QUANTIFICATION OF SUBJECTIVE TASTES IN NEURAL NETWORK-BASED MUSIC PREFERENCE PREDICTION

Charles Brazier[1] **Maxime Convert**[2] **Kyungyun Lee**[3]
Philip Tovstogan[4] **Yasmina Bagar**[5]
[1] Institute of Computational Perception, JKU, Linz, Austria
[2] Graduate School of Information Science and Technology, University of Tokyo, Japan
[3] Music and Audio Computing Lab, KAIST, S.Korea
[4] Music Technology Group, Universitat Pompeu Fabra, Barcelona, Spain

1. INTRODUCTION

Music tastes are both subjective and subtle. Two persons may like the same song, but they will most certainly like it for different reasons. At the same time, a person may like a song in a certain genre while staying indifferent to other songs in the same genre. In order to truly model musical tastes, it seems important for the model to be able to access finer patterns in lower-level audio features.

Articulating music tastes is a challenging task. We may be able to tell what we like about a song, but pinpointing exactly why we have preferences between songs sharing similar high-level characteristics is harder. The degree to which a user can describe what they like about a song also depends on how musically trained they are. Therefore, being able to express musical tastes in a quantifiable manner would allow a better understanding of one's own tastes.

One way to express music preferences would be to use tags describing properties of songs to get an intuitive understanding music tastes in a bottom-up approach. This method has the downside of limiting the description vocabulary to pre-existing keywords in the dataset being used. Ideally, a music taste prediction system should be able to identify the most relevant features for each user independently in a top down manner, including lower-level features to truly reflect subjective preferences. Neural network-based models seem suitable in that regard.

Choi et al. investigated features learnt by Convolutional Neural Networks (CNN) in a music classification task [1]. Our goal in this project is to further discuss about methods to quantify these features, including lower-level ones, so that a user can better understand their own tastes beyond the more easily-understandable high-level musical descriptors like mode, tempo, instruments, etc.

2. CONCEPT

In this section, we introduce an idea to interpret and quantify music features learnt by a neural network model trained to predict user music taste. This model is a hypothetical model that we would like to test in the future. To train this model, we would use a dataset providing preference ratings of users (such a splay count) for each songs they have listened (e.g. the Taste Profile subset from the Million Song Dataset).

2.1 Model

The main idea is to train a unique neural network for each user, as we focus on the variance of individual's taste rather than the inter user variability.

2.1.1 Using hidden layer activations

It is a common practice to describe music with tags, which encompass genre, mood, year of production and basic song styles, such as 'slow' and 'quiet.' Although these tags are very limiting and stereotypical, they can be thought of as the words that are often used to describe music preference. Thus we plan to adopt the music auto-tagging model in a transfer learning setting, shown in Figure 1, by adding a single fully connected layer at the end, where its weights will only be trained uniquely for each user.

We think that the user preference vectors are not only represented by the output of the last hidden layer, but also from the lower layers. For instance, some user tastes may be captured only at the lower layers, which may be relevant to chords and loudness, or at the higher layers, which may correspond to specific genre. Overall, we think that user preference vector is a linear combination of the output from each hidden layers.

After we obtain a single feature vector for a user, we perform dimension reduction. The dimensions that are reduced to hopefully represent the features that are meaningful and important for each user, even though we cannot name the features. We hypothesize that these user-dependent important features correspond to the main aspect of music that the user is sensitive about, thus contributing mostly to the user's preference.

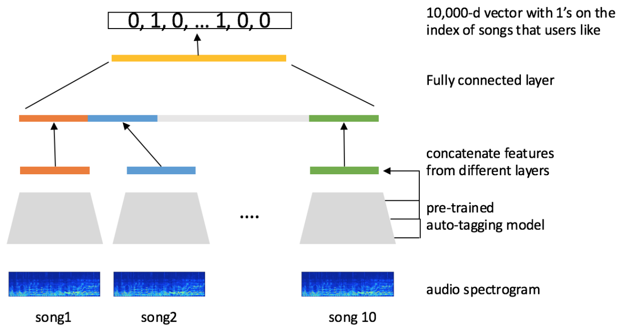

Figure 1. Proposed model for learning user music taste

3. EXPERIMENTAL RESULTS

3.1 Taste identification with audio features

We design a simple demo that uses the pre-computed audio features from the Spotify Web API. Our assumption here is that our proposed neural network will eventually learn features similar to the ones provided by the web API as well as lower-level features. We selected 7 features, such as danceability and instrumentalness, to describe each song (see Table 1). From each user, we ask to select 10 songs they like, and retrieved features for each songs. For each features, we computed the mean and variance across the selected songs. We assume that the features with low variance mean that the user is particularly sensitive about this feature.

