Score Formulation and Parametric Synthesis of Musical Track as a Platform for Big Data in Hit Prediction

Chapter *in* Advances in Intelligent Systems and Computing · January 2018

DOI: 10.1007/978-981-10-8237-5_35

Prathmesh Matodkar
None

Suraj Iyer
University of Toronto

Nirav Gori
University of Mumbai

Score Formulation and Parametric Synthesis of Musical Track as a Platform for Big Data in Hit Prediction

Sunil Karamchandani, Prathmesh Matodkar, Suraj Iyer and Nirav Gori

Abstract In today's entertainment industry which is becoming increasingly competitive, music producers, record labels are striving hard to give the next big hit song and capture the viable music market. We propose to formulate factors and dependency variables which would form the basis of hit prediction in big data environment. The audio features such as pitch and tempo are analyzed in tandem with statistical parameters such as root mean square energy, slope, period frequency, and musical topographies like acousticness, loudness, and instrumentalness. This is a preliminary experiment where the simulated ratings are paralleled with ground truth obtained from Billboard, Spotify, and Radio Mirchi rankings over a period of 5–10 weeks. The paper covers a wide area of tracks from USA, UK, Australia, and India, and proposes to arrive at a consensus to the factors contributing to the success of the track according to their topography. While acousticness plays a vital role in US and India countdowns, British are highly influenced by the danceability and the energy components of the track. The paper provides a cushion for hit prediction classification of musical tracks in big data applications.

S. Karamchandani (✉)
DJSCE, University of Mumbai, Mumbai, India
e-mail: skaramchandani@rediffmail.com

P. Matodkar · S. Iyer · N. Gori
Electronics & Telecommunication Engg, DJSCE,
University of Mumbai, Mumbai, India
e-mail: prathmat@gmail.com

S. Iyer
e-mail: suraj31295@gmail.com

N. Gori
e-mail: gorinirav5@gmail.com

S. Bhattacharyya et al. (eds.), *Advanced Computational and Communication Paradigms*, Advances in Intelligent Systems and Computing 706,
https://doi.org/10.1007/978-981-10-8237-5_35

1 Introduction

Consign to oblivion Anu Malik, Farah Khan for an ear to audit potential talent in song and music, we suggest to foretell the hit potential of a song in proposing a Hit Conceivable Equivalency (HCE) by forking in statistical parameters with musical topographies. Authors Silk et al. [1] have presented a music browsing service to discover emerging styles in UK using four factors: artist, genre, location, and fans. Statistical data from social media like Twitter mentions and number of hits on SoundCloud have been used to determine the same. However, they have restricted their region of research only to UK and have used non-musical attributes for the analysis. "Popular music estimation based on topic model" [2] proposes a technique for music estimation using audio parameters and time information. Relationship between music trends and corresponding time period has been established using machine learning algorithm. Micro-blogging data via tweets was represented as a time series and correlated with Billboard ratings by the authors E. Zangerle in the paper, "Can microblogs predict music charts?" [3] for predicting the music chart ratings. However, the tweets and chart ratings predicted a mild correlation between them. Hence, results do not consider the demographic variation in Twitter and Billboard consumer data. The authors Yang et al. [4] harp on the success of deep learning by conjoining primitive Mel-spectrogram with Convolution Neural Network (CNN). The aim of their experiment suggests more success with deep learning than convolutional NN. Again, they predict only the popularity of songs in Taiwan. Authors Borg and Hokkanen [5] have employed machine learning algorithm which have been fed with only signal features, such as cepstrum and MFCC coefficient and devoid of any musical features. However, their concentration is focused on pop media genre only.

We have accumulated the factors which would be instrumental in predicting the hit potential of a song. Initially, we have assigned weights to these parameters manually with reference to the Billboard rankings and Spearman rank coefficient correlation method. However, our intention is to assign weights to the parameters dynamically by big data and neural network applications in future.

2 Attribute Extraction

The attributes like acousticness, instrumentalness, loudness, and time signature are obtained from echo-nest API [6]. The audio information with the aid of the appropriate parameters can be extracted with an API call. Echo-nest software comes to our rescue by generating the audio analysis when the song is run through it.

2.1 Statistical Attributes

The following numerical values of the attributes have been formulated to exploit the problem of hit prediction.

2.1.1 Slope

For the given audio signal, we have best curve, and then the derivate of this curve is computed, which gives the value of slope.

2.1.2 Period Frequency

The audio signal or the song is divided into number of frames. We calculate the frequency of the maximum periodicity of these frames and this is estimated using auto-correlation function. In the auto-correlation if there are no peaks, its periodicity is zero.

