Hit Song Science Once Again a Science?

Yizhao Ni[1], Raúl Santos-Rodríguez[2], Matt Mcvicar[1] and Tijl De Bie[1]
1. Intelligent Systems Lab, University of Bristol, U. K.
2. Signal Theory and Communications Department, Universidad Carlos III de Madrid, Spain
yizhao.ni@gmail.com, rsrodriguez@tsc.uc3m.es
matt.mcvicar@bris.ac.uk, tijl.debie@gmail.com

1 Introduction

We are interested in the Music Information Retrieval task that aims at predicting whether a given song will be a commercial success prior to its distribution, based on its audio. This is the topic of what is commonly referred to as 'Hit Song Science' (HSS). The underlying assumption behind HSS is that popular songs are similar with respect to a set of features that make them appealing to a majority of people. These features could then be exploited by learning machines in order to predict whether a song will rise to a high position in the chart.

Several attempts to address this question have been proposed in the last few years [2, 3]. The experiments in [3] show that some subjective labels may indeed be reasonably well-learned by these techniques, but not popularity. In this paper, we argue that predicting popularity may well be feasible after all, given a relevant feature set. We support this claim by providing positive results using state-of-the-art machine learning algorithms.

This abstract overviews part of our investigation of the UK top 40 singles chart from the past 50 years. Here, our aim is to distinguish the most popular (peak position top 5) songs from less popular singles (peak position 30 - 40). This split creates nearly equally large classes, making the problem balanced. Mathematically, the hit potential of a song is denoted by $y \in \{+1, -1\}$, with $+1$ indicating hits and -1 non-hits. The strategy we adopted is to learn a binary classifier $f(\mathbf{x}) = \text{sign}(\mathbf{w}^T\mathbf{x})$, trained on the UK top 40 singles dataset $\{y_i, \mathbf{x}_i\}_{i=1}^N$, where $\mathbf{x} \in \mathbb{R}^d$ is the feature vector of a song.

2 Feature Extraction and Learning Agent

The peak UK chart position (popularity) of 5947 unique songs were collected from the Official Charts Company[1], while the features were mainly extracted (with thanks) from the EchoNest API[2]. We used the following musical features from EchoNest: *Tempo* (categorized into 5 bins); *Time Signature* (3 bins); *Song Duration* (3 bins) and *Loudness*. We also computed more detailed summaries of the songs such as *Coefficient of Variance of Loudness* and *Harmonic Simplicity*[3].

Since the hit potential likely depends on the era, it makes sense to express the classifier as a function of time. We thus organized the dataset chronologically and employed the *Shifting Perceptron* [1] as the learning agent. Table 1 outlines the algorithm, where the time shifting weights $[\mathbf{w}_0, \ldots, \mathbf{w}_k]$ show which features are important through time, and the evolving priorities of music listeners.

3 The Evolution of Popularity

The experiments were repeated 100 times to assess variance, using 5000 uniformly sampled songs each time. The parameter λ was optimized as 0.7. Figure 1 depicts the results, where the evolution of the weights for different features is calculated by $\frac{abs(\mathbf{w}_k)}{\|\mathbf{w}_k\|}$. We see a significantly better-than-

[1] http://www.theofficialcharts.com/
[2] http://the.echonest.com/
[3] Due to space constraints, we postpone the description of further details to a subsequent report.

Input of the learner: The samples $\{y_i, \mathbf{x}_i\}_{i=1}^N$ and memory parameter λ.
Initialization: $\mathbf{w}_0 = \mathbf{0}$, $k = 1$
For $t = 1, \ldots, N$
  $a)$ $\hat{y}_t = SGN(\mathbf{w}_k^T \mathbf{x}_t) \in \{+1, -1\}$
  $b)$ **If** $\hat{y}_t \neq y_t$, **then** $\mathbf{w}_{k+1} = (1-\lambda)\mathbf{w}_k + y_t \mathbf{x}_t, \; k = k + 1$
Output of the learner: Predictions $\{\hat{y}_t\}_{t=1}^N$ and feature weights $[\mathbf{w}_0, \ldots, \mathbf{w}_k]$

Table 1: Pseudo-code of the Shifting Perceptron algorithm.

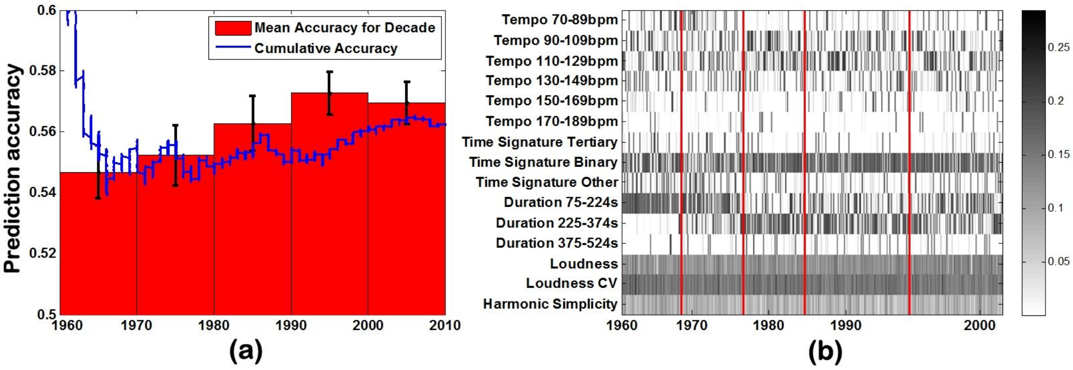

Figure 1: Prediction accuracy with standard deviation for decade (a) and evolution of classifier weights (b). Time progresses on the horizontal axes, decades are grouped for ease of interpretation.

random accuracy ($p = 6.1\text{e}{-125}$), indicating that identifying hits better than randomly from musical features is a feasible. Moreover, we can conclude some interesting trends from the feature evolution:

1. Implied by the importance of *tempo* features, slower songs such as ballads were popular from the 1980s to the 1990s, whilst in the new century music listeners prefer faster songs.
2. In every decade simple songs are hits. This trend is captured by the *Time Signature Binary* and *Harmonic Simplicity* features.
3. There is a clear tendency for music to become relatively longer, indicated by the evolving importance of weights of *duration*.
4. Hits are getting relatively louder. The increasing importance through time shows that the *loudness* of a song is becoming more useful at distinguishing a hit from a non-hit.

4 Discussion

Our results seem to differ from those presented in [3], where it is shown that predicting hits cannot significantly outperform a random oracle. Our more optimistic results can be explained in various ways. The first is that the task description is different - we are trying to distinguish top 5 hits from top 30-40 hits whereas the authors of [3] attempted to learn a label of low, medium or high popularity. Probably more importantly, we used novel audio features not presented in the previously mentioned study. However, a possibly important qualitative difference between the studies is our use of the time-shifting perceptron to account for evolving musical taste. In the future, we will investigate other study designs, including the task mentioned in [3].

References

[1] N. Cesa-Bianchi and C. Gentile. Tracking the best hyperplane with a simple budget perceptron. In *COLT*, volume 4005 of *Lecture Notes in Computer Science*, pages 483–498. Springer, 2006.

[2] F. Pachet. Hit song science. In Tzanetakis & Ogihara Tao, editor, *Music Data Mining*, chapter 10, pages 305–326. Chapman & Hall / CRC Press, 2011.

[3] F. Pachet and P. Roy. Hit song science is not yet a science. In *Proceedings of ISMIR 2008*, pages 355–360, Philadelphia, USA, 2008.