Figure 2 shows the result obtained for each person on our team. The trend reflects the songs that each of us have selected and is interesting to see the difference between the users. For instance, all of Philip's songs have low danceability and acousticness, and Charles prefers songs with high energy, valence and danceability.

Feature	Meaning
Danceability	Tempo, rhythm stability, beat strength
Energy	Dynamic range, perceived loudness
Speechiness	Presence of spoken words in a track
Acousticness	Presence of acoustic instruments
Instrumentalness	Presence of vocals in the track
Liveness	Presence of an audience
Valence	Emotions conveyed (sad-happy)

Table 1. Music Features

3.2 Score prediction on what a user could like

In order to create efficiently a first experience on score prediction, we design a network which would be able to rate a new song seeing the previous songs that a user likes or does not like. The main problem of this experiment is to have an efficient dataset of music which is linked with our personal scores of taste. In that way, we created a dataset of twenty songs where each were previously rated from 0 to 1 (0 means that the user does totally not like the song and 1 means that the user is in love with the song).

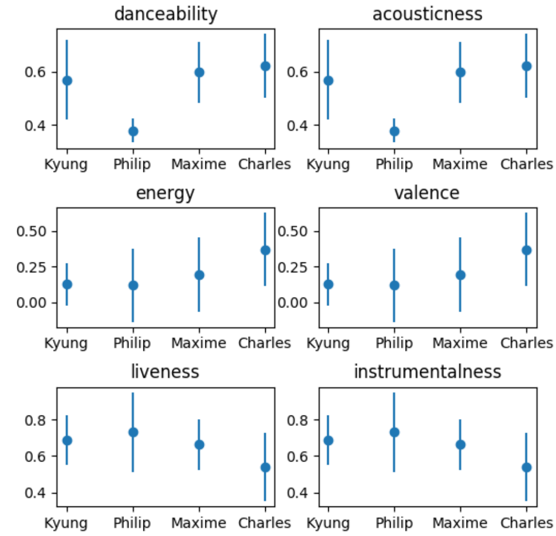

Figure 2. Experimental results

The information which is relevant into songs which could allow to decide if a song will be liked by a user are not evident. In previous music recommendation works [2], it was able to create a similarity world of music. Then two musics from a same album were obviously near. However the taste of a user can change radically even if musics are quite similar. In order to propose an answer at this problem, we build a neural network model based on low level features proposed by the Spotify API where the output is directly the taste score of a user. In that way, a new song will give a prediction score. Obviously the database has to be enlarged to have more relevant results.

The neural network is composed of five fully connected layers where each has respectively 50, 100, 100, 50 and 1 neuron. This type of network allows to create combinations between features and then can learn the idea of putting a score to a song.

4. CONCLUSION

The results outlined in this paper provide a limited music tastes insight of the individuals involved in the experiments. If appropriate time and resources were available to fully conduct the experiments outlined in section 2, the models could potentially identify music taste patterns across genres.

5. REFERENCES

[1] Keunwoo Choi, György Fazekas, and Sandler. Explaining deep convolutional neural network on music classification. In *arXiv:1607.02444v1*, 2016.

[2] Sander Dieleman. Recommending music on spotify with deep learning.

6. APPENDIX: INPUT TRACKS

Kyung:

- Snarky Puppy - Shofukan
- Vulfpeck - Dean Town
- George Ezra - Shotgun
- Weezer - Island in the Sun
- Cigarettes after Sex - K
- Nothing but Thieves - Honey Whiskey
- Twenty One Pilots - Car Radio
- Run River North - Pretender
- Snarky Puppy - Lingus
- Deep Purple - Sometimes I feel like screaming

Philip:

- Nightwish - Ghost Love Score
- Dark Tranquility - Atoma
- Eluveitie - Slania
- Muse - Butterflies and Hurricanes
- Vivienne Mort - Iniy
- Enya - Watermark
- Deep Purple - Wasted Sunsets
- Within Temptation - Reckoning
- The Gathering - Travel
- Dream Theater - In the Name of God

Maxime:

- Clark - Butterfly Prowler
- Nujabes - Sea of Clouds
- Hiatus Kaiyote - Shaolin Monk Motherfunk
- Flying Lotus - Massage Situation
- Birocratic - Boys' Bop
- Tokimonsta - Death by Disco
- BadBadNotGood - Cashmere
- Vendredi - Chemin de Crete
- Smoke City - Devil Mood
- Fakear - Dark Land Song

Charles:

- The Police - Every Breath You Take
- Bowery Electric - Beat
- Pacha - One Kiss
- Raphael Saadiq - Get Involved
- Craig Leon - Nommo
- Cortex - La Bulle
- Lizzy Mercier Desloux - Mais o sont passes les gazelles
- Lord Shorty - Vibrations Groove
- Francis Bebey - Forest Nativity
- The Word - Lobster