

Music Recommendation System Based on Machine Learning

Yunzhe Dong*

University of Nottingham Ningbo China, Ningbo, China

* Corresponding Author Email: ppb0702@163.com

Abstract. Machine learning is a hot topic of research in recent years and has a very wide range of applications in the fields of banking, insurance, transportation, biology, and medicine. A recommendation system is an information filtering system that uses algorithms to pinpoint a user's biased preferences by filtering redundant information from a large dataset and recommending new relevant content to the target user that he or she may like. In this paper, we categorize and outline several mainstream music recommendation methods: content-based and collaborative filtering recommendations, as well as weighted hybrid models of both.

Keywords: Machine learning; recommender systems; algorithms.

1. Introduction

Hundreds of millions of people may now access a wide range of resources thanks to the rapid proliferation of content on the Internet and the introduction of technologies such as audio and video compression. In the age of information explosion, vendors need to solve the problem of how to effectively meet user preferences, and recommendation systems have been created. A recommendation system is an information filtering system. The algorithm of the filtering system can pinpoint the user's bias preferences by obtaining information from a vast number of data sets, and thus (and) develop a matching user profile, and on this basis recommend to the user new relevant content that the user may enjoy. Music platforms such as Netease Cloud Music and QQ Music; shopping platforms such as Taobao and Jingdong; short video platforms such as Jitterbug and Racer; and also some news media platforms, knowledge Q&A platforms, and online communities, all of which satisfy the dual conditions of having a need to shape user profiles for users and having large information resources on the platform itself, are realistic examples of recommendation systems.

The system uses data to project the user's possible preferences based on the user's (historical) choices and provides the user with relevant suggestions to ensure the user's experience, bringing freshness to the user before satisfying the user's needs, and reducing the unnecessary time for the user to find preference data to ensure the user's retention rate is high on this basis. In the current fast-paced society, from movies and music to shopping and food, recommendation system algorithms are all over people's lives and have a real impact on people's lives. Recommendation systems are extremely valuable and meaningful for the current Internet era of exploding knowledge data.

Among them, the music recommendation system is one of the important application scenarios of recommendation systems. Due to the wide distribution of music, consumers have to face the problem of information overload. And the music industry is meeting the challenge of personalized promotion and distribution. The emergence of music recommendation systems can be a better solution to the problem of data explosion, both for users and music companies. As the name suggests, a music recommendation system is a kind of recommendation system based on music data, its purpose is to recommend suitable music or songs to the target users. The goal of this type of system is usually to increase the time users spend using the music sharing platform and therefore has a high commercial value. Therefore, from a business perspective, music recommendation systems are one of the important components of a music-sharing platform. For users, music recommendation systems aim to reduce the unnecessary time spent by users on finding their preferred songs and thus high user experience on the platform.

2. Literature Review of Music Recommendation Systems

The mobile Internet has made digital music widespread, thanks to online music stores and music streaming services. The massive amount of heterogeneous music data exceeds people's basic needs and affordability, and music users have become overwhelmed as a result. In the unprecedented tens of millions of music libraries, it's especially crucial to assist consumers find the music they want fast and accurately, and music recommendation systems have become a major research hotspot in the recommendation field.

