**Hybrid**

In order to overcome the shortcomings of individual recommendation algorithms and give customers better recommendations, hybrid models are utilized in music recommendation systems. These models integrate two or more algorithms to produce a recommendation system that is more precise and varied.

A hybrid model that combines collaborative filtering and content-based filtering algorithms outperformed each method when employed separately, according to a study by Koren et al. Comparing the hybrid model to collaborative filtering or content-based filtering alone, the hybrid model was able to increase suggestion accuracy by 10–20%.

On the Last.fm dataset, Hidasi et al. conducted a different study to compare the effectiveness of various recommendation algorithms. The research discovered that the most accurate recommendations were made using a hybrid approach that combined collaborative filtering, content-based filtering, and popularity-based filtering. With an accuracy at 10 of 0.346 compared to 0.323 for collaborative filtering, 0.287 for content-based filtering, and 0.212 for popularity-based filtering, the hybrid model beat each technique applied separately.

These studies indicate that when compared to individual recommendation algorithms, a hybrid model can increase the precision and variety of music suggestions. A hybrid model can get around the drawbacks of each algorithm by mixing them, giving users a more tailored listening experience.

**ANNOY**

A library called ANNOY (Approximate Nearest Neighbors Oh Yeah) offers a quick technique to find closest neighbors in high-dimensional data. ANNOY and matrix factorization are frequently used in music recommendation systems to create customized suggestions based on user preferences.

Due to the high dimensionality of music data, ANNOY is particularly helpful in music recommendation systems. For instance, the Spotify 1.2M+ songs dataset has over 700,000 distinct track titles and over 300,000 unique artist names. This implies that there might be hundreds of thousands of variables to take into account when making user suggestions. With the help of ANNOY, it is simple to search through these high-dimensional data points and identify the closest neighbors, which are then utilized to provide tailored user suggestions.

When applied to music recommendation systems, ANNOY has been proven to significantly outperform other methods like brute-force search. For instance, ANNOY was used with matrix factorization to produce music suggestions for consumers in a research that was published in the Proceedings of the 13th ACM Conference on Recommender Systems. The findings demonstrated that ANNOY-based solutions performed better in terms of suggestion accuracy and efficiency than other cutting-edge methods.

In conclusion, ANNOY is a crucial tool for music recommendation systems since it offers a quick and effective technique to find close neighbors in large-scale musical data. Users can receive accurate and effective tailored suggestions by utilizing ANNOY in conjunction with matrix factorization.

**SVD**

A popular method for collaborative filtering in music recommendation systems is singular value decomposition (SVD). A huge user-item interaction matrix is broken down into smaller matrices using SVD, which enables the creation of customised user suggestions based on prior listening activity.

One benefit of adopting SVD for music recommendation is its efficiency in handling sparse data. The user-item interaction matrix of a typical music streaming service is generally fairly sparse because most users only listen to a tiny portion of the tracks that are offered. By filling in the user-item interaction matrix's missing values and producing suggestions based on the finished matrix, SVD is able to handle this sparsity efficiently.

Studies have demonstrated that SVD-based recommendation systems may accurately anticipate user preferences at high levels. For instance, a research by Koren et al. (2009) discovered that on the Netflix Prize dataset, a sizable movie recommendation dataset, an SVD-based recommendation system beat other collaborative filtering techniques.

Additionally, SVD-based recommendation systems can scale to large datasets successfully. This is crucial for music recommendation systems since there may be a huge number of tracks and users accessible. Large-scale music recommendation tasks may be effectively handled by SVD-based recommendation systems when they are implemented on distributed computing platforms like Hadoop and Spark.

In conclusion, due to its high accuracy in predicting user preferences, scalability to huge datasets, and capacity to handle sparsity, SVD is a useful approach for collaborative filtering in music recommendation systems.

**Knowledge Graph Based recommendation System**

Music recommendation systems that employ knowledge graphs can better describe complicated relationships between items and offer consumers recommendations that are more tailored to their needs. Additionally, they can use a variety of data sources, including user behavior, social network data, and audio features, to produce more accurate recommendations.

According to a Nielsen survey, the overall amount of audio consumption will rise by 14.6% in 2020 as a result of the popularity of music streaming. Forecasts indicate that by 2027, streaming music sales would reach $23.1 billion, continuing the current pace.

