AUTOMATIC MUSIC RECOMMENDATION SYSTEMS: DO

DEMOGRAPHIC, PROFILING, AND CONTEXTUAL FEATURES

IMPROVE THEIR PERFORMANCE?

ABSTRACT

Traditional automatic music recommendation systems performance typically rely on the accuracy of statistical models learned from past preferences of users on music items. However, additional sources of data such as demographic attributes of listeners, their listening behaviour, and their listening contexts encode information about listeners, and their listening habits, that may be used to improve the accuracy of music recommendation models. In this paper we introduce a large dataset of music listening histories with listeners’ demographic information, and a set of features to characterize aspects of people’s listening behaviour. The longevity of the collected listening histories, covering over two years, allows the retrieval of basic forms of listening context. We use this dataset in the evaluation of accuracy of a music artist recommendation model learned from past preferences of listeners on music items and their interaction with several combinations of people’s demographic, profiling, and contextual features. Our results indicate that using listeners’ self-declared age, country, and gender improve the recommendation accuracy by 8 percent. When a new profiling feature termed exploratoryness was added, the accuracy of the model increased by 12 percent.

1. LISTENING BEHAVIOUR AND CONTEXT

The context in which people listen to music has been the object of study of a growing number of publications, particularly coming from the field of music psychology. Kone˘cni has suggested that the act of music listening has vacated the physical spaces devoted exclusively to music performance and enjoyment long ago, and that music nowadays is listened to in a wide variety of contexts [13]. As music increasingly accompanies our everyday activities, the music and the listener are not the only factors, as the context of listening has emerged as another variable that influences, and is influenced, by the other two factors [11]. It has been also observed that people consciously understand these interactions [6] and use them when choosing music for daily life activities [23]. The context of music listening seems to influence the way in which people chooses music, and so music recommenders should suggest music items to fit the situation and needs of each particular listener. Modelling the user needs was identified by Schedl et al. as one key requirement for developing user-centric music retrieval systems [20]. They noted also that personalized systems customize their recommendations by using additional user information, and context-aware systems use dynamic aspects of the user context to improve the quality of the recommendations. The need for contextual and environmental information was highlighted by Cunningham et al. and others [5, 12, 16]. They hypothesized that listeners’ location, activity, and context were probably correlated with their preferences, and thus should be considered when developing music recommendation systems. As a result, frameworks for abstracting the context of music listening by using raw features such as environmental data have been proposed in the literature [16, 22]. While some researchers have reported that context-aware recommendation systems perform better than traditional ones [15, 22, 24], others have shown only minor improvements [10]. Finally, others have carried out experiments with only the most highly-ranked music items, probably leading to models biased by popularity [15, 25]. We will now discuss the impact of using listeners’ demographic and profiling characteristics— hereafter referred to as user-side features [19]—in improving the accuracy of a music recommendation model. User-side features were extracted from self-declared demographics data and a set of custom-built profiling features characterizing the music listening behaviour of a large amount of users of a digital music service. Their music listening histories were disaggregated into different time spans to evaluate if the accuracy of models changed using different temporal contexts of listening. Finally, models based on latent factors were learned for all listening contexts and all combinations of user-side features. Section 2 presents the dataset collection, Section 3 introduces a set of custom-built features to profile listeners’ listening behaviour, and Section 4 describes the experimental set up and presents the results.