In recent years, music recommendation system research has focused on methods such as content-based recommendation, collaborative filtering-based recommendation, context-aware recommendation, and hybrid recommendation. Item similarity is frequently computed between item feature vectors for content-based recommendations. Any information collected from the audio signal, as well as metadata provided by external sources, is considered content information. Acoustic and musical features computed directly from the audio, as well as semantic annotations inferred or forecasted from these acoustic features using machine learning algorithms, make up the audio content analysis. Since audio content-based recommender systems are not affected by popularity bias and can improve the music "long tail" problem [1], MIR researchers recommends using audio content analysis as a substitute for or in addition to metadata and collaborative filtering techniques [2-3]. Manual annotations, social tags, and annotations automatically mined from the Web using text retrieval techniques are all examples of music metadata. There are already many online communities of edited metadata databases created by music experts or enthusiasts, annotated with genres, rhythms, moods, eras, and emotions. The data source for collaborative filtering-based recommendations usually comes from users' preference behavior for music, by mining their underlying preferences and recommending music based on users with the same interests. Even though the method has problems such as sparsity, scalability, and cold start [4], it is still used by some online music websites. Various elements, such as time, surroundings, emotional state, or the user's current activity, have been demonstrated to influence users' short-term music choices in music psychology studies [5]. Context-aware recommendation based on context takes into account this information that may influence users' appreciation of music. Fully observable, partially observable, and unobservable contexts are distinguished by Adomavicius et al [6], with unobservable contexts described using potential attributes that impact changes in users' short-term preferences [7]. Dey et al [8] propose separating primary and secondary contexts, defining primary contexts as location, identity, activity, and time, and secondary contexts as additional information obtained from primary context components, such as weather. These contexts can be collected from an external information service function or measured by sensors on the user's mobile device. Music liking is a complex and nuanced concept, therefore include several characteristics of music similarity in the suggestion step makes sense. Hybrid music recommenders, or strategies that integrate two or more systems to improve performance while avoiding the disadvantages of any one recommendation technique [9]. There have been related studies that consider any combination of two or more approaches to implement hybrid recommendation strategies to improve user satisfaction, such as combining content with context, collaborative filtering with content, and collaborative filtering with context. There are currently no fully automatic context-aware music recommenders available to the general public, while certain music players allow the user to query their mood or activities.

Furthermore, full access to huge datasets, including not just user ratings but also contextual information and audio content, is a major difficulty in music recommendation research. Unfortunately, public datasets with complete audio access are scarce and generally small. There is no single optimum method for music recommendation in terms of functionality or algorithms, hybrid recommendation approaches especially combined with context-aware music recommendation techniques are a major trend, and group music recommendation techniques may also be a new research idea [10]. To summarize, research on recommendation systems in the music domain is still in its early stages, and many issues require further investigation.

3. Music recommendation algorithm

The recommendation algorithm is at the heart of the recommendation system; its technical merits dictate its performance, and it is also the foundation for distinguishing between different types of recommendation systems. There are three traditional recommendation algorithms widely used in music recommendation, including two methods based on content filtering and collaborative filtering recommendation algorithms based on items, and hybrid recommendation algorithms that combine the above two methods.

3.1. Content-based recommendation algorithms

The idea used in the content-based recommendation algorithm is to match the content of the user's preferred songs with the song resources in the repository by counting the user's preferred songs, sorting the songs in order of similarity from highest to lowest, and recommending similar songs to the user based on the sorting results.

Unlike collaborative filtering, users are independent of each other and do not need additional data from other users, which does not lead to less accurate recommendations due to the sparsity of user data. The "item cold start" problem can be avoided if the item has sufficient content description (audio, track analysis, etc.). The content representation is diverse and can be processed in different ways, e.g., text processing techniques, use of semantic information, inference, etc. The recommendation coverage is high because the recommended products are closely connected with the user's past preferences, and the recommended outcomes are more in line with the user's preferences, and the user's unique niche preferences (the long tail) can be found.

3.2. Collaborative filtering recommendation algorithm

When one wants to recommend certain music to a user, the most logical approach is to find people with similar interests, analyze their behavior, and recommend other music that the user likes with similar interests to the user. Or one can look at the user's preferred songs, such as songs that the user has previously favorited, and if the more similar the audience demographic is between every two songs, recommend the songs to people in their respective demographic who have not yet heard one of the songs. This method of recommendation by finding a collection of similar people/items is called Collaborative Filtering Recommendations (CF). Collaborative filtering is a popular recommendation algorithm with two primary approaches: user-based Collaborative Filtering and item-based Collaborative Filtering [11].

3.3. Combined recommendation algorithm

Because both content-based and collaborative filtering recommendations have their benefits and drawbacks, and each different algorithmic model has its uniqueness, if we simply rely on single-model algorithms, we will lose both sides, a certain algorithm is not perfect in every data set to obtain good results, there is no way to solve the problems of the system from multiple perspectives and levels, and cannot meet the requirements of different users on the recommendation target. Therefore, in practical applications, several different models can be used to fuse to the high accuracy of algorithm prediction, generally using content-based collaborative filtering according to certain rules or strategies combined algorithm, called a hybrid recommendation. In most cases, the recommender system with model fusion has better performance in all aspects than the recommender system using separate models. Hybrid systems are generally blended using weighted, waterfall, transformation, feature combination type, etc [12]. Thus, the commonly used combinatorial recommendation techniques are divided into the following approaches: weighted, hierarchical, waterfall, feature, and recommendation model approaches.