2.1.3 Entropy

We also calculate the Shannon entropy of the audio signal, i.e., the song to be analyzed. Shannon entropy gives us the amount of uncertainty. The higher the entropy, the more is the noise in the song. We use Eq. (1) to calculate the entropy, where b is the base of the logarithmic function used.

$$H(X) = -\sum_{i=1}^{n} p(x_i) * \log_b p(x_i) \tag{1}$$

2.1.4 RMS Energy

An audio spectrum is comprised of a number of peaks. The RMS energy gives us the effective energy of the song, and statistically it is area under the graph. The RMS can be calculated by the following formula. The global energy of the audio can be calculated by taking the root average of the square of the amplitude, also called root mean square (RMS). It is given by Eq. (2).

$$\mathrm{x}_{\mathrm{rms}} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i} = \sqrt{\frac{x_1^2 + x_2^2 + x_3^3 + \cdots + x_n^2}{n}}, \tag{2}$$

where x is the input audio signal. It basically provides us with the energy throughout the song based on its sampled values and also gives the effective power which is independent of the phase of the sample.

Tempo, key, energy, and signature features are extracted from echo-nest API. Let us look for a hit track levels by Avicii. This gives us a song-ID of the track. A profile call is made where we pass on song-ID along with a bucket which characterizes the audio summary parameter [7]. The attribute data of the song is returned in separate variables on extraction from the echo-nest software. Also, the description of each feature is available in echo-nest's documentation as well as below the analysis.

2.2 Musical Topographies

Different musical topographies influence the quality of the sound and are used to enhance the soundtrack. The topographies which dominate the sound quality are considered for the hit prediction analysis and are described in this section.

2.2.1 Danceability

Dance is a vibrant form of expression. It also acts as an indicator of whether the song is likeable or not. Hence, we have considered a factor called the danceability which is given in Eq. (3).

$$\mathrm{Danceability} = \mathrm{Sum}\,(\mathrm{Tempo} * \mathrm{v}\,(\mathrm{t})), \tag{3}$$

where v(t) represents the velocity of the beat at sample time t, i.e., beat strength.

2.2.2 Energy

Energy is a normalized measure representing a perceptual measure of intensity to a value below unity. As we know tracks with high energy have a fast, loud, and noisy feel, death metal on the higher end of the scale with a Bach prelude on the other extreme end is example of the energy variations. Energy value is a consequence of dynamic range, perceived loudness, timbre, onset rate, and entropy calculated as in (4).

$$\mathrm{dx} = 25.6/\mathrm{nx}$$

$$\mathrm{xn} = \mathrm{dx}/2 * \mathrm{n}$$

$$\mathrm{Energy} = \mathrm{Sum}\,(\mathrm{xn})/\mathrm{n} \tag{4}$$

2.2.3 Key

Key signifies the scale or the different pitches that occur along a soundtrack. Each individual scale has been assigned a key, say a key of 1 is allocated to track sung in D-flat.

2.2.4 Liveliness

We have considered this feature which informs us about the involvement of the audience in that particular recording. Tracks when performed live suggest an increase in the liveliness coefficient. A score of 0.8 indicates a lively track. Values below 0.6 most likely suggest studio recordings.

2.2.5 Loudness

Loudness measure is obtained as the amplitudes of the audio waves averaged across the entire time period. It represents a primary psychological correlate of physical strength defined in (5) and (6).

$$\text{Loudness} = (\text{RM} \cdot \text{avg} + \text{Peak} \cdot \text{avg})/2 \tag{5}$$

The typical values of the loudness range from −80 to 0 dB.

2.2.6 Speechiness

The dominance of spoken words in a track is a measure of the speechiness of a track. It is more pronounced in songs which are more like, poetry. The normalized speechiness values above 0.66 are indicative of tracks that may be entirely of spoken words. The median values of speechiness describe tracks which contain equal contribution of both music and speech. Rap music is one such example where speechiness is dominant in layers. Instrumentals can be mapped to lower range of speechiness such as 0.33 and below.

2.2.7 Tempo

Tempo is analogous to the number of beats per minute in a sound track. For a full duration of the song if $V_{inst}(200 \text{ Hz}) > -6$ dB, then the

$$\text{Tempo} = b/6 \tag{6}$$

Range of tempo analysis ranges from 70 to 180 bpm.

