

楽曲宣伝のターゲット決定のためのユーザ推薦システム

高橋滉弥

法政大学情報科学部デジタルメディア学科

E-mail: koya.takahashi.5u@stu.hosei.ac.jp

Abstract

Although there are many services that have introduced music recommendation systems, there are cases where the recommended music does not resonate with the user. In this study, we propose a user recommendation system for songs that can limit the target audience for song recommendation. To predict user preferences, we use 1 million playlists, which contains multiple songs that users prefer, for a total of 60 million songs. Although the playlists can be used to express pattern of preferences in an easy-to-understand manner, the information is sparse due to the huge number of songs. In this study, we use a hybrid filtering method that combines music features and matrix factorization to create a matrix of evaluation values for recommendation. As a result of the recommendation, we were able to present multiple appropriate recommended users for a single song, and this allowed us to limit the target audience for song promotion to about 10% of the total.

1 まえがき

音楽ストーリーミングサービスとは、スマートフォンやPCなどを用いて、インターネット上にある音楽をすぐに再生できるサービスである。近年、音楽ストーリーミングサービスの普及に伴い、音楽はCDなどの媒体からweb上に移行してきており、web上で様々な音楽を視聴できる環境になってきている。このようなサービスは今では種類も多く、利用しているユーザも増えている。音楽ストーリーミングサービスでは、利用している中でアプリケーション側からユーザの好みに合いそうなおすすめ曲を推薦するものが多い。そこで、提示された楽曲に対して、その楽曲を好みそうなユーザを推薦するシステムを提案する。

web上で音楽を聴く際には、多くの場合は検索だけでなく、再生履歴やプレイリストなどに基づいてアプリケーション側から楽曲の推薦が行われる。楽曲推薦システムは多数存在しているが、推薦された楽曲がユーザの心に響かない場合がある。ユーザの好みを先に知ることで、楽曲の推薦や宣伝を効率よく行える可能性がある。

本研究では、楽曲推薦の際にターゲット層を限定することができない問題を解決するため、楽曲に対するユーザ推薦システムを提案する。ユーザ推薦を用いると、新しい曲を宣伝する際にもターゲット層を明確にできる。Spotifyでは、アーティスト向けのダッシュボードを確認すると、楽曲の保存率が平均して全リスナーの3.0~7.0%であり、10%を超えると高いとされている。本研究では、ターゲット層のユーザを10%出力できるような、高性能な推薦システムの作成を目標とする。

2 従来研究

従来の音楽推薦技術では、嗜好推定のために「協調フィルタリング」と「内容ベースフィルタリング」という2つの提案が

されているが、それぞれ相補的な性質を持っている。

協調フィルタリングは、推薦を行う際に他のユーザの楽曲評価を参考にして行う手法である[1]。例として、Amazon.co.jpやiTunes Music Storeなどではこの手法を用いた推薦方法で大きな成果を上げている。しかし、この手法では評価のなされていない楽曲に対して嗜好の推定ができず、推薦することができないという問題がある。これをコールドスタート問題と呼ぶ。これまでコールドスタート問題を解決するような推薦システムの提案はされてきているが、どれも定量的な評価はされていない[2-4]。一方、内容ベースフィルタリングは、音色やリズムといった音楽的特徴に基づいて推薦を行う手法である。この手法では、楽曲の特徴量を抽出することで、楽曲同士の類似度を計算し、類似した楽曲を推薦するため、評価されていないアイテムも推薦可能である。帆足ら[5]は、ベクトル類似度を用いた楽曲検索システムを提案しており、750曲のデータを用いて50%程度の精度を出している。しかし、楽曲の特徴量のみで依存した偏った推薦がされてしまう、楽曲の持つ背景や市場での人気を考慮できないといった問題が存在する。

これまで述べた問題を解決するため、他のユーザの楽曲評価と音楽的特徴量を同時に考慮し推薦を行うハイブリッド型推薦手法を提案する。これにより、評価がされていない楽曲も含め、精度の高い嗜好推定を行えると考える。このような従来の楽曲推薦システムを参考にしながら、楽曲宣伝のためのユーザ推薦システムを完成させる。吉井ら[6]は、ユーザによる評価スコアと音響特徴を結合させたハイブリッド型楽曲推薦手法を提案している。この研究ではハイブリッド型の推薦手法を用いて楽曲のランク付けを行うことにより、5年間分の日本のシングルCDの週間ランキング上位20曲、合計で2500曲を学習データとして用いて、約90%という高い推薦精度を出すことに成功している。しかし、ジャンルや雰囲気が多様な現実的な大規模データベースを使用していないため、この研究で用いられている潜在変数の特性が既存のジャンルと類似しているかどうかを検証することができていないという問題点がある。