4. Deep learning-based recommendation algorithms

Since the high-dimensional and deep-level features of users and items cannot be fully learned in the traditional shallow model, which makes the effectiveness and scalability of the model greatly limited, this has affected the performance of traditional recommendation methods to some extent. Deep learning has produced advancements in a variety of sectors in recent years, and it has become a boom in artificial intelligence, opening up new avenues for study into recommendation systems. The First Workshop on Recommendation Systems Based on Deep Learning was organized in 2016 during the ACM Annual Conference on Recommendation Systems, indicating that recommendation systems based on deep learning will be the next key research topic in the field of recommendation systems.

Three layers make up a deep learning-based recommendation system, where the data in the input layer can be explicit or implicit feedback data from users, such as behavioral data like user ratings, browsing or clicking, user profile and item content data, such as user preferences, age, image and audio content data, and user-generated auxiliary data like annotations and comments. Other deep learning models, such as deep belief networks, convolutional neural networks, recurrent neural networks, and self-encoders, can be used in the model layer. The output layer's job is to generate a list of recommended items for the user using techniques like Softmax classification functions, similarity calculations, and combining implicit user and item representations learned by the model. Deep learning can be integrated into any traditional recommender system and applied to other recommender systems such as social network-based and context-aware recommender systems, as well as content-based recommender systems, collaborative filtering, and hybrid recommender systems, as described below.

4.1. Deep learning-based content recommendation algorithms

The extraction of effective data features is a major challenge in content-based recommender systems, and deep learning is an end-to-end learning system whose main advantage is that it can automatically extract deep features from raw data without the need for human-designed features. As a result, deep neural network models are employed as feature extraction techniques in content-based recommendation systems to extract hidden representations of products and users from their content information, as well as their pictures or past behavioral data, respectively. Then based on the above-hidden representation, similar items to those visited by the user are recommended to the user by calculating the user-item match. The accuracy of the recommender system can be improved by utilizing deep learning to learn the hidden vectors of users and objects.

Deep learning models such as MLP, CNN, RNN, DBN, and others are commonly used in content recommendation algorithms. Elkahky et al. [13] introduced a multi-view deep neural network model based on a deep structured semantic model, which maps two information entities, user and item, into the same hidden space and then uses cosine similarity to calculate the similarity of the two entities. Zheng et al [14] proposed a deep collaborative neural network model, which used two parallel network models to model user preferences based on user review data and item features based on item review data, respectively, based on which an interaction layer was constructed to predict user ratings of items, and the method effectively improved recommendation quality. Gong et al. [15] suggested an attention-based Hashtag recommendation approach in microblogs based on a convolutional neural network model, where CNN was employed to extract microblog features. Zhang et al. [16] introduced a multimodal microblog Hashtag recommendation approach that used CNN to extract picture characteristics and RNN to extract text features.

Deep learning-based content recommendation algorithms combine the extraction of user and item features with the recommendation process, thereby alleviating the difficulties of extracting user and item features as well as the cold start problem of new items. However, it still cannot solve the problem of new users in traditional recommender systems, nor can it discover new items of interest and preference for users.

4.2. Deep learning-based collaborative filtering recommendation algorithm

Deep learning is extensively employed in collaborative filtering recommendation systems because it can generate excellent results in large-scale data processing. The deep learning-based collaborative filtering recommendation algorithm takes explicit or implicit feedback data between users and items, such as a rating matrix, as input, optimizes the parameters of the deep model using a relevant objective optimization function, trains the deep learning model to learn the implicit representation of users or items, and then predicts users' ratings or preferences for items based on the implicit representation, and thus makes recommendations.

