Music Recommendation

Modern Approaches and Problems

T. Shan, M. Finley, R. Jaswal

December 3rd, 2019

Background/Motivation

- Music streaming is most common medium for music listening currently
- Large part of music streaming is content recommendations
- Need for personalized recommendations that incentivize usage
- Research has led to "leaderboard-chasing" resulting in stunted knowledge development

Introduction/Significance

- **Goal**
  - Allows listeners to discover new music that matches their tastes
  - Enables online music stores to target their wares to the right audience
- However, complicated! So many factors...
  - Music content
  - User profile
  - Social and cultural factors...
- Recommend system needed also in other area, like movies, books, scientific articles...

Outline

Methods Overview

- Recommendation systems take the form of:
  - Neural Network architectures
  - K-nearest Neighbors (KNN), collaborative filtering
  - Graph based method

Outline

Summary of Literature

Titles	Authors	Type	Date
Deep Content-Based Music Recommendation	A. van den Oord et al.	Content-based	2013
Learning Content Similarity for Music Recommendation	B. McFee et al.	Content-based	2012
RecGAN: Recurrent Generative Adversarial Networks for Recommendation Systems	H. Bharadhwaj et al.	Neural Network	2018
Neural Collaborative Filtering	X. He et al.	Neural Network	2017
Improving Collaborative Metric Learning with Efficient Negative Sampling	V. Tran et al.	Negative Sampling	2019
Are We Really Making Much Progress? A Worrying Analysis of Recent Neural Recommendation Approaches	M. Ferrari Dacrema et al.	N/A	2019
The Neural Hype and Comparisons Against Weak Baselines	J. Lin	N/A	2019
A Comparison of Offline Evaluations, Online Evaluations, and User Studies in the Context of Research-Paper Recommender Systems	J. Beel et al.	N/A	2015
Investigating the Persuasion Potential of Recommender Systems from a Quality Perspective: An Empirical Study	P. Cremonesi et al	N/A	2012
Improved Neighborhood-based Collaborative Filtering	R. Bell et al.	Neighborhood-Based	2007

Figure: Table 1. Description of literature reviewed.

Collabrative Filtering

- Most prevalent method in the recommendation field
- A method of making automatic predictions (filtering)about the interests of a user by collecting preferences or taste information from many users (collaborating).
-
  1. Look for users who share the same rating patterns with the active user (the user whom the prediction is for).
  2. Use the ratings from those like-minded users found in step 1 to calculate a prediction for the active use

Collabrative Filtering

- **Memory-based** To find the rating R that a user U would give to an item I, the approach includes:
  1. Finding users similar to U who have rated the item I
  2. Calculating the rating R based the ratings of users found in the previous step
- **Model-based** Dimension Reduction

Deep content-based music recommendation

Aaron van den Oord et al, 2013

- **Goal**
  Bridge the semantic gap in music by training a CNN model to predict latent factors from music audio

Content Based Neural Networks

- **Cold start problem of collaborative filtering**
  Only applicable when usage data is available
  New items that have not been consumed before cannot be recommended.
- **Content-based recommendation**
  Predict user preferences from item content and metadata

Content Based Neural Networks

- **Dataset** Million Song Dataset (MSD)
- **Ground truth** Weighted Matrix Factorization (Hu et al, 2008)
  - Dataset contains play counts per song and per user

$$p_{ui} = I(r_{ui} > 0)$$
$$c_{ui} = 1 + \alpha log(1 + \epsilon^{-}1 r_{ui})$$

    where r_{ui} is the play count for user u and song i, and p_{ui} is the preference variable, c_{ui} is the confidence variable
  - WMF

$$min_{x,y} \sum c_{ui}(p_{ui} - x_u^T y_i)^2 + \lambda(\sum_u ||x_u||^2 + \sum_i ||y_i||^2)$$

    where x_u is the latent factor vector for user u and y_i is the latent factor for song i

Content Based Neural Networks

- **Bag-of-words representation**
  Traditionlly used as feature input to classifier or regressor
  - Extract MFCC
  - Vector quantize the MFCCs with K-means algorithm
  - Aggregate them into a bag-of-words representation: Count how many times each cluster was selected.
  - The result vector of counts is a bag-of-words feature representation of the song

  Linear regression, multilayer perceptron (1000 hidden units), metric learning to rank (MLR)
- **Convolutional Neural Network**
  - Log-scale Melspectrogram

