Predicting Musical Sophistication from Music Listening Behaviors: A Preliminary Study

Bruce Ferwerda
Department of Computer Science and Informatics
Jönköping University
Jönköping, Sweden
bruce.ferwerda@ju.se

Mark Graus
Department of Marketing and Supply Chain Management
Maastricht University
Maastricht, the Netherlands
mp.graus@maastrichtuniversity.nl

ABSTRACT

Psychological models are increasingly being used to explain online behavioral traces. Aside from the commonly used personality traits as a general user model, more domain dependent models are gaining attention. The use of domain dependent psychological models allows for more fine-grained identification of behaviors and provide a deeper understanding behind the occurrence of those behaviors. Understanding behaviors based on psychological models can provide an advantage over data-driven approaches. For example, relying on psychological models allow for ways to personalize when data is scarce. In this preliminary work we look at the relation between users' musical sophistication and their online music listening behaviors and to what extent we can successfully predict musical sophistication. An analysis of data from a study with 61 participants shows that listening behaviors can successfully be used to infer users' musical sophistication.

KEYWORDS

Musical sophistication; Gold-MSI; Predictive Modeling; Music Listening Behavior

1 INTRODUCTION & RELATED WORK

There has been an increased interest in understanding online behaviors with psychological models and incorporating them to personalize systems (e.g., [3]). Using psychological models to infer user characteristics from behavior has the advantage that personalization can be done without the need of additional, explicit data collection. Hence, it can be used to mitigate problems where data is scarce (e.g., the cold-start problem).

Most of the research on using psychological models in technological contexts rely on personality traits of users. Personality is

a general model to categorize users and can even be used across domains [2]. However, using more domain dependent psychological models allows for more fine-grained identification of behaviors. Recent research are tapping into these domain dependent psychological models to improve personalizations. For example, Graus et al. [5] showed that personalization based on parenting styles improved the overall user experience of an online parenting library. Hauser et al. [7] exploited the cognitive styles of website users to provide a personalized experience. Germanakos et al. [4] looked at learning styles to adapt the learning environment of students.

In this preliminary work we look at the music domain and explore the influence of musical sophistication and the possibilities to infer musical sophistication of users from their music listening behavior. Müllensiefen et al. [10] created a survey that measures musical sophistication which they define as "a psychometric construct that can refer to musical skills, expertise, achievements, and related behaviors across a range of facets that are measured on different subscales." This translates to that people with a higher degree of music sophistication in general engage more frequently in musical skills and behaviors, and have a greater and more varied repertoire of music behavior patterns. Hence, people's musical sophistication may well be reflected in their music listening patterns. In addition, musical sophistication has been suggested to be related to peoples' needs with regards to music systems [1]. The rise of online music services allows for tracking and analyzing music listening behavior on a larger scale. It provides opportunities to gain deeper insights on the relationships between music sophistication and listening behavior, as well as inferring musical sophistication from listening behaviors.

2 METHOD

The data for this study was collected as part of a larger study that investigated how the order of individual songs affects playlist experience. To study the relationship between musical sophistication and music listening behavior, participants' logged into our app through the Spotify API, which allowed us to retrieve their music listening behavior. In addition, participants completed a survey with items from the Goldsmiths Musical Sophistication Index (Gold-MSI; [10]). The Gold-MSI measures musical sophistication on five subscales: active engagement, emotions, singing abilities, perceptual abilities, and musical training. However, in this preliminary work, we only asked participants to respond on two of these subscales as we believe that they are the most prominent ones reflected in online music listening behaviors:

(1) Active engagement (e.g., how much time and money one spends on music; measured by 9-items).

(2) Emotions (e.g., active behaviors related to emotional responses to music; measured by 6-items).

A total of 61 participants were recruited in December 2017 through a participant pool managed by the Human-Technology Interaction group at Eindhoven University of Technology: 28 male, 33 female (mean age: 23.92 years, SD: 4.57 years). Using Spotify's API we retrieved the participants' top tracks, which resulted in a dataset of 21,080 tracks. For each track we retrieved the audio features through Spotify's API: valence (0-1: negative-positive emotions), liveness, instrumentalness, energy (0-1: calm-energetic), danceability, tempo (BPM), time signature, loudness (dB), track popularity and artist popularity. [1] For each feature we calculated the standard deviation, mean, median, min and max values.

3 RESULTS

We used a learner-based feature selection to select the best features (track properties) to create a model to predict participants' emotions and active engagement scores [10] from their music listening behaviors. A ZeroR classifier was used to create a baseline predictive model. Two different classifiers were used and compared against the baseline model: random forest and radial basis function network (RBF network). Each classifier was applied to the selected features (see Table 1).