そこで、本研究では多種多様な楽曲が含まれている100万のプレイリストを用いて推薦システムを作成することで、従来研究の問題点を解決する。本研究では、内容ベースフィルタリングの役割を担うSpotify Web APIにより求められる特徴量と、協調フィルタリングの役割を担う行列因子分解を組み合わせたハイブリッド型推薦を用いる。

プレイリストの性質について説明する。プレイリストには、ユーザが好む楽曲が複数含まれているため、プレイリスト全体を通して共通点が存在すると考える。プレイリストの自動生成とその評価について行われている研究[7]では、プレイリスト生成のための戦略や必要な情報についてまとめられている。この研究により、プレイリストは何かしらの特徴を持ち、楽曲の出現パターンやプレイリスト内のメタデータなどを含む特徴の類似性が推薦システムに活用できるという利点を持っていることがわかっている。しかし、プレイリストが持つ欠点として、ユーザはプレイリスト内の楽曲を気に入っていると直感的に考えられるため、これらの情報からはユーザのマイナスの評価を確認することができない。本研究では、楽曲宣伝のためのター

ゲット層の限定化が目的であるため、ユーザのマイナスの評価は使用せず、好みの楽曲の情報のみを扱う。よって、プレイリストから得られるユーザの好みの特性を膨大なデータ数から知ることによって推薦のターゲット層を絞り込む。

3 準備

3.1 特異値分解

特異値分解とは行列因子分解の手法の一つであり、2つのユニタリ行列と1つの対角行列の内積で表現できる。

$m \times n$ の行列 C は、特異値分解により以下の形に分解できる。

$$C = U\Sigma V^T \quad (1)$$

ここで、行列 U および V はユニタリ行列、行列 U および V の列ベクトルをそれぞれ左特異ベクトル、右特異ベクトルと呼ぶ。 Σ は対角行列成分が $\sigma_i (i = 1, 2, \dots, k, \sigma_i \geq \sigma_j (i < j))$ となる対角行列であり、この対角成分を特異値と呼ぶ。特異値は0より大きい値である。特異値 σ は、対応する軸の情報の重要度を表す。 σ が小さい箇所の次元まで軸を減らすことで、元の行列を低次元で再現できる。これにより、プレイリストの楽曲の出現パターンを100パターン程度に限定する。

3.2 主成分分析

主成分分析とは、元の情報の損失が最小限となるように新たな指標に要約する方法である。これにより、楽曲の特徴量の次元数を落とし、少ない成分で楽曲の特徴を表現する。

主成分の計算方法について説明する。 m 個のデータ $x_i (i = 1, 2, \dots, m)$ があるとき、データの主成分 z_j は式 (2) のようになる。また、主成分の分散 $V(z_j)$ は式 (3) のようになる。

$$z_j = a_{1j}x_1 + a_{2j}x_2 + \dots + a_{mj}x_m \quad (2)$$

$$V(z_j) = \frac{1}{n} \sum_{i=1}^n (z_{ji} - \bar{z}_j)^2 = \sum_{j=1}^m \sum_{k=1}^m s_{jk} a_j a_k \quad (3)$$

ここで、 a_j は主成分の係数 ($j = 1, 2, \dots, m$)、 s_{jk} は分散・共分散である。

分散 $V(z_j)$ が最大となるように係数 a_j を求めるのだが、これは式 (4) の固有値問題を解くことに等しい。

$$V\mathbf{a} = \lambda\mathbf{a}; V = \begin{bmatrix} s_{11} & s_{12} & \cdots & s_{1m} \\ s_{11} & s_{12} & & s_{1m} \\ \vdots & & \ddots & \vdots \\ s_{m1} & s_{m2} & \cdots & s_{mm} \end{bmatrix}, \mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix} \quad (4)$$

ここで、 λ は行列 V の固有値、 \mathbf{a} は固有値 λ に対応する固有ベクトルである。固有値 λ の中で最大のものに対応する固有ベクトルを用いた主成分が第1主成分となる。式 (2) に実際のデータを代入して値を計算したものを主成分得点という。この値から、各主成分における傾向や関係がわかる。