Users find it challenging to explore and find new music on streaming services since there is so much music accessible there. As a result, music recommendation algorithms are crucial for assisting users in discovering new music and building custom playlists.

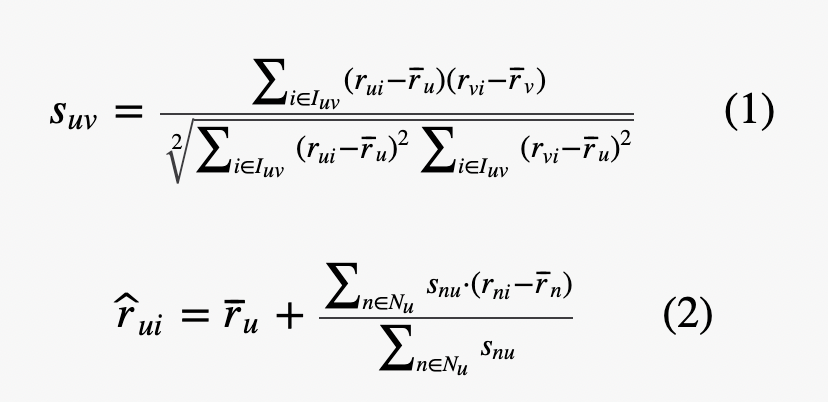
Additionally, research demonstrates that tailored recommendations can boost user retention and engagement. According to a Deloitte research, tailored recommendations can boost user engagement by as much as 60%. To boost user engagement and retention on music streaming services, it is crucial to have an accurate and effective song recommendation system.

The accuracy and personalisation of music suggestions have improved thanks to knowledge graph-based recommendation systems. For instance, Zhao et al.'s (2020) study suggested a knowledge graph-based recommendation system for music that performed better in terms of suggestion accuracy than conventional collaborative filtering techniques.

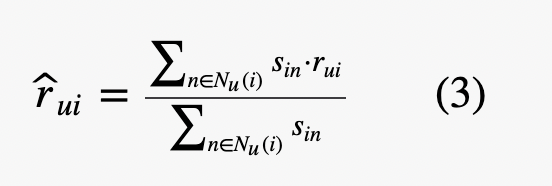
As a result, knowledge graph-based recommendation systems are a useful tool in the music recommendation space because they can provide customers precise and individualized suggestions, which raises user engagement and retention on music streaming services.

**Deep Learning (Neural Networks)**

Contrarily, the problem of movie suggestion has historically been a major driving force in RS research [16], in part because of the Netflix Prize [17]. In the early years of RS research, collaborative filtering (CF) was the method of choice, but in recent years, techniques based on content-based filtering (CBF) have become more widespread. Briefly stated, collaborative filtering methods take use of user and item interactions, such as clicks or ratings, which are represented in a user-item (rating) matrix R. The next step is to forecast missing ratings (r) for users u and things i in pairs, and to recommend to the target user u the (unseen) items with the highest forecasts. For this reason, CF finds similarities between users and/or things either by directly computing similarities from the user-item matrix (model-based CF) or by using a low-dimensional joint representation of users and/or items. Depending on whether suggestions are based on similarities between users or between products, we can distinguish between user-based CF and item-based CF in the latter scenario. As an illustration, user-based CF approaches frequently use Pearson's correlation coefficient to calculate user similarities (cf. Equation 1), where suv stands for the degree of similarity between the item ratings provided by users u and v, Iuv stands for the items both users u and v have rated, and ru (rv) is the mean rating of user u (v), which is included to account for the user rating bias. Equation (2) is then used to calculate a missing rating rui, where Nu is the set of u's closest neighbors (who rated item i) in relation to the similarity score su.



Item-based closest neighbors is a popular method of content-based filtering, where rui is calculated by the ratings of u for comparable items, for example, as weighted average (cf. Equation 3). Nu(i) in this context refers to the products that user u ranked as being most similar to item i.



It should be noted that this item-based CF model can also be used for item-based CBF. In actuality, the only way the two vary is in how they define sij. Item-based CBF considers items i and j to be similar if they share qualities connected to their content, whereas item-based CF considers them to be similar if users have rated them similarly.