Content Based Neural Networks

Objective function

- MSE
- weighted prediction error (WPE)

$$min_{\theta} \sum c_{ui}(p_{ui} - x_u^T y_i')^2$$

where θ is the model parameters, y_i' is the predicted latent vector by the model

Content Based Neural Networks

Experiments

- **Versatility of the latent factor representation**
  Tag classification task: (Logistic regression) Latent factor-based feature has higher AUC than pure audio-based classification
- **Latent factor prediction: quantitative evaluation**

Model	mAP	AUC
MLR	0.01801	0.60608
linear regression	0.02389	0.63518
MLP	0.02536	0.64611
CNN with MSE	0.05016	0.70987
CNN with WPE	0.04323	0.70101

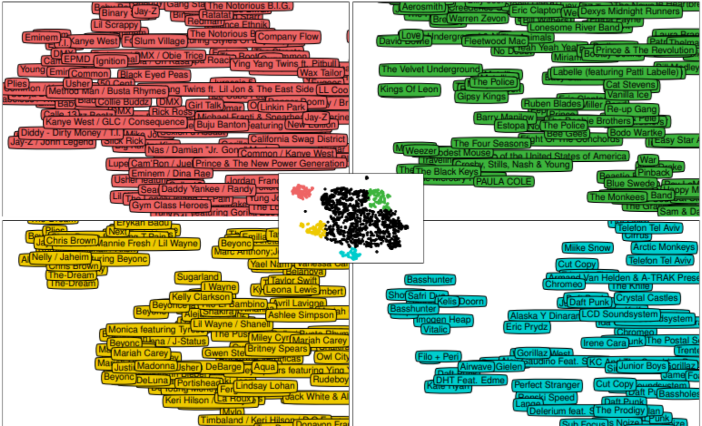
where mAP is mean average precision and AUC is the area under ROC curve

Content Based Neural Networks

Experiments

- **Latent factor prediction: qualitative evaluation**

  t-SNE visualization of the distribution of predicted usage patterns, using latent factors predicted from audio. We can discern hip-hop(red), rock (green), pop (yellow) and electronic music (blue).



Content Based Neural Networks

Conclusion

- Using CNN model to predict latent factor when there is no usage data
- A solution of cold start problem
- Recommend new and unpopular song

Are We Really Making Much Progress? A Worrying Analysis of Recent Neural Recommendation Approaches

Maurizio Ferrari Dacrema et al, 2019

- **Goal**

  Reproduce various recent deep learning approaches to top-n recommender systems and analyze their performances with tuned baseline and heuristic based recommender system.

Deep Learning

Issues

- Influx of deep learning techniques with regards to recommender systems.
- Published in top-level research conferences.
- Many not reproducible.
- Those reproducible beaten by tuned baselines.

Deep Learning

Factors

- Weak baselines or establishment of weak methods as new baselines.
- Difficulties in comparing or reproducing results across papers.
- Differences in dataset, evaluation protocols, performance measure and data preprocessing steps

Deep Learning

Baselines taken

- Non-personalized - TopPopular
- Collaborative based k-nearest neighbour methods - ItemKNN, UserKNN
- Content based k-nearest neighbour methods - ItemKNN-CBF, ItemKNN-CFCBF
- Graph based methods - P^3A, RP^3B

Deep learning methods

- included papers that appeared between 2015 and 2018 in the following four conference series: KDD, SIGIR, TheWebConf (WWW), and RecSys.
- Some of them which were reproducible - CMN, MCRec, CVAE, CDL

Deep Learning

Results

- Many of them were not reproducible
- Some were reproduced using all the parameters listed in the papers
- They were compared with tuned baselines using the databases mentioned in the papers.
- The deep learning methods were outperfomed by the baselines using evaluation metrics mentioned in the papers.

