** Tracks in this pattern become truly popular. They attract a lot of attention since their release, and stay at top-rank for a long time.
- **Slower Rise, Lengthy Popularity (Pattern 3).** These tracks stay "under the radar," but eventually emerge on the top-chart ranking list and attract more listeners.
- **Average Rise, Lengthy Popularity (Pattern 4).** This pattern occurs when a track takes an average amount of time to rise to top-chart ranking, but enjoys lengthy popularity.
- **Faster Rise, Average Popularity (Pattern 5).** Tracks in this common pattern achieve average top-chart ranking popularity duration, but they reach the top-chart listing more quickly.
- **Slower Rise, Average Popularity (Pattern 6).** Similar to Pattern 5, only the artists' tracks take a longer time to achieve a position in the top-chart ranking list.

PATTERN	POPULARITY PATTERNS FOR RISE TO TOP-RANK	*DURATION*	*TIME2 TOPRANK*	*#TRACKS*	*SMOTE*
1	Flash in the Pan, Short Popularity	[1, 3]	All	2,159	1,000
2	Overnight Sensation, Lengthy Popularity	>13	[1,2]	273	819
3	Slower Rise, Lengthy Popularity	>13	>12	232	928
4	Average Rise, Lengthy Popularity	>13	[3,12]	249	996
5	Faster Rise, Average Popularity	[4, 13]	[1, 12]	800	800
6	Slower Rise, Average Popularity	[4, 13]	>12	168	840

Notes. *Duration* and *Time2TopRank* are stated in weeks, and indicate the range of weeks for each pattern as [Lower, Upper] bounds. *#Tracks* measures the number of tracks observed for each popularity pattern. The numbers suggest there is an unbalanced distribution of the data, which is corrected to the number of observations indicated in the *SMOTE* column. Pattern 1, for this dataset, is the *majority pattern*. Its representation is reduced from 2,159 to 1,000 tracks with SMOTE, and stays the majority. Over-sampling was done for the *minority patterns* (2, 3, 4, and 6), while Pattern 5 was sampled without change. This process yielded a balanced dataset with 5,383 tracks, that set up the appropriate conditions for prediction.

Table 6. Music popularity patterns in music social networks.

[5] SMOTE is used for pre-processing data when the classification categories are not equally represented, and exhibit imbalance. Imbalanced data related to the patterns or classes of observations induce an *accuracy paradox* in classification (Chawla et al., 2002). The algorithm *under-samples the majority pattern* as a way to give a balanced classification.

Next, Table 7 shows the results for popularity prediction for 3 machine learning methods.

	Imbalanced Dataset			Balanced Dataset		
	SVM Coef. (SE)	Bagging Coef. (SE)	RF Coef. (SE)	SVM Coef. SE)	Bagging Coef. (SE)	RF Coef. (SE)
Accuracy	0.56 (0.00)	0.67 (0.03)	0.69 (0.03)	0.57 (0.02)	0.73*** (0.01)	0.80*** (0.01)
K	0.00 (0.00)	0.40 (0.04)	0.43 (0.04)	0.48*** (0.02)	0.67*** (0.02)	0.76*** (0.02)
Patterns	**Precision**	**Recall**	**AUC**	**Precision**	**Recall**	**AUC**
1	0.71	0.94	**0.89**	0.69	0.80	0.77
2	0.70	0.47	**0.62**	0.90	0.81	0.92
3	0.57	0.31	0.47	0.88	0.89	0.95
4	0.52	0.18	0.34	0.87	0.84	0.92
5	0.62	0.48	0.59	0.64	0.62	0.66
6	0.47	0.09	0.16	0.90	0.86	0.95

Notes. Significance (*** $p < 0.01$) compares balanced and imbalanced datasets for precision, recall and AUC. There are no significant differences for the imbalanced data-based estimation of patterns, because there was no comparison until the dataset correlations were done. *AUC* indicates the RF-produced *area under the ROC curve* fit metrics for the 6 patterns.

Table 7. Performance of three algorithms for popularity pattern prediction.