2.2.8 Valence

Valence illustrates the musical positivity in a track as shown in (7). As suggested, high valence tracks are represented in qualities of happy and cheerful tones while tracks with sad, depressive, or even angry tones are an indication of low valence.

$$\text{Valence} = ((\text{loudness}(\text{rms}) * \text{Average segment duration}) + -\sigma) * \mu, \tag{7}$$

where the sign of the standard deviation depends on the mode.

3 Observation and Score Formulation

For understanding the listening and likeability patterns of the listeners of UK, USA, and Australia, we have considered the same set of songs for these three nations. To compute the weights, we needed a reliable source of ranking which we found out in the form of Billboard hot 100 charts for the countries USA and UK and Spotify top 20 charts for Australia. Since India is a culturally diverse country, it is only fair to include local Indian songs which would help us correctly gauge the likeability factor and help us recognize the taste of the Indian audience, for which we have used Radio Mirchi's top 200 chart listings as a reference.

3.1 USA

$$\begin{aligned}
&\text{Score (India)} = \\
&(\text{RMS Energy} * 100) + \text{Slope} + \text{Period Freq} + \text{Entropy} + \text{Tempo} \\
&+ \text{BeatIntensity} + \text{Mean} + (\text{Energy} * 100) + \text{Liveness} + (\text{Acousticness} * 1000) \\
&+ \text{Time Signature} + \text{Key} + \text{Instrumentalness} + \text{Speechiness} + (\text{Danceability} * \\
&100) + \text{Valence} + \text{Mode} + \text{Loudness}
\end{aligned}$$

From the above formulation of the score for USA, it can be seen that for a song to attract the audiences, acousticness parameter has the highest weight, which means that the song should have less electronic instruments or tweaking and should contain more acoustic instruments like guitar, piano, etc. which are electronically treated. This is followed by danceability and energy with same weightage. Table 1 provides the hit prediction values with attributes for US song database (Table 2).

Table 1 USA song database and score

	Closer	Heathens	Cold water	Let me love you	Cheap thrills	Brocolli
RMS energy	0.2482	0.23679	0.24521	0.2282	0.23249	0.32611
Slope	0.1283	−0.097	0.0322	−0.2549	0.0502	0.0752
Period freq	0.2781	0.1668	0.3034	1.6684	0.8342	0.1517
Entropy	0.9962	0.9964	0.9707	0.9721	0.9692	0.9716
Tempo	95.01	90.024	92.94	199.864	89.972	145.99
Beat intensity	59	68	59	66	64	72
Mean	154.41	150.89	134.29	125	113.76	141.34
Energy	0.524	0.396	0.798	0.718	0.8	0.525
Liveness	0.111	0.105	0.156	0.122	0.0775	0.057
Acousticness	0.414	0.0841	0.0736	0.0784	0.0561	0.236
Time signature	4	4	4	4	4	4
Key	8	4	6	8	6	8
Instrumentalness	0	0.0000358	0	0.000010	0.0000020	0
Speechiness	0.0388	0.0286	0.0432	0.0576	0.215	0.131
Danceability	0.748	0.732	0.608	0.476	0.592	0.886
Valence	0.635	0.547	0.488	0.143	0.736	0.711
Mode	1	0	0	1	0	1
Loudness	−5.599	−9.438	−5.092	−5.309	−4.931	−7.39
Score	**884.03**	**529.80**	**531.85**	**622.37**	**494.23**	**776.75**

Table 2 USA average rank and analysis rank

Relative ranking:						
Rank (weeks)	Closer	Heathens	Cold water	Let me love you	Cheap thrills	Brocolli
1	1	2	3	4	5	6
2	1	2	3	4	5	6
3	1	2	3	5	4	7
4	1	2	3	6	4	7
5	1	3	2	6	4	7
6	1	4	2	6	3	7
7	1	4	3	6	2	7
8	5	4	2	6	1	7
9	4	5	3	0	1	6
10	0	4	2	0	1	5
Average rank	**1.77**	**3.2**	**2.6**	**4.3**	**3**	**6.5**
Analysis rank	**1**	**6**	**5**	**4**	**7**	**2**

3.2 India

$$\text{Score (India)} = (\text{RMS Energy} * 100) + \text{Slope} + \text{Period Freq} + \text{Entropy} + \text{Tempo} + \text{BeatIntensity} + \text{Mean} + (\text{Energy} * 100) + \text{Liveness} + (\text{Acousticness} * 1000) + \text{Time Signature} + \text{Key} + \text{Instrumentalness} + \text{Speechiness} + (\text{Danceability} * 100) + \text{Valence} + \text{Mode} + \text{Loudness}$$

Indian music industry has come a long way from 1950s to 2017 and it has seen many changes in terms of song production, recording, and distribution. Although there is a rise of usage of electronic instruments, samples, and loops in today's song from the score, we can make out that even today acousticness parameter dominates the most which is followed by energy and danceability. From this, we can conclude that the Americans and Indians tastes of music are analogous (Tables 3, 4, and 5).