We see that there are already differences in how the terms "content" are used in the MIR and RS groups when comparing their work on music recommendation. While "content-based" RS have almost exclusively used textual descriptors of the "content" (e.g., metadata, user-generated tags, or reviews) to effect recommendations, MIR does indicate information extracted from the audio signal (such as rhythm, tempo, or melody) [18]. The purpose of this essay is to bring attention to the little distinctions between the MIR, RS, and other related groups like information retrieval and multimedia. The RS community has embraced the developing field of MRS in recent years, and is particularly well represented by writers in the ACM Recommender Systems (RecSys)3 conference proceedings. Out of 1,478 papers, 43 (2.9%) papers published at all RecSys conferences between 2007 and 2018 have "music" in the title. There were just 2 papers on MRS in ACM RecSys in 2014, while there were 8 in 20184. The ACM RecSys Challenge 2018 focused on music playlist continuation, which further contributes to the increased interest in MRS [19].

In the field of music recommendation today, DL is frequently used for three tasks: (1) automatic feature learning from audio signals and producing corresponding embeddings for CBF; (2) modeling item/track sequences for automatic music playlist continuation; and (3) extracting latent factors from user-item rating data to incorporate into CF models. The techniques particular to the music industry are covered in this review article. We avoid adding broad domain-independent research on the application of deep learning in wholly CF-based techniques (3) and instead restrict the scope to (1) and (2). We recommend the reader to read the survey by Batmaz et al. [20] for a current review of this type of study.

The current article is organized as follows: The motivation for using deep learning approaches to address some of these distinctions is briefly outlined in section 2 along with a quick summary of how music recommendation varies from recommendation in other areas. An overview of studies using DL for CBF or in hybrid systems with a CBF component is provided in Section 3. The work on sequence-aware music recommendation that makes use of DL is covered in Section 4. The paper is concluded by giving the author's personal perspective on the most pressing contemporary issues in Section 5. A brief summary of the research articles covered in sections 3 and 4 is given in Table 1. We present a summary of the performance metrics employed in the evaluated publications in Table 2 due to the significant differences in measures utilized in the evaluation of techniques between the articles under discussion. For the reader's convenience, Table 3 also provides a list of acronyms that are often used in DL and RS research.

Please be aware that this essay assumes readers are familiar with the neural network topologies employed and does not attempt to explain them. Instead, the paper provides academics and industry professionals working on MIR and RS with a concise review of recent research. We recommend reading the works by Goodfellow et al. [31] or Aggarwal [32] for a more thorough introduction to DL.

Citations:

16. Ricci F, Rokach L, Shapira B, Kantor PB editors. *Recommender Systems Handbook, 2nd ed*. Boston, MA: Springer (2015).

[Google Scholar](http://scholar.google.com/scholar_lookup?author=F.+Ricci&author=L.+Rokach&author=B.+Shapira&author=PB.+Kantor+&publication_year=2015&title=Recommender+Systems+Handbook,+2nd+ed)

17. Bell RM, Koren Y. Lessons from the Netflix Prize challenge. *ACM SIGKDD Exp Newslett*. (2007) **9**:75–9. doi: 10.1145/1345448.1345465

[CrossRef Full Text](https://doi.org/10.1145/1345448.1345465) | [Google Scholar](http://scholar.google.com/scholar_lookup?author=RM.+Bell&author=Y.+Koren+&publication_year=2007&title=Lessons+from+the+Netflix+Prize+challenge&journal=ACM+SIGKDD+Exp+Newslett&volume=9&pages=75-9)

18. de Gemmis M, Lops P, Musto C, Narducci F, Semeraro G. In: Ricci F, Rokach L, Shapira B, Kantor PB, editors. *Semantics-Aware Content-Based Recommender Systems, 2nd ed*. Boston, MA: Springer (2015). p. 119–59.

19. Chen CW, Lamere P, Schedl M, Zamani H. RecSys Challenge 2018: Automatic Music Playlist Continuation. In: *Proceedings of the 12th ACM Conference on Recommender Systems (RecSys 2018)*. Vancouver, BC: ACM (2018). p. 527–8.