Our predictive models were trained with the aforementioned classifiers in Weka [6] with a 10-fold cross-validation with 10 iterations. For each classifier used, we report the root-mean-square error (RMSE) in Table 2 to indicate the root mean square difference between predicted and observed values. The RMSE of each music sophistication trait relates to a [1,7] score scale.

Emotions	Active Engagement
Valence std.dev	Track popularity std.dev
Tempo std.dev	Valence mean
Time signature mean	Valence median
Time signature std.dev	Valence max
Time signature min	Tempo std.dev
Liveness mean	Time signature std.dev
Liveness std.dev	Time signature min
Liveness median	Loudness median
Instrumentalness std.dev	Energy max
Energy std.dev	Danceability mean
Energy min	Danceability median
Danceability mean	
Danceability std.dev	
Danceability median	
Danceability min	

Table 1: Selected features for the predictive models.

We first trained a random forest classifier. Random forests have shown to have a reasonable performance when the features consist of high amounts of noise [8]. As the random forest classifier failed to outperform the baseline in the emotions dimension, we used the RBF network classifier. The RBF network is a neural network that has shown to work well on smaller datasets [9].

---

[1]Popularity measures are not explained in detail by Spotify, but range from 0 to 100.

	ZeroR	Random Forest	RBF network
Emotions	0.97	0.99	**0.95**
Active Eng.	0.97	**0.93**	**0.93**

Table 2: RMSE scores ($r \in [1,7]$) of predicting emotions and active musical engagement from listening behavior. Bold-faced numbers indicate an out performance of the baseline.

4 CONCLUSION, LIMITATIONS & OUTLOOK

In this preliminary work we explored the prediction of musical sophistication subscales (i.e., emotions and active engagement) from music listening behavior. Our results show that music listening behavior can be used to infer the musical sophistication of users. We used a random forest classifier and an RBF network classifier to create the predictive models. Although both classifiers were able to outperform the baseline model on active engagement prediction, only the RBF network was able to also outperform the baseline on predicting the emotions subscale.

Although we were able to predict participants' scores on two subscales of Gold-MSI from music listening behavior, performance can likely be improved more. To do this, we plan to extend the analysis in several ways. We aim to expand our dataset by increasing the number of participants in our dataset and the number of measurements per participant. This will allow for a more in-depth investigation of the relationship between behavioral features and the Gold-MSI scores. Furthermore, we plan to explore the prediction of other subscales of the Gold-MSI as well as exploring the predictive value of other music listening behaviors that are available through the Spotify API (e.g., user's playlists and social networks).

5 ACKNOWLEDGEMENTS

REFERENCES

[1] Òscar Celma. 2010. *Music Recommendation and Discovery.* Springer Berlin Heidelberg, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-13287-2
[2] Ignacio Fernández-Tobías and Iván Cantador. 2015. On the use of cross-domain user preferences and personality traits in collaborative filtering. In *International Conference on User Modeling, Adaptation, and Personalization.* Springer, 343–349.
[3] Bruce Ferwerda and Markus Schedl. 2016. Personality-based user modeling for music recommender systems. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases.* Springer, 254–257.
[4] Panagiotis Germanakos, Marios Belk, et al. 2016. *Human-Centred Web Adaptation and Personalization.* Springer.
[5] Mark P Graus, Martijn C Willemsen, and Chris CP Snijders. 2018. Personalizing an Online Parenting Library: Parenting-Style Surveys Outperform Behavioral Reading-Based Models. In *The 23rd International on Intelligent User Interfaces.*
[6] Mark Hall, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, and Ian H. Witten. 2009. The WEKA Data Mining Software: An Update. *SIGKDD Explor. Newsl.* 11, 1 (Nov. 2009), 10–18. https://doi.org/10.1145/1656274.1656278
[7] John R Hauser, Glen L Urban, Guilherme Liberali, and Michael Braun. 2009. Website morphing. *Marketing Science* 28, 2 (2009), 202–223.
[8] Elizabeth M Humston, Joshua D Knowles, Andrew McShea, and Robert E Synovec. 2010. Quantitative assessment of moisture damage for cacao bean quality using two-dimensional gas chromatography combined with time-of-flight mass spectrometry and chemometrics. *Journal of Chromatography A* 1217, 12 (2010).
[9] Lav R Khot, Suranjan Panigrahi, Curt Doetkott, Young Chang, Jacob Glower, Jayendra Amamcharla, Catherine Logue, and Julie Sherwood. 2012. Evaluation of technique to overcome small dataset problems during neural-network based contamination classification of packaged beef using integrated olfactory sensor system. *LWT-Food Science and Technology* (2012).
[10] Daniel Müllensiefen, Bruno Gingras, Jason Musil, and Lauren Stewart. 2014. The musicality of non-musicians: an index for assessing musical sophistication in the general population. *PloS one* 9, 2 (2014), e89642.