The right side of Table 7 shows the results that were obtained. SVM did not produce very accurate predictions for the unbalanced dataset, but there was a significant increase for *K* with the balanced dataset. Bagging and RF yielded improved accuracy 2 times of 3 methods, as suggested by the *K*-values. These improvements are seen from the *areas under the precision-recall curve* (AUC)[6] values over 0.60 for all 6 patterns discovered by RF.

The left side of Table 7 shows the 10-fold CV performance of the models for pattern prediction accuracy for the unbalanced dataset, including SVM, bagging, and RF. The accuracies based on bagging and RF were acceptable. However, their accuracy in classification is due to the strength of the majority pattern, Pattern 1 (Flash in the Pan, Short Popularity). SVM had 56% accuracy for observations assigned to Pattern 1, with 100% recall, while the other 5 patterns had 0% recall. So this was not useful. RF yielded somewhat better performance, but there still were 3 patterns with AUCs of less than 0.50, and the *majority pattern* was retained. (See the left-side lower half of Table 7.)

6.2 Ranking Prediction

Another research question is whether it is possible to predict a specific track's ranking in a music social network – up to real-time prediction. To do this, *ordinal regression* is used to learn the ranking. The dependent variable is *polytomous ordinal* for the marginal effects of changes in TMCV on predicting an improving, stable, or declining weekly rank. (See Table 8.)

A rank correlation of 46.0% between the TMCV constructs and one-period look-ahead weekly track ranks (*RankNextWeek)*. *Awards*, *SocialComments*, and *Other* variables, which had explanatory capability for popularity *Duration*, are useful predictors here. *PriorRankChange* in the past 2 weeks was useful for forecasting *RankNextWeek* for a track. A positive coefficient for *PriorRankChange* indicates decreasing popularity. A negative coefficient suggests increasing popularity. The estimates of the variables for *SimilarTracks* are in line with the outcomes of the recommender system that is at work among the social network members of Last.fm. When there are more similar tracks in the current week (*SimTracks-CurrWeek*: β = -0.012, $p < 0.01$), this seems to have helped the target track to move toward a more favorable top-chart rank. The opposite was true for more similar tracks in the past week though (*SimTracksPastWeek*: β = 0.015, $p < 0.01$): the positive coefficient points to a less favorable rank. Looking back at the data, there is evidence of some oscillating signs in these estimates. So rather than suggest that there is final reading here, there is a need to investigate the effects closely, and not

[6] Davis and Goadrich (2006) showed that AUC is more suitable compared to the area under the *receiver operating characteristics* (ROC) curve for evaluating the classification performance of skewed data.

draw a quick conclusion.

VARIABLES	COEF (S.E.)	VARIABLES	COEF (S.E.)
Awards		**Top-Rank, Past Year**	
AwardsPastMonth	-0.098** (0.047)	*Top 1-50*	-0.005*** (0.000)
NominationsPastMonth	-0.305*** (0.039)	*Top 51-100*	-0.001*** (0.000)
AwardsPast3Years	0.084*** (0.007)	*Top 101-150*	0.010*** (0.000)
NomininationsPast3Years	-0.018*** (0.007)		
SocialComments		**Similar Tracks**	
CommentsCurrWeek	-0.050*** (0.003)	*SimTracksCurrWeek*	-0.012*** (0.003)
CommentsPastWeek	0.010*** (0.002)	*SimTracksPastWeek*	0.015*** (0.004)
CommentsPast2Weeks	0.003** (0.002)	*SimTracksPast2Weeks*	-0.010** (0.004)
CommentsPastMonth	-0.010*** (0.001)	*SimTracksPastMonth*	0.008*** (0.003)
CommentsPastYear	-0.001*** (0.000)	*SimTracksPastYear*	-0.001*** (0.000)
Other		**Regression Metrics**	
PriorRankChange	0.014*** (0.000)	Rank Correlation	46.0%
HolidayDebut	-0.033 (0.022)	Discrimination R^2	19.4%
Notes. Model: Ordinal regression; dep. var. = *RankNextWeek*; obs. = 67,508; Wald Z score used. Signif.: * $Pr\|Z\| < 0.10$, ** $Pr\|Z\| < 0.05$, *** $Pr\|Z\| < 0.01$.			