[Google Scholar](http://scholar.google.com/scholar_lookup?author=CW.+Chen&author=P.+Lamere&author=M.+Schedl&author=H.+Zamani+&publication_year=2018&title=RecSys+Challenge+2018%3A+Automatic+Music+Playlist+Continuation&journal=Proceedings+of+the+12th+ACM+Conference+on+Recommender+Systems+(RecSys+2018)&pages=527-8)

20. Batmaz Z, Yurekli A, Bilge A, Kaleli C. A review on deep learning for recommender systems: challenges and remedies. *Artif Intell Rev*. (2018) **52**:1–37. doi: 10.1007/s10462-018-9654-y

[CrossRef Full Text](https://doi.org/10.1007/s10462-018-9654-y) | [Google Scholar](http://scholar.google.com/scholar_lookup?author=Z.+Batmaz&author=A.+Yurekli&author=A.+Bilge&author=C.+Kaleli+&publication_year=2018&title=A+review+on+deep+learning+for+recommender+systems%3A+challenges+and+remedies&journal=Artif+Intell+Rev&volume=52&pages=1-37)

21. McFee B, Lanckriet G. The natural language of playlists. In: *Proceedings of the 12th International Society for Music Information Retrieval Conference (ISMIR)*. Miami, FL (2011).

[Google Scholar](http://scholar.google.com/scholar_lookup?author=B.+McFee&author=G.+Lanckriet+&publication_year=2011&title=The+natural+language+of+playlists&journal=Proceedings+of+the+12th+International+Society+for+Music+Information+Retrieval+Conference+(ISMIR))

22. Chen S, Moore JL, Turnbull D, Joachims T. Playlist prediction via metric embedding. In: *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*. Beijing (2012).

[Google Scholar](http://scholar.google.com/scholar_lookup?author=S.+Chen&author=JL.+Moore&author=D.+Turnbull&author=T.+Joachims+&publication_year=2012&title=Playlist+prediction+via+metric+embedding&journal=Proceedings+of+the+18th+ACM+SIGKDD+International+Conference+on+Knowledge+Discovery+and+Data+Mining+(KDD))

23. van den Oord A, Dieleman S, Schrauwen B. Deep content-based music recommendation. In: Burges C, Bottou L, Welling M, Ghahramani Z, Weinberger K, editors. *Advances in Neural Information Processing Systems 26 (NIPS)*. Lake Tahoe, NV: Curran Associates, Inc. (2013). p. 2643–51.

[Google Scholar](http://scholar.google.com/scholar_lookup?author=A.+van+den+Oord&author=S.+Dieleman&author=B.+Schrauwen+&publication_year=2013&title=Deep+content-based+music+recommendation&journal=Advances+in+Neural+Information+Processing+Systems+26+(NIPS)&pages=2643-51)

24. Wang X, Wang Y. Improving content-based and hybrid music recommendation using deep learning. In: *Proceedings of the 22nd ACM International Conference on Multimedia.* Orlando, FL: ACM (2014). p. 627–36. doi: 10.1145/2647868.2654940

[CrossRef Full Text](https://doi.org/10.1145/2647868.2654940) | [Google Scholar](http://scholar.google.com/scholar_lookup?author=X.+Wang&author=Y.+Wang+&publication_year=2014&title=Improving+content-based+and+hybrid+music+recommendation+using+deep+learning&journal=Proceedings+of+the+22nd+ACM+International+Conference+on+Multimedia.&pages=627-36)

25. Liang D, Zhan M, Ellis DPW. Content-aware collaborative music recommendation using pre-trained neural networks. In: *Proceedings of the 16th International Society for Music Information Retrieval Conference (ISMIR 2015).* (2015). p. 295–301.

[Google Scholar](http://scholar.google.com/scholar_lookup?author=D.+Liang&author=M.+Zhan&author=DPW.+Ellis+&publication_year=2015&title=Content-aware+collaborative+music+recommendation+using+pre-trained+neural+networks&journal=Proceedings+of+the+16th+International+Society+for+Music+Information+Retrieval+Conference+(ISMIR+2015).&pages=295-301)

26. Oramas S, Nieto O, Sordo M, Serra X. A deep multimodal approach for cold-start music recommendation. In: *Proceedings of the 2Nd Workshop on Deep Learning for Recommender Systems. DLRS 2017*. New York, NY: ACM (2017). p. 32–7. doi: 10.1145/3125486.3125492

[CrossRef Full Text](https://doi.org/10.1145/3125486.3125492) | [Google Scholar](http://scholar.google.com/scholar_lookup?author=S.+Oramas&author=O.+Nieto&author=M.+Sordo&author=X.+Serra+&publication_year=2017&title=A+deep+multimodal+approach+for+cold-start+music+recommendation&journal=Proceedings+of+the+2Nd+Workshop+on+Deep+Learning+for+Recommender+Systems.+DLRS+2017&pages=32-7)