4 提案手法

膨大な数のプレイリストは極端なスパース性を持ち、これにより全体で1,2回しか登場しない楽曲は、他のデータとの関連性のない、データの大きさを増やすだけのノイズとしてシステムの性能を下げってしまう。本研究ではプレイリストの持つデータの大きさという利点を活かしつつ、特異値分解やクラスタリングを行うことによりデータの極端なスパース性を補うことで、膨大な数のプレイリストを用いた推薦システムを実現させる。多くのプレイリストを用いて推薦システムを作成することにより、大量のデータを用いて推薦精度を上げられるだけでなく、楽曲同士の類似性やプレイリストの特性を考慮した推薦システムにできる。

提案するシステムの概要について説明する。まず最初に、100万のプレイリストが含まれる訓練データを用意し、そのプレイリスト内の各楽曲の特徴量を求める。求められた特徴量から、膨大な数の楽曲はクラスタリングにより複数のカテゴリに分類される。また、プレイリストに入っている楽曲情報からスパース行列を作成し、特異値分解によりプレイリストの行列は次元削減が行われ、低次元の行列で表現される。完成したプレイリストと楽曲の2つの低次元行列の積によって評価値を予測し、その数値を基にユーザの嗜好を予測する。最後に、推定された嗜好を基に推薦されるべきユーザを選択し、出力する。推薦されるべきユーザは1人に限定されることはないため、該当する複数のユーザを出力する。

プレイリストを用いたユーザの嗜好をまとめた行列は [8], [9] のように定義する。 m 個のプレイリストが与えられ、それぞれが n 個の曲の一部を含んでいるとする。プレイリスト i に曲 j が含まれていれば $C_{ij} = 1$ 、そうでなければ0とする行列 $C \in \{0,1\}^{m \times n}$ を作成する。これは、行列の要素の99.9%は0で構成されるスパース行列となる。プレイリストを用いて集められたユーザの嗜好情報は、視聴履歴を用いた場合に比べて遙かにスパース性が高く、推薦の精度が低い。そこで、行列因子分解やクラスタリングにより、スパース行列 C をプレイリストと楽曲の2つの行列 A, B に分解する。これにより、プレイリストに含まれる楽曲の情報から、他楽曲に対する評価値の予測を行う。どのプレイリストにも含まれていない新曲が入力された場合は、最初はユーザの評価がなく評価値が全て0となってしまい行列分解によって正しい評価値予測ができないが、クラスタリングを行った後の楽曲についての行列内で学習を行うことで、他楽曲の中から類似している楽曲を探す。

4.1 行列の次元削減

本研究では、膨大な数のプレイリスト、楽曲を扱う。そのため、プレイリストや楽曲同士の類似度を学習するにあたり、関係があまりないデータが多く存在すると、学習に時間がかかるだけでなくそのデータが他のデータとの関係性がないノイズとなってしまう。すると、関係ない情報として干渉してしまい、システムの性能を下げる原因となってしまう。そこで次元削減をすることにより、とても大きなサイズの行列も、低次元の主要な特徴量を持ったデータとして扱えるようにする。

次元削減の手法として、特異値分解 (SVD) や主成分分析 (PCA) などが存在する。本研究では、膨大なサイズのスパース行列を特異値分解することによりプレイリストの低次元行列を求める。また、データセットから入手した楽曲をクラスタリングする際にも、楽曲の特徴量に対して主成分分析を行う。

4.2 訓練データ

本研究では、ユーザがよく聴く複数の楽曲からユーザの特徴量の決定ができるという観点から、プレイリストを採用する。プレイリストの訓練データとして、「Spotify Million Playlist Dataset」を用いる。これは、音楽推薦の研究のために用意されたデータセットであり、実際に入力プレイリストに自然に続くような関連性の高い楽曲のリストを返す研究などに用いられている。

このデータセットは、100万のプレイリストと全体で延べ6000万曲の情報で構成されている大きなデータである。これらのプレイリストは、音楽ストリーミングサービス「Spotify」内での公開プレイリストからサンプリングされたものである。このデータセットはトラック、アルバム、アーティストの名前やURIで構成されている。