Table 8. Ordinal regression results for top-chart ranking increases and decreases

The reader should further note that RF with 10-fold CV predicted *RankNextWeek* with an overall correlation of 81% ($\rho = 0.81$, $p < 0.05$) based only on TMCV. Better performance was achieved with the MCV with 91% ($\rho = 0.91$, $p < 0.01$). This shows the fundamental roles of the artist and social network context to compensate for limitations of the musical constructs for predicting track ranks.

7 Discussion and Conclusion

7.1 What Do the Duration Analysis and Prediction Results Mean?

Explanatory analysis with duration modeling yielded key results on the impacts of musical and non-musical constructs. Genre is most important construct as a popularity driver. A major label cannot guarantee the success of tracks in an album though. People remember only 1 or 2 album tracks, yet this is enough to support interest. For artists, being active drives popularity, especially for new tracks. The variables for social context indicate the importance of the drivers of popularity for the streaming platform on the Internet. The higher the numerical ranking when a track first reached the top-rank chart list, the more likely was the track be able to achieve a longer popularity duration. The cumulative effect that an artist created who achieved more top-chart track ranking also had a positive effect.

Out-of-sample prediction based on what was learned from the explanatory econometrics analysis supported prediction of a track's top-chart rank popularity with ~72-73% accuracy, which is better than the 67% accuracy in the previous works. The learned insights can help a record label or an independent producer to assess an artist's future performance, even for a new artist. This capability can be refined to steer music track outcomes in the market. This will provide a clearer understanding of how much promotion and spending are needed to improve the likelihood that a track will have higher, more sustainable popularity. Popularity and future sales revenue are highly correlated. And, even if artists do not want to sacrifice their creativity or be fenced in by past definitions of what constitutes "good" or "marketable" music, their record labels, producers, and agents will still want to keep their "fingers on the pulse of the market," to be sure that they are doing what they can to maximize value.

Since people remember 1 or 2 tracks at most in albums, the record labels should make strategic choices on which track to select as a title track. This can promote the album to maximize future revenues. With just 1 or 2 popular tracks, the entire album may be successful. So more of the tracks will achieve average popularity duration, a label may wish to release digital singles on a one-by-one basis to generate enough attention before they put the album out, or in parallel with release, to build market value.

7.2 How Can Track Popularity Patterns Be Understood?

The analytics revealed information about popularity, but not how long a track takes to become a top hit. Future superstars may stay "under the radar" before their starpower is recognized. The patterns of popularity may help to understand how popular tracks develop: Overnight Sensation, Lengthy Popularity; Slower Rise, Lengthy Popularity; and Normal Rise, Lengthy Popularity. Yet how tracks and artists manage to reach high popularity with different their own ramp-up processes is not understood.

Popularity patterns of European music artists. 754 tracks with Patterns 2, 3 and 4 were chosen. 39.7% (299) are from European, with 17 females, 17 males, and 39 groups. They include: Adele, Avicii, David Guetta, Coldplay, and Daft Punk. They had all 3 patterns before top-rank popularity occurred, and had Slower Rise if they had 2 hits or less. Coldplay has been active since 1996. Yet only half of their tracks (12 of 27) reached the top ranks with normal (7 of 12) or slower speed (5 of 12).

Another finding on European music artists is that 31 of them became active in producing popular music after 2005, so they had no pre-2005 reputation (e.g., awards, top-rank tracks). Some post-2004 top-chart rank tracks had a Normal Rise (Pattern 4) to long-lived popularity. Just 4 of 31 tracks were Overnight Sensations (Pattern 2) – all British. Superstars do not rise fast apparently – even European artists; they require time to develop. The popularity patterns for music artists add a new dimension to understanding how the successful ones grew their popularity. Hidden information in a social network like Last.fm can play a role in identifying popularity patterns, as a basis for how far an artist can go.