Table 3 Indian song database and score

	Ae Dil Hai Mushkil	Toota Jo Kabhi Tara	Channa Mereya	Tere Sang Yaara	Bulleya	Kala Chashma	Jag Ghoomeya
RMS energy	0.2823	0.292	0.287	0.235	0.29088	0.2846	0.2097
Slope	0.0121	−0.1478	0.026	0.028	−0.296	0.026	−0.0126
Period freq	0.1192	0.175	0.256	0.667	0.1854	0.232	0.3337
Entropy	0.969	0.9755	0.9687	0.972	0.9655	0.970	0.9712
Tempo	123.89	139.9	90.06	76.02	86.95	106.04	82.927
Beat intensity	59	68	64	66	64	72	72
Mean	108.817	113.5266	156.1014	129	134.97	105.9	117.12
Energy	0.654	0.672	0.788	0.6	0.769	0.837	0.577
Liveness	0.165	0.219	0.106	0.075	0.112	0.272	0.103
Acousticness	0.713	0.412	0.237	0.649	0.497	0.107	0.49
Time signature	4	4	4	4	4	4	4
Key	3	9	9	8	2	0	11
Instrumentalness	0	0.000016	0.000024	0	0	0.000026	0
Speechiness	0.0679	0.0369	0.0446	0.033	0.046	0.0774	0.0344
Danceability	0.495	0.437	0.476	0.615	0.552	0.806	0.536
Valence	0.344	0.219	0.745	0.559	0.705	0.911	0.618
Mode	0	0	0	1	0	1	1
Loudness	−6.639	−6.339	−4.281	−6.308	−4.504	−4.273	−7.83
Score	**1149.88**	**881.83**	**713.18**	**1074**	**947.335**	**587.02**	**904.54**

Table 4 India average rank and analysis rank

Rank (Weeks)	Ae Dil Hai Mushkil	Toota Jo Kabhi Tara	Channa Mereya	Tere Sang Yaara	Bulleya	Kala Chashma	Jag Ghoomeya
1	1	2		3	5	4	7
2	1	2		3	5	4	8
3	1	2		3	5	4	10
4	1	2	4	3	6	5	14
5	1	2	3	4	5	8	16
Average rank	**1**	**2**	**3.5**	**3.2**	**5.2**	**5**	**11**
Analysis rank	**1**	**5**	**6**	**2**	**3**	**7**	**4**

Table 5 India relative ranking

Relative ranking							
Rank (Weeks)	Ae Dil Hai Mushkil	Toota Jo Kabhi Tara	Channa Mereya	Tere Sang Yaara	Bulleya	Kala Chashma	Jag Ghoomeya
1	1	2		3	5	4	6
2	1	2		3	5	4	6
3	1	2		3	5	4	6
4	1	2	4	3	6	5	7
5	1	2	3	4	5	6	7
Average rank	**1**	**2**	**3.5**	**3.2**	**5.2**	**4.6**	**6.4**

3.3 United Kingdom

Score (UK) =
(RMS Energy * 100) + Slope + Period Freq + Entropy + Tempo + Beat Intensity + Mean + (Energy * 100) + Liveness + Acousticness + Time Signature + Key + Instrumentalness + Speechiness + (Danceability * 100) + Valen ce + Mode + Loudness

As we analyzed the scores for USA and INDIA, we noticed that the people of these two nations are more acoustically inclined, but what we observe from the score for UK is that here acousticness does not play a major role, and the energy and danceability of the song have equal importance and more prominence over other parameters (Table 6, 7, and 8).