Results

	CiteULike-a			
	HR@5	NDCG@5	HR@10	NDCG@10
TopPopular	0.1803	0.1220	0.2783	0.1535
UserKNN	**0.8213**	**0.7033**	**0.8935**	**0.7268**
ItemKNN	**0.8116**	**0.6939**	0.8878	**0.7187**
$P^3\alpha$	**0.8202**	**0.7061**	0.8901	**0.7289**
$RP^3\beta$	**0.8226**	**0.7114**	**0.8941**	**0.7347**
CMN	0.8069	0.6666	0.8910	0.6942

	PREC@10	REC@10	NDCG@10
TopPopular	0.1907	0.1180	0.1361
UserKNN	0.2913	0.1802	0.2055
ItemKNN	**0.3327**	**0.2199**	**0.2603**
$P^3\alpha$	0.2137	0.1585	0.1838
$RP^3\beta$	0.2357	0.1684	0.1923
MCRec	0.3077	0.2061	0.2363

	REC@50	REC@100	REC@300
TopPopular	0.0044	0.0081	0.0258
UserKNN	0.0683	0.1016	0.1685
ItemKNN	**0.0788**	0.1153	0.1823
$P^3\alpha$	**0.0788**	0.1151	0.1784
$RP^3\beta$	**0.0811**	0.1184	0.1799
ItemKNN-CFCBF	**0.1837**	**0.2777**	**0.4486**
CVAE	0.0772	0.1548	0.3602

	Cutoff 20		Cutoff 60		Cutoff 100	
	REC	MAP	REC	MAP	REC	MAP
TopPopular	**0.1853**	**0.0576**	**0.3335**	**0.0659**	**0.4244**	**0.0696**
UserKNN CF	**0.2881**	**0.1106**	**0.4780**	**0.1238**	**0.5790**	**0.1290**
ItemKNN CF	**0.2819**	**0.1059**	**0.4712**	**0.1190**	**0.5737**	**0.1243**
$P^3\alpha$	**0.2853**	**0.1051**	**0.4808**	**0.1195**	**0.5760**	**0.1248**
$RP^3\beta$	**0.2910**	**0.1088**	**0.4882**	**0.1233**	**0.5884**	**0.1288**
SpectralCF	0.1843	0.0539	0.3274	0.0618	0.4254	0.0656

Heuristic Method

Random Walks in Recommender Systems: Exact Computation and Simulations

Colin Cooper et al.

- **Goal**

  Recommend movie to user based on ranking of nearest three-step random walk along graph of users and movies

Heuristic Method

- $G = (U \cup I, R)$
- G is graph, U is set of users, I is set of movies
- Edges of G show which movies the users have watched
- s-step random walk distribution (P^s) ranks movies user hasn't watched based on the probability distribution of the random walk at step s, if u was the starting vertex

Conclusion

- Music recommendation system will most likely expand and evolve to include recommendations based on more high level abstract musical concepts
- Future research interests include:
  - Improving deep learning methods to outperform tuned baselines consistently
  - Recommendation systems which recommend songs based on ever-changing metrics such as weather, mood.
- Social ethical issue derived from recommendation system

Questions

- Questions?

References I

Homanga Bharadhwaj, Homin Park, and Brian Y. Lim. "RecGAN: Recurrent Generative Adversarial Networks for Recommendation Systems". In: *Proceedings of the 12th ACM Conference on Recommender Systems*. RecSys '18. Vancouver, British Columbia, Canada: ACM, 2018, pp. 372–376. ISBN: 978-1-4503-5901-6. DOI: 10.1145/3240323.3240383. URL: http://doi.acm.org/10.1145/3240323.3240383.

Maurizio Ferrari Dacrema, Paolo Cremonesi, and Dietmar Jannach. "Are We Really Making Much Progress? A Worrying Analysis of Recent Neural Recommendation Approaches". In: *CoRR* abs/1907.06902 (2019). arXiv: 1907.06902. URL: http://arxiv.org/abs/1907.06902.

References II

Yifan Hu, Yehuda Koren, and Chris Volinsky. "Collaborative filtering for implicit feedback datasets". In: *2008 Eighth IEEE International Conference on Data Mining*. Ieee. 2008, pp. 263–272.

Aaron van den Oord, Sander Dieleman, and Benjamin Schrauwen. "Deep content-based music recommendation". In: *Advances in Neural Information Processing Systems 26*. Ed. by C. J. C. Burges et al. Curran Associates, Inc., 2013, pp. 2643–2651. URL: http://papers.nips.cc/paper/5004-deep-content-based-music-recommendation.pdf.

References III

Viet-Anh Tran et al. "Improving Collaborative Metric Learning with Efficient Negative Sampling". In: *Proceedings of the 42Nd International ACM SIGIR Conference on Research and Development in Information Retrieval*. SIGIR'19. Paris, France: ACM, 2019, pp. 1201–1204. ISBN: 978-1-4503-6172-9. DOI: 10.1145/3331184.3331337. URL: http://doi.acm.org/10.1145/3331184.3331337.