7.3 Future Research

In this research, a relatively complete approach to music track popularity analysis was constructed for use with very large datasets. The approach combined machine-based methods with explanatory econometric analysis to understand the sustainability to top-chart rank popularity for music track. There are limitations in this research though. First, deeper analysis is needed of different kinds of music tracks, such as specific genres and artists. Second, modeling rank changes can be done like stock price change, where the *momentum* that price changes have behind them is relevant.

There are other future research directions: (1) prediction methods for the emergence popular artists, especially for specific types of music; (2) more revealing diffusion analysis approaches for streaming music services to understand artists' paths to popularity; and (3) promotion-focused recommendation algorithms that leverage knowledge about the *present value of growth opportunities* (PVGO) in revenue terms for an artist, along with the artist's music catalog.

References

Bhattacharjee, S., Gopal, R.D., Lertwachara, K., Marsden, J.R., and Telang, R. (2007). "The effect of digital sharing technologies on music markets: A survival analysis of albums on ranking charts." *Management Science* 53(9), 1359-1374.

Bischoff, K., Firan, C. S., Georgescu, M., Nejdl, W., and Paiu, R. (2009). "Social knowledge-driven music hit prediction." In *International Conference on Advanced Data Mining and Applications* (ADMA), Berlin, Germany: Springer, 43-54.

Blei, D.M., Ng, A.Y., and Jordan, M.I. (2003). "Latent dirichlet allocation." *Journal of Machine Learning Research* 3(1), 993-1022.

Borg, N., and Hokkanen, G. (2011). "What makes for a hit pop song? What makes for a pop song?" Stanford University, Stanford, CA.

Bronson, F. (2012). "Hot 100 55th anniversary: The all-time top 100 songs." *Billboard*, August 2.

Chang, R., Kauffman, R.J., and Kwon, Y. (2014). Understanding the paradigm shift to computational social science in the presence of big data. *Decision Support Systems* 63, 67-80.

Chawla, N.V., Bowyer, K.W., Hall, L.O., and Kegelmeyer, W.P. (2002). "SMOTE: Synthetic minority over-sampling technique." *Journal of Artificial Intelligence Research* 16, 321-357.

Chen, H., Chiang, R., and Storey, V. C. (2012). "Business intelligence and analytics: From big data to impact." *Management Information Systems Quarterly* 36(4), 1165-1188.

Cheng, Z., and Shen, J. (2016). "On effective location-aware music recommendation." *ACM Transaction on Information Systems* 34(2), 13.

Chon, S.H., Slaney, M., and Berger, J. (2006). "Predicting success from music sales data: A statistical and adaptive approach." In *Proceedings of the 1st ACM Workshop on Audio and Music Computing Multimedia*, New York: ACM Press, 83-88.

Chung, K.H., and Cox, R.A. (1994). "A stochastic model of superstardom: An application of the Yule distribution." *The Review of Economics and Statistics* 76(4), 771-775.

Davis, J., and Goadrich, M. (2006). "The relationship between precision-recall and ROC curves." In *Proceedings of the 23rd International Conference on Machine Learning* (ICML), New York: ACM Press, 233-240.

Dhanaraj, R., and Logan, B. (2005). "Automatic prediction of hit songs." In *Proceedings of the 6th International Society for Music Information Retrieval* (ISMIR), London, UK, 488-491.

Fernández-Delgado, M., Cernadas, E., Barro, S., and Amorim, D. (2014). "Do we need hundreds of classifiers to solve real world classification problems." *Journal of Machine Learning Research* 15(1), 3133-3181.

Flacy, M. (2012). "Unlimited listening on Spotify will vanish for U.S. early adopters next week." DigitalTrends.com, Portland, OR, January 6.

Forbes. (2016). "Adele: Musician."

Frieler, K., Jakubowski, K., and Müllensiefen, D. (2015). "Is it the song and not the singer? Hit song prediction using structural features of melodies." *Yearbook German Society for Music Psychology,* Göttingen, COUNTRY: Hogrefe Verlag, 24, 41-54.

Greenburg, J.O (2016). Beyonce's net worth: $265 million in 2016. *Forbes*, June 1.