Table 6 UK song database and score

	Say you won't let go	Starboy	Closer	Side to side	Let me love you	Cold water	The greatest
RMS energy	0.188	0.3244	0.2482	0.269	0.2282	0.2452	0.26347
Slope	0.0739	0.2108	0.1283	0.029	−0.2549	0.0322	0.06619
Period freq	0.256	0.2086	0.2781	0.208	1.6684	0.3034	0.8342
Entropy	0.970	0.9704	0.9962	0.967	0.9721	0.9707	0.9689
Tempo	99.26		95.01	159.1	199.86	92.943	192.024
Beat intensity	68	70	59	69	66	59	66
Mean	130.9	155.67	154.4	124.9	125	134.29	139.37
Energy	0.564	0.594	0.524	0.738	0.718	0.798	0.723
Liveness	0.086	0.134	0.111	0.292	0.122	0.156	0.0507
Acousticness	0.693	0.159	0.414	0.040	0.0784	0.0736	0.0109
Time signature	4	4	4	4	4	4	4
Key	10	7	8	6	8	6	1
Instrumentalness	0	0.00002	0	0	0.00001	0	0.0004
Speechiness	0.052	0.28	0.038	0.247	0.0576	0.0432	0.268
Danceability	0.4	0.608	0.748	0.648	0.476	0.608	0.666
Valence	0.449		0.635	0.606	0.143	0.488	0.731
Mode	1	1	1	0	1	0	1
Loudness	−7.444	−7.021	−5.599	−5.883	−5.309	−5.092	−6.164
Score	**423.5**	**385.5**	**470.4**	**525.1**	**544**	**458.33**	**565.40**

Table 7 UK average rank and analysis rank-I

Rank (weeks)	Say you won't let go	Starboy	Closer	Side to side	Let me love you	Cold water	The greatest
1	1	3	2	6	5	7	8
2	1	2	3	5	6	7	8
3	1	3	2	4	7	8	5
4	2	4	3	7	9	11	6
5	2	4	3	9	12	13	8
6	2	4	5	12	14	16	9
7	3	4	8	14	15	18	10
8	2	5	9	14	16	19	11
9	4	8	9	13	15	22	12
10	4	3	10	20	19	24	16
Average rank	**2.2**	**4**	**5.4**	**10.4**	**11.8**	**14.5**	**9.3**
Analysis rank	**6**	**7**	**4**	**3**	**2**	**5**	**1**

Table 8 UK average rank and analysis rank-II

Relative ranking:							
Rank (weeks)	Say you won't let go	Starboy	Closer	Side to side	Let me love you	Cold water	The greatest
1	1	3	2	5	4	6	7
2	1	2	3	4	5	6	7
3	1	3	2	4	6	7	5
4	1	3	2	5	6	7	4
5	1	3	2	5	6	7	4
6	1	2	3	5	6	7	4
7	1	2	3	5	6	7	4
8	1	2	3	5	6	7	4
9	1	2	3	5	6	7	4
10	2	1	3	6	5	7	4
Average rank	**1.1**	**2.3**	**2.6**	**4.9**	**5.6**	**6.8**	**4.7**
Analysis rank	**6**	**7**	**4**	**3**	**2**	**5**	**1**

4 Verdict and Conclusion

We have considered a number of parameters that goes into making of song and based on the rakings of the respective songs we analyzed and formulated a score that defines the position of the song on the chart list, which indirectly gives us an idea about the taste of the audience of these three nations. Hence, we can conclude that the audience from India and USA like songs which are more acoustic in nature as well which are highly energetic and have danceability quotient in it, whereas on the other hand the UK audience wants only high energetic and highly danceability quotient songs where acousticness does not play a major role. Also, we plan to create and feed these results which include the parameters and the weights obtained to a big data neural network as a future scope.

References

1. Silk, H., Santos-Rodriguez, R., Mesnage, C., DeBie, T.: Data science for the detection of emerging music styles. In: Extended abstracts for the Late-Breaking Demo Session of the 17th International Society for Music Information Retrieval Conference (2016)
2. Kinoshita, S., Ogawa, T., Haseyama, M.: Popular music estimation based on topic model using time information and audio features. In: IEEE 3rd Global Conference on Consumer Electronics (GCCE), Tokyo, pp. 102–103 (2014)
3. Hupfauf, B., Pichl, M., Specht, G., Zangerle, E.: Can microblogs predict music charts? An analysis of the relationship between #nowplaying tweets and music charts. In: ISMIR (2016)

4. Yang, L.-C., Chou, S.-Y., Liu, J.-Y., Yang, Y.-H., Chen, Y.-A.: Revisiting the problem of audio-based hit song prediction using convolutional neural networks. In: 2017 IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP), pp. 621–625 (2017). ISSN 2379-190X
5. Borg, N., Hokkanen, G.: What makes for a hit pop song? What makes for a pop song? http://cs229.stanford.edu/proj2011/BorgHokkanen-WhatMakesForAHitPopSong.pdf (2011)
6. Echonest's developer.echonest.com and phyechonest python library
7. Herremans, D., Martens, D., Sorensen, K.: Dance hit song prediction. J. New Music Res. **43**(3), 291–302 (2014)