Literature Review: Personality-Based Music Recommendation using Machine Learning

Xue Ling Teh

Student ID: 12690175

Research Methods and Professional Practice: Unit 7

Contents

Introduction

Recently, recommender systems have become popular due to their appeal to the general audience, such as bringing about a personalised experience that caters to their interests and dynamic customisation possibilities. Many modern recommender systems make recommendations based on ranking. In the music industry, household names such as Spotify, YouTube Music and Apple Music are several examples of globally used music streaming services and utilise Music Recommender Systems (MRS) based on users' music preferences. Research shows that MRS relies on a range of factors from human-related emotions and personality (Schedl et al., 2018) to musical approaches such as song similarity rankings (Fessahaye et al., 2019). The following literature review aims to explore the aspects of MRS in the music industry and evaluate Machine Learning (ML) techniques used in implementing personality-based MRS.

Music Recommendation Systems

Multiple factors would be considered in the design of an MRS, particularly in determining each of their importance and relevance to the recommendation results. Collaborative Filtering (CF), a term introduced by the authors of the first recommender system, Tapestry (Goldberg et al., 1992), includes user actions and opinions in the recommendation decision. Other than the generic issue of possible inaccurate recommendations, one of the bigger challenges posed by MRS is known as the cold start problem (Schedl et al., 2018). This happens typically during the initial stages as there is insufficient data about a new user, or if a user uses the service without logging in, for example. Similarly, when a new song is added but there is minimal association to make follow-up reliable recommendations.

One of the MRS approaches is modelling based on user personality. The study (Ferwerda & Schedl, 2016) references the five-factor model (FFM) to classify personality as OCEAN, namely openness to experience (O), conscientiousness (C), extraversion (E), agreeableness (A), and neuroticism (N). User listening needs, meta information of music and user interface behaviour were observed to understand and map user personality to their preferences. User personality traits were gathered through the social media sites Facebook, and Twitter – both of which have since been rebranded Meta and X, respectively – and Instagram. The Root Mean Square Error (RMSE) was chosen as the measure for personality trait prediction, which is presumably a simple regression model. The formula for RMSE, which measures the average difference between the predicted and actual values is: $RMSE = sqrt\ [(\Sigma(Pi - Oi)^2) / n]$

Table 1: RMSE Results of FFM on Facebook (Meta), Twitter (X) and Instagram

RMSE \ Social Media	**Facebook (Meta)**	**Instagram**	**Instagram and Twitter (X)**
Openness to experience (O)	.73	.68	.51
Conscientiousness (C)	.73	.66	.67
Extraversion (E)	.99	.90	.71
Agreeableness (A)	.73	.69	.50
Neuroticism (N)	.83	.95	.73

Table 1 shows a comparison of RMSE across the three social media sources. According to the study results, user personality could be inferred despite sparse data undisclosed by users from Facebook (Meta). However, it is apparent that combining multiple sources leads to an improvement in predicting results.

On the other hand, there are MRS designed using ML algorithms, which is elaborated on in the next section.

Machine Learning Algorithms

Designed by Fessahaye et al. (2019), the Tunes Recommendation System (T-RECSYS), uses Content-Based Filtering (CBF) and CF as input vectors to train a neural network using deep learning, which is a subset of ML. The model takes in data input from nine songs in order to predict whether the listener likes the tenth song. CBF takes in six song metadata, which were genre, release year, artist type, artist era, tempo and mood. These categorical metadata were transformed via one-hot encoding into Boolean representation values (e.g. Hip-hop genre had a binary value of 0000001000) for model training. On the other hand, CF measures the pair similarity between the tenth song and each song the user favoured.

In the experiments reported by Moscato et al. (2020), five different MRS classifiers, one per OCEAN trait, were trained on two datasets – MyPersonality (E_1) and PsychoFlickr (E_2) – to evaluate the effectiveness of user personality recognition. There were 50 participants between the ages of 22 and 55, with differing education levels. As there was an overabundance of 47 features, the method applied was Forward Features Selection in tandem with 10-fold Cross Validation.