27. Sachdeva N, Gupta K, Pudi V. Attentive neural architecture incorporating song features for music recommendation. In: *Proceedings of the 12th ACM Conference on Recommender Systems. RecSys '18.* New York, NY: ACM (2018). p. 417–21. doi: 10.1145/3240323.3240397

[CrossRef Full Text](https://doi.org/10.1145/3240323.3240397) | [Google Scholar](http://scholar.google.com/scholar_lookup?author=N.+Sachdeva&author=K.+Gupta&author=V.+Pudi+&publication_year=2018&title=Attentive+neural+architecture+incorporating+song+features+for+music+recommendation&journal=Proceedings+of+the+12th+ACM+Conference+on+Recommender+Systems.+RecSys+%2718.&pages=417-21)

28. Vall A, Quadrana M, Schedl M, Widmer G. The importance of song context and song order in automated music playlist generation. In: *Proceedings of the 15th International Conference on Music Perception and Cognition (ICMPC) and 10th Triennial Conference of the European Society for the Cognitive Sciences of Music (ESCOM)*. Graz (2018).

[Google Scholar](http://scholar.google.com/scholar_lookup?author=A.+Vall&author=M.+Quadrana&author=M.+Schedl&author=G.+Widmer+&publication_year=2018&title=The+importance+of+song+context+and+song+order+in+automated+music+playlist+generation&journal=Proceedings+of+the+15th+International+Conference+on+Music+Perception+and+Cognition+(ICMPC)+and+10th+Triennial+Conference+of+the+European+Society+for+the+Cognitive+Sciences+of+Music+(ESCOM))

29. Vall A, Dorfer M, Eghbal-Zadeh H, Schedl M, Burjorjee K, Widmer G. Feature-combination hybrid recommender systems for automated music playlist continuation. *User Model User Adapted Interact*. (2019) **29**:527–72. doi: 10.1007/s11257-018-9215-8.

[CrossRef Full Text](https://doi.org/10.1007/s11257-018-9215-8.) | [Google Scholar](http://scholar.google.com/scholar_lookup?author=A.+Vall&author=M.+Dorfer&author=H.+Eghbal-Zadeh&author=M.+Schedl&author=K.+Burjorjee&author=G.+Widmer+&publication_year=2019&title=Feature-combination+hybrid+recommender+systems+for+automated+music+playlist+continuation&journal=User+Model+User+Adapted+Interact&volume=29&pages=527-72)

30. Lin Q, Niu Y, Zhu Y, Lu H, Mushonga KZ, Niu Z. Heterogeneous knowledge-based attentive neural networks for short-term music recommendations. *IEEE Access*. (2018) **6**:58990–9000. doi: 10.1109/ACCESS.2018.2874959

[CrossRef Full Text](https://doi.org/10.1109/ACCESS.2018.2874959) | [Google Scholar](http://scholar.google.com/scholar_lookup?author=Q.+Lin&author=Y.+Niu&author=Y.+Zhu&author=H.+Lu&author=KZ.+Mushonga&author=Z.+Niu+&publication_year=2018&title=Heterogeneous+knowledge-based+attentive+neural+networks+for+short-term+music+recommendations&journal=IEEE+Access&volume=6&pages=58990-9000)

31. Goodfellow I, Bengio Y, Courville A. *Deep Learning.* Cambridge, MA: MIT Press (2016).

[Google Scholar](http://scholar.google.com/scholar_lookup?author=I.+Goodfellow&author=Y.+Bengio&author=A.+Courville+&publication_year=2016&title=Deep+Learning.)

32. Aggarwal CC. *Neural Networks and Deep Learning: A Textbook*. Springer International Publishing (2018).

33. Yang X, Dong Y, Li J. Review of data features-based music emotion recognition methods. *Multimedia Syst*. (2018) **24**:365–89. doi: 10.1007/s00530-017-0559-4.

[CrossRef Full Text](https://doi.org/10.1007/s00530-017-0559-4.) | [Google Scholar](http://scholar.google.com/scholar_lookup?author=X.+Yang&author=Y.+Dong&author=J.+Li+&publication_year=2018&title=Review+of+data+features-based+music+emotion+recognition+methods&journal=Multimedia+Syst&volume=24&pages=365-89)