4.3 楽曲特徴量

本研究では、推薦システムのために、テンポや調といった耳でわかる特徴量の他に、楽曲を専門用語などを用いずに表現し比較するために楽曲の雰囲気や数値化した特徴量を用いる。そこで、Spotify が提供している膨大な数の楽曲を扱うにあたり、データセットに含まれるプレイリストの特徴量は

Spotify の Web API を用いて取得する。これは、Spotify Data Catalogue からトラック、アルバム、アーティスト、およびトラックに関する JSON メタデータを返す Web API である。また、Spotify Web API では、楽曲のいくつかの属性を機械的に数値化した解析データが存在する。これらを取得することにより、楽曲の特徴量をまとめたデータベースを構築する。

本研究で扱う特徴量は、表 1 の 11 次元の特徴量である。主成分分析により 4 次元まで特徴量を落とし、4 次元の特徴量によって楽曲のクラスタリングを行う。

表 1. 取得される楽曲特徴量

acousticness	楽曲のアコースティック感
danceability	リズムの安定性、ビートの強さに基づく値
energy	強度と活動性の知覚的な尺度
instrumentalness	楽曲にボーカルが含まれていないかの予測
liveness	レコーディング中の観客の存在を検出
loudness	楽曲の全体的な物理的な強さ (db)
speechiness	楽曲に含まれる話し言葉の存在を検出
mode	楽曲の調
valence	音楽的なポジティブ度
tempo	楽曲の BPM
time signature	楽曲の拍子

5 実験

5.1 行列 C の作成方法

行列 C は、プレイリスト i に曲 j が含まれていれば $C_{ij} = 1$ 、そうでなければ 0 とする行列 $C \in \{0,1\}^{m \times n}$ である。データセットを参照して全曲の ID に番号を振った辞書を作成し、そこから辞書に存在する楽曲の行列番号を列挙することでスパース行列を作成していく。

データセットのプレイリストは全部で 100 万個、楽曲数は重複しているものを除くと全部で 2262292 曲である。よって、この手順により作成される行列 C は 1000000×2262292 のスパース行列となる。

5.2 行列 A,B の作成方法

行列 A,B は、プレイリスト間の距離や曲の類似度を表す低次元の行列である。行列 A,B はそれぞれ、特異値分解やクラスタリングによって次元削減が行われることで完成する。 1000000×2262292 のスパース行列 C を特異値分解し、特異値が大きい順番にプロットすると図 1 のようになる。特異値は $k=100$ 付近までは急激に減少し、 $k=100$ 以降は緩やかに減少し続けるような値を取っていることがわかる。これより、 $k=100$ で特異値分解を行い生成された 1000000×100 の行列 A、 100×2262292 の行列 B の積で、元のスパース行列 C に近い行列を表現できるといえる。

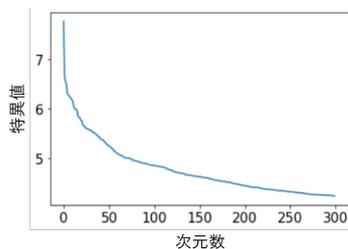


図 1. 特異値の推移

これにより、列数が k に落とされた行列 A は、 k 個のパターンを持ったプレイリストの特性を表していると言える。

5.2.1 楽曲のクラスタリング

3.3 章で述べた、Spotify により提供されている特徴量を用いて楽曲のクラスタリングは行われる。まず、11 次元の特徴量を主成分分析により 4 次元まで落とす。特徴量を正規化し、累積寄与率を確認すると図 3 のようになる。第 1 主成分の寄与率は 40% であり、第 4 主成分までで元の特徴量の 80% 以上

を表すことができている。この 4 次元の特徴量を用いて、1 回しか登場しない楽曲を除いたおよそ 100 万曲をクラスタリングし、100 個のクラスタに分類した。データセット内の全ての楽曲は、100 パターンのいずれかのカテゴリに分類される。

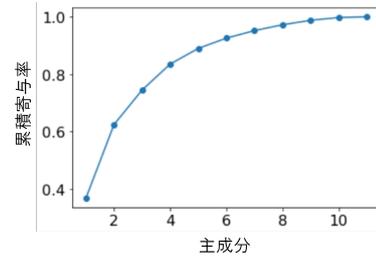


図 2. 各主成分の累積寄与率

6 評価

提案した手法の有効性を確認するため、全く評価がされていない新曲を用いてユーザ推薦を行うコールドスタート問題と、わずかな評価値のみが用意されている準コールドスタート問題のような状況を用意し、評価実験を行った。