Hamlen, W.A. (1991). "Superstardom in popular music." *The Review of Economics and Statistics* (4), 729-733.

Herremans, D., Martens, D., and Sörensen, K. (2014). "Dance hit song prediction." *Journal of New Music Research* 43(3), 291-302.

Hu, P., Liu, W., Jiang, W., and Yang, Z. (2014). "Latent topic model for audio retrieval." *Pattern Recognition* 47(3), 1138-1143.

IFPI (International Federation of the Phonographic Industry) (2015). Digital music report. Recording industry in numbers, London.

Janani, S., Iyswarya, K., Priya, K., and Visuwasam, L. (2012) "Combining spectral features to identify the musical instruments and recognize the emotion from music." *International Journal on Computer Science and Engineering* 4, 1253-1259.

Karydis I., Gkiokas A., and Katsouros V. (2016) "Musical track popularity mining dataset." In L. Iliadis Maglogiannis I. (eds.), In *Proceedings of the Artificial Intelligence Applications and Innovations Conference*, IFIP Advances in Information and Communication Technology, 475, Cham, Switzerland: Springer, 562-572.

Kim, Y., Schmidt, E., Migneco, R., Morton, P., Richardson, J., Speck, J., and Turnbull, D. (2010). "Music emotion recognition: A state of the art review." In *Proceedings of the 11th International Society for Music Information Retrieval* (ISMIR), Utrecht, Netherlands, 255-266.

Kim, Y., Suh, B., and Lee, K. (2014). "# now playing the future Billboard: Mining music listening behaviors of Twitter users for hit song prediction." In *Proceedings of the 1st International Workshop on Social Media Retrieval and Analysis*, New York: ACM Press, 51-56.

Kleinbaum, D.G., and Klein, M. (2006). *Survival Analysis: A Self-Learning Text.* Berlin: Springer.

Koenigstein, N., and Shavitt, Y. (2009). "Song ranking based on piracy in peer-to-peer networks." In *Proceedings of the 10th International Society for Music Information Retrieval* (ISMIR), Kobe, Japan, 633-638.

Lee, J., and Lee, J. S. (2015). "Predicting music popularity patterns based on musical complexity and early stage popularity." In *Proceedings of the 3rd Edition Workshop on Speech, Language & Audio in Multimedia*, New York: ACM Press, 3-6.

Napiorkowski, S. (2015). "Music mood recognition: State of the art." RWTH Aachen University.

Ni, Y., Santos-Rodriguez, R., McVicar, M., and De Bie, T. (2011). "Hit song science once again a science." In *Proceedings of the 4th International Workshop on Machine Learning and Music*, Sierra Nevada, Spain.

Nunes, J. C., and Ordanini, A. (2014). "I like the way it sounds: The influence of instrumentation on a pop song's place in the charts." *Musicae Scientiae* 18(4), 392-409.

Nunes, J. C., Ordanini, A., and Valsesia, F. (2015). "The power of repetition: Repetitive lyrics in a song increase processing fluency and drive market success." *Journal of Consumer Psychology* 25(2), 187-199.

Pachet, F., and Roy, P. (2008). "Hit song science is not yet a science." In *Proceedings of the 9th International Society for Music Information Retrieval* (ISMIR), Philadelphia, 355-360.

Schedl, M. (2011). "Analyzing the potential of microblogs for spatio-temporal popularity estimation of music artists." In *Proceedings of the 22nd International Joint Conference on Artificial Intelligence* (IJCAI), Barcelona, Spain, 539-553.

Singhi, A., and Brown, D.G. (2015). "Can song lyrics predict hits?" In *Proceedings of the 11th International Symposium on Computer Music Multidisciplinary Research* (CMMR), Plymouth, UK.

Strobl, E.A., and Tucker, C. (2000). "The dynamics of chart success in the U.K. pre-recorded popular music industry." *Journal of Cultural Economics* 24(2), 113-134.

Zhang, B., Shen, J., Xiang, Q., and Wang, Y. (2009). "CompositeMap: A novel framework for music similarity measure." In *Proceedings of the 32nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, New York: ACM Press, 403-410.