Table 2: Accuracy Results for Dataset E_1 (Moscato et al., 2020)

		Openess	Conscientiousness	Extroversion	Agreeableness	Neuroticism
Naive–Bayes	**Train**	0.6142	0.6583	0.7255	0.5969	0.6311
	Test	0.6507	0.6594	0.7325	0.6194	0.6456
	#features	6	10	10	25	4
Logistic Regression	**Train**	0.7187	0.7240	0.7378	0.7706	0.7395
	Test	0.6972	0.7163	0.7438	0.7007	0.6853
	#features	25	20	16	27	18
SVM	**Train**	0.6929	0.7257	0.6979	0.6788	0.6679
	Test	0.6840	0.7059	0.6900	0.6548	0.6346
	#features	17	20	21	21	14
Random Forest	**Train**	0.7905	0.7706	0.7725	0.7543	0.7357
	Test	0.7110	0.7007	0.7507	0.7313	0.7124
	#features	16	30	13	25	22
GBT	**Train**	0.8700	0.7782	0.7324	0.7176	0.7567
	Test	0.6884	0.6853	0.6679	0.6435	0.6976
	#features	21	20	20	16	15

Table 3: Accuracy Results for Dataset E_2 (Moscato et al., 2020)

		Openess	Conscientiousness	Extroversion	Agreeableness	Neuroticism
Alexnet	**Train**	0.5805	0.5931	0.5974	0.5813	0.5465
	Test	0.5504	0.5401	0.5491	0.5553	0.5234
Resnet50	**Train**	0.6233	0.5731	0.5698	0.6234	0.6432
	Test	0.6051	0.5668	0.5621	0.5973	0.6123
Squeezenet	**Train**	0.6051	0.5432	0.5876	0.5575	0.5567
	Test	0.5960	0.5411	0.5323	0.5438	0.5365
Densenet	**Train**	0.6009	0.5991	0.5787	0.5643	0.5434
	Test	0.5617	0.5512	0.5384	0.5243	0.5564
Vgg	**Train**	0.6322	0.6102	0.5543	0.5776	0.5576
	Test	0.6022	0.5898	0.5221	0.5349	0.5234

Table 2 displays the accuracy of the training and test sets of MyPersonality and the number of chosen features by the Forward Features Selection algorithm per classifier. In contrast, Table 3 summarises the accuracy of the training and test sets of PsychoFlickr. An identified research gap was the missing combination of multiple variable inputs with the incorporation of real-time updates as data sources (Fang et al., 2017). Broadly speaking, the results of the research studies reflect this notion.

Conclusion

To sum up, the role of user personality and evaluation of ML models in creating an MRS have been discussed, including the collection of user metadata sourced from various social media and ML modelling personality based on OCEAN traits. While there is no straightforward strategy to assess reliability in music recommendations due to their subjective nature, ML models demonstrate the improvement in a personalised musical experience that is highly relevant to the user. Despite challenges persisting, such as the cold start issue, certain ML models have the ability to address the reliability of recommendations with sparse data to an acceptable extent.

References

Fang, J., Grunberg, D., Lui S. & Wang, Y. (2017) Development of a music recommendation system for motivating exercise. *2017 International Conference on Orange Technologies (ICOT)*, Singapore 83-86. DOI: https://doi.org/10.1109/ICOT.2017.8336094

Ferwerda, B. & Schedl, M. (2016) Personality-based user modeling for music recommender systems. *Machine Learning and Knowledge Discovery in Databases* 9853(3): 254-257. DOI: https://doi.org/10.1007/978-3-319-46131-1_29

Fessahaye et al. (2019) January. T-recsys: A novel music recommendation system using deep learning. *2019 IEEE international conference on consumer electronics (ICCE)* 1-6. DOI: https://doi.org/10.1109/ICCE.2019.8662028

Goldberg, D., Nichols, D., Oki, B.M. & Terry, D. (1992) Using collaborative filtering to weave an information tapestry. *Communications of the ACM* 35(12): 61-70.

Moscato, V., Picariello, A. & Sperli, G. (2020) An emotional recommender system for music. *IEEE Intelligent Systems* 36(5): 57-68.

Schedl, M., Zamani, H., Chen, CW., Deldjoo, Y. & Elahi, M. (2018) Current challenges and visions in music recommender systems research. *International Journal of Multimedia Information Retrieval* 7: 95-116. DOI: https://doi.org/10.1007/s13735-018-0154-2