コールドスタート問題に関しては、訓練データとして用意したプレイリストから全体の中で 13205 回登場する楽曲を抜き出し、それらの楽曲を入力としてシステムを動かす。出力は、ターゲット層を 10% に限定する目標のもとで、吉井らの研究 [6] の実験で行われている人数の選び方と同じように、評価値が高いユーザをそれぞれ上位 1000 人、10000 人、100000 人に分けて確認し、入力楽曲が元々含まれていたプレイリストを保有するユーザが出力結果として表示されるかどうか確認した。

次に、入力楽曲を複数選択し、欠損値が 100% で協調フィルタリングにより解決が難しいコールドスタート問題の状態と、評価値の欠損値を作成する準コールドスタート問題に関して、入力楽曲が元々含まれていたプレイリストを保有するユーザが出力結果として表示されるかどうか確認した。欠損値を作る楽曲はランダムに 10 曲選択し、その楽曲の列の要素数によって推薦精度がどのように変化するか確認した。

7 実行結果

データセット内で 13205 回以上出現している楽曲を入力楽曲とした時、出力ユーザに含まれる正解ユーザの割合を適合率 (precision)、正解ユーザのうち出力できた正解ユーザの割合を再現率 (recall) とすると、それぞれ表 2、表 3 のようになった。

表 2. precision

Rankings	Our method(Frequent songs)	Our method	Random
1000	2.9%	2.7%	0.9%
10000	1.8%	1.7%	1.1%
100000	1.5%	1.4%	1.2%

表 3. recall

Rankings	Our method(Frequent songs)	Our method	Random
1000	0.2%	0.2%	0.1%
10000	1.4%	1.2%	1.0%
100000	11.2%	10.6%	9.2%

Spotify では、楽曲の保存率が平均して全リスナーの 3.0~7.0% であるのに対して、提案手法ではランダム約 2 倍の精度が出ているものの、1.5~3.0% と低い正答率である。実験により出力されたユーザは該当楽曲を好む、すなわち楽曲を保存すると予測されたユーザであり、平均的な楽曲の保存率を下回っていることから、現状のシステムは良いものとは言えない。これは、データの極端なスパース性や楽曲の特徴量の問題があると考えられる。

次に、欠損値の割合と入力楽曲を変化させた時の評価値の予想精度は図 3 のようになった。

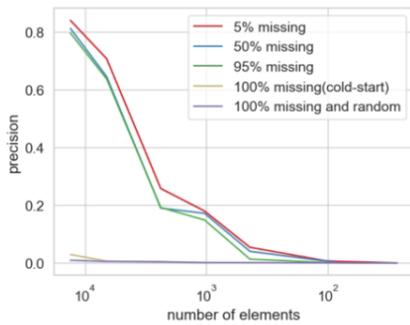


図 3. 予測精度の変化

50% の欠損値を持つ楽曲と 95% の欠損値を持つ楽曲を比較すると、欠損値が少ないほどわずかに高い精度を出す、大きな違いは見られない。しかし、どちらも約 1 万個の要素を持つ楽曲を入力とした時、80% に近い精度でユーザの推薦を行えている。また、楽曲の要素数が多いほど、すなわち列から得られる情報が多いほど精度は高くなっている。

この実験から、全く評価されていない新規アイテムに対しても先に少しの評価値を与えることが、コールドスタート問題の解決へのアプローチとなると言える。しかし、要素数がとても少ない楽曲に関しては、50% の評価値がわかっているような楽曲でも正しいユーザを返すことができていないことから、スパース性が極端に高い行列を用いて評価値を予測することは、新規アイテムを用いる場合に限らず困難である。

8 考察

8.1 特徴量の有効性

システムの性能が良くならない原因の一つとして、扱う特徴量の曖昧さが考えられる。クラスタリングを行った後で同クラスタ内の類似楽曲を実際に聴いてみると、全く類似していないケースが多く見られた。そこで、精度の向上のために正しくクラスタリングが行われるための別な特徴量の提案が必要である。例として、本研究で特徴量として用いなかったメタデータのような指標も活用できるようなシステムにすることが必要だと考えられる。その特徴量を有効な内容ベースのフィルタリングとして協調フィルタリングと組み合わせ、新規データの持ちうる情報から類似データを探し出すことで、コールドスタート問題の解決にも繋がると考えている。