The algorithm can usually utilize deep learning models such as AE, NADE, CNN, RNN, GAN, etc. Salakhutdinov et al [17] were the first to employ deep learning in recommendation systems, proposing a collaborative filtering recommendation model based on the RBM model for recommending movies. For solving the TopN recommendation problem, Sedhain et al [18] and Wu et al [19] developed a self-encoder-based collaborative filtering approach and a noise-reducing self-encoder-based collaborative filtering method, respectively. Convolutional neural networks were utilized by Geng et al [20] to learn a uniform representation of people and photos in a vast, sparse, and heterogeneous social network, and to conduct image suggestions based on the similarity of individuals and images computed in the same space. Wang et al [21] proposed an IRGAN model that integrates the generative and discriminative retrieval models in information retrieval modeling, drawing on the idea that generators and discriminators confront each other in generative adversarial networks. To model the user's time-series behavior and achieve prediction, Wu et al [22] suggested a recurrent recommendation network model based on recurrent neural networks.

Deep learning-based collaborative filtering algorithms introduce nonlinear feature transformations in learning the hidden representations of users and items and can achieve nonlinear feature modeling. Compared with traditional collaborative filtering algorithms, the system is more automated and has richer recommendation results, but still cannot solve the data sparsity and cold start problems.

4.3. Hybrid recommendation algorithms based on deep learning

Traditional hybrid recommendation algorithms combine explicit or implicit user feedback data with auxiliary data such as user portraits, item content, and social annotations, effectively alleviating the problem of data sparsity in collaborative filtering methods. However, the biggest challenge with this approach is figuring out how to represent the auxiliary data. Deep learning with automatic feature extraction solves this difficulty precisely, and deep learning can effectively obtain the auxiliary data representation. The deep learning-based hybrid recommendation technique blends content-based recommendation methods with collaborative filtering methods, bringing the feature learning and item recommendation processes together in one system architecture. Various deep learning models are first used to learn the implicit representations of users or items, while an optimization function is constructed for parameter training in combination with traditional recommendation methods, and finally, the trained models are used to achieve item recommendations for users.

The algorithm can usually utilize deep learning models such as AE, CNN, RNN, MLP, etc., among which the AE model is the most widely used. Wang et al [23] introduced a collaborative deep learning hybrid recommendation technique based on the AE model, which learns the hidden representation of text-like auxiliary data of items using a Bayesian stacked noise reduction AE model. Zhang et al [24] used a deep learning approach to learn the hidden representation of items based on a probabilistic matrix decomposition model based on a collaborative knowledge base embedding model that learns the hidden representation of items using three unsupervised representation learning methods. By combining measure learning and collaborative filtering, Hsieh et al [25] established a collaborative measure learning approach that can increase the accuracy of similarity learning. To increase the accuracy of rating prediction, Kim et al [26] created a convolutional neural network based on a convolutional matrix factorization model that combines the textual auxiliary data of products into a recommendation system. Bansal et al [27] proposed an RNN-based text recommendation method that utilizes both rating data and textual content and labels data for a recommendation.

Deep learning-based hybrid recommendation algorithms make extensive use of auxiliary data such as user information, item content, and social annotation to successfully address data sparsity and cold start issues. However, current research is only limited to adopting different deep learning models for different auxiliary data, and further research is needed on how to build a hybrid recommendation framework that can be applied to all data.

References

- [1] Celma, òscar. music Recommendation and Discovery [J].2010,10. 1007/978-3-642-13287-2.
- [2] Barrington L, Oda R, Lanckriet G. SMARTER THAN GENIUS? HUMAN EVALUATION OF MUSIC RECOMMENDER SYSTEMS[C]// International Society for Music Information Retrieval Conference. dblp, 2009.
- [3] Casey M A, Veltkamp R, Goto M, et al. Content-Based Music Information Retrieval: Current Directions and Future Challenges[J]. Proceedings of the IEEE, 2008, 96(4): 668-696.
- [4] Liu JD, Qi LN. User cold start problem in music recommendation system[J]. Journal of Jishou University (Natural Science Edition), 2017, 38(4): 34-37.
- [5] North A C, Hargreaves D. The social and applied psychology of music [M]. Oxford University Press, 2008.
- [6] Verbert K , Duval E , Lindstaedt S N , et al. Context-aware Recommender Systems[C]// Acm Conference on Recommender Systems. acm, 2008.
- [7] Hariri N, Mobasher B, Burke R. Context-Aware Music Recommendation Based on Latent Topic Sequential Patterns[C]// Acm Conference on Recommender Systems. acm, 2012: 131-138.
- [8] Dey A, Abowd G. Towards a better understanding of context and context-awareness [J]. CHI 2000 workshop on the what, who, where, when, and how of context-awareness, 2000, 4: 1-6.
- [9] Burke, R. Hybrid recommender systems: survey and experiments [J]. User Modeling and UserAdapted Interaction, 2002, 12(4): 331-370.
- [10] Tan X. Q., He S.. A review of research on personalized recommendation systems for music [J]. Modern Library and Intelligence Technology, 2014, 30(9): 22-32.
- [11] Meng Xiaolong. Research on recommendation technology based on machine learning [D]. Jilin:Jilin University,2016.
- [12][12] Zhong W. Design and implementation of a hybrid music recommendation system based on music genes [D]. Anhui:Anhui University,2014.
- [13]Elkahky A M, Song Y , He X. A Multi-View Deep Learning Approach for Cross Domain User Modeling in Recommendation Systems[C]// the 24th International Conference on World Wide Web. Florence, Italy, 2015: 278-288.
- [14]Zheng L, Noroozi V, Yu P S. Joint Deep Modeling of Users and Items Using Reviews for Recommendation[C]// Tenth ACM International Conference on Web Search and Data Mining. 2017: 425-434.
- [15]Gong Y, Zhang Q. Hashtag Recommendation Using Attention-Based Convolutional Neural Network[C]// International Joint Conference on Artificial Intelligence, New York, USA, 2016: 2782-2788.
- [16]Zhang Q, Wang J, Huang H, et al. Hashtag Recommendation for Multimodal Microblog Using Co-Attention Network[C]// Twenty-Sixth International Joint Conference on Artificial Intelligence. 2017: 3420-3426.
- [17]Salakhutdinov R, Mnih A, Hinton G. Restricted Boltzmann machines for collaborative filtering [C]. Proceedings of the 24th international conference on Machine learning Corvalis, USA, 2007: 791-798.

- [18] Suvash S, Menon A K, Sanner S, et al. AutoRec: Autoencoders Meet Collaborative Filtering [C]. Proceedings of the 24th International Conference on World Wide Web. Florence, Italy, 2015: 111-112.
- [19] Wu Y, Dubois C, Zheng A X, et al. Collaborative Denoising Auto-Encoders for Top-N Recommender Systems[C]// Proceedings of the 9th ACM International Conference on Web Search and Data Mining. San Francisco, USA, 2016: 153-162.
- [20] Geng X, Zhang H, Bian J, et al. Learning Image and User Features for Recommendation in Social Networks[C]// IEEE International Conference on Computer Vision. IEEE Computer Society, 2015: 4274-4282.
- [21] Wang J, Yu L, Zhang W, et al. IRGAN: A Minimax Game for Unifying Generative and Discriminative Information Retrieval Models[C]. Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval. Tokyo, Japan, 2017: 515-524.
- [22] Wu C Y, Ahmed A, Beutel A, et al. Recurrenet recommender networks [C]. Proceedings of the 10th ACM International Conference on Web Search and Data Mining. Cambridge, UK, 2017: 495-503.
- [23] Wang H, Wang N, Yeung D Y. Collaborative Deep Learning for Recommender Systems[C]. Proceedings of the 21st ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Sydney, Australia, 2014: 1235-1244.
- [24] Zhang F, Yuan N J, Lian D, et al. Collaborative Knowledge Base Embedding for Recommender Systems[C]// Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. San Francisco, USA, 2016: 353-362.
- [25] Hsieh C K, Yang L, Cui Y, et al. Collaborative metric learning [C]. Proceedings of the 26th International Conference on World Wide Web. Perth, Australia, 2017: 193-201.
- [26] Kim D, Park C, Oh J, et al. Convolutional Matrix Factorization for Document Context-Aware Recommendation [C]// Proceedings of the 10th ACM Conference on Recommender Systems. ACM, 2016: 233-240.
- [27] Bansal T, Belanger D, McCallum A. Ask the GRU: Multi-task Learning for Deep Text Recommendations [J]. 2016: 107-114.