8.2 システムの性能確認

評価実験では、入力楽曲に対して特徴量を用いて類似している楽曲を探し出すことで評価値の予測をしていた。ここでは、類似している評価列を用いることによってユーザ推薦ができているかどうかを確認するため、楽曲の特徴量を考えず、スパース行列の列情報だけを見て要素の値が類似している楽曲を探し出し、同じ手法で正解ユーザの適合率と再現率を確認した。

表 4. precision

Rankings	Similar Row	Random
1000	3.1%	0.9%
10000	1.8%	1.1%
100000	1.5%	1.2%

表 5. recall

Rankings	Similar Row	Random
1000	0.2%	0.1%
10000	1.4%	1.0%
100000	11.0%	9.2%

結果は表 4,5 のようになり、これもランダムに推薦した場合の約 2 倍の精度は出ているものの、目標の 10% と比べると低い精度である。また、クラスタリングにより類似楽曲を探し出した場合 (表 2,3) と比べてもほとんど変わらない結果となった。これより、類似した評価値を新規アイテムの評価値として代わりに用いても、目標の精度まで上げることが困難であることがわかった。

9 あとがき

本研究で提案した手法によってユーザの限定化はできたが、入力楽曲に対するターゲット層を目標としていた全体の 10% 提示することができなかった。また、評価実験により性能を確認すると、ランダムにユーザ推薦を行った場合の 2,3 倍の精度は出ているが、行列の評価値を用いた従来研究 [8,9] に比べて良い結果を残すことができなかった。協調フィルタリングにより実験データに存在しないアイテムを推薦するようなコールドスタート問題解決のための推薦システムを提案している研究は多く存在するが、いずれも高い精度でアイテムを推薦することができていない。本研究により、従来研究の 100 倍以上の大きさを持つ、スパース性が極めて高いプレイリスト情報を用いて新規アイテムの評価値を予測することが困難であることがわかった。

今後は、プレイリストの性質をより一層活用できるような手法を提案することにより、精度の高い推薦システムにできるように努めたい。また、本研究で用いた特徴量とは違う、有効性のある特徴量の提案をすることも、推薦システムの性能向上に繋がると考えている。

参考文献

- [1] J. Breese, D. Heckerman, and C. Kadie, "Empirical analysis of predictive algorithms for collaborative filtering", in Proc. 14th Conf. Uncertainty Artif. Intell. (UAI), 1998, pp. 43-52.
- [2] R.Zhou, S.Khemmarat, L.Gao, "The impact of YouTube recommendation system on video views", Proceedings of the 10th ACM SIGCOMM Conference on Internet Measurement, ACM, 2010, pp.404-410.
- [3] D.-C. Nie, Z.-K. Zhang, J.-L. Zhou, Y. Fu, K. Zhang, "Information filtering on coupled social networks", PLoS One 9, 2014, e101675.
- [4] J.-H. Liu, T. Zhou, Z.-K. Zhang, Z. Yang, C. Liu, W.-M. Li, "Promoting cold-start items in recommender systems", PLoS One 9, 2014, e113457.
- [5] K. Hoashi, K. Matsumoto, and N. Inoue, "Personalization of user profiles for content-based music retrieval based on relevance feedback", Proc. ACM Multimedia, pp. 110-119, 2003.
- [6] Kazuyoshi Yoshii, Masataka Goto, Kazunori Komatani, Tetsuya Ogata and Hiroshi G. Okuno, "An Efficient Hybrid Music Recommender System Using an Incrementally Trainable Probabilistic Generative Model", IEEE Trans. on Audio, Speech, and Language Processing, Vol.16, No.2, pp.435-447.
- [7] Geoffroy Bonnin and Dietmar Jannach, "Automated Generation of Music Playlists: Survey and Experiments", ACM Computing Surveys, Vol.147, Issue.2, No.26, pp.1-35.
- [8] Kirell Benzi, Vassilis Kalofolias, Xavier Bresson and Pierre Vandergheynst, "Song recommendation with non-negative matrix factorization and graph total variation", IEEE International Conference on Acoustics, Speech and Signal Processing, 2016, pp.2439-2443.
- [9] S.Rendle, C.Freudenthaler, Z.Gantner, and L.Schmidt-Thieme, "BPR: Bayesian Personalized Ranking from Implicit Feedback", Proceedings of Conference on Uncertainty in Artificial Intelligence, 2009, pp.452-461.